



HAL
open science

Etude sur la compétition intra-spécifique entre souches bactériennes dans le contexte de la vaccination de masse : approches théorique et statistique

Florian Lecorvaisier

► To cite this version:

Florian Lecorvaisier. Etude sur la compétition intra-spécifique entre souches bactériennes dans le contexte de la vaccination de masse : approches théorique et statistique. *Ecologie, Environnement*. Université Claude Bernard - Lyon I, 2024. Français. NNT : 2024LYO10285 . tel-04918453

HAL Id: tel-04918453

<https://theses.hal.science/tel-04918453v1>

Submitted on 29 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE DE DOCTORAT DE L'UNIVERSITÉ CLAUDE BERNARD LYON 1

École Doctorale N°341
Évolution, Écosystèmes, Microbiologie et Modélisation

Discipline : Éco-épidémiologie

Soutenue publiquement le 06/12/2024 par
Florian Lecorvaisier

ÉTUDE SUR LA COMPÉTITION INTRA-SPÉCIFIQUE ENTRE SOUCHES BACTÉRIENNES DANS LE CONTEXTE DE LA VACCINATION DE MASSE : APPROCHES THÉORIQUES ET STATISTIQUES

Devant le jury composé de :

Gandon, Sylvain
Blanquart, François
Soubeyrand, Samuel
Lopes, Christelle
Pontier, Dominique
Fouchet, David

Directeur de Recherche CNRS
Chargé de Recherche CNRS
Directeur de Recherche INRAE
Maîtresse de Conférences UCBL
Professeure des Universités UCBL
Maître de Conférences UCBL

Rapporteur
Rapporteur
Président du jury
Examinatrice
Directrice de thèse
Codirecteur de thèse



Cette thèse a été rédigée au sein du Laboratoire de Biométrie et Biologie Évolutive

UMR CNRS 5558 – LBBE

Université Claude Bernard Lyon 1 – Bâtiment Gregor Mendel

43 boulevard du 11 novembre 1918

69622 Villeurbanne cedex

Ce qui est simple est toujours faux. Ce qui ne l'est pas est inutilisable.

Paul Valéry, *Mauvaises pensées et autres*, 1942

RÉSUMÉ EN FRANÇAIS

Aujourd'hui, les activités humaines comme l'exploitation de l'environnement (déforestation, chasse, agriculture...) et la mondialisation des échanges conduisent à l'émergence continuelle de nouvelles maladies dont la récente pandémie de Covid-19 n'est finalement qu'un exemple parmi une multitude d'autres. Au même moment, l'humanité développe de nouvelles méthodes de lutte contre les agents infectieux, comme les antibiotiques et les vaccins, qui conduisent à l'émergence de nouveaux variants de pathogènes séculaires. La vaccination a par exemple conduit à la sélection de nouvelles souches de *Corynebacterium diphtheriae* et de *Bordetella pertussis*, les agents étiologiques de la diphtérie et de la coqueluche, respectivement.

Dans mes travaux de thèse, je me suis dans un premier temps intéressé à la manière dont l'utilisation de vaccins à anatoxine contre *C. diphtheriae* pouvait conditionner l'évolution de cette bactérie pathogène. Pour ce faire, j'ai développé un modèle dynamique de type SIR. Ce modèle m'a permis de mettre en évidence que ce type de vaccin favorise la sélection de souches ne produisant pas la toxine diphtérique, principal facteur de virulence de *C. diphtheriae*. Je montre également que la compétition entre souches toxinogènes et souches non toxinogènes facilite l'éradication des souches toxinogènes, donc virulentes. Ainsi, la vaccination, combinée à la compétition entre souches, permet de réduire la circulation des souches virulentes de *C. diphtheriae* et de diminuer la prévalence de la diphtérie.

Dans un deuxième temps, je me suis intéressé à identifier les facteurs environnementaux ayant un impact sur la vitesse d'émergence de nouvelles souches de pathogènes dans un contexte de vaccination de masse. Faute de données sur la diphtérie, j'ai étudié un jeu de données de souches de *B. pertussis* aux États-Unis. Plus précisément, je me suis intéressé à la vitesse de remplacement de souches sauvages de la bactérie par des souches mutantes déficientes en pertactine (PRN-), un antigène de surface présent dans une grande partie des vaccins en circulation. J'ai

dû développer une nouvelle méthode pour analyser ces données, les approches traditionnelles se montrant inadaptées pour étudier un phénomène d'invasion par définition hautement autocorrélé dans le temps. Les résultats de cette analyse semblent indiquer que la température moyenne de l'environnement dans lequel ces souches circulent pourrait avoir une incidence sur la vitesse d'invasion des souches PRN-. Cependant, des analyses complémentaires ont montré que le jeu de données était finalement limité pour mettre en évidence un effet d'un facteur environnemental sur cette invasion.

Dans un troisième temps, j'ai développé un modèle informatique et mathématique pour simuler la circulation de deux souches d'un pathogène type *B. pertussis* au sein d'une population mimant celle des États-Unis. Le but était *in fine* de produire des jeux de données artificiels similaires à celui étudié dans la partie précédente, pour vérifier comment pouvait fluctuer la qualité de l'échantillonnage et des estimations des cofacteurs telles que calculées dans le modèle précédent. Les résultats préliminaires de cette partie montrent que la dynamique de propagation d'un pathogène est un phénomène complexe, la rendant particulièrement difficile à étudier de manière rigoureuse. L'estimation des cofacteurs est biaisée par la qualité et la quantité des échantillons, et leur effet est alors difficile à mettre en évidence. Des travaux supplémentaires sur ces aspects sont nécessaires.

En conclusion, au cours de cette thèse, j'ai employé diverses approches mathématiques et informatiques pour étudier la compétition entre souches de pathogènes dans un contexte vaccinal. J'ai pu mettre en évidence un rôle de la vaccination dans l'émergence de souches non virulentes, développé une méthode statistique pour étudier la vitesse d'émergence d'une telle souche, et étudié les forces et les limites d'une telle approche.

Mots-clés : relation hôte-pathogène, émergence de souches, compétition intra-spécifique, épidémiologie, statistiques, modélisation

RÉSUMÉ EN ANGLAIS

Nowadays, human activities such as the exploitation of the environment (deforestation, hunting, agriculture...) and the globalization of trade lead to the continual emergence of new diseases, of which the recent Covid-19 pandemic is just one example. At the same time, humanity is developing new methods of combating infectious agents, such as antibiotics and vaccines, leading to the emergence of new variants of age-old pathogens. Vaccination, for example, has led to the selection of new strains of *Corynebacterium diphtheriae* and *Bordetella pertussis*, the etiological agents of diphtheria and pertussis, respectively.

In my thesis work, I first investigated how the use of toxoid vaccines against *C. diphtheriae* could condition the evolution of this pathogenic bacterium. To do this, I developed a dynamic SIR-like model. This model enabled me to demonstrate that this type of vaccine favors the selection of strains that do not produce the diphtheria toxin, the main virulence factor of *C. diphtheriae*. I have also shown that competition between toxigenic and non-toxigenic strains facilitates the eradication of toxigenic, and therefore virulent, strains. Thus, vaccination, combined with competition between strains, reduces the circulation of virulent strains of *C. diphtheriae* and lowers the prevalence of diphtheria.

Secondly, I set out to identify the environmental factors affecting the rate of emergence of new pathogen strains in the context of mass vaccination. In the absence of data on diphtheria, I studied a dataset of *B. pertussis* strains in the United States. More specifically, I was interested in the rate of replacement of wild strains of the bacterium by mutant strains deficient in pertactin (PRN⁻), a surface antigen present in a large proportion of vaccines in circulation. I had to develop a new method for analyzing these data, as traditional approaches proved unsuitable for studying an invasion phenomenon that is, by definition, highly autocorrelated over time. The results of this analysis suggest that the average temperature of the environment in which these strains

circulate could have an impact on the invasion rate of PRN- strains. However, further analysis showed that the dataset was ultimately limited to highlight an effect of an environmental factor on this invasion.

Thirdly, I developed a mathematical and computer model to simulate the circulation of two strains of a *B. pertussis*-like pathogen within a population mimicking that of the USA. The ultimate aim was to produce artificial datasets similar to the one studied in the previous section, in order to check how the quality of sampling and cofactor estimates calculated in the previous model might fluctuate. The preliminary results of this section show that pathogen propagation dynamics is a complex phenomenon, making it particularly difficult to study rigorously. The estimation of cofactors is biased by the quality and quantity of the samples, and their effect is therefore difficult to demonstrate. More work is needed on these aspects.

In conclusion, in the course of this thesis I have employed various mathematical and computational approaches to study competition between pathogen strains in a context of mass vaccination. I was able to highlight the role of vaccination in the emergence of non-virulent strains, develop a statistical method to study the rate of emergence of such a strain, and study the strengths and limitations of such an approach.

Mots-clés (*keywords*) : host-pathogen relationship, strains emergence, intra-specific competition, epidemiology, statistics, modelling

REMERCIEMENTS

Si le résultat d'une thèse est la production d'un document signé d'un nom unique, il convient de reconnaître que son accomplissement ne peut se faire dans l'individualité la plus totale. Ainsi, je n'aurais jamais réussi à mener à bien ce travail sans l'aide, le soutien et l'amitié d'un certain nombre de personnes auxquelles je voudrais adresser ici quelques mots de remerciement.

Tout d'abord, je tiens à remercier ma directrice de thèse, Dominique Pontier, et mon codirecteur, David Fouchet. Ils m'ont accordé leur confiance et ont accepté de m'encadrer sur un sujet de thèse pourtant éloigné de ma formation académique. Je leur suis reconnaissant pour leur aide et leurs conseils au cours de ces trois années de thèse qui m'ont permis de produire un travail au maximum de mes capacités. Je les remercie également de m'avoir accordé une grande liberté dans les directions prises dans ma thèse, tout en me réorientant lorsque je courais droit vers un mur – ce qui est arrivé plusieurs fois. Je voudrais également remercier Frank Sauvage qui, bien que n'étant pas officiellement membre de ma direction de thèse, s'est fortement impliqué dans ma thèse pour mes deuxième et troisième axes, et dont l'aide a été précieuse.

Je tiens également à remercier Benoît Soubeyrand et Nicole Guiso, qui m'ont apporté leur expertise tout au long de ma thèse. Benoît, vaccinologue, a été à l'origine de mon sujet de thèse à travers la rédaction d'un mémoire sur les vaccins imparfaits. Il m'a beaucoup aidé pendant mon stage et pour les deux premiers axes de ma thèse. Nicole, experte internationale de la coqueluche, nous a rejoint pour le deuxième axe de ma thèse en nous fournissant ses connaissances sur cette maladie et son pathogène, apportant un éclairage considérable sur une pathologie dont j'ignorais presque tout.

Je voudrais également adresser mes remerciements aux membres de mon comité de suivi qui, à deux reprises, ont accepté de se réunir pour m'écouter parler de mon sujet de thèse et me donner leurs conseils. Merci donc à Yannis Michalakis, Matthieu Domenech de Cellès, Thibaud

Porphyre pour leur temps, leur savoir et leurs conseils. Merci également à Bernard Kaufmann, qui a accepté d'être mon tuteur à l'école doctorale alors que mon sujet de thèse est foncièrement éloigné de ses thématiques de recherche, et qui a animé les réunions du comité de suivi avec une totale bienveillance.

Je tiens de même à exprimer mes remerciements aux membres du jury qui, sans me connaître, ont accepté de m'accorder leur temps pour évaluer mon travail de thèse. Merci à Sylvain Gandon, François Blanquart, Samuel Soubeyrand et Christelle Lopes. Je remercie plus particulièrement encore Sylvain et François qui ont accepté le rôle chronophage d'être rapporteurs de mon manuscrit. Les travaux de Sylvain ont été à l'origine de mon sujet de thèse, et je le remercie d'autant plus d'avoir accepté d'être rapporteur. Je voudrais également adresser mes remerciements à Florence Débarre et Samuel Alizon qui, n'ayant pas pu faire partie de mon jury, m'ont néanmoins témoigné leur intérêt pour mon travail et m'ont aidé à trouver d'autres membres pour mon jury.

Encore merci à toutes ces personnes pour m'avoir accordé confiance, temps et conseils et avoir ainsi rendu cette thèse possible. Pendant mes trois années de thèse, j'ai également eu l'occasion de rencontrer une multitude de personnes qui m'ont donné chaque jour envie de venir au laboratoire et de donner le meilleur de moi-même au service de mon sujet de recherche. Ces personnes sont trop nombreuses pour toutes être nommées et remerciées personnellement, alors je voudrais commencer par exprimer un grand merci général à tous les membres du LBBE. Pour leur aide dans le bon déroulé de ma thèse, je voudrais plus particulièrement remercier les membres du pôle informatique et du pôle administratif du laboratoire. Enfin, pour les moments conviviaux, les repas du midi, les activités hors du laboratoire, les encouragements, il y a quelques personnes que je voudrais remercier personnellement. Merci à Lisa Nicvert, Emma Acacia, Blandine Charrat, Alice Genestier, Léa Keurinck, Marie Morel, Alexia Nguyen-Trung, Rémi Tuffet et Barbara Vuillaume ; et plus globalement merci aux stagiaires, doctorants et postdoctorants qui rendent le laboratoire si vivant et si plaisant.

Si j'ai décidé de faire une thèse, c'est notamment grâce à plusieurs enseignants que j'ai pu rencontrer au cours de ma formation universitaire, en licence puis en master. Pour cette raison, je voudrais en remercier quelques-uns. J'ai déjà nommé mes directrice et directeur, Dominique et David, mais je voudrais les citer à nouveau ici. Leurs cours m'ont donné envie de faire de l'épidémiologie et de la modélisation, respectivement. En plus d'eux, de nombreux intervenants

croisés au détour d'un unique cours ou au contraire des dizaines de fois en cinq années d'étude m'ont transmis le goût de la modélisation, des sciences de la biodiversité, ou plus simplement du travail de chercheur. Parmi elles et eux, je voudrais nommer Isabelle Amat, Sylvain Charlat, François-Xavier Dechaume-Moncharmont, Christophe Douady, Gilles Escarguel, Emmanuel Desouhant, Anne-Béatrice Dufour, Marie Fablet, Clémentine François, Bernard Kaufmann et Christelle Lopes que j'ai déjà cités, Tristan Lefébure, Jean-Paul Léna, Frédéric Menu, Sylvain Mousset, Marc Philippe, Laure Segurel, Fabrice Vavre et Yann Voituron. J'espère n'en oublier aucun.

En plus de ces enseignants, j'ai pu faire la rencontre en licence et en master d'étudiants et d'étudiantes sans qui je n'en serais pas là aujourd'hui. Je voudrais exprimer ma gratitude générale à la promotion du master 1 BEE de l'année 2019-2020 et à la promotion du master 2 EEG de l'année 2020-2021. Merci pour les discussions – scientifiques ou non –, les soirées, les heures de révisions à la BU, les écritures de rapports et autres travaux collectifs. Il serait long de donner tous les noms, mais je voudrais tout de même remercier quelques personnes en particulier pour leur présence pendant le master et depuis. Merci donc à Anthony Antos, Nans Barthélémy, Romain « Romso » Dupuy-Jandard, Thibault Froger, Céline Jannas, Ninon Larcher-Orlando, Sarah Pouly et Hugues Saint-Pierre. Merci également aux « parasites » du master, Théo Fumagalli et Thomas Sagredo, ainsi qu'à Eddie Pérochon, qui, bien que ne faisant pas partie du master, ont été assez présents pour qu'on le croie.

Merci également à ces quelques personnes qui, en licence, ont accepté le gamin asocial que j'étais et qui m'ont donné le goût de la vie étudiante : Léa Backory, Nils Dumarski, Maureen Konrad, Claire Ménival et Livia-Maria Venturini. Un grand merci également à Mélodie Bastian qui, la première et dès la licence, m'a motivé à faire une thèse, et qui depuis lors n'a pas cessé de me pousser à me dépasser intellectuellement.

Je terminerai sur les amitiés en remerciant les quelques personnes qui me suivent depuis le lycée et qui ont supporté, pendant huit ans, de m'entendre raconter mes études puis ma thèse. Merci donc à Antoine Archier, Alexandre Poisson et Etienne Royan pour leur indéfectible amitié ; et merci également à Joshua Gonnet pour la stimulation intellectuelle apportée pendant ma dernière année de lycée.

Je remercie également ma famille pour leur soutien apporté tout au long de mes – longues –

études. Merci à mes parents, Véronique et Yannick Lecorvaisier, ainsi qu'à mon frère Kévin. Puisque ceux qui m'ont soutenu ne sont pas tous humains, je voudrais également remercier les chats de mes parents, Galika et surtout Hiro, qui nous a tragiquement quittés le 10 juillet de cette année.

Enfin, passionné de lecture, je voudrais citer quelques auteurs et autrices scientifiques qui, sans le savoir, grâce leurs ouvrages, m'ont continuellement inspiré et motivé à apprendre toujours plus, à étudier un maximum, et à poursuivre une carrière académique. Je voudrais ainsi remercier Samuel Alizon, déjà mentionné, pour son livre *Évolution, écologie et pandémies* qui m'a fait comprendre que l'épidémiologie représente bien plus que de compter des malades ; Marylène Patou-Mathis, pour son livre *L'homme préhistorique est aussi une femme*, qui m'a donné l'envie de lier science et histoire dans ce manuscrit ; Hope Jahren, pour son livre *Lab Girl*, qui m'a sorti d'un moment difficile de ma thèse en me montrant la beauté de la recherche à travers le plus beau livre traitant de science que j'aie lu ; Lluís Quintana-Murci, pour son livre *Le Peuple des humains*, qui m'a simplement passionné ; et enfin Steve Brusatte, pour ses livres *Le Triomphe et la Chute des dinosaures* et *Le Triomphe et le Règne des mammifères*, qui, comme *Lab Girl*, m'ont montré la beauté de la recherche.

PRÉAMBULE

Mon intérêt pour les sciences commence au collège. Je me souviens qu'en 4^{ème} je voulais faire des études pour devenir astrophysicien, pour finalement m'orienter vers les sciences de la vie au lycée, en terminale. J'ai commencé une licence de SVT sans trop savoir où cela me mènerait, puis en deuxième année, j'ai commencé à vouloir m'orienter vers la biologie de la conservation. En troisième année, je me suis convaincu que j'allais faire une thèse. J'ai obtenu en 2019 ma licence en Sciences de la Biodiversité, laquelle me permit de commencer un master Biodiversité, Écologie, Évolution.

La variété des champs disciplinaires présentés dans ce master m'a permis de découvrir de nombreux objets d'études, et de réveiller chez moi un intérêt pour une discipline qui en réalité attisait ma curiosité depuis le lycée : l'épidémiologie. Dans le cadre de ma première année de master, je devais effectuer un stage, et je me suis donc rapproché de Dominique Pontier, dont les travaux de recherche m'intéressaient. Elle m'a redirigé vers David Fouchet, qui accepta de m'encadrer sur un stage dont j'avais choisi le sujet : modéliser la propagation de la chytridiomycose, une maladie responsable de la disparition de populations entières d'amphibiens et à laquelle je m'intéressais depuis quelques années déjà. Malheureusement, cette année 2020 allait être marquée par le début de la pandémie de Covid-19, et les stages de première année de master furent annulés.

Je décidai de travailler sur mon sujet de stage de chez moi. L'avantage d'un travail de modélisation, c'est qu'il ne demande que peu de matériel : un ordinateur, des feuilles et un crayon sont suffisants. J'envoyais régulièrement mes avancées à David, qui acceptait patiemment et sans obligation de relire mon travail. À la fin de la période qui aurait dû être celle du stage, David me parla d'un projet de stage de M2 que Dominique avait depuis quelques années. Je la rencontrai pendant l'été pour en discuter, et elle accepta de me prendre en stage de janvier à mai 2021.

Ce stage consistait à créer un modèle pour déchiffrer les contextes dans lesquels la vaccination contre *Corynebacterium diphtheriae*, l'agent étiologique de la diphtérie, sélectionne des souches moins virulentes du pathogène. Après ce stage, j'obtins une bourse ministérielle de l'école doctorale E2M2 pour commencer une thèse en octobre de la même année, et mon travail de stage fut amélioré et terminé pour devenir le premier axe de ma thèse.

Dans ce manuscrit, je présente les résultats de ces trois années de thèse. Trois années de travail, de doutes, de collaborations, de discussions, d'écriture et de ratures. Trois années parfois plaisantes, parfois beaucoup moins. Mais trois années que ne regrette pas.

TABLE DES MATIÈRES

Résumé en français	5
Résumé en anglais (<i>abstract</i>)	7
Remerciements	9
Préambule	13
Table des figures	20
Table des tableaux	21
Table des encadrés	22
1 Introduction générale	23
1.1 Histoire naturelle des parasites	23
1.1.1 Origine évolutive des pathogènes et des parasites	23
1.1.2 Histoire des maladies infectieuses et des épidémies	27
1.2 Histoire et principe de la vaccination	34
1.2.1 De Lady Montagu aux vaccins contre le paludisme	34
1.2.2 Principe de la vaccination et types de vaccins	36
1.2.3 Histoire des vaccins anticoquelucheux	38
1.3 Évolution de la virulence	41
1.3.1 Définir et mesurer la virulence	41
1.3.2 Théories et observations sur l'évolution de la virulence	44
1.3.3 Manipulation expérimentale de la virulence	51
	15

1.3.4	Cas d'école : l'exemple de la myxomatose	54
1.4	Impact des vaccins sur l'écologie et l'évolution des pathogènes	56
1.4.1	Généralités	56
1.4.2	Le cas de <i>Corynebacterium diphtheriae</i>	60
1.4.3	Le cas de <i>Bordetella pertussis</i>	61
1.5	Modélisation en (éco-)épidémiologie	69
1.5.1	Brève histoire de la modélisation en épidémiologie	69
1.5.2	Exemples de modèles et de questions	71
1.5.3	Modéliser l'évolution de la virulence	74
1.5.4	Analyse de données	78
1.6	Conclusion	80
2	Modélisation théorique : Vaccination anatoxinique et évolution de la virulence	103
2.1	Introduction	103
2.1.1	Retours sur l'impact des vaccins sur l'écologie et l'évolution des pathogènes	103
2.1.2	La diphtérie comme modèle d'étude	106
2.2	Description du modèle	107
2.2.1	Formulation mathématique	107
2.2.2	Paramétrisation	111
2.2.3	Analyse d'invasion	113
2.3	Résultats des simulations	116
2.3.1	Prévalence de la souche 1	117
2.3.2	Prévalence de la souche 2	122
2.3.3	Invasion par une unique souche d'une population saine	126
2.3.4	Invasion endémique, compétition et coexistence	126
2.4	Discussion générale	129
2.5	Conclusion	134

3	Modélisation statistique : Facteurs environnementaux et propagation	
	de souches mutantes	141
3.1	Introduction	141
3.1.1	De la modélisation dynamique à l'analyse de données	141
3.1.2	Contexte et mise en situation	143
3.1.3	Étudier l'invasion d'une souche mutante	144
3.2	Modèle mathématique	146
3.3	Collection des données	148
3.3.1	Données de prévalence de souches	148
3.3.2	Choix et acquisition des cofacteurs	153
3.4	Transformation des cofacteurs	164
3.5	Approche fréquentiste	167
3.5.1	Modèle linéaire généralisé	167
3.5.2	Modèle linéaire généralisé mixte	169
3.6	Analyse de robustesse	171
3.6.1	Principe	171
3.6.2	Modèle linéaire généralisé	173
3.6.3	Modèle linéaire généralisé mixte	175
3.7	Approche bayésienne	177
3.7.1	Principe	177
3.7.2	Résultats	180
3.8	Analyse de puissance	182
3.8.1	Principe	182
3.8.2	Résultats	184
3.9	Discussion	188
3.10	Conclusion	191
4	Modélisation informatique : Propagation et échantillonnage de souches	
	mutantes	199
4.1	Introduction	199

4.1.1	Des données <i>in natura</i> aux données <i>in silico</i>	199
4.1.2	Utilisation des données <i>in silico</i> en sciences de la vie	200
4.2	Modèle épidémique	202
4.2.1	Description du modèle	202
4.2.2	Matrice de transition	203
4.2.3	Paramétrisation de la matrice de transition	205
4.2.4	Calcul des cofacteurs	210
4.2.5	Calcul des forces d'infection	211
4.2.6	Mise en place des simulations	212
4.3	Dynamique épidémiologique d'une souche unique	213
4.4	Compétition entre souches	218
4.5	Échantillonnage, estimation et qualité des cofacteurs	220
4.5.1	Simuler des données réalistes	220
4.5.2	Estimation des cofacteurs	222
4.5.3	Estimation des coefficients	226
4.6	Discussion	230
4.7	Conclusion	234
5	Discussion générale et perspectives	241
5.1	Retour sur les chapitres précédents	241
5.2	Limites des approches de modélisation	242
5.2.1	De l'utilité des modèles	242
5.2.2	Limites aux modèles dynamiques	243
5.2.3	Limites aux modèles statistiques	246
5.3	La pluridisciplinarité de l'(éco-)épidémiologie	249
5.3.1	Retour aux origines	249
5.3.2	L'approche de terrain	251
5.3.3	L'approche historique	251
5.3.4	L'approche moléculaire	252
5.3.5	L'approche mathématique	253

5.3.6	L'approche éco-évolutive	253
5.4	Quel rôle pour le chercheur et la chercheuse ?	255
5.4.1	Étymologie du terme	255
5.4.2	Interagir avec la société	255
5.4.3	Penser l'avenir	259
5.5	Questionner la recherche	265
5.6	Pour finir : quelques considérations personnelles	268
5.6.1	Généralités	268
5.6.2	La modélisation face la réalité du terrain	272
5.6.3	Être (dé)connecté de son sujet de recherche	274
	Acronymes	289
	Annexes	291
A	Articles publiés	291
B	Suppléments au Chapitre 2	308
B.1	Critère d'invasion mutuelle	308
B.2	Durée des simulations	309
B.3	Couverture vaccinale et proportion de vaccinés	310
C	Suppléments au Chapitre 3	312
C.1	Discretisation du modèle SIR	312
C.2	Formulation du modèle logit	313
C.3	Code du modèle bayésien	315

TABLE DES FIGURES

1.1	Relation entre le type d'interaction antagoniste et la taille des acteurs	24
1.2	Illustration du principe de la théorie du compromis entre virulence et transmission	47
1.3	Relation entre la perte de fitness induite par le parasite et la dépendance à la mobilité de l'hôte pour sa transmission	49
1.4	Schéma simplifié de l'évolution du virus de la maladie de Marek	59
2.1	Modèle épidémiologique	108
2.2	Prévalence de la souche 1 1	118
2.3	Prévalence de la souche 1 2	119
2.4	Prévalence de la souche 2 1	123
2.5	Prévalence de la souche 2 2	124
2.6	Nombre de reproduction effectif de la souche 2 1	127
2.7	Nombre de reproduction effectif de la souche 1	128
2.8	Nombre de reproduction effectif de la souche 2 2	129
2.9	Prévalence de la souche 1 ou de la souche 2	130
2.10	Nombre de reproduction effectif de la souche 1 ou de la souche 2	131
2.11	Nombre de reproduction effectif de la souche 2 3	132
3.1	Distribution des isolats de <i>B. pertussis</i> dans le jeu de données	150
3.2	Évolution temporelle dans la proportion de souches PRN-	153
3.3	Distribution des isolats de <i>B. pertussis</i> conservés dans les analyses	154
3.4	Localisation des stations météorologiques	158
3.5	Fluctuations des <i>z-scores</i> des différents cofacteurs	160
3.6	Illustration du calcul de $\chi_{i,s}(k)$	166
3.7	Résultats de l'analyse de robustesse du modèle linéaire généralisé	174

3.8	Résultats de l'analyse de robustesse du modèle linéaire généralisé mixte	176
3.9	Diagramme acyclique dirigé du modèle bayésien	179
3.10	Puissance du modèle bayésien 1	185
3.11	Puissance du modèle bayésien 2	186
3.12	Puissance du modèle bayésien 3	187
4.1	Influence de σ_y sur la prévalence de l'infection	214
4.2	Influence de σ_i sur la prévalence de l'infection	215
4.3	Influence de φ sur la prévalence de l'infection	216
4.4	Extinction locale et persistance de l'infection	217
4.5	Influence de μ_2 sur la prévalence de l'infection	218
4.6	Proportion de simulations pour lesquelles les souches se maintiennent dans la population	220
4.7	Valeurs obtenues pour les différents critères d'ajustement des simulations	223
4.8	Corrélation entre la valeur réelle et la valeur estimée des cofacteurs 1	224
4.9	Corrélation entre la valeur réelle et la valeur estimée des cofacteurs 2	225
4.10	Valeurs médiane de a_0 et valeurs de P 1	228
4.11	Valeurs médianes de a_f et valeurs de P 1	229
4.12	Valeurs médianes de a_0 et valeurs de P 2	230
4.13	Valeurs médianes de a_f et valeurs de P 2	231
5.1	Vision personnelle de la pluridisciplinarité de l'éco-épidémiologie	250
5.2	Calligraphie	256
S1	Proportion d'individus vaccinés en fonction de la couverture vaccinale	311

TABLE DES TABLEAUX

2.1	Paramètres du modèle	112
3.1	Résultats de l'analyse par le GLM 1	168

3.2	Résultats de l'analyse par le GLM 2	169
3.3	Résultats de l'analyse par le GLMM 1	171
3.4	Résultats de l'analyse par le GLMM 2	172
3.5	Résultats de l'analyse par le modèle bayésien 1	181
3.6	Résultats de l'analyse par le modèle bayésien 2	182
4.1	Formules des différentes probabilités de transition	204
4.2	Valeurs arbitraires des paramètres épidémiologiques	222

TABLE DES ENCADRÉS

1.1	Zoonoses, maladies émergentes et pandémies	30
1.2	Les mouvements antivax	35
1.3	Définitions en sciences de la biodiversité	42
1.4	Vaccins imparfaits	60
5.1	Éthique de la recherche	248
5.2	Taxonomie et problèmes éthiques	267

CHAPITRE I

INTRODUCTION GÉNÉRALE

Dans cette introduction générale, je commence par discuter de l'histoire naturelle des parasites, incluant les bactéries et virus, et l'histoire des maladies infectieuses. Ensuite, je présente l'histoire de la première grande arme de l'humanité contre les pathogènes : la vaccination et discute brièvement son principe. À partir de là, je traite de l'évolution de la virulence des pathogènes, puis je mets en lien l'usage massif de la vaccination au XX^e siècle et l'évolution des pathogènes en circulation, notamment au niveau de la virulence de ceux-ci. Finalement, je traite de la manière dont la modélisation mathématique permet de répondre à des questions de première importance en éco-épidémiologie.

1.1 – Histoire naturelle des parasites

1.1.1 – Origine évolutive des pathogènes et des parasites

Interroger l'origine des parasites revient à questionner l'origine du parasitisme. Le parasitisme est par définition une interaction entre deux organismes, l'un profitant de la relation (gain de valeur sélective, également appelée *fitness*) tandis que l'autre en souffre (perte de *fitness*). La différence entre parasitisme et prédation est mince, et tient plus de la différence de degré que de nature (Figure 1.1). Généralement, un parasite est beaucoup plus petit que son hôte, alors qu'un prédateur est plus grand que sa proie ou d'une taille comparable. De plus, les prédateurs vont généralement tuer leur proie immédiatement tandis que les parasites vont exploiter leur hôte sur un temps qui peut être considérablement long et ne vont pas nécessairement le tuer. Dans

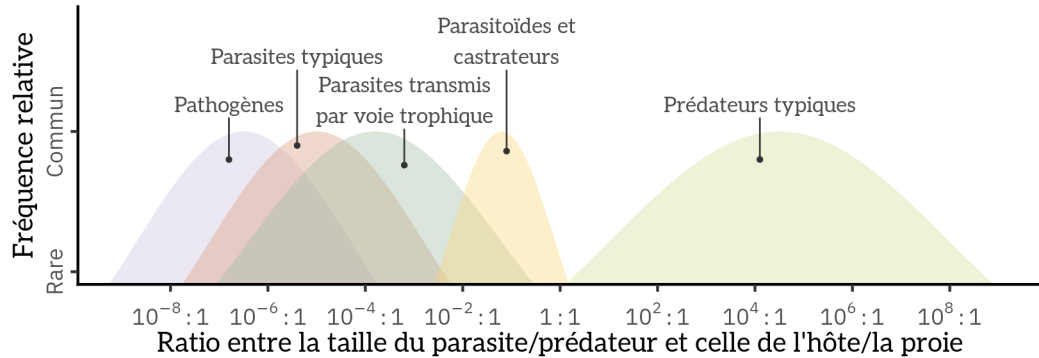


FIGURE 1.1 – **Relation entre le type d’interaction antagoniste et la taille des acteurs.** L’axe des abscisses indique le ratio de taille entre les antagonistes et l’axe des ordonnées indique une indication schématique de la fréquence du type d’interaction pour le ratio considéré. Adapté de la Figure 2 de [Lafferty et Kuris \(2002\)](#).

le langage courant, nous avons tendance à regrouper sous l’appellation de « pathogène » des organismes microscopiques tels que des bactéries et des virus, tandis que le terme « parasite » est plutôt réservé à des organismes macroscopiques et pluricellulaires tels les nématodes, les plathelminthes, etc. En écologie, la notion de parasite recouvre ces deux catégories, et j’utiliserai donc préférentiellement ce terme.

Premiers pas : les virus

Si les virus accompagnent l’humanité depuis ses origines (Section 1.1.2), à l’échelle de l’histoire de la vie, ceux-ci sont apparus bien longtemps avant que l’espèce humaine n’émerge. Les méthodes d’observation directe de virus permettent de remonter de quelques dizaines de millions d’années leur évolution. L’une de ces méthodes est l’observation d’insectes conservés dans l’ambre. Avec cette approche [Poinar et Poinar \(2005\)](#) ont pu observer la présence de virus chez plusieurs diptères et un lépidoptère dans des échantillons d’ambre datant de quelques millions d’années¹.

Pour des virus plus anciens, les méthodes directes ne fonctionnent pas à cause de l’absence de traces fossiles laissées par les virus. Il est néanmoins possible d’utiliser des méthodes indirectes pour détecter leur présence dans ces temps reculés. Par exemple, en étudiant des rétrovirus endogènes, c’est-à-dire des virus à ARN dont le génome s’est intégré dans celui d’un hôte, les chercheurs peuvent dater le moment de l’intégration, et donc estimer la date où ces virus ont

1. Pour plus d’informations sur la détection de virus anciens, voir [Poinar \(2021\)](#).

commencé à interagir avec leurs hôtes (Katzourakis, 2013). De cette manière, Lee *et al.* (2013) ont identifié un rétrovirus endogène ayant infecté l'ancêtre commun des mammifères il y a entre 110 et 104 millions d'années et Suh *et al.* (2013) ont identifié que le virus de l'hépatite B s'est endogénisé chez l'ancêtre commun des Neoves il y a plus de 82 millions d'années.

Finalement, ces différentes méthodes ne permettent pas de remonter très loin dans le temps à l'échelle du vivant. Pourtant, il est vraisemblable que les virus soient beaucoup plus anciens que ces ordres de grandeur. Il est probable que les premiers parasites aient été des virus, et ces premiers virus auraient émergé entre l'apparition des premières cellules et la différenciation entre bactéries, archées et eucaryotes, voire avant l'apparition des premières cellules – ce qui impliquerait l'existence de virus archaïques non parasitaires (Forterre, 2006 ; Holmes, 2011 ; Nasser *et al.*, 2020). Ainsi, le dernier ancêtre commun universel (LUCA) des bactéries, archées et eucaryotes, ayant émergé il y a quatre milliards d'années (Moody *et al.*, 2024), était vraisemblablement déjà infecté par des virus (Krupovic *et al.*, 2020). Les chercheurs avançant une origine aussi primitive que la vie aux virus se basent sur des indices indirects, comme l'ubiquité des virus dans le règne du vivant et l'homologie de certaines protéines comme l'ARN polymérase ARN-dépendante, partagée par tous les virus à ARN (Wolf *et al.*, 2018).

Premières bactéries parasites

Avec la diversification des formes de vie cellulaires, de nombreux parasites ont pu émerger, mais également des formes de vie symbiotiques et commensales. Une symbiose désigne une interaction dans laquelle la relation est bénéfique pour chacun des participants (gain mutuel de *fitness*) tandis que le commensalisme désigne une interaction neutre pour l'un des participants et positive pour l'autre. Il est parfois difficile d'évaluer la nature d'une interaction, et celle-ci peut avoir changé au cours du temps.

Il est donc particulièrement difficile d'identifier l'origine des bactéries parasites. Il est plus simple de dater l'apparition des interactions entre les bactéries et d'autres organismes. L'exemple typique d'association entre une bactérie et une autre forme de vie est la mitochondrie. Selon l'hypothèse la plus largement acceptée, cet organite proviendrait de l'endosymbiose d'une alphaprotéobactérie proche du genre *Rickettsia* – comprenant de nombreuses espèces pathogènes et symbiotiques – par une cellule eucaryote ; cette endosymbiose daterait de plus de deux milliards

d'années (Feng *et al.*, 1997; Andersson *et al.*, 1998).

L'ancienneté de l'endosymbiose mitochondriale laisse supposer une origine primitive des bactéries parasitaires, ou en tout cas des bactéries capables de s'introduire dans le cytosol d'autres cellules comme les rickettsies. De telles bactéries ont nécessairement émergé entre l'endosymbiose, datée de deux milliards d'années, et l'apparition de LUCA il y a quatre milliards d'années.

Diversification des eucaryotes et multiplication des parasites

La diversification des formes parasitaires suit celle des hôtes et notamment celle liée à l'apparition de la multicellularité². Quoique celle-ci soit possiblement apparue de manière précoce il y a deux milliards d'années (El Albani *et al.*, 2010), c'est au cours de l'Édiacarien (-635 à -541 Ma) puis du Cambrien (-539 à -485 Ma) que les formes de vie multicellulaires vont grandement se diversifier.

C'est du début du Cambrien que datent les plus anciens parasites multicellulaires connus à ce jour. Il s'agit d'organismes tubulaires – leur taxonomie est inconnue – parasitant la coquille de bivalves de l'espèce *Neobolus wulongqingensis* (Zhang *et al.*, 2020). Ce cas serait une forme de cleptoparasitisme plutôt que de parasitisme *stricto sensu*, c'est-à-dire que ces « tubes » n'exploitaient pas directement leur hôte mais s'approprièrent sa nourriture. Au Cambrien moyen, d'autres traces de parasites tubulaires sont retrouvées, cette fois sur le brachiopode *Linnarssonina constans* (Bassett *et al.*, 2004).

D'autres indices placent également l'origine des parasites multicellulaires au Cambrien, notamment des études phylogénétiques chez les Myxozoa (Kodádková *et al.*, 2015) et les Pentatomides (Sanders et Lee, 2010). Au cours des époques géologiques suivantes, les traces fossiles de parasites vont se multiplier, preuve du succès continu de ce mode de vie (Leung, 2017).

De nos jours, les parasites se retrouvent dans tous les grands clades du vivant, que ce soit les animaux (par exemple chez les insectes, les isopodes, les nématodes...), les végétaux, les bactéries, les virus ou les protozoaires. Le mode de vie parasite est nécessairement avantageux dans une grande variété de contextes pour avoir ainsi émergé de manière convergente parmi de si nombreux groupes du vivant, et il a été estimé que la moitié des espèces vivantes seraient des parasites,

2. Ou plutôt aux apparitions, la multicellularité étant vraisemblablement apparue plusieurs fois de manière indépendante (Knoll, 2011)

encore qu’il soit difficile de produire des estimations fiables (Windsor, 1998 ; Poulin et Morand, 2000 ; Dobson *et al.*, 2008 ; Poulin, 2014).

Par définition, les parasites induisent une perte de *fitness* chez leur hôte et leur sont donc nuisibles. Ainsi, un grand nombre d’entre eux sont responsables de pathologies, qu’on nommera maladies infectieuses.

1.1.2 – Histoire des maladies infectieuses et des épidémies

Une relation ancienne

L’histoire humaine est parsemée d’épidémies. La plus célèbre d’entre elles est certainement la peste noire, due à la bactérie *Yersinia pestis*, ayant ravagé l’Europe – mais également une partie de l’Asie et de l’Afrique du Nord – vers le milieu du XIV^e siècle. Mais si cette épidémie est certainement l’une des plus meurtrières de l’histoire, elle est loin d’être la plus ancienne. La peste de Justinien (c. 540 – c. 750), la peste de Cyprien (c. 248 – c. 270), la peste antonine (c. 165 – c. 188) et la peste d’Athènes (c. -430 – c. -426) l’ont précédée³ (Stange, 2021). Si elles sont parmi celles ayant fait couler le plus d’encre, que ce soit par les historiens contemporains des événements ou ceux d’aujourd’hui, ces épidémies sont en réalité récentes dans l’histoire humaine. En remontant le cours du temps, on retrouve la trace d’autres maladies et épidémies moins documentées. L’une des plus anciennes maladies infectieuses dont on ait une trace historique est la variole : des indices médicaux et documentaires laissent à supposer que cette maladie aurait décimé la famille du pharaon Ramsès V au XII^e siècle avant notre ère⁴ (Strouhal, 1996). Il y a plus longtemps encore, des épidémies de tularémie (causées par la bactérie *Francisella tularensis*) sont suspectées dans l’empire hittite au XIV^e siècle avant notre ère (Trevisanato, 2007) et en Égypte au XIX^e siècle avant notre ère (Trevisanato, 2004).

Si ce sont là les plus anciennes épidémies documentées, il est certain que l’humanité a souffert de maladies infectieuses bien avant que l’histoire commence à en garder des traces écrites. Ainsi, le célèbre Ötzi, momifié naturellement dans les Alpes il y a 5 300 ans, était infecté par *Helicobacter pylori*, une bactérie pouvant provoquer des ulcères (Maixner *et al.*, 2016) et, à la

3. Si toutes portent le nom de « peste », il est probable que certaines aient été causées par d’autres pathogènes que le bacille de la peste

4. « Avant notre ère » signifiera dans tout le texte « avant l’an 1 ».

même période, un chasseur-cueilleur d'Europe du Nord portait le bacille de la peste ([Susat et al., 2021](#)). La génomique permet aujourd'hui de retracer l'histoire évolutive de maladies pré-datant les premières traces écrites ([Spyrou et al., 2019](#)).

Il est possible de remonter encore plus loin que les traces historiques et fossiles. L'étude génomique du virus JC montre que celui-ci a commencé à infecter la lignée humaine avant sa sortie d'Afrique, ([Pavesi, 2003, 2004, 2005](#)) comme le virus de l'hépatite G ([Pavesi, 2001](#))⁵. On sait également que la lignée humaine subissait le joug de parasites bien avant l'apparition de *Homo sapiens*. C'est par exemple le cas des herpèsvirus : ces virus à ADN, dont les différentes espèces ont coévolué avec les vertébrés depuis les 200 derniers millions d'années, accompagnent la lignée humaine depuis ses origines, il y a environ six millions d'années⁶ ([McGeoch et Cook, 1994](#); [McGeoch et al., 1995](#); [Wertheim et al., 2014](#)).

L'un des événements les plus importants dans l'histoire des maladies humaines est certainement la néolithisation ([Barrett et al., 1998](#); [Harper et Armelagos, 2013](#)). Ce terme désigne un changement radical dans le mode de vie de nombreux peuples humains qui passent de chasseurs-cueilleurs nomades à agriculteurs sédentaires. Ce changement de mode vie a eu deux conséquences majeures. Premièrement, les densités des populations humaines augmentent fortement, favorisant l'émergence et le maintien de maladies. Deuxièmement, la domestication entraîne une forte proximité entre les humains et de nombreuses espèces animales, dont les bovidés, qui transmettent de nouveaux pathogènes et maladies aux humains ([Brüssow, 2009](#)). Ces maladies issues d'animaux non humains, ou zoonoses, ont continué d'émerger pendant le reste de l'histoire humaine et ont été à l'origine de plusieurs pandémies (Encadré 1.1). [Barrett et al. \(1998\)](#) considèrent que la néolithisation constitue la première « transition épidémiologique »⁷. Parmi les maladies dont l'émergence a été reliée à l'invention de l'agriculture, on peut citer la peste, certainement apparue 3 400 ans avant notre ère ([Susat et al., 2021](#)), ou la variole, qui serait apparue à la même époque ou plus tôt encore mais dont la date d'émergence reste fortement débattue ([Shchelkunov, 2009](#); [Babkin et Babkina, 2015](#)).

5. L'étude de ces pathogènes anciens permet d'étudier les grandes migrations humaines. Voir la très belle thèse d'Alexia [Nguyen-Trung \(2024\)](#) à ce sujet.

6. Date proposée par [Kumar et al. \(2005\)](#) pour la divergence entre la lignée humaine et la lignée des chimpanzés. D'autres auteurs avancent d'autres dates.

7. La deuxième transition épidémiologique est le recul de la mortalité des maladies infectieuses à la faveur des maladies chroniques au cours de la révolution industrielle, et la troisième correspond à l'émergence de nouveaux pathogènes ou leur réémergence (et l'apparition des résistances aux antibiotiques) à l'ère de la mondialisation.

Par leur histoire parfois ancienne, les maladies ont un rôle majeur dans le façonnement des populations humaines. En augmentant la mortalité, elles réduisent l'espérance de vie des individus et ralentissent la croissance des populations, jouant ainsi un rôle clé dans la démographie humaine. Par exemple, il a été estimé que la variole a été responsable de 8 à 20 % de la mortalité du XVIII^e siècle (Mercer, 1985). De même, les multiples épidémies de peste présentées plus tôt ont, par leur forte mortalité, contribué à réduire la croissance humaine mais aussi perturbé l'équilibre des sociétés les ayant subies – amorçant même la chute de l'Empire Romain (Stange, 2021). Mais les maladies ont également des conséquences sur le plus long terme. En coévoluant avec leurs hôtes, les parasites sont responsables de fortes pressions de sélection exercées sur leurs génomes ; ainsi, avec la diversification des pathogènes au néolithique, les populations humaines ont développé une résistance génétique à certaines infections (Kerner *et al.*, 2023).

La malaria – aussi appelée paludisme – est l'exemple typique de maladie qui est à la fois une zoonose, une maladie de la néolithisation, et un acteur clé de l'évolution humaine. La maladie est provoquée par des protozoaires du genre *Plasmodium*, dont *Plasmodium falciparum* et *P. vivax*. *P. falciparum* aurait été transmis à l'humain par le gorille (*Gorilla gorilla*) (Liu *et al.*, 2010), tandis que *P. vivax* proviendrait d'un macaque⁸ (genre *Macaca*) (Mu *et al.*, 2005). Le plus virulent de ces parasites, *P. falciparum*, aurait commencé à se propager dans les populations humaines il y a 2 500 ans, alors que la néolithisation de l'Afrique de l'Ouest s'est accomplie il y a 4 000 à 5 000 ans (Carter et Mendis, 2002 ; Michel *et al.*, 2024). Le paludisme a représenté une telle pression sur les populations humaines au cours de l'histoire qu'aujourd'hui encore des mutations persistent dans le génome de certaines populations conférant un avantage contre la maladie, mais qui sont pourtant délétères en son absence (Kwiatkowski, 2005). Autrement dit, certaines mutations ayant un effet négatif en absence du pathogène ont été sélectionnées parce qu'elles confèrent une protection contre ce pathogène dont le bénéfice contre celui-ci est supérieur au coût à porter ces mutations.

Dans mes travaux de thèse, je me suis intéressé à deux pathologies humaines d'origine bactérienne : la diphtérie et la coqueluche. Dans les sections suivantes, je présente brièvement l'histoire à la fois sociale et évolutive de ces deux maladies infantiles.

8. Cela implique une origine asiatique de *P. vivax*, contestée par certaines études (voir Michel *et al.*, 2024).

ENCADRÉ 1.1 – Zoonoses, maladies émergentes et pandémies

Dans la nature, il arrive qu'un pathogène circulant habituellement au sein d'une population animale non humaine commence à se propager chez l'humain. La maladie qui en résulte est alors appelée une zoonose. La grippe espagnole est un exemple typique de zoonose devenue pandémie. Avant de commencer à se propager dans la population humaine, le virus responsable de la grippe espagnole circulait vraisemblablement chez un hôte aviaire (Taubenberger *et al.*, 2005 ; Reid *et al.*, 2004). Il se serait ensuite adapté aux hôtes mammifères peu de temps avant la déclaration de la pandémie (Reid *et al.*, 1999, 2000). Une fois la barrière d'espèce franchie – vraisemblablement en mars 1918 aux alentours d'un camp militaire aux États-Unis (Berche, 2022) – la maladie débarque rapidement en Europe, transportée par des soldats américains. Elle arrive ensuite en Asie via les colonies européennes avant de finalement atteindre l'Afrique et l'Océanie ^a.

La grippe espagnole, comme le Covid-19 ou le SIDA, sont des exemples assez spectaculaires de zoonoses devenues des pandémies. Cependant, elles ne représentent que la partie émergée d'un immense iceberg éco-épidémiologique. De très nombreux pathogènes extrêmement virulents ont été transmis à l'humain depuis des populations animales sauvages : le virus Ebola est suspecté provenir de différentes espèces de chauve-souris (Marí Saéz *et al.*, 2015 ; Leroy *et al.*, 2005 ; Pourrut *et al.*, 2009), comme le virus Marburg (Pourrut *et al.*, 2009 ; Amman *et al.*, 2012), le virus Nipah (Chua *et al.*, 2002) et le virus Hendra (Halpin *et al.*, 2000) ; le VIH aurait été transmis aux humains par la consommation de viande de chimpanzés (Peeters *et al.*, 1989 ; Keele *et al.*, 2006) ; les hantavirus circulent chez plusieurs espèces de rongeurs (Escutenaire *et al.*, 2000 ; Heyman *et al.*, 2004, 2009 ; Plyusnina *et al.*, 2011)... ^b

D'autres maladies plus anciennes sont suspectées d'avoir été transmises à l'humanité via les espèces animales domestiquées. C'est par exemple le cas de la rougeole ^c qui aurait une origine bovine. Traditionnellement, on pensait que la maladie était apparue au cours du Néolithique tardif, c'est-à-dire rapidement après la domestication des bovins, lorsque les populations humaines sont devenues assez importantes pour permettre la subsistance de la maladie (Brüssow, 2009 ; Black, 1966). Cependant, des études moléculaires estiment que la divergence entre le virus de la rougeole et son plus proche cousin, le *rinderpest virus* RPV, aurait eu lieu plus tardivement. Mais les dates proposées divergent beaucoup : VI^e siècle avant notre ère (Düx *et al.*, 2020) ou XII^e siècle de notre ère (Furuse *et al.*, 2010) ^d.

Une maladie infectieuse émergente est définie comme « une maladie inconnue qui apparaît subitement dans la population, ou qui existait mais dont l'incidence et l'aire de distribution augmentent soudainement » (Stephen Morse cité dans Robin, 2021). À l'époque moderne, la plupart des maladies émergentes sont causées par des pathogènes franchissant

la barrière d'espèce (Taylor *et al.*, 2001). Les activités humaines comme les déplacements, les interactions avec la faune sauvage (incluant la chasse et la consommation de chair) et la destruction de l'environnement facilitent le passage d'une espèce à l'autre et d'une région à une autre (Weiss et McMichael, 2004).

- a. Exposition « Épidémies, prendre soin du vivant », du 12 avril 2024 au 16 février 2025, Musée des Confluences, Lyon.
- b. On peut également citer le SARS-CoV, le MERS-CoV, de nombreux virus grippaux...
- c. Ou encore de la variole.
- d. Les données historiques coïncideraient plutôt avec une émergence au XII^e siècle, la maladie n'ayant jamais été décrite plus tôt.

Le cas de la diphtérie

La diphtérie est une maladie touchant principalement les voies respiratoires causée par la bactérie *Corynebacterium diphtheriae*⁹, quoique d'autres espèces du genre *Corynebacterium* comme *C. ulcerans* puissent également provoquer des symptômes proches de la diphtérie (Tiwari et Wharton, 2012). Plus de détails sur la bactérie responsable de la diphtérie et sur la maladie elle-même sont donnés dans la Section 2.1.

Comme pour la variole, la plus ancienne trace d'une maladie ressemblant à la diphtérie nous vient d'Égypte. Une tête momifiée datant de l'époque du Nouvel Empire (période s'étalant du XVI^e au XI^e siècle avant notre ère) présente de l'ADN du genre *Corynebacterium*, quoique l'espèce précise n'ait pas pu être identifiée (Zink *et al.*, 2001). Plus tôt encore, des manuscrits datés entre le XIX^e et le XVI^e siècle avant notre ère font la description d'une maladie évoquant la diphtérie (Zink *et al.*, 2001 ; Tiwari et Wharton, 2012). Par la suite, Hippocrate décrit au VII^e siècle avant notre ère une maladie ressemblant à la diphtérie, puis Arétée de Cappadoce au II^e siècle de notre ère apporte ce qui est généralement accepté comme étant la première description non ambiguë de la maladie (Guilfoile, 2009). La diphtérie est par la suite régulièrement décrite dans des traités médicaux jusqu'au VI^e siècle avant de disparaître des écrits pendant presque un millénaire, pour finalement réapparaître en Europe à la fin du XVI^e siècle (Berche, 2007 ; Guilfoile, 2009). À partir de ce moment, la maladie se maintient en Europe et déclenche de nombreuses épidémies.

Quoique la diphtérie puisse occasionnellement toucher des animaux non humains, les données épidémiologiques indiquent que la transmission de la maladie se fait presque exclusivement entre personnes (Guilfoile, 2009). Il est cependant intéressant de noter que *C. ulcerans*, espèce sœur

⁹. Pour l'anecdote : avant que sa nature bactérienne ne soit découverte, la diphtérie a été suspectée d'être causée par un champignon (Taylor, 1881).

de *C. diphtheriae* et également pathogène, possède de nombreuses espèces réservoir comme le cochon domestique (*Sus domesticus*) (Schuhegger *et al.*, 2009), le rakali (*Hydromys chrysogaster*) (Eisenberg *et al.*, 2015), le sanglier (*Sus scrofa*) (Contzen *et al.*, 2011) ou encore le castor européen (*Castor fiber*) (Sting *et al.*, 2023).

Finalement, peu de choses sont actuellement connues sur l'origine évolutive de la diphtérie. L'existence de multiples espèces de *Corynebacterium* provoquant différentes pathologies chez plusieurs espèces de mammifères suggère l'existence d'une relation ancienne entre les *Corynebacterium* et les mammifères (Nagaraja, 2022). Au cours de cette coévolution, *C. diphtheriae* aurait acquis un gène issu d'un bactériophage, *tox*, lui permettant de produire la toxine diphtérique (TD), une protéine lui permettant d'exploiter plus efficacement son hôte (Pappenheimer et Gill, 1973), tout en se spécialisant pour devenir un parasite quasiment exclusif de *Homo sapiens* (Tiwari et Wharton, 2012; Guilfoile, 2009).

Le cas de la coqueluche

La coqueluche est finalement assez semblable à la diphtérie. Elle est provoquée par la bactérie *Bordetella pertussis* et touche également les voies respiratoires. Elle provoque dans sa forme classique une toux longue (parfois pendant plusieurs semaines) avec des inspirations produisant un son aigu d'où lui vient son nom anglais de *whooping cough*.

Son histoire débute plus tardivement que celle de la diphtérie. De nombreux auteurs citent Guillaume de Baillou (1538-1616) traitant de l'épidémie de Paris en 1578 comme étant la première description de la coqueluche aisément reconnaissable comme telle (Weston, 2012; Aslanabadi *et al.*, 2015; Rohani et Scarpino, 2018). Il existe cependant des mentions plus précoces de maladies ressemblant à la coqueluche. En Perse, à la fin du XV^e siècle et au début du XVI^e siècle, ce sont trois épidémies ressemblant à la coqueluche qui sont décrites : deux entre 1484 et 1495 (les dates précises sont inconnues) et une en 1501 (Aslanabadi *et al.*, 2015). Plus tôt encore, en Corée, une encyclopédie médicale publiée en 1433, le *Hyangyak Chipseong pang*, décrit une maladie évoquant la coqueluche (Magner, 1993). La maladie va être largement décrite en Europe à partir du XVII^e siècle mais surtout aux XVIII^e et XIX^e siècles dans plusieurs traités médicaux (Weston, 2012).

Les connaissances actuelles sur l'origine évolutive de la coqueluche sont plus fournies que celles sur la diphtérie. La majorité des souches de *B. pertussis* en circulation aujourd'hui seraient

issues d'une lignée ayant subi un considérable goulot d'étranglement il y a environ 500 ans, qui correspondrait à la période où cette lignée de *B. pertussis* se serait spécialisée comme pathogène humain, et qui coïncide avec les premières mentions de la maladie (Bart *et al.*, 2014).

Si *B. pertussis* est un pathogène exclusivement humain, elle appartient à un genre comprenant plusieurs espèces parasites infectant une grande étendue d'hôtes. Parmi ces espèces, *B. avium* infecte de nombreux oiseaux (Raffel *et al.*, 2002), *B. bronchiseptica* infecte une grande variété de mammifères et quelques oiseaux (Farrington et Jorgenson, 1976; Goodnow, 1980; Musser *et al.*, 1987) et *B. parapertussis* est séparée en deux lignages, l'un spécifique de l'humain et l'autre spécifique du mouton (*Ovis aries*) (Van Der Zee *et al.*, 1997). Les espèces *B. pertussis*, *B. parapertussis* et *B. bronchiseptica* sont regroupées sous l'appellation d'espèces « classiques » de *Bordetella*, et leur étude indique une évolution de *B. pertussis* à partir d'un ancêtre généraliste mais peu virulent, ressemblant à *B. bronchiseptica*, vers une bactérie spécialiste et plus virulente il y a entre 0,3 et 3,5 millions d'années (Parkhill *et al.*, 2003; Diavatopoulos *et al.*, 2005; Linz *et al.*, 2016). Cette spécialisation se serait faite principalement par perte de gènes, un phénomène qui serait également responsable de la spécialisation de *B. holmesii* pour l'hôte humain et de *B. avium* pour les oiseaux.

Il est possible de remonter encore plus loin l'histoire évolutive du genre *Bordetella*. Contrairement à la plupart des espèces du genre, *B. petrii* n'est pas un parasite obligatoire, et peut se reproduire dans l'environnement (von Wintzingerode *et al.*, 2001). L'étude de *B. petrii* a mené à postuler l'hypothèse que l'ancêtre des *Bordetella* était une bactérie environnementale ayant subi une lente radiation évolutive jusqu'à obtenir la diversité de pathogènes observée aujourd'hui (Gross *et al.*, 2008; Soumana *et al.*, 2017).

En conclusion, l'histoire évolutive de *B. pertussis* peut être résumée de la manière suivante : une bactérie initialement environnementale, ressemblant à *B. petrii* aurait évolué vers plusieurs espèces pathogènes généralistes. L'une de ces espèces, pathogène des mammifères et ressemblant à *B. bronchiseptica*, se serait par la suite spécialisée pour devenir un pathogène exclusif de *Homo sapiens*. Cependant, l'histoire évolutive de *B. pertussis* est loin d'être terminée. Évidemment, toute espèce évolue en continu, mais une force sélective importante, créée par l'humanité, dirige l'évolution de ce pathogène et bien d'autres : la vaccination.

1.2 – Histoire et principe de la vaccination

1.2.1 – De Lady Montagu aux vaccins contre le paludisme

L'histoire de la vaccination commence par celle de la variolisation. Le procédé est répandu en Orient, notamment en Afrique et en Inde, mais également en Chine où il est documenté dès le X^e siècle (Dinc et Ulman, 2007; Gross et Sepkowitz, 1998). La variolisation est importée de Turquie vers le Royaume-Uni par Mary Wortley Montagu (1689 – 1762) en 1721 (Dinc et Ulman, 2007). Elle consiste à appliquer du pus d'un malade atteint de la variole et de l'appliquer sur la peau entaillée de la personne à immuniser¹⁰. Le procédé sera adopté notamment chez les nobles européens, malgré les réticences de nombreux savants de l'époque (Encadré 1.2).

Quoiqu'efficace, le procédé est dangereux, beaucoup plus que son successeur : la vaccination. Celle-ci est inventée – ou plutôt documentée – par le médecin britannique Edward Jenner¹¹ (1749 – 1823) à la fin du XVIII^e siècle. Il observe que les paysannes trayant les vaches ont souvent des pustules sur les mains mais ne contractent pas la variole. Il émet alors l'hypothèse que les vachères gagnent une forme de protection contre la variole en contractant la vaccine, *cowpox* en anglais, maladie bénigne. Il va vérifier son hypothèse en vaccinant – c'est-à-dire en inoculant la vaccine – un premier enfant en 1796 puis en lui inoculant plusieurs fois la variole¹² (Jenner, 1798). De là, le procédé va se généraliser au Royaume-Uni, remplaçant peu à peu la variolisation.

Il faudra ensuite attendre presque 80 ans pour que de nouveaux progrès majeurs aient lieu dans le domaine de la vaccination, avec la théorie des germes du chimiste français Louis Pasteur (1822 – 1895), fondateur de la bactériologie, et le développement de nouveaux vaccins efficaces¹³. Pasteur développe d'abord des vaccins à destination des animaux (choléra des poules, maladie du charbon) puis c'est au tour du vaccin contre la rage d'être mis au point. Ce vaccin a la

10. En Chine, la méthode traditionnelle consiste à se placer du coton imbibé de pus de malade séché dans les narines.

11. En réalité, Benjamin Jesty (1736 – 1816) a été le premier, en 1774, à inoculer la vaccine à lui-même ainsi qu'à sa femme et ses fils – l'histoire ne nous dit pas s'ils étaient consentants – pour tester l'hypothèse que cette méthode protégeait contre la variole (Gross et Sepkowitz, 1998). C'est cependant Jenner qui étudiera scientifiquement la méthode et la popularisera.

12. L'éthique en sciences expérimentales était en ce temps un concept presque inexistant.

13. D'autres avant lui, comme le Belge Louis Willem (1822 – 1907), ont tenté de produire de nouveaux vaccins, mais à part le vaccin de Willem contre la péripneumonie du bétail, tous ont été des échecs (Berche, 2007).

particularité d'être à visée curative, là où les autres vaccins sont à visée prophylactique. Le XX^e siècle sera celui des grandes innovations médicales, avec la découverte des antibiotiques et de très nombreux vaccins, par exemple contre la tuberculose, la poliomyélite ou encore la coqueluche (Section 1.2.3). Le 8 mai 1980, l'OMS annonce officiellement que la variole a été éradiquée sur toute la planète, conclusion de plus de 250 ans de lutte vaccinale contre un pathogène qui accompagnait l'humanité depuis l'aube des civilisations.

Au XXI^e siècle la lutte continue contre les maladies infectieuses et de nouveaux vaccins sont continuellement développés. Par exemple, la pandémie de Covid-19 a conduit au rapide développement et à la mise sur le marché de plusieurs vaccins et depuis 2022 deux vaccins contre le paludisme sont entrés en circulation dans plusieurs pays d'Afrique. Néanmoins, l'espoir de voir un jour disparaître les maladies infectieuses grâce aux vaccinés est ténu, notamment à cause des mouvements de contestation de la vaccination (Encadré 1.2), de l'évolution des pathogènes en réponse à l'utilisation des vaccins (Section 1.4) ou encore du mode d'action de certains vaccins, qui n'empêchent pas l'infection et la transmission des pathogènes mais protègent de leurs symptômes (Section 1.4).

ENCADRÉ 1.2 – Les mouvements antivax

Le contenu de cet Encadré est essentiellement adapté du livre *Antivax : la résistance aux vaccins du XVIII^e siècle à nos jours* de [Salvadori et Vignaud \(2019\)](#).

Peu de sujets médico-scientifiques font autant couler d'encre que la vaccination. Ce rejet des vaccins n'est pas une particularité de l'époque moderne, bien au contraire : il est aussi vieux que la variolisation. Lorsque Lady Montagu importe la méthode en Europe, elle fait face à de nombreuses critiques à la fois sexistes (le procédé est présenté par une femme) et xénophobes (le procédé vient d'Orient). Lorsque Edward Jenner propose la vaccination comme alternative plus sûre à la variolisation, il fait à son tour face à des critiques contre son procédé qu'on traite d'impie car il implique de s'injecter du pus de vache.

Pourtant, cela ne va pas empêcher le progrès de se faire et les politiques vaccinales de se construire. La vaccination contre la variole devient obligatoire dans de nombreuses sociétés au XIX^e siècle, puis ce sera le tour d'autres vaccins au siècle suivant. Après l'annonce de l'éradication de la variole le 8 mai 1980 par l'OMS, les « antivax » comme on les appelle aujourd'hui obtiennent une victoire : l'arrêt de la vaccination contre cette maladie.

L'arrêt de la vaccination contre la variole suit une logique simple : les accidents vaccinaux existent pour tous les vaccins ; or si la maladie n'existe plus, alors les accidents vaccinaux

deviennent impardonnables. C'est l'un des principaux arguments rationnels utilisés dans le passé comme aujourd'hui par les groupes antivax : si le coût à la vaccination est supérieur au coût engendré par la maladie, alors la vaccination n'a aucun intérêt. Mais il existe également d'autres arguments contre la vaccination, notamment philosophiques ou religieux. Par exemple, certaines personnes considèrent la vaccination contre-nature ou blasphématoire.

Finalement, le débat sur la vaccination est loin d'être clos. Les parents français doivent faire vacciner leurs enfants pour de plus en plus de maladies qu'ils n'ont souvent jamais croisées de leur vie. Il est légitime de se poser la question du risque que l'on prend volontairement (celui de soumettre son enfant aux effets indésirables du vaccin) face à celui que l'on subit contre son gré (celui que l'enfant tombe malade).

1.2.2 – Principe de la vaccination et types de vaccins

Le principe général de la vaccination est d'introduire tout ou partie d'un corps étranger (pathogène) chez un individu pour déclencher une réponse immunitaire, idéalement sans provoquer de symptômes ou d'effets indésirables. Le but est ainsi de protéger des futures infections par le pathogène, l'individu dut-il le rencontrer dans la nature. Ce principe est exploité avant même que les bases de l'immunologie ne soient posées.

Au temps de la variolisation, la méthode consiste à introduire l'agent pathogène – et tout ce qui l'accompagne dans le pus des malades – directement dans le corps de l'inoculé. À l'époque de Jenner, la méthode n'évolue globalement pas. C'est Pasteur, grâce à la découverte des microbes, qui commence à produire des vaccins de manière « propre ». Son premier vaccin, contre le choléra des poules, est constitué de bactéries *Pasteurella multocida* atténuées simplement en laissant les colonies vieillir à l'air libre. Elles perdent ainsi leur virulence tout en gardant leur capacité à induire une protection contre une infection naturelle. Lors de ses travaux sur la maladie du charbon, provoquée par *Bacillus anthracis*, Pasteur chauffe ses cultures bactériennes pour atténuer plus efficacement la virulence du pathogène, et ses acolytes Émile Roux (1853 – 1933) et Charles Chamberland (1851 – 1908) améliorent le procédé grâce à l'utilisation d'agents chimiques : le bichromate de potasse et l'acide phénique. Pour une troisième maladie, le rouget du porc, Pasteur atténue le pathogène (*Erysipelothrix rhusiopathiae*) en le faisant passer plusieurs fois par un hôte auquel il n'était pas adapté, le lapin. Il réutilise cette technique, ainsi que celle de la dessiccation à l'air libre, lors de la confection de son vaccin contre la rage qui le fera passer à la postérité (Berche, 2007).

Ainsi, les premiers « vaccins vrais » sont des vaccins où le germe (ou dans le cas de la rage, le virus) est entier et vivant mais atténué. Ce sont Émile Roux et Alexandre Yersin ¹⁴ (1863 – 1943) qui découvrent la première toxine bactérienne : la toxine diphtérique, produite par *Corynebacterium diphtheriae*. Ils émettent alors l'idée que cette toxine puisse être utilisée pour développer une immunité contre la diphtérie. Ces résultats sont publiés en 1888 dans la thèse de Yersin (Berche, 2007).

À la même période, l'immunologie fait ses débuts. En 1884, Élie Metchnikoff (1845 – 1916), naturaliste russe spécialiste des vers et des échinodermes, fait fortuitement la découverte dans le corps des organismes qu'il étudie de cellules capables de migrer pour « dévorer » les microbes lors d'une infection : les phagocytes. Il établit ainsi la première théorie de l'immunité cellulaire et obtient le prix Nobel de médecine en 1908. Ensuite, en 1890, le médecin prussien Emil Adolf von Behring (1854-1917, prix Nobel de médecine en 1901) et le japonais Kitasato Shibasaburō ¹⁵ (1853 – 1931) publient les premiers travaux sur les anticorps, ajoutant la théorie de l'immunité humorale à la discipline naissante de l'immunologie (Berche, 2007). Alors que la variolisation puis la vaccination sont utilisées depuis des siècles, la science commence à comprendre la manière dont ces procédés permettent de protéger contre les maladies.

Au cours de ses expériences sur les anticorps, von Behring s'intéresse particulièrement à la diphtérie. Il cherche, sans succès, à produire un vaccin à partir de la toxine diphtérique mais celle-ci est trop dangereuse pour être utilisée pour produire une immunité chez l'humain. Il parvient cependant, à partir d'injections de doses croissantes de toxine chez le mouton, à produire un sérum à faible capacité immunogène (il ne protège pas plus longtemps que deux semaines) mais hautement efficace pour contrer les effets de la toxine chez les malades, réduisant de manière drastique la mortalité (Berche, 2007).

C'est ensuite au tour de Gaston Ramon (1886 – 1963) de tenter la production d'un vaccin antidiphtérique. Comme la toxine est trop dangereuse pour être administrée, il cherche un moyen de la rendre inoffensive. La solution est finalement similaire à celle qu'utilisait Pasteur pour atténuer certaines de ses bactéries : de la chaleur et de la chimie, en l'occurrence l'utilisation de formol. Ramon produit ainsi la première anatoxine, ouvrant la voie à nombre de nouveaux

14. Découvreur du bacille de la peste, qui porte désormais son nom : *Yersinia pestis*.

15. Il n'obtint pas le prix Nobel, mais, petite consolation, son portrait apparaît sur les nouveaux billets japonais de 1 000 yens mis en circulation en juillet 2024.

vaccins dont les vaccins antitétaniques (Berche, 2007).

À ce moment de l'histoire, les deux grands types de vaccins ont été découverts : les vaccins cellulaires, composés de pathogènes atténués ou tués, et les vaccins moléculaires ou acellulaires, composés de certains éléments de ces pathogènes. L'innovation continuera pour ces derniers, avec notamment l'arrivée des vaccins à ARN messager (ARNm), rendus célèbres pour leur utilisation au cours de la pandémie de Covid-19 (Verbeke *et al.*, 2019). Les différents types de vaccins ont tous des avantages et des inconvénients, notamment en termes de réactogénicité, de coût et de capacité de production, comme je le présente dans la section suivante avec l'exemple des vaccins anticoquelucheux.

1.2.3 – Histoire des vaccins anticoquelucheux

L'histoire des vaccins anticoquelucheux débute en 1906 avec la première mise en culture de *Bordetella pertussis* par Jules Bordet et Gustave Gengou (Bordet et Gengou, 1906). Les deux bactériologistes belges vont développer un proto-vaccin en immunisant des animaux avec *B. pertussis* pour produire un sérum qui sera ensuite injecté à des patients. Néanmoins, cette première tentative de vaccin sera un échec, le sérum ne parvenant pas à protéger les patients des effets délétères de l'infection (Madsen, 1925).

Au printemps 1913, à l'occasion d'une épidémie de coqueluche survenue à Tunis, le célèbre médecin Charles Nicolle, habitué de la production de vaccins, développe un vaccin anticoquelucheux dans un but thérapeutique, c'est-à-dire pour soigner des individus déjà malades. Les résultats sont plutôt mauvais, malgré un nombre parfois élevé d'inoculations – jusqu'à neuf! – puisque seulement le tiers des individus inoculés guérissent (Nicolle et Conon, 1913).

Le vaccin tel que développé par Nicole va ensuite commencer à être utilisé comme moyen de protection prophylactique, notamment en Amérique du Nord, en Angleterre et en Italie (Madsen, 1925). Plusieurs pays, comme les États-Unis en 1914 et le Danemark en 1916, vont se donner les moyens de produire à grande échelle le vaccin anticoquelucheux. En 1923/1924 et en 1929, deux épidémies de coqueluche frappent les Îles Féroé. À cette occasion, le médecin danois Thorvald Madsen¹⁶ va lancer un programme de vaccination de la population de l'archipel. Il en conclura

16. Souvent crédité comme le découvreur du vaccin anticoquelucheux en 1923, on voit qu'en réalité de nombreux acteurs ont contribué à développer le vaccin au cours des premières décennies du XX^e siècle.

que le vaccin anticoquelucheux a en réalité une faible efficacité prophylactique mais une (relativement) bonne efficacité thérapeutique, notamment dans la réduction de la mortalité induite par la maladie (Madsen, 1933).

La prochaine grande étape de l'histoire des vaccins anticoquelucheux a lieu aux États-Unis dans les années 1930. Alors que plusieurs vaccins sont commercialisés depuis des années, aucun ne semble en réalité satisfaisant (Council on Pharmacy and Chemistry, 1931). En 1932, deux chercheuses américaines, Pearl Kendrick et Grace Eldering, commencent un programme de recherche pour développer un nouveau vaccin qui soit à la fois sûr et efficace. Entre 1934 et 1938 les deux chercheuses mettent en place une série d'essais cliniques à large échelle pour tester un vaccin et vérifier son efficacité prophylactique, le dosage nécessaire, ainsi que la possibilité de le conjuguer avec un vaccin antidiphtérique (Kendrick *et al.*, 1939; Kendrick, 1940; Kendrick et Weiss, 1942; Kendrick, 1943). Leur vaccin étant un succès, il sera approuvé et mis sur le marché aux États-Unis, où il contribuera à réduire dramatiquement la prévalence de la maladie. Par la suite, Kendrick deviendra consultante pour l'OMS dans de nombreux pays pour aider à la mise en place de programmes de vaccination (Shapiro-Shapin, 2010). Les vaccins contre la coqueluche entrent dans les calendriers vaccinaux de nombreux pays dans les années 1940 et 1950.

Le vaccin développé par Kendrick et Eldering consiste en une solution de *B. pertussis* tuées chimiquement (Kendrick *et al.*, 1939), ce qui le fait entrer dans la catégorie des vaccins à germes entiers, ou WCV (pour *whole-cell vaccine*). Au cours des années 1960 et 1970, les populations et les autorités médicales de plusieurs pays commencent à s'inquiéter des effets secondaires de ces vaccins. Ces effets secondaires sont déjà remarqués par Madsen, qui cite la mort de deux nourrissons consécutive à une deuxième injection de vaccin (Madsen, 1933). Au cours de leur premier essai clinique, Kendrick et Eldering vont notifier un cas de convulsion post-vaccination ainsi que deux cas de forte fièvre et de vomissements (Kendrick *et al.*, 1939).

L'un des effets secondaires les plus discutés de la vaccination anticoquelucheuse est le risque d'encéphalopathie. Dès 1958, une revue de la littérature fait état d'une centaine de cas d'encéphalopathies post-vaccinales, dont les conséquences sur le long terme ont parfois été dramatiques : convulsions récurrentes, hémiplégie, perte des facultés cognitives, et dans certains cas décès de l'enfant (Cockburn, 1958). Dans les deux décennies suivantes, de nombreux auteurs vont essayer de quantifier la fréquence des accidents et de lister les risques post-vaccinaux. Stewart (1977)

avance une incidence des encéphalopathies de 1/54 000 enfants complètement vaccinés (trois doses) à Glasgow en Écosse, en précisant que cette incidence est certainement sous-estimée. [Ström \(1960, 1967\)](#) avance une incidence de 1/6 000 en Suède sur la période 1955-1958 et de 1/3 600 sur la période 1959-1965, avec une augmentation progressive de l'incidence sur cette seconde période¹⁷. D'autres auteurs encore font état d'encéphalopathies suivant directement une inoculation de vaccin anticoquelucheux (*e.g.* [Kulenkampff et al., 1974](#); [Ehregut, 1978](#)). Finalement, les inquiétudes croissantes vont mener à une réduction drastique de la couverture vaccinale, comme au Royaume-Uni dans les années 1970, voire un abandon total de la vaccination, comme en Suède de 1979 à 1986, ou au Japon.

En décembre 1974 et janvier 1975, deux décès suivant l'injection de vaccin conduisent le ministère de la Santé japonais à arrêter l'administration du vaccin en février 1975. Le vaccin est finalement à nouveau utilisé dès avril de la même année, mais avec un âge de primo-vaccination décalé à deux ans. Cependant, les populations et professionnels de santé refusent en masse la vaccination, conduisant à une diminution drastique de la couverture vaccinale et à une augmentation des cas. C'est alors que les autorités du pays lancent un programme de développement d'un nouveau vaccin, moins dangereux, dans le but de regagner la confiance du public et de remonter la couverture vaccinale. En 1981, les recherches aboutissent sur l'autorisation de mise sur le marché d'un nouveau type de vaccin, premier de sa génération : un vaccin acellulaire, ou ACV (pour *acellular vaccine*) ([Kimura et Kuno-Sakai, 1988](#); [Nakayama, 2013](#)). Ces nouveaux vaccins seront ensuite adoptés dans la majorité des pays à fort revenu dans les années 1990 et 2000.

Le principe des vaccins acellulaires contre la coqueluche est de n'inclure dans leur formule qu'un nombre réduit d'antigènes de *B. pertussis*, permettant ainsi d'exclure ceux responsables des effets indésirables des vaccins à germe entier. Les ACV en circulation aujourd'hui contiennent un à cinq antigènes : la toxine pertussique (TP) avec possiblement une adhésine (hémagglutinine filamenteuse ou FHA), deux adhésines (FHA et pertactine ou PRN) ou quatre adhésines (FHA, PRN et fimbriae de type 2 et de type 3 ou Fim-2 et Fim-3) ([Guiso et al., 2022](#)). Ces vaccins sont appelés respectivement ACV1, ACV2, ACV3 et ACV5. Quoique les ACV soient plus sûrs que les WCV, ces derniers restent aujourd'hui utilisés dans la majeure partie des pays du monde pour

17. À noter que Ström a été critiqué pour ne pas avoir exclu les cas de prédisposition aux encéphalopathies, surestimant la fréquence des accidents post-vaccination ([Kulenkampff et al., 1974](#)).

une raison simple : ils sont beaucoup moins chers à produire que les ACV, du fait des étapes de purification nécessaires à la production des vaccins acellulaires (Edwards et Decker, 2012).

Comme présenté dans la Section 1.4.3, l'utilisation massive des vaccins, et notamment des ACV, a conduit à l'émergence de souches de plus forte virulence et capables d'échapper au moins partiellement à l'immunité induite par les vaccins. De plus, certains auteurs suggèrent que la durée de l'immunité conférée par les vaccins acellulaires serait fortement réduite par rapport à celle conférée par les vaccins à germe entier (Szwejszer-Zawislak *et al.*, 2022), tandis que d'autres estiment que la durée de l'immunité est comparable (*e.g.* Wendelboe *et al.*, 2005; Domenech de Cellès *et al.*, 2019). Ces différents éléments ont conduit à l'identification d'un besoin : celui de développer une nouvelle génération de vaccins (Poolman, 2014; Rumbo et Hozbor, 2014; Locht, 2018; Dewan *et al.*, 2020). Parmi les solutions proposées, on retrouve le développement d'un vaccin vivant atténué ou l'ajout de nouveaux antigènes aux vaccins acellulaires. Mais à ce jour ces vaccins sont toujours en développement ou en test, et *B. pertussis* continue d'évoluer.

1.3 – Évolution de la virulence

1.3.1 – Définir et mesurer la virulence

Ce que la médecine nomme « maladie » correspond à l'expression de la virulence d'un pathogène, une notion centrale en épidémiologie ainsi qu'en écologie des parasites. Mais le mot « virulence » fait partie d'une longue liste de termes en biologie dont la définition est débattue, imprécise, voire manquante (Encadré 1.3). Une définition générale de la virulence serait « la capacité d'un pathogène à se multiplier et à causer des dommages chez son hôte »¹⁸ (Poulin et Combes, 1999). La virulence se définirait donc à la fois par la valeur sélective du pathogène (sa capacité à se multiplier) et par son impact sur celle de l'hôte (les dommages induits, généralement mesurés au niveau de la survie ou de la reproduction).

Toutefois, cette définition est souvent réduite à sa deuxième partie lorsqu'on étudie la virulence empiriquement ou par modélisation. Ainsi, la virulence se quantifie ou se qualifie par les dommages infligés à l'hôte. Le paradigme général de la recherche est donc que « la virulence est

18. « *the ability of a pathogen to multiply and cause harm to its host* »

un trait du parasite [ou plutôt, de l'interaction hôte-parasite] mais sa valeur est un paramètre démographique de son hôte »¹⁹ (Alizon et van Baalen, 2008a).

Puisque la virulence est – généralement – une valeur quantitative, elle peut être mesurée. Cependant, il existe une multitude de manières de considérer les dommages subis par l'hôte. Dans les travaux de modélisation (voir Section 1.5.3), on considère généralement les dégâts par le prisme d'un excédent de mortalité (*e.g.*, Levin et Pimentel, 1981; Lipsitch et Nowak, 1995; Day et Proulx, 2004), quoique d'autres métriques puissent être considérées comme la quantité de particules virales produites (Berngruber *et al.*, 2013). Dans les travaux expérimentaux ou observatoires, il existe une plus grande diversité des mesures de la virulence, comme la charge virale/parasitaire (*e.g.*, Locher *et al.*, 2003; Spence *et al.*, 2013; Cornet *et al.*, 2014; Owolabi *et al.*, 2024), la densité des globules rouges de l'hôte, (*e.g.*, Ferguson *et al.*, 2003; Spence *et al.*, 2013; Barclay *et al.*, 2014; Cornet *et al.*, 2014; Owolabi *et al.*, 2024), sa réduction de poids (*e.g.*, Ferguson *et al.*, 2003; Spence *et al.*, 2013; Cornet *et al.*, 2014; Owolabi *et al.*, 2024) ou, simplement, son taux de mortalité associée à l'infection (*e.g.*, Ferguson *et al.*, 2003).

ENCADRÉ 1.3 – Définitions en sciences de la biodiversité

La biologie est une science de définitions approximatives. Ceci est particulièrement vrai pour les sciences de la biodiversité (écologie et biologie de l'évolution). Prenons l'exemple de la définition de la vie sur le Wiktionary, pour comprendre le problème. La vie y est définie comme « le statut possédé par un nombre quelconque d'entités [...] ayant les propriétés de réplication et de métabolisme »^a. La réplication y est définie comme étant « le processus de production de répliques de molécules d'ADN ou d'ARN »^b. Nous avons déjà là un premier problème à la définition de la vie, qui la confinerait à des structures à base d'ADN ou d'ARN. Mais le vrai problème est le métabolisme, défini comme « les processus chimiques qui se produisent au sein d'un organisme vivant pour maintenir la vie »^c. Est donc vivante une structure... capable de se maintenir en vie.

Évidemment, cette définition parfaitement circulaire de la vie provient d'un dictionnaire « grand public », il ne s'agit pas d'une proposition de définition scientifique. De nombreux auteurs scientifiques ont proposé des définitions à la vie, et peut-être plus encore ont débattu de la pertinence des définitions proposées. Une définition scientifique courante de la vie, celle utilisée par la NASA dans sa recherche de vie extraterrestre, est que la vie est « un système autoentretenu capable d'évolution darwinienne »^d.

19. « *the virulence is a trait of the parasite, but its value is a demographic parameter of its host* »

Les problèmes de cette définition sont cependant multiples. Le premier est que cette définition est centrée sur ce que l'on connaît de la vie *ici* et *maintenant*, et pourrait ainsi être trop contraignante pour la découverte de formes de vie extraterrestre, ou encore pour identifier quand et comment la vie est apparue sur Terre (Luisi, 1998). Un autre problème de cette définition est qu'elle est fondamentalement population-centrée, et ne permet pas d'identifier si un individu unique est vivant ou non. La non-reconnaissance de ce deuxième point a conduit certains auteurs à faire émerger le « problème de la mule » (Chodasewicz, 2014). Voici son énoncé sous la forme d'un syllogisme : la vie se définit par l'évolution darwinienne, or l'évolution darwinienne implique la reproduction, donc ce qui ne peut pas se reproduire - par exemple, une mule - n'est pas vivant. On remarque cependant que la mule est issue de populations capables de reproduction, et est donc vivante. Félicitations à elle.

Quoique soluble, le problème de la mule illustre à quel point les définitions du vivant sont débattables et sujettes à caution. Prenons un exemple inverse et donnons une définition trop large à la vie. Définissez la vie comme un système chimique qui convertit de la matière en énergie pour se multiplier et vous voilà à passer vos pauses méridiennes à expliquer à vos collègues incrédules que le feu est un organisme vivant.

J'ai utilisé la définition de la vie comme exemple, mais de nombreuses autres définitions posent des difficultés en biologie, comme celle d'espèce (Mallet, 1995), de conscience (Seth et Bayne, 2022), de taille efficace de population (Waples, 2022), de communauté écologique (Looijen et van Andel, 1999)... Parfois, le problème n'est pas de *produire* une définition, mais de la *comprendre*, notamment dans l'enseignement des sciences de la vie (Zukswert *et al.*, 2019 ; Wandersee, 1988).

Toutefois, le problème des définitions variables ou imprécises en est-il vraiment un ? Certains auteurs estiment que l'absence de définition permet à la biologie de progresser. Ainsi, ne pas définir la vie permettrait plus de liberté de recherche, notamment dans les disciplines émergentes de l'exobiologie, de l'étude de systèmes de vie artificielle, ou encore de la recherche de l'origine terrestre de la vie (Cleland, 2012). Finalement, la quête de définitions universelles serait vaine. Ce qui est nécessaire en recherche est plutôt d'avoir des termes correctement définis *au sein de chaque projet de recherche*. Si par exemple un projet de recherche s'intéresse aux caractéristiques du vivant, ce projet doit être capable de produire une définition du vivant cohérente pour ce projet (Hodges, 2008).

a. « *the status possessed by any of a number of entities, [...] which have the properties of replication and metabolism* », disponible sur <https://en.wiktionary.org/wiki/life> (dernier accès le 04/06/2024).

b. « *the process of producing replicas of DNA or RNA molecules* », disponible sur <https://en.wiktionary.org/wiki/replication> (dernier accès le 04/06/2024).

c. « *the chemical processes that occur within a living organism in order to maintain life* », disponible sur <https://en.wiktionary.org/wiki/metabolism> (dernier accès le 04/06/2024).

d. « *a self-sustaining chemical system capable of Darwinian evolution* », disponible sur <https://astrobiology.nasa.gov/research/life-detection/about/> (dernier accès le 04/06/2024).

1.3.2 – Théories et observations sur l'évolution de la virulence

Sélection naturelle chez les pathogènes

Depuis la publication de *L'Origine des espèces* par Charles Darwin (1809 – 1882), une idée fondamentale sert de clé de voûte à la théorie de la sélection naturelle : ce sont les caractères des individus qui se reproduisent le plus qui sont sélectionnés. Autrement dit, la sélection naturelle favorise les individus maximisant la taille de leur descendance. Certes, on comprend facilement qu'un individu ne peut pas avoir un nombre illimité de descendants, puisque les ressources ne le sont pas et que la durée pendant laquelle l'individu peut se reproduire est limitée, au moins par sa durée de vie. Des stratégies sont donc apparues au cours de l'évolution pour faire face à ces contraintes de ressources et de temps.

Lorsque Darwin publie ses travaux, il pense d'abord aux animaux comme étant les entités soumises à la sélection naturelle. À cette époque, la science ignore tout des germes. Pourtant, la sélection naturelle s'applique également à ces organismes microscopiques. Pour un pathogène bactérien ou viral, discuter de descendance est complexe. Un pathogène peut produire une progéniture importante, mais si celle-ci n'est pas transmise à un nouvel hôte alors le pathogène sera voué à voir ses gènes disparaître. C'est pour cela que la valeur sélective d'un pathogène est généralement mesurée à travers son nombre de reproduction de base, noté \mathcal{R}_0 . Quoique sa définition exacte soit plus rigoureuse, j'utiliserai pour l'instant une définition simple du \mathcal{R}_0 : le nombre d'individus auquel l'hôte d'un pathogène va transmettre celui-ci au cours de son infection. Ainsi, si les animaux sont sélectionnés pour maximiser leur nombre de descendants, les pathogènes sont sélectionnés pour maximiser leur \mathcal{R}_0 – au moins dans les modèles hôte-pathogène simples, négligeant notamment l'hétérogénéité au sein de la population hôte. En effet, il existe plusieurs situations dans lesquelles une souche pathogène avec un plus faible \mathcal{R}_0 peut être sélectionnée (Kao, 2006 ; Roberts, 2007 ; Lion et Metz, 2018).

Conventional wisdom

Des années 1930 aux années 1970, la théorie la plus largement acceptée sur l'évolution de la virulence des pathogènes est celle que May et Anderson (1990) nomment la *conventional wisdom*. Cette théorie présente la virulence comme un artefact de la relation hôte-pathogène,

puisque celle-ci nuit aux deux parties. L'idée derrière cette théorie est simplement que si le pathogène tue son hôte, alors ce dernier ne pourra plus transmettre le pathogène, diminuant alors mécaniquement son \mathcal{R}_0 par rapport à une situation où le pathogène laisserait son hôte survivre indéfiniment. La *conventional wisdom* stipule ainsi que les systèmes hôte-pathogène doivent nécessairement converger vers une forme de commensalisme tandis que les acteurs en interaction coévoluent : l'hôte développe une résistance au pathogène alors que le parasite devient moins virulent (Johnson, 1986). La virulence observée aujourd'hui chez de nombreux pathogènes notamment humains serait alors due à la récente émergence de ceux-ci chez notre espèce et au manque de temps qu'auraient eues les deux espèces pour s'adapter l'une à l'autre.

Cette théorie du commensalisme comme finalité évolutive repose sur les observations de plusieurs maladies, comme la maladie de Lyme ou la légionellose, provoquées par des bactéries circulant sans effet négatif ou presque chez des espèces animales non humaines, et dont l'humain ne représente qu'un hôte « accidentel » voire indésirable, finalement un cul-de-sac épidémiologique et évolutif (Johnson, 1986 ; Levin, 1996). Dans son célèbre article « *The Units of Selection* », Richard Lewontin utilise l'exemple de la myxomatose pour expliquer comment la sélection à l'échelle populationnelle justifier la sélection de pathogènes avirulents (Lewontin, 1970). Cependant, cette théorie d'une sélection mécanique vers l'avirulence est aujourd'hui mise au ban, et quoique les arguments de Lewontin soient logiques, nous verrons dans la Section 1.3.4 que le virus de la myxomatose n'a pas évolué vers une avirulence, mais vers une virulence modérée.

Trade-off theory

L'évolution des pathogènes vers une virulence modérée est mise en lumière dans les années 1980. Elle serait le résultat d'une relation entre la capacité de reproduction du pathogène et les dommages subis par l'hôte associés à cette reproduction (Anderson et May, 1982 ; May et Anderson, 1990). Ainsi, un parasite se reproduisant beaucoup se transmettra beaucoup, mais incapacitera son hôte ou le tuera rapidement, tandis qu'un parasite se reproduisant peu ne se transmettra presque pas mais ne handicapera que peu son hôte.

Cette dualité trouve son origine dans le double niveau de sélection auquel sont soumis les parasites, et surtout les microparasites. Le premier niveau est l'intra-hôte. Au sein d'un hôte, la souche parasitaire favorisée est celle qui se multiplie le plus, comme cela a pu être montré ex-

périmentalement chez *Plasmodium chabaudi*, protozoaire responsable de la malaria des rongeurs (de Roode *et al.*, 2005). La virulence du parasite n'a alors pas d'incidence dans le processus de sélection. Le deuxième niveau est l'inter-hôte. Pour continuer à se multiplier, le parasite doit finalement se transmettre à un nouvel individu. Or, s'il se multiplie trop au sein de son hôte précédent, le parasite risque de le tuer ou de le handicaper, réduisant ses chances de transmission. À ce niveau, la virulence devient un facteur important. La théorie du compromis entre sélection intra-hôte et sélection inter-hôte, ou *trade-off theory*, stipule ainsi qu'il doit exister une virulence optimale, intermédiaire, maximisant la transmission d'un pathogène. Anderson et May (1982) reprennent l'exemple de la myxomatose pour illustrer cette théorie, s'opposant ainsi à Lewontin (1970) avec le même exemple.

En plus de produire un modèle mathématique (nous y reviendrons), Anderson et May (1982) s'appuient sur quatre observations pour défendre leur théorie du *trade-off* en opposition à la *conventional wisdom* :

- La transmission est souvent corrélée avec la virulence des parasites.
- Le parasitisme est un mode de vie extrêmement commun, là où le commensalisme (prédit comme finalité évolutive par la *conventional wisdom*) est rare. Les estimations de la proportion d'espèces parasites varient selon la définition donnée au parasitisme et les taxa considérés, mais la moitié de toutes les espèces vivantes seraient des parasites (Windsor, 1998 ; Poulin et Morand, 2000 ; Dobson *et al.*, 2008 ; Poulin, 2014).
- Il existe un polymorphisme extrêmement fort pour les gènes impliqués dans la réaction anti-parasitaire chez les populations d'hôtes. Ce polymorphisme serait lié à une coévolution ancienne et continue entre les hôtes et les parasites, à mettre en lien avec la théorie de la Reine Rouge (Van Valen, 1973).
- Le système immunitaire est une machinerie tellement complexe qu'il est plus que vraisemblable que les relations avec les parasites ont une incidence majeure sur la *fitness* des espèces animales.

Dans sa formulation originale, la théorie du *trade-off* stipule un compromis entre le taux de transmission d'un pathogène et sa virulence – quantifiée par la mortalité induite par celui-ci (Figure 1.2A). Le résultat est qu'il existe une virulence optimale pour laquelle la reproduction du pathogène (son \mathcal{R}_0) est maximisée (Figure 1.2B). Néanmoins, il est possible d'envisager d'autres

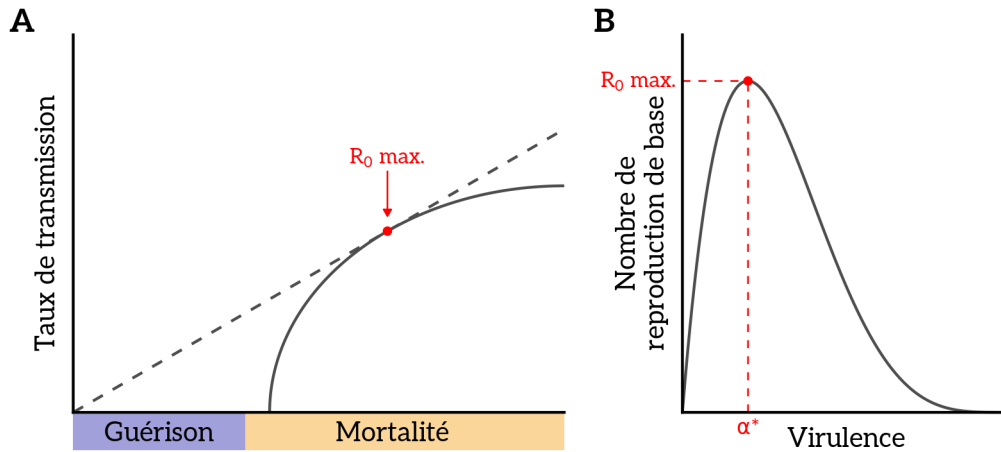


FIGURE 1.2 – **Illustration du principe de la théorie du compromis entre virulence et transmission.** (A) Exemple d’une forme que peut prendre le compromis (*trade-off*) entre le taux de transmission du parasite et sa virulence. Le nombre de reproduction de base (\mathcal{R}_0) est le produit du taux de transmission et de la durée de l’infection, cette dernière étant l’inverse de la somme de la virulence (ici résumée à la mortalité) et du taux de guérison. La valeur maximale du \mathcal{R}_0 est égale à la pente de la droite passant par l’origine et tangentielle à la courbe du *trade-off*. On remarque que si on continue à augmenter la mortalité, le taux de transmission augmente mais pas le nombre de reproduction de base. (B) Exemple de relation entre le nombre de reproduction de base et la virulence d’un parasite. Comme attendu, le nombre de reproduction maximal est atteint pour une virulence optimale (α^*) non nulle. Panel A adapté de la Figure 1 de [Bull et Lauring \(2014\)](#) et de la Figure 1 de [Bull et Antia \(2022\)](#) et panel B adapté de la Figure 2 de [Anderson et May \(1982\)](#).

compromis, par exemple entre le taux de transmission et le taux de guérison. Il s’agit donc plus d’un cadre de réflexion qu’une théorie universelle, et il est possible que d’autres facteurs – comme l’immunité de l’hôte ou les ressources à disposition – limitent à la fois la transmission et la virulence des pathogènes ([Bull et Lauring, 2014](#)).

Il est intéressant de noter que certains auteurs nuancent le lien théorisé par Anderson et May entre taux de transmission et virulence. Au cours d’une expérience, [Tardy *et al.* \(2019\)](#) ont infecté des roselins familiers (*Haemorrhous mexicanus*) avec des souches de *Mycoplasma gal-lisepticum*, bactérie responsable de la mycoplasmosse aviaire, collectées entre 1994 et 2015. En utilisant différentes mesures de virulence (baisse de la masse des oiseaux, sévérité de l’infection et taux de mortalité), les auteurs ont mis en évidence une hausse de la virulence des souches au fil des années, sans que cette hausse soit associée à une augmentation du taux de réplication des

souches testées.

Rôle du mode de transmission

Johnson (1986) et Ewald (1995) font le postulat que l'expression du compromis entre transmission du pathogène et virulence dépend de la nature de la relation entre l'hôte et le parasite et du mode de transmission du pathogène (Figure 1.3). Plus généralement, la formulation classique de la théorie du *trade-off* ne fonctionnerait que pour les pathogènes transmis directement (Alizon et Michalakis, 2015). Par exemple, le tétanos est une maladie à forte virulence – son taux de mortalité est estimé entre 35,0 et 54,3 % selon les pays (Li *et al.*, 2023) – or la bactérie responsable, *Clostridium tetani*, est capable de survivre et se reproduire hors de l'hôte. Dans cet exemple, la relation entre l'hôte et le pathogène est donc accidentelle, et il n'y a pas de pression de sélection favorisant des souches de *C. tetani* de virulence plus tempérée. Il existe également des maladies, comme la peste ou le kuru²⁰ pour lesquelles le décès de l'hôte peut faciliter la transmission du pathogène, favorisant ainsi les souches de forte virulence.

De la même manière, Ewald (2004a) propose que les maladies transmises par un vecteur évoluent vers une plus forte virulence puisque l'immobilisation de l'hôte n'est pas un frein à la transmission du pathogène. Plusieurs études portant sur *P. chabaudi* ont montré que les clones naturellement de forte virulence (comprendre : virulence pour le rongeur, pas pour le moustique) se transmettent mieux aux moustiques que les clones de faible virulence (Mackinnon et Read, 1999 ; Ferguson *et al.*, 2003 ; Spence *et al.*, 2012). Cependant, la formation de clones hypervirulents par passages sanguins (voir Section 1.3.3) conduit, lorsque la transmission par moustique est rétablie, à la sélection de clones de virulence atténuée (Spence *et al.*, 2012, 2015), un résultat qui coïnciderait avec la théorie du *trade-off*. Ce phénomène pourrait s'expliquer par le fait que les moustiques ne piquent que des hôtes vivants. Finalement, l'existence d'un *trade-off* entre transmission et virulence chez les maladies à transmission vectorielle reste floue et pourrait, comme souvent en biologie, demander une étude au cas par cas. Une revue sur les maladies végétales à transmission vectorielle estime que les indices pour l'existence d'un *trade-off* dans de tels cas sont minces (Froissart *et al.*, 2010).

20. Le kuru était une encéphalopathie, provoquée par un prion et généralement associée au cannibalisme dans les populations Fore de Nouvelle-Guinée.

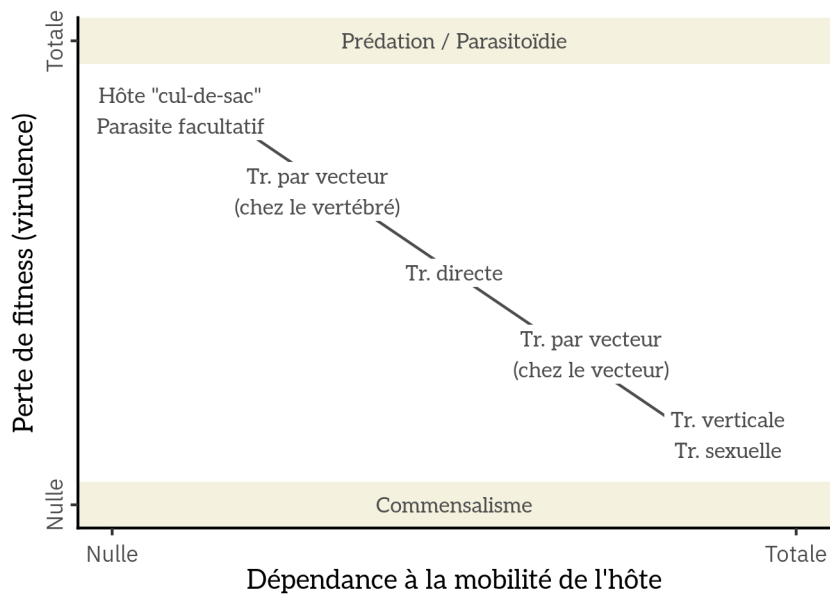


FIGURE 1.3 – **Relation entre la perte de fitness induite par le parasite et la dépendance à la mobilité de l'hôte pour sa transmission.** Plus un parasite a besoin que son hôte conserve sa mobilité pour assurer sa transmission, plus la virulence du parasite doit être basse. Les positions des catégories de parasitisme sont indicatives d'une prédiction moyenne et des parasites appartenant à une catégorie peuvent en réalité se retrouver ailleurs sur le gradient. Les axes ne doivent pas être interprétés comme étant linéaires. Adapté de la Figure 1 de [Ewald \(2004a\)](#). Tr. : transmission.

A l'inverse, un parasite nécessitant un hôte vivant pour se transmettre devrait subir des pressions favorisant la sélection de souches de plus faible virulence – mais pas nécessairement des souches avirulentes. C'est par exemple le cas de la myxomatose introduite en Australie, dont je discute plus loin dans la Section 1.3.4, ou des infections à *Chlamydia pecorum* chez différents mammifères ([Mohamad et al., 2014](#)). Le cas des maladies à transmission sexuelle est particulièrement intéressant : pour que le pathogène puisse se transmettre, il est nécessaire que l'hôte soit en suffisamment bonne santé pour s'engager dans une activité sexuelle et, s'il appartient à une espèce avec un fort choix du partenaire, que ses symptômes soient suffisamment discrets pour ne pas être repoussé par le partenaire potentiel. De plus, il doit maintenir son hôte en bonne santé suffisamment longtemps pour que ce dernier ait l'occasion d'avoir une activité sexuelle car, pour citer [Ewald \(2004a\)](#), « une personne a généralement des rapports sexuels avec beaucoup moins de personnes par semaine que sur lesquelles elle peut éternuer »²¹. La situation du VIH

21. « a person generally has sex with many fewer people per week than he can sneeze on »

au Sénégal dans les années 1980 et 1990 illustre cette sélection pour une plus faible virulence lorsque les opportunités de transmission sont réduites. Dans ce pays, le sous-type de VIH-1 circulant le plus est le A et ce sous-type est moins virulent que les autres en circulation en Afrique de l'Ouest, réduisant d'un facteur huit le risque de contracter le SIDA par rapport aux autres sous-types (Kanki *et al.*, 1999). Cette circulation privilégiée du sous-type A par rapport à tout autre sous-type a été reliée aux efforts de prévention mis en place dans le pays pour lutter contre la transmission du VIH et donc à sa faible transmission (Meda *et al.*, 1999). Il est cependant important de nuancer ce résultat. Ewald (2004a) écrit, en citant Toure-Kane *et al.* (2000), qu'au Sénégal le type VIH-1 n'a pas été remplacé par le type VIH-2, reconnu moins virulent et moins transmissible (Marlink, 1996; Nyamweya *et al.*, 2013); or Toure-Kane *et al.* (2000) les auteurs précisent à l'inverse que le VIH-2 a récemment été remplacé par le VIH-1 au Sénégal²². La situation du VIH au Sénégal n'est donc pas non plus totalement claire.

De l'écologie à l'évolution des parasites

Quoique la notion d'évolution de la virulence se comprenne comme un processus long et multigénérationnel (du point de vue du pathogène), il est également intéressant d'étudier les facteurs écologiques modulant la virulence au temps présent. Ces facteurs peuvent contribuer à définir la stratégie évolutive adoptée par les parasites et leurs hôtes. Par exemple, il a été théorisé que la qualité des hôtes joue un rôle sur la capacité d'infection et de reproduction des parasites, ainsi que sur leur virulence. La grande limnée (*Lymnaea stagnalis*) est un escargot d'eau douce parasité par le trématode *Diplostomum spathaceum*. Il a été montré que lorsque l'hôte est affamé, le parasite est plus virulent, induisant une plus forte mortalité chez son hôte (Seppälä *et al.*, 2008). L'idée générale derrière ce phénomène est qu'un hôte de plus faible qualité ou en état de stress sera moins capable de résister aux dommages induits par le parasite. Dans une expérience comparable portant sur des serins des Canaries (*Serinus canaria*) infectés par l'agent étiologique du paludisme aviaire, *P. relictum*, les auteurs ont obtenu des résultats similaires : les individus moins nourris résistaient moins à l'infection, ce qui implique une virulence réalisée plus forte des parasites (Cornet *et al.*, 2014).

22. Marlink (1996) montre même que le Sénégal est en réalité l'un des pays d'Afrique de l'Ouest avec la plus faible séroprévalence du VIH-2, tandis que la séroprévalence du VIH-1 serait plus élevée qu'estimée par Meda *et al.* (1999). Malheureusement, des données récentes manquent pour une conclusion définitive.

D'autres facteurs écologiques comme la compétition entre souches peuvent également influencer l'évolution de la virulence des pathogènes. Comme expliqué plus tôt, la théorie du *trade-off* implique que les souches de plus forte virulence se répliquent plus au sein de leur hôte. Lors d'une infection multiple, il est donc attendu que la compétition favorise les souches de plus forte virulence (van Baalen et Sabelis, 1995). Plusieurs études expérimentales donnent du crédit à cette théorie, mais d'autres la contredisent, laissant supposer que la compétition intra-hôte est un phénomène finalement complexe qui implique de multiples acteurs et notamment le système immunitaire de l'hôte (Alizon et van Baalen, 2008a).

Le système immunitaire de l'hôte est justement un des acteurs majeurs de l'évolution des pathogènes. Par exemple, Wallace (1989) propose que les hôtes immunodéficients pourraient représenter des portes d'entrée pour de nouveaux pathogènes, mais pourraient également contribuer à l'évolution de pathogènes circulant déjà au sein d'une population. De leur côté, Gandon *et al.* (2001) proposent que les vaccins stimulant une immunité bloquant l'effet de la virulence des pathogènes sur l'hôte (*i.e.*, l'augmentation de la mortalité) favorisent l'émergence de souches de plus forte virulence. Cela s'est vérifié dans le cas de la maladie de Marek (Section 1.4.1) mais pas dans le cas de la diphtérie (Section 1.4.2).

1.3.3 – Manipulation expérimentale de la virulence

Pasteur est le premier à expérimenter sur la virulence de pathogènes. Comme souvent en science, ses premières découvertes sur le sujet se font par sérendipité : il observe que des cultures de *Pasteurella multocida* oubliées au laboratoire pendant ses vacances d'été sont devenues moins virulentes²³. Pasteur mène ensuite d'autres expériences sur différents pathogènes dont il parvient à augmenter ou diminuer la virulence en faisant passer les pathogènes d'un hôte à l'autre (Berche, 2007).

Dans les décennies suivantes, de nombreuses expériences sont menées au cours desquelles un pathogène est transmis artificiellement d'un hôte au suivant par l'expérimentateur, une méthode appelée « passage ». Cette transmission artificielle altère la balance coût-bénéfice de la virulence puisque le pathogène n'a plus besoin de maintenir son hôte dans un état de santé suffisant pour

23. De la même manière, en 1928, Alexander Fleming découvre la pénicilline après avoir oublié des cultures bactériennes pendant ses vacances. Ainsi, la science progresse en manquant d'organisation et en partant en vacances.

qu'il puisse de lui-même transmettre le pathogène. Ainsi, dans ces expériences, le coût à tuer son hôte est rendu nul et, dans plusieurs systèmes, ce type d'expérience conduit à la sélection de souches de plus forte virulence (Ebert, 1998). C'est par exemple ce qui a été observé dans un système où des babouins (*Papio cynocephalus*) ont été infectés par le VIH-2 (Locher *et al.*, 2003). Grâce à ce type de manipulation, Barclay *et al.* (2014) ont pu vérifier la théorie de Wallace (1989) sur l'impact de l'état de santé immunitaire de l'hôte sur l'évolution des pathogènes : le passage du parasite *P. chabaudi* entre souris immunodéficientes conduit à l'évolution vers une plus forte virulence que le passage entre hôtes immunocompétents. Il est intéressant de noter qu'en plus du coût nul à la transmission, la diversité génétique des hôtes dans ces expériences de passage est généralement faible, créant une pression uniforme pour sélectionner un seul phénotype et génotype du pathogène (Ebert, 1998).

Cependant, il est intéressant de noter que Jansen *et al.* (2015) ont observé, dans un système mêlant l'hôte nématode *Caenorhabditis elegans* et la bactérie pathogène *Pseudomonas aeruginosa*, une sélection pour des souches commensales de la bactérie plutôt que des souches de forte virulence. Quoique cette étude soit plutôt un cas à part, elle laisse supposer que la direction de l'évolution de la virulence pourrait dépendre du système hôte-pathogène étudié – même indépendamment du mode de transmission du pathogène – et appelle à de plus amples recherches à ce sujet.

Pasteur, en plus de produire des souches de plus forte virulence, réussit également à réduire la virulence de plusieurs pathogènes par la méthode du passage. Il le fait notamment avec la bactérie responsable du rouget du porc (*Erysipelothrix rhusiopathiae*). La différence principale avec l'expérience de Jansen *et al.* (2015) est que Pasteur fait passer la bactérie par un nouvel hôte, le lapin. Pasteur ne le savait pas à l'époque, mais en changeant d'hôte le pathogène acquiert des mutations facilitant sa reproduction chez le nouvel hôte mais le rendant moins adapté à son hôte originel, ce qui diminue sa virulence chez celui-ci (Berche, 2007).

En plus du passage, d'autres méthodes existent pour modifier artificiellement la virulence d'un pathogène. L'une d'elles est la surexpression/inhibition (*knockout*) de gènes. Fu *et al.* (2008) ont manipulé l'expression de plusieurs gènes de la levure *Candida albicans* et ont ainsi pu faire varier la virulence de ce pathogène responsable de la candidose²⁴. De la même manière, Scorpio

24. L'objectif des auteurs était d'identifier les « gènes de virulence » de *Candida albicans*, pas de produire des

et al. (2010) ont augmenté l'expression du gène codant la capsule dépolymérase de *B. anthracis*, bactérie responsable de la maladie du charbon, et ainsi observé une diminution de la virulence de la bactérie. Il est également possible de directement ajouter des facteurs de virulence dans le génome de pathogènes. *Harjai et al.* (1994) ont introduit un plasmide contenant des gènes incomplets codant un facteur de virulence – en l'occurrence, un sidérophore – chez *Escherichia coli* et ont observé une augmentation de la cytotoxicité de la souche mutante.

Aujourd'hui, les expériences de manipulation de la virulence, et plus globalement les expériences de modification de pathogènes, font polémique. En juin 2012, *Science* publie un article dans lequel les auteurs rapportent avoir produit, par la technique du passage, un variant du virus de la grippe aviaire H5N1 capable de se transmettre entre mammifères, en l'occurrence des furets (*Herfst et al.*, 2012). Il ne s'agit pas directement d'une expérience de manipulation de la virulence – d'ailleurs les auteurs n'ont pas quantifié la virulence de leur souche mutée – cependant il est évident que si une souche de H5N1 transmissible entre humains devait émerger, elle ferait beaucoup plus de dégâts qu'une souche uniquement transmissible de l'oiseau à l'humain, fût-elle moins virulente.

Cette expérience a été le déclencheur de la polémique mondiale, dont deux camps ont émergé : ceux qui soutiennent que les expériences de manipulation de pathogènes sont essentielles, notamment pour la production de vaccins et pour aider à prévenir de futures pandémies ; et ceux qui estiment que les risques associés à ces recherches – comme les fuites de laboratoires et les attaques bioterroristes – sont trop grands (*Casadevall et Imperiale*, 2014). La récente pandémie de Covid-19 a amplifié ces craintes : le laboratoire P4 de l'Institut de virologie de Wuhan, ville de départ de la pandémie, connu pour ses manipulations sur les coronavirus, s'est retrouvé suspecté d'être à l'origine du virus responsable de millions de morts à travers la planète (*Berche*, 2023).

En plus des manipulations directes et volontaires de la virulence, il existe des cas où les actions humaines ont conduit indirectement à la sélection de souches altérées de la virulence de certains pathogènes. Dans la section suivante, je présente le cas d'école de l'évolution de la virulence : la myxomatose en Australie.

souches de forte virulence.

1.3.4 – Cas d'école : l'exemple de la myxomatose

La myxomatose est une maladie touchant principalement les lagomorphes (lapins et lièvres) causée par le *myxoma virus* (MYXV), un virus à ADN de la famille des poxvirus. Historiquement, le MYXV est un pathogène du tapeti (*Sylvilagus brasiliensis*) et du lapin de Bechmann (*Sylvilagus bachmani*), deux espèces endémiques des Amériques chez qui l'infection est bénigne (Regnery et Marshall, 1971).

En 1788, le lapin européen (*Oryctolagus cuniculus*) est introduit pour la première fois en Australie par des colons, mais c'est en 1859 qu'une petite population est relâchée dans la nature. À partir de là, le lapin européen envahira en quelques décennies une importante partie du territoire. Dans les années 1930 et 1940, de nombreuses expériences sont menées pour tester le potentiel du MYXV comme moyen de lutte biologique contre le lapin invasif. Ces différentes études montrent que les souches testées sont extrêmement létales pour le lapin européen, atteignant jusqu'à 100 % de mortalité en laboratoire et 99,7 % sur le terrain (Fenner *et al.*, 1953). Cependant, Bull et Mules (1944) estiment que la maladie est un moyen peu efficace pour réduire les populations de lapin, celle-ci se transmettant difficilement d'une colonie à l'autre.

Enfin, au cours de l'été 1950-1951²⁵, de multiples introductions du virus sont effectuées et la maladie parvient à s'établir, notamment grâce à l'acquisition du virus par des insectes vecteurs (mouches et surtout moustiques) qui permettent les contaminations entre colonies et par le transport de lapins morts de la myxomatose à travers le territoire australien (Nichols, 1951 ; Ratcliffe *et al.*, 1952).

Bien que la maladie soit alors en voie de devenir endémique, des doutes sont très rapidement émis sur sa capacité à réduire voire éradiquer le lapin européen d'Australie. Ratcliffe *et al.* (1952) préviennent que la coévolution entre les populations de lapin et le virus risque de conduire à la sélection de souches moins virulentes tandis que les lapins deviennent plus résistants à l'infection. Dès les premiers temps de la propagation, on observe que les lapins survivant à l'infection développent une immunité et que cette immunité les protège contre les formes graves de la myxomatose en cas de réinfection (Fenner *et al.*, 1953)²⁶.

25. L'été australien a lieu pendant les mois correspondant à l'hiver dans l'hémisphère nord.

26. Frank Fenner, que je cite beaucoup dans cette section, est non seulement l'un des principaux scientifiques impliqués dans l'étude de la myxomatose dans les années 1950 à 1980, mais également l'un des grands noms à qui on doit l'éradication de la variole.

Au cours de l'été 1951-1952 (deuxième épidémie), il est estimé que la maladie est encore extrêmement létale, avec 90-99 % de létalité, quoique ce pourcentage soit parfois plus bas dans certaines localités (Fenner *et al.*, 1953). Cependant, on remarque déjà une baisse de la létalité par rapport à l'épidémie précédente, et Fenner *et al.* (1953) propose trois mécanismes qui pourraient expliquer cet état de fait :

- La génération de lapins de 1951-1952 est issue de mères ayant survécu à l'épidémie précédente et transmis leurs anticorps à leur descendance.
- Seuls les lapins ayant un patrimoine génétique leur permettant de résister à la maladie ont survécu à l'épidémie précédente et ont donc transmis leurs gènes à leur descendance.
- Un variant moins virulent du MYXV a émergé et est devenu dominant.

Pendant l'été 1952-1953, la létalité du virus diminue encore, et les premiers tests sont menés démontrant qu'un nouveau variant moins virulent est bel est bien entré en circulation (Fenner, 1953). Dans les trois décennies suivantes, la virulence des souches en circulation en Australie est systématiquement évaluée. Il apparaît alors une évidente sélection de souches de virulence intermédiaire (autour de 60-70 % de létalité) plutôt que de souches avirulentes et une disparition presque complète des souches de très forte virulence (Marshall et Fenner, 1960 ; Edmonds *et al.*, 1975 ; Fenner, 1983).

Le cas de la myxomatose est certainement le meilleur exemple permettant d'illustrer la théorie du *trade-off*. Pour résumer : le MYXV, virus endémique des Amériques, est introduit dans une population d'hôtes qui ne lui sont pas naturels mais chez lesquels il est capable de se transmettre. Mal adapté, il provoque une extrêmement forte mortalité, telle que pendant plusieurs années les essais d'introduction de la maladie dans les populations sauvages de lapins en Australie échouent²⁷. La *conventional wisdom* énonce qu'on devrait avoir sélection pour des souches avirulentes du MYXV, or nous avons vu que ce n'est pas le cas. Comme noté plus haut, la transmission se fait essentiellement – du moins en Australie – via les moustiques. Or, les moustiques ne piquent que des hôtes vivants, et les souches de très forte virulence tuent leur hôte avant qu'il ait eu le temps de se faire piquer. A l'inverse, les souches de très faible virulence ne produisent pas assez de matériel viral pour infecter les moustiques les piquant (Fenner, 1983).

27. D'autres facteurs ont été proposés pour justifier ces échecs, comme les conditions environnementales ou la présence de prédateurs retirant les carcasses infectées des colonies.

Il est intéressant de noter que le MYXV a continué d'évoluer en Australie, où il a été observé que dans certaines conditions des souches de forte virulence peuvent réémerger (Kerr *et al.*, 2022). Cette découverte montre qu'il reste encore beaucoup de choses à apprendre sur la manière dont la virulence des pathogènes évolue. De plus, si le pathogène évolue, l'hôte également. Il a ainsi été montré que les lapins vivant dans les régions où le MYXV circule possèdent une plus grande résistance génétique à l'infection par le virus (Alves *et al.*, 2019). Pour de plus amples informations sur la myxomatose, l'évolution du MYXV et également son introduction en Europe, contemporaine de celle en Australie, je conseille les *reviews* de Kerr (2012) et Kerr *et al.* (2015).

1.4 – Impact des vaccins sur l'écologie et l'évolution des pathogènes

1.4.1 – Généralités

Les vaccins ont une multitude de répercussions à la fois sur le pathogène contre lesquels ils sont dirigés et sur l'hôte ou la population d'hôtes les utilisant. Ces répercussions existent à l'échelle individuelle, populationnelle et évolutive.

À l'échelle individuelle, le premier effet des vaccins, celui pour lequel ceux-ci sont produits et distribués, est de protéger les individus contre les effets délétères des pathogènes. Cet effet peut être atteint soit en protégeant directement les individus contre l'infection par le pathogène, soit en empêchant les symptômes de l'infection d'apparaître. Ces deux modes d'action des vaccins ont des conséquences différentes sur l'écologie et l'évolution des pathogènes, comme discuté plus loin et notamment dans la Section 2.1.1. Les vaccins peuvent également avoir des conséquences délétères sur les individus, comme c'est le cas par exemple des vaccins anticoquelucheux (Edwards et Decker, 2012, voir également la Section 1.2.3).

À l'échelle populationnelle, l'utilisation des vaccins a deux conséquences majeures. La première, au cœur des stratégies de vaccination de masse, est de générer ce qu'on appelle une immunité de groupe (*herd immunity* en anglais). L'immunité de groupe consiste en la réduction de la circulation d'un pathogène par l'augmentation de la part immunisée de la population (Fine, 1993). Pour une maladie à transmission interhumaine, le pathogène peut se transmettre tant que

son réservoir d'hôtes susceptibles est suffisamment grand (*i.e.*, tant qu'il parvient à entrer en contact avec un nouvel hôte susceptible); or, ce réservoir se réduit lorsqu'une partie des hôtes acquièrent une immunité – naturellement ou par vaccination – et ainsi l'immunité de groupe est atteinte lorsqu'une part suffisamment importante de la population est immunisée et donc retirée du réservoir d'hôtes susceptibles pour que la balance entre le nombre de nouveaux infectés et le nombre d'infections guéries devienne négative (Fine, 1993). L'un des buts de la vaccination de masse est donc de réduire le réservoir d'hôtes susceptibles pour atteindre cette immunité de groupe. Lorsque le pathogène ne peut plus se transmettre faute d'hôtes susceptibles il est voué à disparaître – quoique cette éradication puisse prendre un temps plus ou moins long selon les caractéristiques du système hôte-pathogène – et ainsi même la part non vaccinée de la population se retrouve au moins partiellement protégée lorsque l'immunité de groupe est atteinte.

L'autre effet que les vaccins peuvent avoir sur les populations est d'induire des changements démographiques et épidémiologiques. En premier lieu, l'utilisation de vaccins réduit la circulation de maladies et donc la morbidité, la mortalité mais aussi les coûts de santé associés à ces maladies. Il a été estimé que depuis l'introduction du Programme Élargi de Vaccination (PEV) par l'OMS en 1974, 154 millions de morts ont pu être empêchées par la vaccination – sachant que cette estimation omet de nombreux vaccins (Shattock *et al.*, 2024). En réduisant la mortalité associée aux maladies infectieuses, la vaccination perturbe également la croissance des populations. Comme présenté plus tôt, il a été estimé qu'au cours du XVIII^e siècle la variole a été responsable de 8 à 20 % des morts en Europe : la découverte de la vaccination à la fin du siècle, en réduisant le fardeau de la variole, a contribué à l'expansion démographique qui a eu lieu au XIX^e siècle et ensuite (Mercer, 1985).

Ensuite, la vaccination peut perturber la distribution en âge et la dynamique des infections, comme cela a été décrit par exemple pour la rougeole et la coqueluche. Historiquement, ces deux maladies touchent principalement les enfants en bas âge et les stratégies vaccinales cherchent à immuniser en priorité les jeunes individus. Ainsi, l'immunité des enfants est rapidement renforcée pour réduire le risque d'infection contre ces pathogènes. Avec le temps, l'immunité s'estompe et n'est souvent pas renouvelée faute de rappels vaccinaux. Ce sont alors les classes d'âges suivantes qui deviennent susceptibles de contracter ces maladies. En Chine, il a été montré qu'avant 1990 la majorité des cas de rougeole sont distribués autour d'un unique pic entre 5 et 15 ans alors

qu'après 2000 les cas sont distribués autour de deux pics : un chez les enfants de moins de deux ans et un chez les adultes de 15 à 35 ans (Li *et al.*, 2017). En Suède, la réintroduction de la vaccination contre la coqueluche en 1996 après 17 années d'interruption a conduit à un rapide changement dans la distribution en classes d'âge des cas. Avant la réintroduction, plus de 80 % des cas de coqueluche sont observés chez les enfants de moins de cinq ans. Une dizaine d'années après la réintroduction, environ 60 % des cas sont observés chez les enfants de six ans et plus, avec également un pic chez les moins d'un an (Rohani *et al.*, 2010).

À l'échelle évolutive, l'un des effets principaux de la vaccination est l'émergence de souches de pathogènes capables d'éviter l'immunité induite par la vaccination. On parle d'échappement vaccinal (*vaccine escape* en anglais). L'échappement vaccinal survient lorsque le pathogène accumule des mutations sur un antigène présent dans le vaccin, réduisant avec le temps l'affinité entre les anticorps de l'hôte et les antigènes du virus ou de la bactérie. L'échappement vaccinal est donc un produit de la sélection naturelle : sous la pression du système immunitaire des hôtes renforcé par le vaccin, les mutants sont sélectionnés car ils sont moins efficacement reconnus par le système immunitaire de l'hôte, donc moins rapidement éliminés et se reproduisent donc plus. Ce phénomène a largement été médiatisé au cours de la pandémie de Covid-19, avec l'apparition continue de nouveaux mutants capables d'échapper à l'immunité induite par les vaccins (Wang *et al.*, 2021 ; Chakraborty *et al.*, 2022). Il a également été documenté chez d'autres pathogènes comme le virus de l'hépatite B (dos Santos *et al.*, 2017 ; Wang *et al.*, 2017), le pneumocoque (Brueggemann *et al.*, 2007) ou le virus de la grippe (Carrat et Flahault, 2007).

Une autre conséquence évolutive de la vaccination de masse est la sélection de souches avec une virulence altérée. L'effet de la vaccination sur la virulence des souches de pathogènes en circulation est principalement étudiée par le prisme de la modélisation (voir Section 1.5.3). Il existe néanmoins quelques exemples *in situ* où un changement de virulence a été observé et lié à l'utilisation massive de la vaccination. L'un de ces exemples est le virus de la maladie de Marek (MDV). Cette maladie se transmet à la faveur des fortes concentrations d'individus dans les élevages industriels de volaille et son virus était initialement de virulence modérée. À partir des années 1950, la virulence des souches de MDV en circulation augmente progressivement, d'abord en lien avec l'industrialisation de l'élevage – et l'augmentation de la densité d'hôtes – puis avec la vaccination à partir de 1968 (Witter, 1997 ; Gimeno, 2008). Suite à l'utilisation d'une première

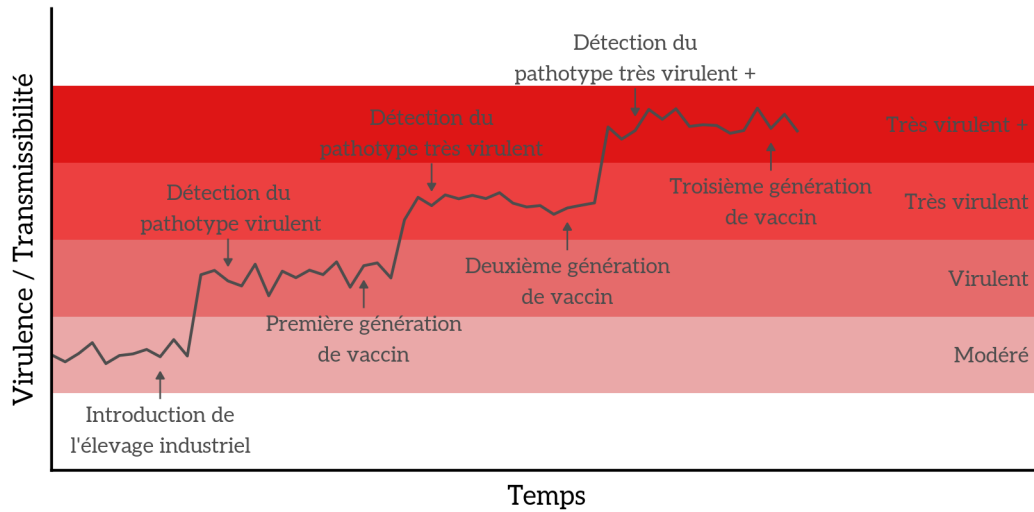


FIGURE 1.4 – **Vision simplifiée de l'évolution du virus de la maladie de Marek** à travers le temps, en lien avec l'industrialisation de l'élevage aviaire et le développement de nouveaux vaccins. La virulence est corrélée positivement à la transmission, et la vaccination réduit le coût à la virulence, sélectionnant ainsi des souches qui deviennent fortement létales pour les hôtes non vaccinés. Les axes sont donnés à titre indicatif et ne représentent pas la réalité. La virulence fluctue entre pas de temps pour illustrer l'émergence et la disparition de mutants de virulence altérée.

génération de vaccins, des souches de plus forte virulence encore vont apparaître, nécessitant le déploiement d'une deuxième génération de vaccins. Suivant l'introduction de cette deuxième génération de vaccin, la virulence du virus va encore augmenter, nécessitant le développement d'une troisième génération pour contrer ces virus qualifiés de « très virulents + » (Figure 1.4) (Geoghegan et Holmes, 2018). Il est intéressant de noter que l'évolution vers une plus forte virulence correspond dans le cas du MDV à une convergence évolutive entre plusieurs souches plutôt qu'à la propagation d'un clone de forte virulence (Trimpert *et al.*, 2017). L'apparition de souches de MDV de plus en plus grande virulence s'explique en grande partie par la nature des vaccins en circulation (mais également par de mauvaises pratiques d'élevage et de vaccination). Ce sont des vaccins appelés « imparfaits » (Encadré 1.4), bloquant le développement de la maladie mais favorisant les souches de plus forte virulence (Gandon *et al.*, 2003 ; Bull et Antia, 2022), ces souches ayant un taux de transmission plus élevé que des souches moins virulentes (Read *et al.*, 2015).

Le virus de la maladie de Marek (MDV) est certainement l'exemple le plus souvent cité

pour illustrer le rôle de la vaccination dans l'évolution de la virulence des pathogènes. Dans les sections suivantes, l'évolution de *C. diphtheriae* et de *B. pertussis* à l'ère vaccinale est discutée, notamment à travers des altérations dans la production de facteurs de virulence que sont la toxine diphtérique et la toxine pertussique.

ENCADRÉ 1.4 – Vaccins imparfaits

Là où les vaccins traditionnels bloquent l'infection ou la transmission des pathogènes, les vaccins imparfaits (*imperfect* ou *leaky vaccines* en anglais) bloquent les symptômes de l'infection sans interrompre la circulation des maladies – quoiqu'ils puissent la réduire (Halloran *et al.*, 1989; Gandon *et al.*, 2001; Bull et Antia, 2022). Il a été théorisé que de tels vaccins conduisent à la sélection de souches de plus forte virulence (voir Section 1.5.3) en perturbant le compromis entre taux de transmission des pathogènes et virulence (Gandon *et al.*, 2001).

Le principe est le suivant : si le vaccin protège contre les symptômes d'une maladie, c'est qu'il protège contre les effets de la virulence du pathogène qui lui est associé. Ainsi, sous l'hypothèse d'une corrélation positive entre taux de transmission et virulence, un pathogène circulant au sein d'une population d'hôtes vaccinés sera favorisé s'il a une forte virulence, puisqu'il aura en même temps un fort taux de transmission et que les désavantages à avoir une forte virulence (*i.e.*, la mort ou l'immobilisation de l'hôte) sont annulés par l'effet du vaccin.

Ainsi, l'utilisation massive d'un vaccin imparfait devrait sélectionner des souches présentant une virulence exacerbée, comme cela a été montré chez le virus de la maladie de Marek (voir texte principal). Ces souches de forte virulence feraient alors des dégâts considérables chez les individus non vaccinés, notamment si elles devaient atteindre des populations où aucun vaccin ne circule ou si la couverture vaccinale devait diminuer.

1.4.2 – Le cas de *Corynebacterium diphtheriae*

En Écosse, le suivi continu des infections par *C. diphtheriae* dans la région du Grampian entre 2002 et 2010 a permis d'isoler 27 infections, toutes dues à des souches non toxigènes (Edwards *et al.*, 2011). En France, les souches toxigènes ont été remplacées par des souches non toxigènes à la fin du XX^e siècle (Farfour *et al.*, 2013). En Angleterre et au Pays de Galles, les souches non toxigènes deviennent largement majoritaires dès le début des années 1990 et représentent presque toutes les souches isolées à la fin de la décennie (Reacher *et al.*, 2000; Wren et Shetty, 2005). Le cas le plus fréquemment cité sur le lien entre vaccination et émergence

des souches non toxigènes est la Roumanie. Entre 1958 et 1972, la couverture immunitaire augmente rapidement de 60 % à 97 %. En même temps, on observe une réduction progressive du pourcentage de souches toxigènes, passant de 86 % à 3 %, et la maladie disparaît du pays (Ewald, 2004b).

Tous ces exemples aident à formuler l'idée que la pression exercée par les vaccins antidiphthériques, visant particulièrement la TD, favorise l'émergence et la propagation de souches dépourvues de cet unique antigène vaccinal (Guiso *et al.*, 2022). La production de la TD représente un coût métabolique important pour la bactérie. En population non vaccinée, ce coût est aisément compensé par le bénéfice que la toxine apporte à la bactérie, mais en population vaccinée ce coût est rendu nul par l'action du système immunitaire qui inactive la TD (Soubeyrand et Plotkin, 2002).

On peut également penser que l'arrêt de production de la TD participe à une stratégie d'échappement vaccinal. Cependant, la toxine est sécrétée à distance de la bactérie. Une réponse immunitaire dirigée contre la toxine ne devrait donc pas compromettre la bactérie. L'hypothèse d'un coût élevé à la production de la toxine semble donc aujourd'hui être la meilleure hypothèse pour expliquer la sélection de souches mutantes ne la produisant pas.

1.4.3 – Le cas de *Bordetella pertussis*

L'impact de la vaccination sur l'évolution de *B. pertussis* est beaucoup plus largement documenté que pour *C. diphtheriae*, notamment parce que les vaccins anti-coquelucheux sont plus complexes que les vaccins anti-diphthériques par leur composition et leur histoire, et que leurs conséquences évolutives sont à leur tour plus complexes. Il existe trois effets majeurs de la vaccination sur l'évolution de la bactérie : la sélection de nouveaux allèles, le changement d'expression de la toxine pertussique (TP) et la perte de la production de la pertactine (PRN), une adhésine de surface. Un quatrième effet existe, le changement dans le sérotype des fimbriae, mais ne sera pas discuté ici.

Sélection de nouveaux allèles

La majorité des études portant sur l'évolution de *B. pertussis* se concentre sur les gènes *prn*, *fim2*, *fim3* et *ptxA*, codant respectivement la PRN, le Fim2, le Fim3 et le domaine A de la TP (*i.e.*, sa partie active). Comme présenté dans la Section 1.2.3, cela correspond aux antigènes inclus dans les ACV utilisés aujourd'hui dans la majorité des pays à fort revenu, à l'exception de la FHA²⁸.

Ces gènes sont dispersés sur le chromosome bactérien plutôt que sur des plasmides ou autres éléments génétiques mobiles. Ce chromosome est d'une longueur d'environ 3,8 millions de paires de bases variable entre souches REF. Les gènes *prn* et *fim2* sont voisins tandis que le gène *fim3* est isolé des autres gènes sur le chromosome, séparé par 0,5 millions de paires de bases du gène *fim2*. Les différents gènes utilisés pour la production de la toxine pertussique, *ptxA-E*, qui forment l'opéron *ptx*, sont fortement rapprochés sur le chromosome, et largement séparés des autres gènes codant les antigènes vaccinaux (Stibitz et Garletts, 1992; Andrade *et al.*, 2014).

La majorité des WCV et ACV historiquement utilisés sont produits à partir de souches possédant les allèles *fim2-1*, *fim2-3*, *prn1* et *ptxA2*, quoiqu'il existe une certaine variabilité, notamment pour le gène *ptxA* où l'allèle *ptxA2* n'est représenté que dans la moitié des souches vaccinales (Bart *et al.*, 2014). C'est également le cocktail d'allèles que l'on retrouve chez la souche Tohama I, utilisée pour produire la majorité des ACV. Au cours de l'ère pré-vaccinale, le génotype des souches en circulation correspond globalement à celui des souches utilisées dans les vaccins. À la suite de l'introduction des vaccins, une proportion de plus en plus grande de souches présentant de nouveaux allèles est observée, menant à une forte discordance entre souches vaccinales et souches en circulation.

Comme présenté plus haut, la majorité des vaccins sont basés sur des souches *prn1*, mais il existe quelques vaccins pour lesquels ont été utilisées des souches *prn7* ou *prn10*. Depuis quelques années, l'émergence et la propagation d'un allèle présent dans aucun vaccin, *prn2*, sont documentées dans de très nombreux pays comme en Australie (Octavia *et al.*, 2012), aux États-Unis (Schmidtke *et al.*, 2012), en Finlande (Elomaa *et al.*, 2005), en Italie (Loconsole *et al.*, 2018), au Japon (Miyaji *et al.*, 2013; Zomer *et al.*, 2018), en Pologne (Polak *et al.*, 2019), au

²⁸. Le gène de la FHA est moins étudié du fait de sa grande taille, qui le rend difficile à séquencer (Nicole Guiso, communication personnelle).

Royaume-Uni (Fry *et al.*, 2001), en Russie (Borisova *et al.*, 2007) ou encore en Suède (Hallander *et al.*, 2005). De nombreux indices lient l'émergence de *prn2* à la vaccination. La comparaison de plusieurs pays utilisant soit un (ou plusieurs) ACV, soit un WCV, soit les deux types de vaccin, conduit à émettre l'hypothèse que les ACV comme les WCV conduisent à la sélection de souches *prn2*, quoiqu'il ait été précédemment proposé que ce soit la vaccination acellulaire qui soit à l'origine de l'émergence des souches *prn2* (van Gent *et al.*, 2015).

Il semblerait donc que l'émergence de souches *prn2* soit un cas typique d'évolution par échappement vaccinal (Section 1.4.1). Les cas danois et serbe sont particulièrement intéressants pour illustrer comment l'intensité de la pression vaccinale module la sélection de souches présentant un gène *prn* différent de celui présent dans les vaccins.

Au Danemark, l'allèle *prn2* est détecté pour la première fois au début des années 1970. Il devient largement majoritaire dans les années 1990 et, à la fin du siècle dernier, l'allèle *prn1* disparaît de la population danoise. Ce n'est qu'en 2003 que le Danemark remplace son WCV contre un vaccin acellulaire et, exception parmi les autres pays, c'est ACV1 que le Danemark commence à utiliser ; c'est-à-dire un vaccin ne contenant pas de PRN. Or, à la fin des années 2000, l'allèle *prn1* réapparaît dans le pays (Petersen *et al.*, 2012). On a donc là un important indice d'une sélection dirigée par le WCV, et qui s'annule une fois la pression contre l'allèle *prn1* relâchée par l'utilisation d'un ACV dépourvu de PRN. Il serait intéressant de disposer de données plus récentes pour vérifier si la réémergence de *prn1* observée à la fin des années 2000 est anecdotique ou représente réellement une tendance au retour de cet allèle.

La Serbie est également bien différente des autres pays dans son utilisation des vaccins antioquelucheux. Un WCV y est utilisé jusqu'en 2014 et sa composition a changé au cours de l'histoire. Avant 1985, le vaccin serbe est composé de deux souches porteuses de l'allèle *prn1* mais en 1985 deux nouvelles souches, dont une *prn2*, sont ajoutées au vaccin. Il en résulte une émergence lente de l'allèle *prn2* dans ce pays, et surtout l'émergence de deux autres allèles, *prn3* et *prn11* (Dakic *et al.*, 2010 ; Pljesa *et al.*, 2016). Quoiqu'il y ait eu en Serbie une émergence partielle de l'allèle *prn2*, son retard ainsi que l'émergence des allèles non vaccinaux *prn3* et *prn11* contribuent à l'idée d'une sélection par échappement vaccinal de souches présentant de nouveaux allèles du gène *prn*.

Pour le gène *ptxA* en revanche, les choses sont plus complexes et il est difficile de voir une

forme d'échappement vaccinal. Comme pour l'émergence de *prn2*, de nombreux pays font état de l'émergence de souches portant l'allèle *ptxA1*, différent de l'allèle vaccinal majoritaire *ptxA2*. Ces émergences sont documentées par exemple en Russie (Borisova *et al.*, 2007), aux États-Unis (Schmidtke *et al.*, 2012) ou en Pologne (Polak *et al.*, 2019) avant l'introduction de la vaccination acellulaire, excluant un rôle particulier des ACV mais semblant indiquer ici encore un phénomène d'échappement vaccinal.

Cependant, les exemples serbe et britannique nuancent cette hypothèse. En Serbie, avant le changement de vaccin de 1985, le vaccin contient uniquement l'allèle *ptxA2* et l'allèle *ptxA1* se propage au cours des années 1960 et 1970. Ensuite, après l'ajout d'une paire de souches *ptxA1* dans le vaccin, l'allèle *ptxA2* réémerge dans le pays. Au Royaume-Uni, l'émergence de l'allèle *ptxA2* est également documentée, passant de la moitié des isolats collectés dans les années 1940 à la totalité dans les années 1990 (Fry *et al.*, 2001). Or, durant toute cette période le pays a utilisé un WCV composé de deux souches : une *ptxA2* et une *ptxA1*. Ainsi, dans ces deux pays, l'utilisation de vaccins composés à la fois de souches *ptxA1* et *ptxA2* a modifié la structure des souches en circulation sans que l'on observe l'émergence de nouveaux allèles non vaccinaux. Il est donc possible que le gène *ptxA* subisse des pressions autres que vaccinales maintenant les allèles *ptxA1* et *ptxA2* au sein des populations humaines.

Changement d'expression de la toxine pertussique

Un deuxième phénomène évolutif en lien avec la vaccination est le changement de production de la toxine pertussique par les souches de *B. pertussis* à l'ère vaccinale. On désigne sous le nom *ptxP* le promoteur de l'opéron *ptx* codant les différentes sous-unités de la TP (Nicosia et Rappuoli, 1987). Historiquement, les souches en circulation ainsi que les souches vaccinales portent l'allèle *ptxP1*, or plusieurs pays comme l'Afrique du Sud (Moosa *et al.*, 2023a), l'Australie (Lam *et al.*, 2014), le Danemark (Petersen *et al.*, 2012), l'Italie (Loconsole *et al.*, 2018), le Japon (Miyaji *et al.*, 2013; Zomer *et al.*, 2018), les Pays-Bas (Mooi *et al.*, 2009) et la Suède (Advani *et al.*, 2011) font état à l'ère vaccinale de l'émergence d'un nouvel allèle, *ptxP3*.

Cette sélection pour un nouvel allèle de promoteur est particulièrement intéressante. Un promoteur régule la transcription d'un gène, il ne code pas directement une protéine et n'est donc pas reconnaissable par le système immunitaire de l'hôte. La sélection de souches *ptxP3*

ne résulte donc pas d'un phénomène d'échappement vaccinal. Les souches *ptxP3* produisent vraisemblablement une plus grande quantité de TP que les souches *ptxP1* (Mooi *et al.*, 2009; Gouw *et al.*, 2014), or la TP serait responsable des symptômes de la coqueluche et donc de la virulence de *B. pertussis* (Scanlon *et al.*, 2019a,b). Plusieurs hypothèses ont été proposées pour expliquer cette sélection de souches de plus forte virulence :

- Les ACV protègent contre les formes graves de la coqueluche mais échouent à bloquer l'infection et la transmission de *B. pertussis* (Warfel *et al.*, 2014), or un modèle théorique montre que des vaccins bloquant les symptômes d'une maladie sans interrompre la circulation du pathogène (les vaccins imparfaits présentés plus tôt) devraient sélectionner des souches de plus forte virulence puisqu'ils réduisent le coût chez l'hôte à la production d'une toxine avantageuse pour la bactérie (voir Section 1.5.3).
- La TP a plusieurs effets sur l'hôte de *B. pertussis*, dont la modulation de sa réponse immunitaire (Scanlon *et al.*, 2019a,b). Un système immunitaire amélioré par la vaccination pourrait représenter un obstacle supplémentaire pour la bactérie qui serait sélectionnée pour surexprimer la TP pour « déborder » les défenses de l'hôte (Mooi *et al.*, 2009).
- Comme la TD chez *C. diphtheriae*, la TP est sécrétée à distance de la surface bactérienne, donc il est possible que la réponse immunitaire dirigée contre la toxine n'affecte pas la bactérie (Ma *et al.*, 2021). Plutôt qu'une hypothèse indépendante, celle-ci constitue un complément des deux précédentes.

L'émergence des souches *ptxP3* est un phénomène documenté majoritairement dans des pays utilisant des ACV. De plus, il semblerait que la formulation de l'ACV ait une incidence sur la vitesse de propagation des souches *ptxP3*. Lors de l'introduction de la vaccination acellulaire en Suède, le pays a été divisé en deux régions. Dans la région de Göteborg, un ACV1 était utilisé tandis que le reste du pays employait un ACV3. Il a été observé que l'allèle *ptxP3* a émergé et s'est répandu plus vite dans la région de Göteborg que dans le reste de la Suède (Advani *et al.*, 2011). L'explication proposée est que l'ACV1 a une concentration plus forte que l'ACV3 en TP, déclenchant une plus forte réponse immunitaire chez l'hôte et induisant ainsi une plus grande pression sélective pour produire une plus forte quantité de toxine afin de déborder le système immunitaire. Il est également possible que la présence d'un unique antigène dans le vaccin, en réduisant la diversité des pressions exercées contre les souches de *B. pertussis*, ait permis une

plus rapide sélection des souches *ptxP3* dans la région de Göteborg.

L'émergence de souches *ptxP3* dans des pays utilisant des WCV est assez peu renseignée et les résultats sont parfois ambigus, compliquant les interprétations du rôle des WCV dans la sélection de souches *ptxP3*. En Pologne, l'allèle *ptxP3* a largement remplacé l'allèle *ptxP1*, mais uniquement une fois qu'un vaccin acellulaire a commencé à circuler en parallèle du vaccin à germe entier, et il est donc difficile de conclure à un effet des WCV (Polak *et al.*, 2019). Aux États-Unis, l'émergence de *ptxP3* commence avant l'introduction de la primo-vaccination par ACV mais après l'introduction du vaccin pour les rappels chez l'enfant. Il est donc possible que le WCV ait participé à l'émergence de ce nouvel allèle (Schmidtke *et al.*, 2012). De la même manière, en Afrique du Sud, la vaccination acellulaire a remplacé le WCV en 2009, et la totalité des souches isolées entre 2015 et 2019 portent l'allèle *ptxP3* (Moosa *et al.*, 2023a). La rapidité de propagation du nouvel allèle laisse supposer un effet précédent du WCV. En Argentine, les indices d'une sélection par les WCV sont plus forts. Comme en Pologne, des vaccins acellulaires sont commercialisés mais leur utilisation reste très marginale du fait de leur coût. Pourtant, la prévalence des souches *ptxP3* est extrêmement élevée, cet allèle étant retrouvé dans la totalité des échantillons entre 2005 et 2017 et représentant déjà 40 % des isolats en 2000 (Carrquiriborde *et al.*, 2019).

Finalement, la vitesse de sélection de souches *ptxP3* serait possiblement liée à la « précision » du vaccin : moins celui-ci comporte d'antigènes, plus il cible précisément la TP et plus le nouvel allèle se propage rapidement au sein de la population. Quoique la sélection de souches *ptxP3* soit largement documentée et acceptée, l'hypothèse d'une plus forte virulence de ces souches est cependant controversée, la surexpression de la toxine ayant pour l'instant été observée uniquement chez la souris et pas au sein de populations humaines.

Arrêt d'expression de la pertactine

Un troisième effet de la vaccination sur l'évolution de *B. pertussis* est la sélection de souches ne produisant plus l'un des antigènes vaccinaux : la pertactine. Ces souches seront appelées déficientes en pertactine, ou PRN⁻, en opposition aux souches productrices de pertactine, ou PRN⁺.

Cette sélection pour des souches PRN⁻ est largement documentée dans les pays utilisant des

ACV comme l’Australie (Lam *et al.*, 2014; Xu *et al.*, 2019), le Canada (Tsang *et al.*, 2014, 2019), les États-Unis (Pawloski *et al.*, 2014; Vodzak *et al.*, 2017), ou la Nouvelle-Zélande (Ring *et al.*, 2022). En revanche, dans des pays utilisant encore des WCV comme l’Argentine (Carrquiriborde *et al.*, 2019), le Brésil (Leite *et al.*, 2021) ou encore la Pologne (Polak *et al.*, 2019) les souches PRN⁻ sont rares. Cependant, en Pologne, la proportion de souches PRN⁻ semble augmenter dans les années récentes, mais comme énoncé plus haut dans ce pays un vaccin acellulaire circule grandement aujourd’hui.

Ces indices permettent de formuler l’hypothèse que ce sont les ACV qui conduisent à la sélection de souches PRN⁻. En Europe, une corrélation a été trouvée entre la proportion de souches PRN⁻ et la durée depuis laquelle la vaccination acellulaire a été introduite dans le pays (Barkoff *et al.*, 2019). Mais plus que simplement les ACV, ce seraient les ACV contenant de la PRN qui seraient responsables de cette sélection, donc les ACV3 et ACV5. Ainsi, en Afrique du Sud, où un ACV2 est en circulation, aucune sélection pour des souches PRN⁻ n’est observée (Moosa *et al.*, 2023a). Au Japon, les ACV en circulation ont changé au cours du temps, des vaccins contenant ou non la PRN étant régulièrement ajoutés sur le marché, perturbant la sélection de souches PRN⁻ et faisant ainsi fluctuer leur prévalence (Otsuka *et al.*, 2012; Miyaji *et al.*, 2013; Hiramatsu *et al.*, 2017; Zomer *et al.*, 2018).

L’inactivation du gène *prn*, conduisant au phénotype PRN⁻, est obtenue par différentes mutations, par exemple l’introduction d’éléments transposables dans la séquence codante du gène *prn*, des mutations changeant un codon en codon stop, des délétions de séquences de plusieurs milliers de bases de longueur, ou encore une perte totale du gène *prn* (Otsuka *et al.*, 2012; Pawloski *et al.*, 2014; Barkoff *et al.*, 2019). Cette multiplicité des mutations indique que l’émergence globale de souches déficientes en pertactine n’est pas due à la propagation d’un même clone de *B. pertussis* mais plutôt à un phénomène de convergence où de très nombreuses souches PRN⁻ sont sélectionnées à travers le monde.

La sélection de souches au phénotype PRN⁻ s’explique facilement d’un point de vue immunologique. Contrairement à la toxine pertussique, la pertactine s’exprime en surface de la bactérie. Cela signifie qu’une réponse immunitaire dirigée contre la PRN atteindra la bactérie, la mettant directement en danger. Ne plus produire la PRN permet ainsi à la bactérie d’être protégée contre le système immunitaire de l’hôte stimulé par la vaccination (Ma *et al.*, 2021). L’avantage

des souches déficientes en pertactine a pu être vérifié chez la souris, où il a été montré que les souches PRN⁻ se développent mieux que les souches PRN⁺ chez les souris vaccinées par un ACV (Hegerle *et al.*, 2014). Ces résultats ont été confirmés chez l'humain où il a été montré que les individus vaccinés ont de plus grandes chances d'être infectées par une souche PRN⁻ que par une souche PRN⁺ (Martin *et al.*, 2015).

Il est notable que la pertactine est le seul antigène pour lequel une sélection claire pour son arrêt de production est documentée, alors qu'on s'attendrait à une sélection contre tous les antigènes vaccinaux (Guiso *et al.*, 2022). Certes, il existe quelques détections de souches ne produisant pas la TP aux États-Unis (Williams *et al.*, 2016; Weigand *et al.*, 2017) et en France (Bouchez *et al.*, 2009, 2021) et quelques souches ne produisant plus la FHA, également en France (Bouchez *et al.*, 2021), en Iran (Safarchi *et al.*, 2021) et au Japon (Wakimoto *et al.*, 2022), mais les isolats collectés se comptent sur les doigts d'une main. Comme noté plus tôt, la non-sélection de souches dépourvues de TP pourrait s'expliquer par le fait que la toxine est sécrétée loin de la bactérie, réduisant la pression exercée par le système immunitaire de l'hôte contre cet antigène; de plus la durée de l'immunité induite par les vaccins contre la TP serait relativement courte, relâchant rapidement toute pression qui s'exercerait contre la production de la toxine (Ma *et al.*, 2021). Pour la FHA, l'explication proposée est que la molécule est partiellement sécrétée (mais également produite en surface de la bactérie), comme la TP, et qu'elle serait en réalité peu immunogène (Ma *et al.*, 2021). Par ailleurs, son rôle d'adhésine est similaire à celui de la PRN, ce qui pourrait expliquer la sélection des souches PRN⁻ : la redondance fonctionnelle de la molécule diminue la nécessité pour la bactérie de maintenir sa production.

Il est également possible que le maintien observé des antigènes vaccinaux TP et FHA soit en réalité dû à un biais d'échantillonnage. Comme expliqué dans la Section 1.4.3, la TP est le principal facteur de virulence de *B. pertussis* et est donc responsable de la forme clinique de la coqueluche. Or, la bactérie est généralement isolée chez les patients présentant des symptômes coquelucheux, donc produisant la toxine pertussique. Il est donc peu probable de détecter par cette méthode des souches déficientes en TP. L'Afrique du Sud est l'un des rares pays à avoir initié un programme de dépistage de *B. pertussis* en population générale. Il en a résulté que 0,3 % de la population serait infectée par *B. pertussis*. Parmi ces infectés, le gène *ptxA*²⁹ n'a été

29. Nommé *ptxS1* dans l'article, il s'agit bien du même gène.

détecté que chez 20,3 % d’entre eux, laissant supposer que seulement un cinquième des souches en circulation produisent en réalité la TP (Moosa *et al.*, 2023b). Des études plus approfondies sont nécessaires pour confirmer cette hypothèse.

1.5 – Modélisation en (éco-)épidémiologie

La modélisation, pouvant prendre les formes les plus variées, est au cœur de l’épidémiologie et de l’éco-épidémiologie, c’est-à-dire l’épidémiologie reliant les caractéristiques des écosystèmes à la santé. Le terme « modélisation », et plus généralement le terme « modèle », peuvent revêtir de nombreux sens en science. Dans cette partie, je me concentrerai sur les aspects statistiques et mathématiques des modèles, laissant de côté les modèles biologiques employés en épidémiologie.

1.5.1 – Brève histoire de la modélisation en épidémiologie

L’un des premiers modèles mathématiques utilisés en épidémiologie est celui présenté en 1760 à l’Académie des Sciences par Daniel Bernoulli (1700 – 1782). Bernoulli développe un système d’équations différentielles et se sert de tables de mortalité pour démontrer l’utilité de l’inoculation contre la variole en termes d’années de vie gagnées par les individus acceptant de subir l’inoculation (Bernoulli et Chapelle, 2023 ; Bacaër, 2011a). Plus que la conclusion de ce travail, la contribution principale de Bernoulli à l’épidémiologie est d’y avoir apporté l’approche mathématique. C’est à partir de ce travail pionnier que les mathématiques vont progressivement s’imposer comme une clé de voûte de cette science.

Une des applications principales des mathématiques – et surtout du calcul différentiel – en épidémiologie est de décrire la manière dont un pathogène se répand au sein d’une population. Cette utilisation des mathématiques apparaît en 1911, lorsque Ronald Ross (1857 – 1932) propose dans la deuxième édition de son livre *The Prevention of Malaria* plusieurs modèles mathématiques décrivant la dynamique de la malaria (Ross, 1911 ; Bacaër, 2011c). Grâce à l’un de ces modèles, Ross a montré qu’il était possible d’éradiquer localement la malaria en réduisant les populations de moustiques porteurs (Bacaër, 2011c). Ce modèle, résumé et modernisé par Bacaër (2011c), consiste en un duo d’équations différentielles ordinaires (ODE). L’une représente le flux d’humains infectés, l’autre le flux de moustiques infectés. Ces systèmes d’ODE décrivant

une dynamique temporelle deviendront la base de la modélisation en épidémiologie. Ross restera toute sa vie, et à travers ses très nombreux écrits, convaincu de l'importance capitale des mathématiques en épidémiologie : « toute épidémiologie, intéressée qu'elle est par les variations d'une maladie dans le temps ou l'espace, doit être étudiée mathématiquement, peu importe le nombre de variables impliquées, pour être un tant soit peu considérée comme scientifique »³⁰ (Ross, 1911, p. 651). Il inventa même un terme pour cette application des mathématiques : la pathométrie, terme qui tombera cependant dans l'oubli (Ross, 1908, p. 30).

En 1926, Anderson Gray McKendrick (1876 – 1943) propose un modèle dans lequel une population est divisée en compartiments représentant différents stades infectieux et dont les individus peuvent passer d'un compartiment à l'autre selon des flux caractérisés par un jeu de paramètres. Il s'agit là du premier modèle que l'on appellera plus tard « SIR », pour susceptible-infecté-guéri (*removed*), du nom des trois compartiments qui le composent. Le modèle reprend l'idée proposée par Ross de décrire la dynamique épidémique par des flux dérivés du temps (McKendrick, 1925 ; Bacaër, 2011b). L'année suivante, William Ogilvy Kermack (1898 – 1970) se joint aux recherches de McKendrick et ils continuent ensemble à développer le modèle SIR. Ils mettent alors en évidence l'existence d'une valeur seuil de densité de population en dessous de laquelle une maladie ne peut pas se maintenir et qui dépend à la fois de la transmissibilité du pathogène, du taux de guérison et du taux de mortalité, une notion qui mènera des années plus tard à la formalisation du nombre de reproduction de base, \mathcal{R}_0 (Kermack et McKendrick, 1927 ; Bacaër, 2011b).

Si les travaux de Bernoulli, Ross, McKendrick et Kermack sont les fondations de l'épidémiologie mathématique, de nombreux auteurs et nombreuses autrices ont contribué à développer cette science au cours des XX^e et XXI^e siècles. Entre autres exemples (il serait difficile de toutes et tous les citer), Diekmann *et al.* (1990) ont contribué à formaliser le nombre de reproduction de base, van den Driessche et Watmough (2002) (voir également Diekmann *et al.*, 2010) ont proposé une méthode pour le calculer pour des modèles complexes...

30. « *all epidemiology, concerned as it is with the variation of disease from time to time or from place to place, must be considered mathematically, however many variables are implicated, if it is to be considered scientifically at all* »

1.5.2 – Exemples de modèles et de questions

La modélisation en (éco-)épidémiologie est utilisée pour répondre à un nombre croissant de questions. De la simple description des épidémies aux conséquences écologiques et évolutives de celles-ci, ce sont des livres entiers qui pourraient être écrits pour balayer toutes les interrogations auxquelles cherchent à répondre la modélisation – et tous les modèles qui ont été construits pour y parvenir. Le but de cette partie n'est donc pas de présenter un paysage exhaustif des modèles produits jusqu'à aujourd'hui, mais plutôt de présenter la diversité des questions posées à travers quelques exemples.

Le modèle compartimental proposé par Kermack et McKendrick est à la base d'un grand nombre de modèles épidémiologiques développés depuis maintenant presque un siècle. Ce modèle a été largement modifié, en ajoutant des paramètres et des compartiments, et d'autres modèles mathématiques complètement différents ont également été développés. Une emphase particulière est mise ici sur l'utilisation des modèles compartimentaux (modèles SIR ou dérivés) mais quelques travaux utilisant des modèles d'autre nature sont également présentés.

La plus simple utilisation qui puisse être faite d'un modèle épidémiologique est de décrire la dynamique d'une épidémie. Par exemple, [Rachah et Torres \(2015\)](#) ont utilisé un modèle SIR sous sa forme la plus basique pour modéliser l'évolution du nombre d'individus infectés par le virus Ebola au cours de l'épidémie au Libéria en 2014, et ont réussi à obtenir des estimations du nombre de cas très proches des observations. [Eisenberg *et al.* \(2015\)](#) ont complexifié le modèle SIR en y ajoutant un compartiment, W (pour *water*) pour l'adapter au mode de transmission du choléra (le pathogène est véhiculé par l'eau) et ont ainsi produit un modèle correctement ajusté aux données de l'épidémie ayant touché l'Angola en 2006.

En plus de décrire, les modèles épidémiologiques sont également utilisés pour prédire le déroulement des épidémies en amont de celles-ci. Cette application a été largement retrouvée au cours de la pandémie récente de Covid-19, où de nombreuses équipes se sont évertuées à produire des prédictions sur le cours de la pandémie, à une échelle globale ou plus locale. Par exemple, [Anand *et al.* \(2020\)](#) ont développé un modèle type-SIR incluant une mise en quarantaine pour inférer le moment où l'épidémie devrait s'arrêter dans l'état de Kerala, en Inde. Cependant, la capacité des modèles épidémiologiques à donner des prédictions correctes est remise en question à la fois

à cause de l'importance de la stochasticité intervenant en début d'épidémie et de la méconnaissance des valeurs des paramètres des modèles, les données étant nécessairement limitées en début d'épidémie (Roda *et al.*, 2020 ; Melikechi *et al.*, 2022). De plus, certains auteurs estiment que la complexité des modèles peut, dans certains cas, pénaliser leurs capacités prédictives à cause de l'ajout de paramètres dont les valeurs sont difficiles à estimer (Roda *et al.*, 2020).

Quoique intéressants d'un point de vue pédagogique, les modèles visant uniquement à décrire la dynamique d'une épidémie sont finalement assez peu instructifs, surtout lorsqu'il s'agit d'une épidémie déjà passée. Dans leur article sur le choléra déjà mentionné, Eisenberg *et al.* (2015) sont allés plus loin : la comparaison de leur modèle SIWR avec un modèle SIR (*i.e.*, sans transmission par l'eau) et avec un modèle SWR (*i.e.*, sans transmission par les infectés) leur a permis de mettre en évidence un rôle important du double mode de transmission du pathogène. Ainsi, les modèles épidémiologiques permettent de mettre en évidence des caractéristiques essentielles de la dynamique hôte-pathogène comme les voies de transmission, ce qui permet par la suite la mise en place de mesures de contrôle des maladies.

Un important facteur façonnant la dynamique d'une maladie est la manière dont les individus de la population entrent en contact les uns avec les autres. De nombreux modèles ont été développés pour prendre en considération des contacts hétérogènes au sein d'une population. Pour ce faire, ces modèles divisent généralement les compartiments S, I et R (et les compartiments supplémentaires s'il y en a) en sous-compartiments correspondant à des classes de la population. Ces classes représentent très souvent des classes d'âge, l'âge étant considéré comme un moteur puissant de la structure des contacts au sein d'une population (Anderson et May, 1985 ; Del Valle *et al.*, 2007 ; Mossong *et al.*, 2008). Les taux de transmission interclasses sont alors supposés différents de ceux intra-classes, et les taux ne sont pas nécessairement symétriques (*i.e.*, le taux d'infection de la classe A vers la classe B n'est pas nécessairement le même que celui de la classe B vers la classe A). La méthode classique pour renseigner ces taux de transmission est l'utilisation d'une matrice WAIFW (*Who Acquires Infection From Whom*) dont les éléments correspondent aux taux de transmission d'une classe à l'autre (Anderson et May, 1985). Cette hétérogénéité a des implications fortes sur la dynamique des maladies et les conséquences des moyens de luttés mis en place. Schenzle (1984) a développé un modèle structuré en âge et montré qu'un tel modèle fonctionne beaucoup mieux qu'un modèle « classique » sans classes pour décrire les épidémies de

rougeole, et [Hethcote \(1997\)](#) a développé un modèle comparable – quoique divisé en beaucoup plus de compartiments – pour la coqueluche et tiré des conclusions comparables.

Dans une idée similaire, plusieurs modèles de différentes natures ont été développés pour étudier l’effet de pratiques de contrôle sur la dynamique des maladies infectieuses. Une forme de contrôle souvent étudiée est la fermeture des écoles, dont l’effet a été étudié par le prisme de nombreux types de modèles différents : *agent-based model* ([Milne et al., 2008](#) ; [Lee et al., 2010](#)), modèle de réseau³¹ ([Glass et al., 2006](#)), ou modèle SIR amélioré ([Towers et al., 2012](#)). Ces modèles concluent généralement à un effet bénéfique de la fermeture des écoles sur la dynamique des épidémies (souvent, des épidémies de grippe), mais la saisonnalité de l’épidémie pourrait rendre une telle politique contre-productive dût-elle être déclenchée au mauvais moment ([Towers et al., 2012](#)). L’exemple de la fermeture des écoles montre l’importance et la complexité de la question des mesures de contrôle dans les travaux de modélisation.

Une autre problématique majeure pour laquelle sont développés des modèles épidémiologiques est de comprendre l’impact – réalisé ou potentiel – d’un vaccin sur la dynamique d’une maladie. Notamment, la modélisation épidémiologique permet d’estimer l’impact de la vaccination sur le nombre de vies gagnées. À partir d’un ensemble de modèles mathématiques, [Shattock et al. \(2024\)](#) ont pu produire une estimation du nombre de morts évitées grâce à la vaccination ces 50 dernières années. Dans leur modèle présenté en début de section, [Rachah et Torres \(2015\)](#) se sont intéressés à la manière dont la vaccination aurait pu perturber la dynamique de l’épidémie de virus Ebola. Tout en prenant en compte le coût d’implémentation d’une stratégie vaccinale, ils ont quantifié la stratégie optimale et ont montré comment la mise en place de celle-ci aurait permis de réduire le nombre d’infectés et de réduire la durée de l’épidémie.

La modélisation permet également, au-delà des vies sauvées, de comprendre comment les vaccins vont altérer la dynamique hôte-pathogène. J’ai présenté plus tôt (Section 1.4.1) le concept d’immunité de groupe. L’utilisation de modèles SIR et dérivés permet d’estimer la couverture vaccinale nécessaire pour atteindre une immunité de groupe ([Fine, 1993](#)). Un modèle proposé par [Hethcote \(1997\)](#) a permis de mettre en évidence qu’une vaccination contre la coqueluche ciblée sur les enfants ne permettrait pas d’atteindre une immunité de groupe mais permettrait uniquement de protéger les jeunes classes d’âge, plus sensibles à la maladie. Son modèle a éga-

31. Pour les modèles de réseau en épidémiologie, voir [Keeling et Eames \(2005\)](#)

lement permis de mettre en évidence que la vaccination ciblée sur les jeunes populations décale la distribution des cas vers les classes d'âge plus élevées, un résultat retrouvé par ailleurs dans un travail de modélisation centré sur la rougeole (Magpantay *et al.*, 2019). La modélisation permet également de comprendre comment des altérations dans les calendriers vaccinaux, des fluctuations dans l'efficacité vaccinale ou des altérations du taux de transmission d'un pathogène peuvent changer l'âge et le nombre des individus infectés (Pesco *et al.*, 2014, 2015); ou encore comment différents types de vaccins altèrent différemment la circulation des pathogènes. Gandon *et al.* (2003) montrent que les vaccins bloquant la transmission d'un pathogène conduisent à une diminution de la circulation du pathogène tandis que les vaccins bloquant sa multiplication intra-hôte devrait conduire à une augmentation de sa circulation, en lien avec une augmentation de sa virulence (voir Section 1.5.3).

La modélisation est aussi un outil de premier choix pour étudier les conséquences sur le long terme de la vaccination, c'est-à-dire les conséquences évolutives. Par exemple, la question de l'échappement vaccinal est au cœur de nombreux modèles épidémiologiques. Wilson *et al.* (1998) montrent que l'émergence d'un variant capable d'échapper au vaccin contre l'hépatite B nécessite à la fois une couverture vaccinale forte et une immunité croisée faible contre les variants. Ils prédisent également que si un tel variant devait émerger, sa propagation prendrait des dizaines d'années. McLeod et Gandon (2022) montrent quant à eux que la sélection de nouveaux variants capables d'échapper à la pression vaccinale est conditionnée au type de vaccin utilisé. Une autre question cruciale à laquelle de nombreux travaux sont dédiés est de comprendre la manière dont les vaccins exercent une pression évolutive altérant la virulence des pathogènes.

1.5.3 – Modéliser l'évolution de la virulence

Au temps où la théorie du *trade-off* n'était pas encore complètement établie, Levin et Pimentel (1981) ont formalisé mathématiquement la théorie de Lewontin (1970) d'une sélection de souches avirulentes de MYXV en Australie. Leur modèle, basé sur trois compartiments d'hôtes (non infecté, infecté par une souche avirulente et infecté par une souche virulente), omet le lien entre virulence et transmission, ne regardant que le taux de mortalité plus élevé de la souche virulente. Selon ce modèle, la souche virulente du pathogène devrait, sur le long terme, disparaître du

système.

Quelque temps plus tard, lorsque Roy Anderson et Robert May formalisent la théorie du *trade-off*, ils utilisent des modèles mathématiques pour le faire (Anderson et May, 1982; May et Anderson, 1990). Ils établissent alors les relations entre taux de transmission et virulence des pathogènes, pour conclure que les systèmes hôte-pathogène devraient converger vers des virulences intermédiaires. Par la suite, ce sont de très nombreux modèles proposés par presque autant d’auteurs qui sont développés pour tenter de déchiffrer la manière dont la virulence des parasites évolue. Ces modèles sont utilisés pour décrire de nombreux systèmes hôte-pathogène, parfois d’une grande complexité, là où les premiers modèles d’Anderson et May étaient relativement simples.

Un aspect des systèmes hôte-pathogène non étudié par Anderson est May est l’existence d’infections multiples. À l’échelle d’un hôte, plusieurs souches d’un pathogène peuvent être en compétition pour exploiter les ressources de l’hôte. Comme présenté dans la Section 1.3.2, ce sont les souches se multipliant le plus, donc les souches de plus forte virulence, qui sont avantagées et donc sélectionnées à cette échelle. Cependant, les pathogènes devant se transmettre à de nouveaux hôtes pour perpétuer leur transmission, la dynamique évolutive est en réalité plus compliquée. Ainsi, van Baalen et Sabelis (1995) ont montré que lorsque les infections multiples sont fréquentes (limitées à une double infection dans leur modèle), une plus forte virulence est favorisée. En revanche, lorsque la virulence augmente, la transmission d’un hôte au suivant diminue. Finalement, le niveau optimal de virulence dépendrait de la compétition intra-hôte et de la transmission inter-hôtes. En utilisant un modèle plus complexe, Alizon et van Baalen (2008a) développent cette idée d’un lien complexe entre compétition intra-hôtes et évolution de la virulence. Ils montrent notamment que la compétition entre souches peut mener à une diversification du pathogène, avec l’émergence d’une souche hypovirulente adaptée aux hôtes sains et d’une souche hypervirulente adaptée aux hôtes infectés par une souche hypovirulente.

Comme évoqué dans la Section 1.3.2, un aspect souvent considéré comme important pour étudier l’évolution de la virulence est le mode de transmission des pathogènes. Par exemple, les maladies à transmission sexuelle ont une dynamique très particulière. Lipsitch et Nowak (1995) se sont intéressés à l’évolution de la virulence du VIH. Ils ont pris en considération la démographie de la population d’hôte ainsi que le comportement sexuel des hôtes (en termes

d'acquisition de partenaires) et développé un système où deux souches circulent, une fortement virulente et une faiblement virulente. Ils ont mis en évidence un lien complexe entre la croissance de la population, la vitesse d'acquisition des partenaires et l'évolution de la virulence, montrant que celle-ci peut varier selon l'état de la population. [Alizon et van Baalen \(2008b\)](#) ont produit un modèle s'intéressant à l'évolution de la virulence dans le cas d'une maladie transmise par un vecteur (la malaria). Ils ont montré que dans ce système il n'y a pas de sélection pour une très forte virulence chez l'hôte humain, puisque la virulence chez l'humain n'est pas positivement corrélée à la transmission du pathogène et pourrait, au contraire, freiner sa propagation.

Il a également souvent été montré que l'hétérogénéité dans la population d'hôte avait un impact important sur l'évolution de la virulence des pathogènes. Dans leur modèle présenté précédemment, [Alizon et van Baalen \(2008a\)](#) montrent que c'est la coexistence d'hôtes susceptibles et d'hôtes infectés qui conduit à la diversification des souches en circulation. [Torstenson et Shaw \(2024\)](#) montrent que lorsque deux populations entrent en contact (*e.g.*, une population résidente et une migrante), des différences de caractéristique des populations, notamment dans leur réponse à un pathogène, peuvent impacter à la fois l'évolution de la virulence du pathogène et la taille de la population. [Gandon \(2004\)](#) montre que dans les systèmes multihôtes (*i.e.*, les systèmes où le pathogène peut infecter plusieurs types d'hôtes, comme par exemple plusieurs espèces) l'évolution de la virulence est largement conditionnée par les structures des populations d'hôtes ainsi que par la manière dont les traits du pathogène interagissent entre eux et dont le pathogène est transmis entre types ou au sein d'un type d'hôte.

L'impact de la vaccination sur l'évolution des pathogènes est au cœur de mon sujet de thèse. [Gandon et al. \(2001, 2003\)](#) ont développé des modèles généralistes montrant que les vaccins imparfaits (Encadré 1.4), et notamment les vaccins anatoxiques, conduisent à la sélection de souches de plus forte virulence en réduisant le coût associé à la virulence du pathogène. Par la suite, d'autres modèles ont été développés pour approfondir cette question du rôle des vaccins imparfaits. Par exemple, [Walter et Lion \(2021\)](#) montrent que les effets des vaccins (parfaits et imparfaits) sur l'évolution de la virulence peuvent être mitigés par des changements périodiques dans la couverture vaccinale et par la valeur moyenne de cette couverture vaccinale. D'autres modèles ont été développés sur des systèmes hôtes-pathogènes particuliers. Ainsi, [van Boven et al. \(2005\)](#) montrent pour la coqueluche que la sélection de souches de plus forte virulence

lorsqu'un vaccin imparfait circule dépend en réalité de la manière dont les différentes souches sont capables d'exploiter les hôtes susceptibles et vaccinés. Sur le même système biologique, [Bernhauerová \(2016\)](#) propose un modèle structuré en âge et montre que cette structure n'a pas d'impact sur l'évolution de la virulence, mais conclut comme [Gandon *et al.* \(2001, 2003\)](#) que l'utilisation d'un vaccin imparfait conduit à l'évolution vers une plus forte virulence. À l'inverse, pour la diphtérie, j'ai montré que l'utilisation d'un vaccin imparfait favorise la propagation de souches de plus faible virulence à cause du coût associé à la production d'un facteur de virulence ([Lecorvaisier *et al.*, 2024](#)). Ce travail de recherche est présenté dans le Chapitre 2.

La modélisation permet également de déchiffrer la manière dont des interactions complexes influencent l'évolution de la virulence. [Wood et Ashby \(2023\)](#) ont développé un modèle pour étudier la manière dont les interactions hôte-parasite-hyperparasite (un hyperparasite étant un parasite de parasite) pouvaient moduler la virulence des parasites. Leur modèle montre que de faibles changements dans les caractéristiques des hyperparasites peuvent mener à de grandes perturbations dans la virulence des parasites (augmentation ou réduction de la virulence voire coexistence de souches de forte et de faible virulence). Notamment, les chercheurs montrent que les parasites évoluent vers une plus forte virulence lorsque l'hyperparasite réduit le taux de transmission du parasite. Également, lorsque l'hyperparasite réduit la virulence effective du parasite, il sélectionne des parasites de plus forte virulence, de manière analogue aux vaccins imparfaits dans le modèle de [Gandon *et al.* \(2001\)](#).

Pour terminer cette section, il est intéressant de citer la critique de [Day et Proulx \(2004\)](#) sur les méthodes classiques de modélisation de l'évolution de la virulence, qu'ils désignent comme étant « *game-theoretic* ». Cette appellation implique que ces modèles se concentrent sur la stratégie évolutivement stable de la virulence; autrement dit, ces modèles ne considèrent la virulence qu'à l'équilibre et ne tiennent pas compte des états transitoires par lesquels elle peut passer. Ils proposent de modéliser directement les fluctuations de la virulence au cours du temps. Cette approche permet, par exemple, de découvrir que la virulence peut initialement augmenter dans un système avant de converger vers une valeur plus faible, en lien notamment avec la structure changeante de la population d'hôtes ([Day et Proulx, 2004](#); [Berngruber *et al.*, 2013](#); [Benhamou *et al.*, 2024](#)).

1.5.4 – Analyse de données

L'épidémiologie est en grande partie une science des données : lorsqu'il présente les bénéfices de la variolisation à l'Académie des sciences, Bernoulli utilise les tables de mortalité de la variole pour montrer que l'inoculation permet de sauver bien plus de vies qu'elle n'en prend (Bernoulli et Chapelle, 2023) et depuis plus de deux siècles les autorités publiques collectent des données sur la manière dont les épidémies affectent les populations. D'abord principalement descriptives, les statistiques évoluent lorsque les mathématiques se développent. Aujourd'hui, l'analyse de données permet de répondre à une multitude de questions dans le domaine de l'épidémiologie.

Pour reprendre la célèbre doctrine du statisticien britannique George Box (1919 – 2013), « tous les modèles sont faux mais certains sont utiles »³² (Box, 1976, 1979). Dans la Section 1.5.2, j'ai présenté une certaine variété de modèles dynamiques, souvent très théoriques mais parfois plus appliqués. La collecte et l'analyse de données permet de rendre ces modèles « moins faux » en ajustant leurs paramètres pour rendre leurs estimations plus proches de la réalité. Par exemple, dans l'étude de Rachah et Torres (2015) sur l'épidémie d'Ebola au Libéria en 2014, la collecte de données a permis de paramétrer convenablement leur modèle pour qu'il prédise une dynamique épidémique proche de la réalité. Dans l'étude de Eisenberg *et al.* (2015) sur l'épidémie de choléra en Angola en 2006, c'est la comparaison statistique entre les données et les résultats de leurs modèles qui leur a permis de conclure à un effet à la fois de la transmission par l'eau et entre individus dans la dynamique de l'épidémie.

L'analyse de données est également essentielle pour vérifier les prédictions des modèles (Bull et Luring, 2014). Au cours d'une même étude, Berngruber *et al.* (2013) produisent à la fois un modèle théorique et une expérience d'évolution *in vitro* pour prédire la manière dont la virulence d'une population de bactériophages devrait évoluer au cours du temps et pour vérifier que les données biologiques concordent avec les prédictions du modèle. De la même manière, Tuffet *et al.* (2024) ont couplé une expérience de manipulation du génome de la bactérie *Acinetobacter baumannii* avec un modèle dynamique stochastique pour comprendre les mécanismes menant à l'acquisition et à la propagation de gènes de résistance aux antibiotiques chez ce pathogène et vérifier que les conclusions obtenues mathématiquement coïncidaient avec celles obtenues

32. « *all models are wrong but some are useful* »

expérimentalement.

Plus habituellement, les études de modélisation théorique et de vérification expérimentale/de terrain sont découplées. Dans la Section 1.5.2, j’ai expliqué que plusieurs modèles avaient été développés pour prédire l’impact que pourraient avoir des mesures de santé publique comme la fermeture des écoles sur les épidémies. Plusieurs études statistiques montrent que cette mesure est effectivement efficace pour réduire les contacts entre individus et donc la transmission des pathogènes (*e.g.*, [Cauchemez et al., 2008](#); [Litvinova et al., 2019](#)). Ce type de vérification par les données permet de crédibiliser l’adoption de politiques de santé publique parfois coûteuses (économiquement ou socialement). De la même manière, au cours de la récente pandémie de Covid-19, plusieurs études se sont intéressées à la manière dont les confinements ont permis de réduire le nombre de cas en utilisant des méthodes statistiques et des sources de données variées (*e.g.*, [Alfano et Ercolano, 2020](#); [Moris et Schizas, 2020](#); [Vinceti et al., 2020](#)).

On se rend facilement compte que la modélisation et l’analyse de données sont complémentaires, rendant parfois la frontière entre les deux approches assez floue. Par exemple, [Yeung et al. \(2017\)](#) ont amélioré un modèle publié précédemment par [Crowcroft et al. \(2003\)](#) pour estimer le nombre de cas et de morts dus à la coqueluche en 2014. Le modèle repose sur l’utilisation de données démographiques et épidémiologiques fournies notamment par l’OMS ainsi que sur l’utilisation d’un certain nombre de paramètres (efficacité vaccinale, taux de mortalité...), mélangeant ainsi des éléments propres à l’analyse statistique et des éléments généralement rapprochés à la modélisation dynamique.

Tout comme les modèles sont par nature faux, les données collectées – et subséquemment analysées – ne représentent que partiellement la réalité. L’obtention des données est généralement biaisée, et c’est particulièrement vrai en épidémiologie, où celle-ci est peu contrôlée à cause de la nature même des épidémies. Par exemple, les individus asymptomatiques sont rarement connus, réduisant les estimations de la prévalence des pathogènes. Il a été estimé que les cas de coqueluche rapportés aux autorités sanitaires aux États-Unis entre 1985 et 1988 ne représentaient que 11,6 % des cas véritables ([Sutter et Cochi, 1992](#)). Ce biais d’échantillonnage représente également un problème dans les travaux s’intéressant à l’évolution de la virulence des pathogènes ; il est beaucoup plus probable d’isoler une souche de forte virulence provoquant une forme symptomatique claire d’une maladie qu’une souche de faible virulence provoquant une forme asymptomatique.

Par exemple, quoique la majorité des souches isolées de *B. pertussis* produisent la toxine pertussique – facteur de virulence principal de la bactérie – j’ai présenté dans la Section 1.4.3 une étude sud-africaine montrant qu’en réalité environ quatre individus sur cinq seraient infectés par une souche ne produisant pas la toxine, donc des souches de faible virulence. Ainsi, bien qu’elle soit un outil essentiel, la collecte et l’analyse de données est, comme la modélisation, un outil fatalement sujet à certaines limites (Section 5.2).

1.6 – Conclusion

Les pathogènes nous accompagnent depuis la naissance de l’humanité. Quand *Homo sapiens* quitte l’Afrique et part à la conquête du monde, il emmène avec lui de nombreux virus et bactéries – et autres parasites. Depuis, il a coévolué avec eux pendant des dizaines de milliers d’années, voire des millions d’années si on remonte au-delà de la lignée humaine. En exploitant à déraison son environnement, *Homo sapiens* s’est exposé à de nouveaux pathogènes, de plus en plus nombreux, ravageant des populations de plus en plus grandes. Finalement, il a appris à combattre ces ennemis ancestraux : il a utilisé la variolisation puis la vaccination pour se protéger avant même d’être infecté. Mais la sélection naturelle a fait son œuvre, et désormais de nouvelles souches émergent : plus virulentes, capables d’échapper à la protection des vaccins...

À travers le monde, une multitude de chercheurs et chercheuses cherchent à déchiffrer les mécanismes et les règles selon lesquels les pathogènes évoluent. Les approches sont variées : collecte sur le terrain, expériences de laboratoire, modélisation mathématique... Cette dernière approche est celle que j’ai développée au cours de ma thèse. Dans le Chapitre 2, je propose un modèle dynamique pour expliquer la manière dont la vaccination contre *Corynebacterium diphtheriae*, l’agent étiologique de la diphtérie, a provoqué l’émergence et la propagation de souches de plus faible virulence. Ensuite, dans le Chapitre 3, je change de sujet biologique pour m’intéresser à *Bordetella pertussis*, l’agent étiologique de la coqueluche. Je développe plusieurs analyses statistiques pour tenter de comprendre comment des facteurs environnementaux influencent la propagation de nouvelles souches en lien avec la vaccination. Ces analyses conduisent au Chapitre 4 dans lequel je développe un modèle informatique pour simuler une épidémie dans le but d’étudier la manière dont la méthode développée au chapitre précédent peut être appliquée.

BIBLIOGRAPHIE

- ADVANI A., GUSTAFSSON L., ÅHRÉN C., MOOI F. R. et HALLANDER H. O. (2011) : « Appearance of Fim3 and *ptxP3-Bordetella pertussis* strains, in two regions of Sweden with different vaccination programs » *Vaccine*, vol. 29, n°18, p./id. 3438–3442.
- ALFANO V. et ERCOLANO S. (2020) : « The Efficacy of Lockdown Against COVID-19: A Cross-Country Panel Analysis » *Applied Health Economics and Health Policy*, vol. 18, n°4, p./id. 509–517.
- ALIZON S. et MICHALAKIS Y. (2015) : « Adaptive virulence evolution: the good old fitness-based approach » *Trends in Ecology & Evolution*, vol. 30, n°5, p./id. 248–254.
- ALIZON S. et VAN BAALEN M. (2008a) : « Multiple Infections, Immune Dynamics, and the Evolution of Virulence » *The American Naturalist*, vol. 172, n°4, p./id. E150–E168.
- ALIZON S. et VAN BAALEN M. (2008b) : « Transmission–virulence trade-offs in vector-borne diseases » *Theoretical Population Biology*, vol. 74, n°1, p./id. 6–15.
- ALVES J. M., CARNEIRO M., CHENG J. Y., LEMOS DE MATOS A., RAHMAN M. M., LOOG L., CAMPOS P. F., WALES N., ERIKSSON A., MANICA A., STRIVE T., GRAHAM S. C., AFONSO S., BELL D. J., BELMONT L., DAY J. P., FULLER S. J., MARCHANDEAU S., PALMER W. J., QUENEY G., SURRIDGE A. K., VIEIRA F. G., MCFADDEN G., NIELSEN R., GILBERT M. T. P., ESTEVES P. J., FERRAND N. et JIGGINS F. M. (2019) : « Parallel adaptation of rabbit populations to myxoma virus » *Science*, vol. 363, n°6433, p./id. 1319–1326.
- AMMAN B. R., CARROLL S. A., REED Z. D., SEALY T. K., BALINANDI S., SWANEPOEL R., KEMP A., ERICKSON B. R., COMER J. A., CAMPBELL S., CANNON D. L., KHRISTOVA M. L., ATIMNEDI P., PADDOCK C. D., KENT CROCKETT R. J., FLIETSTRA T. D., WARFIELD K. L., UNFER R., KATONGOLE-MBIDDE E., DOWNING R., TAPPERO J. W., ZAKI S. R., ROLLIN P. E., KSIAZEK T. G., NICHOL S. T. et TOWNER J. S. (2012) : « Seasonal Pulses of Marburg Virus Circulation in Juvenile *Rousettus aegyptiacus* Bats Coincide with Periods of Increased Risk of Human Infection » *PLOS Pathogens*, vol. 8, n°10, p./id. e1002877.
- ANAND N., SABARINATH A., GEETHA S. et SOMANATH S. (2020) : « Predicting the Spread of COVID-19 Using *SIR* Model Augmented to Incorporate Quarantine and Testing » *Transactions of the Indian National Academy of Engineering*, vol. 5, n°2, p./id. 141–148.
- ANDERSON R. M. et MAY R. M. (1982) : « Coevolution of hosts and parasites » *Parasitology*, vol. 85, n°2, p./id. 411–426.
- ANDERSON R. M. et MAY R. M. (1985) : « Age-related changes in the rate of disease transmission: implications for the design of vaccination programmes » *Journal of Hygiene*, vol. 94, n°3, p./id. 365–436.
- ANDERSSON S. G. E., ZOMORODIPOUR A., ANDERSSON J. O., SICHERITZ-PONTÉN T., ALSMARK U. C. M., PODOWSKI R. M., NÄSLUND A. K., ERIKSSON A.-S., WINKLER H. H. et KURLAND C. G. (1998) : « The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria » *Nature*, vol. 396, n°6707, p./id. 133–140.

-
- ANDRADE B. G. N., MARIN M. F. A., CAMBUY D. D., FONSECA E. L., SOUZA N. F. et VICENTE A. C. P. (2014) : « Complete genome sequence of a clinical *Bordetella pertussis* isolate from Brazil » *Memórias do Instituto Oswaldo Cruz*, vol. 109, n°7, p./id. 972–974.
- ASLANABADI A., GHABILI K., SHAD K., KHALILI M. et SAJADI M. M. (2015) : « Emergence of whooping cough: notes from three early epidemics in Persia » *The Lancet Infectious Diseases*, vol. 15, n°12, p./id. 1480–1484.
- BABKIN I. et BABKINA I. (2015) : « The Origin of the Variola Virus » *Viruses*, vol. 7, n°3, p./id. 1100–1112.
- BACAËR N. (2011a), *Daniel Bernoulli, d’Alembert and the inoculation of smallpox (1760)* p. 21–30, Springer Londres, Royaume-Uni.
- BACAËR N. (2011b), *McKendrick and Kermack on epidemic modelling (1926–1927)* p. 89–96, Springer Londres, Royaume-Uni.
- BACAËR N. (2011c) : « Ross and malaria (1911) » *A Short History of Mathematical Population Dynamics*, p. 65–69, Springer (Londres, Royaume-Uni).
- BARCLAY V. C., KENNEDY D. A., WEAVER V. C., SIM D., LLOYD-SMITH J. O. et READ A. F. (2014) : « The Effect of Immunodeficiency on the Evolution of Virulence: An Experimental Test with the Rodent Malaria *Plasmodium chabaudi* » *The American Naturalist*, vol. 184, n°S1, p./id. S47–S57.
- BARKOFF A.-M., MERTSOLA J., PIERARD D., DALBY T., HOEGH S. V., GUILLOT S., STEFANELLI P., VAN GENT M., BERBERS G., VESTRHEIM D., GREVE-ISDAHL M., WEHLIN L., LJUNGMAN M., FRY N. K., MARKEY K. et HE Q. (2019) : « Pertactin-deficient *Bordetella pertussis* isolates: evidence of increased circulation in Europe, 1998 to 2015 » *Eurosurveillance*, vol. 24, n°7, p./id. 1700832.
- BARRETT R., KUZAWA C. W., MCDADE T. et ARMELAGOS G. J. (1998) : « EMERGING AND RE-EMERGING INFECTIOUS DISEASES: The Third Epidemiologic Transition » *Annual Review of Anthropology*, vol. 27, p./id. 247–271.
- BART M. J., HARRIS S. R., ADVANI A., ARAKAWA Y., BOTTERO D., BOUCHEZ V., CASSIDAY P. K., CHIANG C.-S., DALBY T., FRY N. K., GAILLARD M. E., VAN GENT M., GUISSO N., HALLANDER H. O., HARVILL E. T., HE Q., VAN DER HEIDE H. G. J., HEUVELMAN K., HOZBOR D. F., KAMACHI K., KARATAEV G. I., LAN R., LUTYŃSKA A., MAHARJAN R. P., MERTSOLA J., MIYAMURA T., OCTAVIA S., PRESTON A., QUAIL M. A., SINTCHENKO V., STEFANELLI P., TONDELLA M. L., TSANG R. S. W., XU Y., YAO S.-M., ZHANG S., PARKHILL J. et MOOI F. R. (2014) : « Global Population Structure and Evolution of *Bordetella pertussis* and Their Relationship with Vaccination » *mBio*, vol. 5, n°2, p./id. e01074–14.
- BASSETT M. G., POPOV L. E. et HOLMER L. E. (2004) : « THE OLDEST-KNOWN METAZOAN PARASITE? » *Journal of Paleontology*, vol. 78, n°6, p./id. 1214–1216.
- BENHAMOU W., BLANQUART F., CHOISY M., BERNGRUBER T. W., CHOQUET R. et GANDON S. (2024) : « Evolution of virulence in emerging epidemics: from theory to experimental evolution and back » *Virus Evolution*, vol. 10, n°1, p./id. veae069.
- BERCHE P. (2007) : *Une Histoire des microbes*, John Libbey Eurotext (Montrouge, France).
- BERCHE P. (2022) : « The Spanish flu » *La Presse Médicale*, vol. 51, n°3, p./id. 104127.
- BERCHE P. (2023) : « Gain-of-function and origin of Covid19 » *La Presse Médicale*, vol. 52, n°1, p./id. 104167.
- BERNGRUBER T. W., FROISSART R., CHOISY M. et GANDON S. (2013) : « Evolution of Virulence in Emerging Epidemics » *PLOS Pathogens*, vol. 9, n°3, p./id. e1003209.

- BERNHAUEROVÁ V. (2016) : « Vaccine-driven evolution of parasite virulence and immune evasion in age-structured population: the case of pertussis » *Theoretical Ecology*, vol. 9, n°4, p./id. 431–442.
- BERNOULLI D. et CHAPPELLE D. (2023), « Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir ».
- BLACK F. L. (1966) : « Measles Endemicity in Insular Populations: Critical Community Size and Its Evolutionary Implication » *Journal of Theoretical Biology*, vol. 11, n°2, p./id. 207–211.
- BORDET J. et GENGOU O. (1906) : « LE MICROBE DE LA COQUELUCHE » *Annales de l'Institut Pasteur*, vol. 20, n°9, p./id. 731–741.
- BORISOVA O., KOMBAROVA S. Y., ZAKHAROVA N. S., VAN GENT M., ALESHKIN V. A., MAZUROVA I. et MOOI F. R. (2007) : « Antigenic Divergence between *Bordetella pertussis* Clinical Isolates from Moscow, Russia, and Vaccine Strains » *Clinical and Vaccine Immunology*, vol. 14, n°3, p./id. 234–238.
- BOUCHEZ V., BRUN D., CANTINELLI T., DORE G., NJAMKEPO E. et GUIISO N. (2009) : « First report and detailed characterization of *B. pertussis* isolates not expressing pertussis toxin or pertactin » *Vaccine*, vol. 27, n°43, p./id. 6034–6041.
- BOUCHEZ V., GUILLOT S., LANDIER A., ARMATYS N., MATCZAK S., THE FRENCH PERTUSSIS MICROBIOLOGY STUDY GROUP, TOUBIANA J. et BRISSE S. (2021) : « Evolution of *Bordetella pertussis* over a 23-year period in France, 1996 to 2018 » *Eurosurveillance*, vol. 26, n°37, p./id. 2001213.
- BOX G. E. P. (1976) : « Science and Statistics » *Journal of the American Statistical Association*, vol. 71, n°356, p./id. 791–799.
- BOX G. E. P. (1979) : « Robustness in the Strategy of Scientific Model Building » *Robustness in Statistics*, p. 201–236, Academic Press (New York, États-Unis).
- BRUEGGEMANN A. B., PAI R., CROOK D. W. et BEALL B. (2007) : « Vaccine Escape Recombinants Emerge after Pneumococcal Vaccination in the United States » *PLOS Pathogens*, vol. 3, n°11, p./id. e168.
- BRÜSSOW H. (2009) : « Europe, the bull and the Minotaur: the biological legacy of a Neolithic love story » *Environmental Microbiology*, vol. 11, n°11, p./id. 2778–2788.
- BULL J. J. et ANTIA R. (2022) : « Which 'imperfect vaccines' encourage the evolution of higher virulence? » *Evolution, Medicine, and Public Health*, vol. 10, n°1, p./id. 202–213.
- BULL J. J. et LAURING A. S. (2014) : « Theory and Empiricism in Virulence Evolution » *PLOS Pathogens*, vol. 10, n°10, p./id. e1004387.
- BULL L. B. et MULES M. W. (1944) : « An Investigation of *Myxomatosis cuniculi* with Special Reference to the Possible Use of the Disease to Control Rabbit Populations in Australia » *Journal of the Council for Scientific and Industrial Research, Australia*, vol. 17, n°2, p./id. 79–93.
- CARRAT F. et FLAHAULT A. (2007) : « Influenza vaccine: The challenge of antigenic drift » *Vaccine*, vol. 25, n°39-40, p./id. 6852–6862.
- CARRIQUIRIBORDE F., REGIDOR V., AISPURU P. M., MAGALI G., BARTEL E., BOTTERO D. et HOZBOR D. (2019) : « Rare Detection of *Bordetella pertussis* Pertactin-Deficient Strains in Argentina » *Emerging Infectious Diseases*, vol. 25, n°11, p./id. 2048–2054.
- CARTER R. et MENDIS K. N. (2002) : « Evolutionary and Historical Aspects of the Burden of Malaria » *Clinical Microbiology Reviews*, vol. 15, n°4, p./id. 564–594.

-
- CASADEVALL A. et IMPERIALE M. J. (2014) : « Risks and Benefits of Gain-of-Function Experiments with Pathogens of Pandemic Potential, Such as Influenza Virus: a Call for a Science-Based Discussion » *mBio*, vol. 5, n°4, p./id. e01730–14.
- CAUCHEMEZ S., VALLERON A.-J., BOËLLE P.-Y., FLAHAULT A. et FERGUSON N. M. (2008) : « Estimating the impact of school closure on influenza transmission from Sentinel data » *Nature*, vol. 452, n°7188, p./id. 750–754.
- CHAKRABORTY C., SHARMA A. R., BHATTACHARYA M. et LEE S.-S. (2022) : « A Detailed Overview of Immune Escape, Antibody Escape, Partial Vaccine Escape of SARS-CoV-2 and Their Emerging Variants With Escape Mutations » *Frontiers in Immunology*, vol. 13, p./id. 801522.
- CHODASEWICZ K. (2014) : « Evolution, reproduction and definition of life » *Theory in Biosciences*, vol. 133, n°1, p./id. 39–45.
- CHUA K. B., LEK KOH C., HOOI P. S., WEE K. F., KHONG J. H., CHUA B. H., CHAN Y. P., LIM M. E. et LAM S. K. (2002) : « Isolation of Nipah virus from Malaysian Island flying-foxes » *Microbes and Infection*, vol. 4, n°2, p./id. 145–151.
- CLELAND C. E. (2012) : « Life without definitions » *Synthese*, vol. 185, n°1, p./id. 125–144.
- COCKBURN W. C. (1958) : « INCIDENTS ET ACCIDENTS CONSECUTIFS A LA VACCINATION ANTICOQUELUCHEUSE » *Bulletin of the World Health Organization*, vol. 19, n°1, p./id. 109–122.
- CONTZEN M., STING R., BLAZEY B. et RAU J. (2011) : « *Corynebacterium ulcerans* from Diseased Wild Boars » *Zoonoses and Public Health*, vol. 58, n°7, p./id. 479–488.
- CORNET S., BICHET C., LARCOMBE S., FAIVRE B. et SORCI G. (2014) : « Impact of host nutritional status on infection dynamics and parasite virulence in a bird-malaria system » *Journal of Animal Ecology*, vol. 83, n°1, p./id. 256–265.
- COUNCIL ON PHARMACY AND CHEMISTRY (1931) : « PERTUSSIS VACCINES OMITTED FROM N. N. R. » *JAMA*, vol. 96, n°8, p./id. 613.
- CROWCROFT N. S., STEIN C., DUCLOS P. et BIRMINGHAM M. (2003) : « How best to estimate the global burden of pertussis? » *The Lancet Infectious Diseases*, vol. 3, n°7, p./id. 413–418.
- DAKIC G., KALLONEN T., ELOMAA A., PLJESA T., VIGNJEVIC-KRASTAVCEVIC M. et HE Q. (2010) : « *Bordetella pertussis* vaccine strains and circulating isolates in Serbia » *Vaccine*, vol. 28, n°5, p./id. 1188–1192.
- DAY T. et PROULX S. R. (2004) : « A General Theory for the Evolutionary Dynamics of Virulence » *The American Naturalist*, vol. 163, n°4, p./id. E40–E63.
- DE ROODE J. C., PANSINI R., CHEESMAN S. J., HELINSKI M. E. H., HUIJBEN S., WARGO A. R., BELL A. S., CHAN B. H. K., WALLIKER D. et READ A. F. (2005) : « Virulence and competitive ability in genetically diverse malaria infections » *Proceedings of the National Academy of Sciences*, vol. 102, n°21, p./id. 7624–7628.
- DEL VALLE S. Y., HYMAN J. M., HETHCOTE H. W. et EUBANK S. G. (2007) : « Mixing patterns between age groups in social networks » *Social Networks*, vol. 29, n°4, p./id. 539–554.
- DEWAN K. K., LINZ B., DEROCO S. E. et HARVILL E. T. (2020) : « Acellular Pertussis Vaccine Components: Today and Tomorrow » *Vaccines*, vol. 8, n°2, p./id. 217.
- DIAVATOPOULOS D. A., CUMMINGS C. A., SCHOULS L. M., BRINIG M. M., RELMAN D. A. et MOOI F. R. (2005) : « *Bordetella pertussis*, the Causative Agent of Whooping Cough, Evolved from a Distinct, Human-Associated Lineage of *B. bronchiseptica* » *PLOS Pathogens*, vol. 1, n°4, p./id. e45.

- DIEKMANN O., HEESTERBEEK J. A. P. et METZ J. A. J. (1990) : « On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations » *Journal of Mathematical Biology*, vol. 28, n°4.
- DIEKMANN O., HEESTERBEEK J. A. P. et ROBERTS M. G. (2010) : « The construction of next-generation matrices for compartmental epidemic models » *Journal of The Royal Society Interface*, vol. 7, n°47, p./id. 873–885.
- DINC G. et ULMAN Y. I. (2007) : « The introduction of variolation ‘A La Turca’ to the West by Lady Mary Montagu and Turkey’s contribution to this » *Vaccine*, vol. 25, n°21, p./id. 4261–4265.
- DOBSON A., LAFFERTY K. D., KURIS A. M., HECHINGER R. F. et JETZ W. (2008) : « Homage to Linnaeus: How many parasites? How many hosts? » *Proceedings of the National Academy of Sciences*, vol. 105, n°supplement_1, p./id. 11482–11489.
- DOMENECH DE CELLÈS M., ROHANI P. et KING A. A. (2019) : « Duration of Immunity and Effectiveness of Diphtheria-Tetanus–Acellular Pertussis Vaccines in Children » *JAMA Pediatrics*, vol. 173, n°6, p./id. 588–594.
- DOS SANTOS M. I. M. A., PACHECO S. R., STOCKER A., SCHINONI M. I., PARANÁ R., REIS M. G. et SILVA L. K. (2017) : « Mutations associated with drug resistance and prevalence of vaccine escape mutations in patients with chronic hepatitis B infection » *Journal of Medical Virology*, vol. 89, n°10, p./id. 1811–1816.
- DÛX A., LEQUIME S., PATRONO L. V., VRANCKEN B., BORAL S., GOGARTEN J. F., HILBIG A., HORST D., MERKEL K., PREPOINT B., SANTIBANEZ S., SCHLOTTERBECK J., SUCHARD M. A., ULRICH M., WIDULIN N., MANKERTZ A., LEENDERTZ F. H., HARPER K., SCHNALKE T., LEMEY P. et CALVIGNAC-SPENCER S. (2020) : « Measles virus and rinderpest virus divergence dated to the sixth century BCE » *Science*, vol. 368, n°6497, p./id. 1367–1370.
- EBERT D. (1998) : « Experimental Evolution of Parasites » *Science*, vol. 282, n°5393, p./id. 1432–1436.
- EDMONDS J. W., NOLAN I. F., SHEPHERD R. C. H. et GOCS A. (1975) : « Myxomatosis: the virulence of field strains of myxoma virus in a population of wild rabbits (*Oryctolagus cuniculus* L.) with high resistance to myxomatosis » *Journal of Hygiene*, vol. 74, n°3, p./id. 417–418.
- EDWARDS B., HUNT A. C. et HOSKISSON P. A. (2011) : « Recent cases of non-toxigenic *Corynebacterium diphtheriae* in Scotland: justification for continued surveillance » *Journal of Medical Microbiology*, vol. 60, n°4, p./id. 561–562.
- EDWARDS K. M. et DECKER M. D. (2012) : « Pertussis vaccines » *Vaccines* 6^{ème} édition, p. 447–492, Elsevier (Amsterdam, Pays-Bas).
- EHRENGUT W. (1978) : « WHOOPING-COUGH VACCINATION: Comment on Report from Joint Committee on Vaccination and Immunisation » *The Lancet*, vol. 311, n°8060, p./id. 370–371.
- EISENBERG T., MAUDER N., CONTZEN M., RAU J., EWERS C., SCHLEZ K., ALTHOFF G., SCHAUERTE N., GEIGER C., MARGOS G., KONRAD R. et SING A. (2015) : « Outbreak with clonally related isolates of *Corynebacterium ulcerans* in a group of water rats » *BMC Microbiology*, vol. 15, p./id. 42.
- EL ALBANI A., BENGTON S., CANFIELD D. E., BEKKER A., MACCHIARELLI R., MAZURIER A., HAMMARLUND E. U., BOULVAIS P., DUPUY J.-J., FONTAINE C., FÛRSICH F. T., GAUTHIER-LAFAYE F., JANVIER P., JAVAUX E., OSSA F. O., PIERSON-WICKMANN A.-C., RIBOULLEAU A., SARDINI P., VACHARD D., WHITEHOUSE M. et MEUNIER A. (2010) : « Large colonial organisms with coordinated growth in oxygenated environments 2.1 Gyr ago » *Nature*, vol. 466, n°7302, p./id. 100–104.

-
- ELOMAA A., ADVANI A., DONNELLY D., ANTILA M., MERTSOLA J., HALLANDER H. et HE Q. (2005) : « Strain Variation among *Bordetella pertussis* Isolates in Finland, Where the Whole-Cell Pertussis Vaccine Has Been Used for 50 Years » *Journal of Clinical Microbiology*, vol. 43, n°8, p./id. 3681–3687.
- ESCUTENAIRE S., CHALON P., VERHAGEN R., HEYMAN P., THOMAS I., KARELLE-BUI L., AVSIC-ZUPANC T., LUNDKVIST Å., PLYUSNIN A. et PASTORET P.-P. (2000) : « Spatial and temporal dynamics of Puumala hantavirus infection in red bank vole (*Clethrionomys glareolus*) populations in Belgium » *Virus Research*, vol. 67, n°1, p./id. 91–107.
- EWALD P. W. (1995) : « THE EVOLUTION OF VIRULENCE: A UNIFYING LINK BETWEEN PARASITOLOGY AND ECOLOGY » *The Journal of Parasitology*, vol. 81, n°5, p./id. 659.
- EWALD P. W. (2004a) : « Evolution of virulence » *Infectious Disease Clinics of North America*, vol. 18, n°1, p./id. 1–15.
- EWALD P. W. (2004b) : « Imperfect Vaccines and the Evolution of Pathogen Virulence » *Novel Vaccination Strategies*, p. 547–565, Wiley (Weinheim, Allemagne).
- FARFOUR E., BADELL E., DINU S., GUILLOT S. et GUISSO N. (2013) : « Microbiological changes and diversity in autochthonous non-toxicogenic *Corynebacterium diphtheriae* isolated in France » *Clinical Microbiology and Infection*, vol. 19, n°10, p./id. 980–987.
- FARRINGTON D. O. et JORGENSON R. D. (1976) : « PREVALENCE OF *Bordetella bronchiseptica* IN CERTAIN WILD MAMMALS AND BIRDS IN CENTRAL IOWA » *Journal of Wildlife Diseases*, vol. 12, n°4, p./id. 523–525.
- FENG D.-F., CHO G. et DOOLITTLE R. F. (1997) : « Determining divergence times with a protein clock: Update and reevaluation » *Proceedings of the National Academy of Sciences*, vol. 94, n°24, p./id. 13028–13033.
- FENNER F. (1953) : « CHANGES IN THE MORTALITYRATE DUE TO MYXOMATOSIS IN THE AUSTRALIAN WILD RABBIT » *Nature*, vol. 172, n°4371, p./id. 228–230.
- FENNER F. (1983) : « The Florey Lecture, 1983: Biological control, as exemplified by smallpox eradication and myxomatosis » *Proceedings of the Royal Society of London. Series B. Biological Sciences*, vol. 218, n°1212, p./id. 259–285.
- FENNER F., MARSHALL I. D. et WOODROOFE G. M. (1953) : « STUDIES IN THE EPIDEMIOLOGY OF INFECTIOUS MYXOMATOSIS OF RABBITS: I. RECOVERY OF AUSTRALIAN WILD RABBITS (*ORYCTOLAGUS CUNICULUS*) FROM MYXOMATOSIS UNDER FIELD CONDITIONS » *Journal of Hygiene*, vol. 51, n°2, p./id. 225–244.
- FERGUSON H. M., MACKINNON M. J., CHAN B. H. et READ A. F. (2003) : « MOSQUITO MORTALITY AND THE EVOLUTION OF MALARIA VIRULENCE » *Evolution*, vol. 57, n°12, p./id. 2792–2804.
- FINE P. E. M. (1993) : « Herd Immunity: History, Theory, Practice » *Epidemiologic Reviews*, vol. 15, n°2, p./id. 265–302.
- FORTERRE P. (2006) : « The origin of viruses and their possible roles in major evolutionary transitions » *Virus Research*, vol. 117, n°1, p./id. 5–16.
- FROISSART R., DOUMAYROU J., VUILLAUME F., ALIZON S. et MICHALAKIS Y. (2010) : « The virulence–transmission trade-off in vector-borne plant viruses: a review of (non-)existing studies » *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 365, n°1548, p./id. 1907–1918.
- FRY N. K., NEAL S., HARRISON T. G., MILLER E., MATTHEWS R. et GEORGE R. C. (2001) : « Genotypic Variation in the *Bordetella pertussis* Virulence Factors Pertactin and Pertussis Toxin in Historical and Recent Clinical Isolates in the United Kingdom » *Infection and Immunity*, vol. 69, n°9, p./id. 5520–5528.

- FU Y., LUO G., SPELLBERG B. J., EDWARDS, Jr J. E. et IBRAHIM A. S. (2008) : « Gene Overexpression/Suppression Analysis of Candidate Virulence Factors of *Candida albicans* » *Eukaryotic Cell*, vol. 7, n°3, p./id. 483–492.
- FURUSE Y., SUZUKI A. et OSHITANI H. (2010) : « Origin of measles virus: divergence from rinderpest virus between the 11th and 12th centuries » *Virology Journal*, vol. 7, n°1, p./id. 52.
- GANDON S. (2004) : « EVOLUTION OF MULTIHOST PARASITES » *Evolution*, vol. 58, n°3, p./id. 455–469.
- GANDON S., MACKINNON M., NEE S. et READ A. (2003) : « Imperfect vaccination: some epidemiological and evolutionary consequences » *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 270, p./id. 1129–1136.
- GANDON S., MACKINNON M. J., NEE S. et READ A. F. (2001) : « Imperfect vaccines and the evolution of pathogen virulence » *Nature*, vol. 414, n°6865, p./id. 751–756.
- GEOGHEGAN J. L. et HOLMES E. C. (2018) : « The phylogenomics of evolving virus virulence » *Nature Reviews Genetics*, vol. 19, n°12, p./id. 756–769.
- GIMENO I. M. (2008) : « Marek's disease vaccines: A solution for today but a worry for tomorrow? » *Vaccine*, vol. 26, n°s3, p./id. C31–C41.
- GLASS R., GLASS L., BEYELER W. et MIN H. (2006) : « Targeted Social Distancing Designs for Pandemic Influenza » *Emerging Infectious Diseases*, vol. 12, n°11, p./id. 1671–1681.
- GOODNOW R. A. (1980) : « Biology of *Bordetella bronchiseptica* » *Microbiological Reviews*, vol. 44, n°4, p./id. 722–738.
- GOUW D. D., HERMANS P. W. M., BOOTSMA H. J., ZOMER A., HEUVELMAN K., DIAVATOPOULOS D. A. et MOOI F. R. (2014) : « Differentially Expressed Genes in *Bordetella pertussis* Strains Belonging to a Lineage Which Recently Spread Globally » *PLOS ONE*, vol. 9, n°1, p./id. e84523.
- GROSS C. P. et SEPKOWITZ K. A. (1998) : « The Myth of the Medical Breakthrough: Smallpox, Vaccination, and Jenner Reconsidered » *International Journal of Infectious Diseases*, vol. 3, n°1, p./id. 54–60.
- GROSS R., GUZMAN C. A., SEBAIHIA M., MARTINS DOS SANTOS V. A. P., PIEPER D. H., KOEBNIK R., LECHNER M., BARTELS D., BUHRMESTER J., CHOUDHURI J. V., EBENSEN T., GAIGALAT L., HERRMANN S., KHACHANE A. N., LARISCH C., LINK S., LINKE B., MEYER F., MORMANN S., NAKUNST D., RÜCKERT C., SCHNEIKER-BEKEL S., SCHULZE K., VORHÖLTER F.-J., YEVS A T., ENGLE J. T., GOLDMAN W. E., PÜHLER A., GÖBEL U. B., GOESMANN A., BLÖCKER H., KAISER O. et MARTINEZ-ARIAS R. (2008) : « The missing link: *Bordetella petrii* is endowed with both the metabolic versatility of environmental bacteria and virulence traits of pathogenic Bordetellae » *BMC Genomics*, vol. 9, p./id. 449.
- GUILFOILE P. (2009) : *Deadly Diseases and Epidemics: Diphtheria*, Chelsea House (New York, États-Unis).
- GUISSO N., SOUBEYRAND B. et MACINA D. (2022) : « Can vaccines control bacterial virulence and pathogenicity? *Bordetella pertussis*: the advantage of fitness over virulence » *Evolution, Medicine, and Public Health*, vol. 10, n°1, p./id. 363–370.
- HALLANDER H. O., ADVANI A., DONNELLY D., GUSTAFSSON L. et CARLSSON R.-M. (2005) : « Shifts of *Bordetella pertussis* Variants in Sweden from 1970 to 2003, during Three Periods Marked by Different Vaccination Programs » *Journal of Clinical Microbiology*, vol. 43, n°6, p./id. 2856–2865.

-
- HALLORAN M. E., STRUCHINER C. J. et SPIELMAN A. (1989) : « Modeling Malaria Vaccines II: Population Effects of Stage-Specific Malaria Vaccines Dependent on Natural Boosting » *Mathematical Biosciences*, vol. 94, n°1, p./id. 115–149.
- HALPIN K., YOUNG P. L., FIELD H. E. et MACKENZIE J. S. (2000) : « Isolation of Hendra virus from pteropid bats: a natural reservoir of Hendra virus » *Journal of General Virology*, vol. 81, n°8, p./id. 1927–1932.
- HARJAI K., CHHIBBER S., RAO BHAI L. N. et SHARMA S. (1994) : « Introduction of plasmid carrying an incomplete set of genes for aerobactin production alters virulence of *Escherichia coli* HB101 » *Microbial Pathogenesis*, vol. 17, n°4, p./id. 261–270.
- HARPER K. N. et ARMELAGOS G. J. (2013) : « Genomics, the Origins of Agriculture, and Our Changing Microbe-Scape: Time to Revisit Some Old Tales and Tell Some New Ones » *American Journal of Physical Anthropology*, vol. 152, n°S57, p./id. 135–152.
- HEGERLE N., DORE G. et GUISSO N. (2014) : « Pertactin deficient *Bordetella pertussis* present a better fitness in mice immunized with an acellular pertussis vaccine » *Vaccine*, vol. 32, n°49, p./id. 6597–6600.
- HERFST S., SCHRAUWEN E. J. A., LINSTER M., CHUTINIMITKUL S., DE WIT E., MUNSTER V. J., SORRELL E. M., BESTEBROER T. M., BURKE D. F., SMITH D. J., RIMMELZWAAN G. F., OSTERHAUS A. D. M. E. et FOUCHIER R. A. M. (2012) : « Airborne Transmission of Influenza A/H5N1 Virus Between Ferrets » *Science*, vol. 336, n°6088, p./id. 1534–1541.
- HETHCOTE H. W. (1997) : « An Age-Structured Model for Pertussis Transmission » *Mathematical Biosciences*, vol. 145, n°2, p./id. 89–136.
- HEYMAN P., BAERT K., PLYUSNINA A., COCHEZ C., LUNDKVIST Å., ESBROECK M. V., GOOSSENS E., VANDENVELDE C., PLYUSNIN A. et STUYCK J. (2009) : « Serological and genetic evidence for the presence of Seoul hantavirus in *Rattus norvegicus* in Flanders, Belgium » *Scandinavian Journal of Infectious Diseases*, vol. 41, n°1, p./id. 51–56.
- HEYMAN P., PLYUSNINA A., BERNY P., COCHEZ C., ARTOIS M., ZIZI M., PIRNAY J. P. et PLYUSNIN A. (2004) : « Seoul hantavirus in Europe: first demonstration of the virus genome in wild *Rattus norvegicus* captured in France » *European Journal of Clinical Microbiology & Infectious Diseases*, vol. 23, n°9.
- HIRAMATSU Y., MIYAJI Y., OTSUKA N., ARAKAWA Y., SHIBAYAMA K. et KAMACHI K. (2017) : « Significant Decrease in Pertactin-Deficient *Bordetella pertussis* Isolates, Japan » *Emerging Infectious Diseases*, vol. 23, n°4, p./id. 699–701.
- HODGES K. E. (2008) : « Defining the problem: terminology and progress in ecology » *Frontiers in Ecology and the Environment*, vol. 6, n°1, p./id. 35–42.
- HOLMES E. C. (2011) : « What Does Virus Evolution Tell Us about Virus Origins? » *Journal of Virology*, vol. 85, n°11, p./id. 5247–5251.
- JANSEN G., CRUMMENERL L. L., GILBERT F., MOHR T., PFEFFERKORN R., THÄNERT R., ROSENSTIEL P. et SCHULENBURG H. (2015) : « Evolutionary Transition from Pathogenicity to Commensalism: Global Regulator Mutations Mediate Fitness Gains through Virulence Attenuation » *Molecular Biology and Evolution*, vol. 32, n°11, p./id. 2883–2896.
- JENNER E. (1798) : *An Inquiry into the Causes and Effects of the Variolæ Vaccinæ, a Disease Discovered in Some of the Western Counties of England, Particularly Gloucestershire, and Known by the Name of the Cow Pox*, n. d. (Londres, Royaume-Uni).
- JOHNSON R. B. (1986) : « Human Disease and the Evolution of Pathogen Virulence » *Journal of Theoretical Biology*, vol. 122, n°1, p./id. 19–24.

- KANKI P. J., HAMEL D. J., SANKALÉ J.-L., HSIEH C.-c., THIOR I., BARIN F., WOODCOCK S. A., GUÈYE-NDIAYE A., ZHANG E., MONTANO M., SIBY T., MARLINK R., NDOYE I., ESSEX M. E. et MBOUP S. (1999) : « Human Immunodeficiency Virus Type 1 Subtypes Differ in Disease Progression » *The Journal of Infectious Diseases*, vol. 179, n°1, p./id. 68–73.
- KAO R. R. (2006) : « Evolution of pathogens towards low R_0 in heterogeneous populations » *Journal of Theoretical Biology*, vol. 242, n°3, p./id. 634–642.
- KATZOURAKIS A. (2013) : « Paleovirology: inferring viral evolution from host genome sequence data » *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 368, n°1626, p./id. 20120493.
- KEELE B. F., VAN HEUVERSWYN F., LI Y., BAILES E., TAKEHISA J., SANTIAGO M. L., BIBOLLET-RUCHE F., CHEN Y., WAIN L. V., LIEGEOIS F., LOUL S., NGOLE E. M., BIENVENUE Y., DELAPORTE E., BROOKFIELD J. F. Y., SHARP P. M., SHAW G. M., PEETERS M. et HAHN B. H. (2006) : « Chimpanzee Reservoirs of Pandemic and Nonpandemic HIV-1 » *Science*, vol. 313, n°5786, p./id. 523–526.
- KEELING M. J. et EAMES K. T. D. (2005) : « Networks and epidemic models » *Journal of The Royal Society Interface*, vol. 2, n°4, p./id. 295–307.
- KENDRICK P., ELDERING G. et BOROWSKI A. J. (1939) : « A STUDY IN ACTIVE IMMUNIZATION AGAINST PERTUSSIS » *American Journal of Epidemiology*, vol. 29-SectionB, n°3, p./id. 133–153.
- KENDRICK P. L. (1940) : « SECONDARY FAMILIAL ATTACK RATES FROM PERTUSSIS IN VACCINATED AND UNVACCINATED CHILDREN » *American Journal of Epidemiology*, vol. 32-SectionA, n°3, p./id. 89–91.
- KENDRICK P. L. (1943) : « A FIELD STUDY OF ALUM-PRECIPITATED COMBINED PERTUSSIS VACCINE AND DIPHTHERIA TOXOID FOR ACTIVE IMMUNIZATION » *American Journal of Epidemiology*, vol. 38, n°2, p./id. 193–202.
- KENDRICK P. L. et WEISS E. S. (1942) : « Use of Alum-Treated Pertussis Vaccine, and of Alum-Precipitated Combined Pertussis Vaccine and Diphtheria Toxoid, for Active Immunization » *American Journal of Public Health and the Nations Health*, vol. 32, n°6, p./id. 615–626.
- KERMACK W. O. et MCKENDRICK A. G. (1927) : « 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*, vol. 115, n°772, p./id. 700–721.
- KERNER G., NEEHUS A.-L., PHILIPPOT Q., BOHLEN J., RINCHAI D., KERROUCHE N., PUEL A., ZHANG S.-Y., BOISSON-DUPUIS S., ABEL L., CASANOVA J.-L., PATIN E., LAVAL G. et QUINTANA-MURCI L. (2023) : « Genetic adaptation to pathogens and increased risk of inflammatory disorders in post-Neolithic Europe » *Cell Genomics*, vol. 3, n°2, p./id. 100248.
- KERR P. J. (2012) : « Myxomatosis in Australia and Europe: A model for emerging infectious diseases » *Antiviral Research*, vol. 93, n°3, p./id. 387–415.
- KERR P. J., CATTADORI I. M., SIM D., LIU J., HOLMES E. C. et READ A. F. (2022) : « Divergent Evolutionary Pathways of Myxoma Virus in Australia: Virulence Phenotypes in Susceptible and Partially Resistant Rabbits Indicate Possible Selection for Transmissibility » *Journal of Virology*, vol. 96, n°20, p./id. e00886–22.
- KERR P. J., LIU J., CATTADORI I., GHEDIN E., READ A. F. et HOLMES E. C. (2015) : « Myxoma Virus and the Leporipoxviruses: An Evolutionary Paradigm » *Viruses*, vol. 7, n°3, p./id. 1020–1061.
- KIMURA M. et KUNO-SAKAI H. (1988) : « Pertussis Vaccines in Japan » *Pediatrics International*, vol. 30, n°2, p./id. 143–153.

-
- KNOLL A. H. (2011) : « The Multiple Origins of Complex Multicellularity » *Annual Review of Earth and Planetary Sciences*, vol. 39, p./id. 217–239.
- KODÁDKOVÁ A., BARTOŠOVÁ-SOJKOVÁ P., HOLZER A. S. et FIALA I. (2015) : « *Bipteria vetusta* n. sp. – an old parasite in an old host: tracing the origin of myxosporean parasitism in vertebrates » *International Journal for Parasitology*, vol. 45, n°4, p./id. 269–276.
- KRUPOVIC M., DOLJA V. V. et KOONIN E. V. (2020) : « The LUCA and its complex virome » *Nature Reviews Microbiology*, vol. 18, n°11, p./id. 661–670.
- KULENKAMPPF M., SCHWARTZMAN J. S. et WILSON J. (1974) : « Neurological complications of pertussis inoculation » *Archives of Disease in Childhood*, vol. 49, n°1, p./id. 46–49.
- KUMAR S., FILIPSKI A., SWARNA V., WALKER A. et HEDGES S. B. (2005) : « Placing confidence limits on the molecular age of the human–chimpanzee divergence » *Proceedings of the National Academy of Sciences*, vol. 102, n°52, p./id. 18842–18847.
- KWIATKOWSKI D. P. (2005) : « How Malaria Has Affected the Human Genome and What Human Genetics Can Teach Us about Malaria » *The American Journal of Human Genetics*, vol. 77, n°2, p./id. 171–192.
- LAFFERTY K. D. et KURIS A. M. (2002) : « Trophic strategies, animal diversity and body size » *Trends in Ecology & Evolution*, vol. 17, n°11, p./id. 507–513.
- LAM C., OCTAVIA S., RICAFORT L., SINTCHENKO V., GILBERT G. L., WOOD N., MCINTYRE P., MARSHALL H., GUISO N., KEIL A. D., LAWRENCE A., ROBSON J., HOGG G. et LAN R. (2014) : « Rapid Increase in Pertactin-deficient *Bordetella pertussis* Isolates, Australia » *Emerging Infectious Diseases*, vol. 20, n°4, p./id. 626–633.
- LECORVAISIER F., PONTIER D., SOUBEYRAND B. et FOUCHET D. (2024) : « Using a dynamical model to study the impact of a toxoid vaccine on the evolution of a bacterium: The example of diphtheria » *Ecological Modelling*, vol. 487, p./id. 110569.
- LEE A., NOLAN A., WATSON J. et TRISTEM M. (2013) : « Identification of an ancient endogenous retrovirus, predating the divergence of the placental mammals » *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 368, n°1626, p./id. 20120503.
- LEE B. Y., BROWN S. T., COOLEY P., POTTER M. A., WHEATON W. D., VOORHEES R. E., STEBBINS S., GREFFENSTETTE J. J., ZIMMER S. M., ZIMMERMAN R. K., ASSI T.-M., BAILEY R. R., WAGENER D. K. et BURKE D. S. (2010) : « Simulating School Closure Strategies to Mitigate an Influenza Epidemic » *Journal of Public Health Management and Practice*, vol. 16, n°3, p./id. 252–261.
- LEITE D., CAMARGO C. H., KASHINO S. S., POLATTO R., MARTINS L. M., PEREIRA J. C., PAWLOSKI L., TONDELLA M. L., DE OLIVEIRA R. S. et DE ANDRADE VAZ DE LIMA L. R. (2021) : « Prevalence and characterization of pertactin deficient *Bordetella pertussis* strains in Brazil, a whole-cell vaccine country » *Vaccine: X*, vol. 8, p./id. 100103.
- LEROY E. M., KUMULUNGUI B., POURRUT X., ROUQUET P., HASSANIN A., YABA P., DÉLICAT A., PAWESKA J. T., GONZALEZ J.-P. et SWANEPOEL R. (2005) : « Fruit bats as reservoirs of Ebola virus » *Nature*, vol. 438, n°7068, p./id. 575–576.
- LEUNG T. L. F. (2017) : « Fossils of parasites: what can the fossil record tell us about the evolution of parasitism?: Parasitism and the fossil record » *Biological Reviews*, vol. 92, n°1, p./id. 410–430.
- LEVIN B. R. (1996) : « The Evolution and Maintenance of Virulence in Microparasites » *Emerging Infectious Diseases*, vol. 2, n°2, p./id. 93–102.

- LEVIN S. et PIMENTEL D. (1981) : « Selection of Intermediate Rates of Increase in Parasite-Host Systems » *The American Naturalist*, vol. 117, n°3, p./id. 308–315.
- LEWONTIN R. C. (1970) : « THE UNITS OF SELECTION » *Annual Review of Ecology and Systematics*, vol. 1, p./id. 1–18.
- LI J., LIU Z., YU C., TAN K., GUI S., ZHANG S. et SHEN Y. (2023) : « Global epidemiology and burden of tetanus from 1990 to 2019: A systematic analysis for the Global Burden of Disease Study 2019 » *International Journal of Infectious Diseases*, vol. 132, p./id. 118–126.
- LI S., MA C., HAO L., SU Q., AN Z., MA F., XIE S., XU A., ZHANG Y., DING Z., LI H., CAIRNS L., WANG H., LUO H., WANG N., LI L. et FERRARI M. J. (2017) : « Demographic transition and the dynamics of measles in six provinces in China: A modeling study » *PLOS Medicine*, vol. 14, n°4, p./id. e1002255.
- LINZ B., IVANOV Y. V., PRESTON A., BRINKAC L., PARKHILL J., KIM M., HARRIS S. R., GOODFIELD L. L., FRY N. K., GORRINGE A. R., NICHOLSON T. L., REGISTER K. B., LOSADA L. et HARVILL E. T. (2016) : « Acquisition and loss of virulence-associated factors during genome evolution and speciation in three clades of *Bordetella* species » *BMC Genomics*, vol. 17, p./id. 767.
- LION S. et METZ J. A. J. (2018) : « Beyond R_0 maximisation: on pathogen evolution and environmental dimensions » *Trends in Ecology & Evolution*, vol. 33, n°6, p./id. 458–473.
- LIPSITCH M. et NOWAK M. A. (1995) : « The Evolution of Virulence in Sexually Transmitted HIV/AIDS » *Journal of Theoretical Biology*, vol. 174, n°4, p./id. 427–440.
- LITVINOVA M., LIU Q.-H., KULIKOV E. S. et AJELLI M. (2019) : « Reactive school closure weakens the network of social interactions and reduces the spread of influenza » *Proceedings of the National Academy of Sciences*, vol. 116, n°27, p./id. 13174–13181.
- LIU W., LI Y., LEARN G. H., RUDICELL R. S., ROBERTSON J. D., KEELE B. F., NDJANGO J.-B. N., SANZ C. M., MORGAN D. B., LOCATELLI S., GONDER M. K., KRANZUSCH P. J., WALSH P. D., DELAPORTE E., MPOUDI-NGOLE E., GEORGIEV A. V., MULLER M. N., SHAW G. M., PEETERS M., SHARP P. M., RAYNER J. C. et HAHN B. H. (2010) : « Origin of the human malaria parasite *Plasmodium falciparum* in gorillas » *Nature*, vol. 467, n°7314, p./id. 420–425.
- LOCHER C. P., WITT S. A., HERNDIER B. G., ABBEY N. W., TENNER-RACZ K., RACZ P., KIVIAT N. B., MURTHY K. K., BRASKY K., LELAND M. et LEVY J. A. (2003) : « Increased Virus Replication and Virulence after Serial Passage of Human Immunodeficiency Virus Type 2 in Baboons » *Journal of Virology*, vol. 77, n°1, p./id. 77–83.
- LOCHT C. (2018) : « Will we have new pertussis vaccines? » *Vaccine*, vol. 36, n°36, p./id. 5460–5469.
- LOCONSOLE D., DE ROBERTIS A. L., MOREA A., METALLO A., LOPALCO P. L. et CHIRONNA M. (2018) : « Resurgence of Pertussis and Emergence of the *Ptxp3* Toxin Promoter Allele in South Italy » *Pediatric Infectious Disease Journal*, vol. 37, n°5, p./id. e126–e131.
- LOOLIJEN R. C. et VAN ANDEL J. (1999) : « Ecological communities: conceptual problems and definitions » *Perspectives in Plant Ecology, Evolution and Systematics*, vol. 2, n°2, p./id. 210–222.
- LUISI P. L. (1998) : « ABOUT VARIOUS DEFINITIONS OF LIFE » *Origins of Life and Evolution of the Biosphere*, vol. 28, n°4-6, p./id. 613–622.
- MA L., CAULFIELD A., DEWAN K. K. et HARVILL E. T. (2021) : « Pertactin-Deficient *Bordetella pertussis*, Vaccine-Driven Evolution, and Reemergence of Pertussis » *Emerging Infectious Diseases*, vol. 27, n°6, p./id. 1561–1566.

-
- MACKINNON M. J. et READ A. F. (1999) : « GENETIC RELATIONSHIPS BETWEEN PARASITE VIRULENCE AND TRANSMISSION IN THE RODENT MALARIA *PLASMODIUM CHABAUDI* » *Evolution*, vol. 53, n°3, p./id. 689–703.
- MADSEN T. (1925) : « WHOOPING COUGH: Its Bacteriology, Diagnosis, Prevention and Treatment » *The Boston Medical and Surgical Journal*, vol. 192, n°2, p./id. 50–60.
- MADSEN T. (1933) : « VACCINATION AGAINST WHOOPING COUGH » *JAMA*, vol. 101, n°3, p./id. 187.
- MAGNER L. N. (1993) : « Diseases of the Premodern Period in Korea » *The Cambridge World History of Human Disease*, p. 392–400, Cambridge University Press (Cambridge, Royaume-Uni).
- MAGPANTAY F. M. G., KING A. A. et ROHANI P. (2019) : « Age-structure and transient dynamics in epidemiological systems » *Journal of The Royal Society Interface*, vol. 16, n°156, p./id. 20190151.
- MAIXNER F., KRAUSE-KYORA B., TURAEV D., HERBIG A., HOOPMANN M. R., HALLOWS J. L., KUSEBAUCH U., VIGL E. E., MALFERTHEINER P., MEGRAUD F., O’SULLIVAN N., CIPOLLINI G., COIA V., SAMADELLI M., ENGSTRAND L., LINZ B., MORITZ R. L., GRIMM R., KRAUSE J., NEBEL A., MOODLEY Y., RATTEI T. et ZINK A. (2016) : « The 5300-year-old *Helicobacter pylori* genome of the Iceman » *Science*, vol. 351, n°6269, p./id. 162–165.
- MALLET J. (1995) : « A species definition for the Modern Synthesis » *Trends in Ecology & Evolution*, vol. 10, n°7, p./id. 294–299.
- MARÍ SAÉZ A., WEISS S., NOWAK K., LAPEYRE V., ZIMMERMANN F., DÜX A., KÜHL H. S., KABA M., REGNAUT S., MERKEL K., SACHSE A., THIESEN U., VILLÁNYI L., BOESCH C., DABROWSKI P. W., RADONIĆ A., NITSCHKE A., LEENDERTZ S. A. J., PETTERSON S., BECKER S., KRÄHLING V., COUACY-HYMAN E., AKOUA-KOFFI C., WEBER N., SCHAADE L., FAHR J., BORCHERT M., GOGARTEN J. F., CALVIGNAC-SPENCER S. et LEENDERTZ F. H. (2015) : « Investigating the zoonotic origin of the West African Ebola epidemic » *EMBO Molecular Medicine*, vol. 7, n°1, p./id. 17–23.
- MARLINK R. (1996) : « Lessons from the second AIDS virus, HIV-2 » *AIDS*, vol. 10, p./id. 689–699.
- MARSHALL I. D. et FENNER F. (1960) : « Studies in the epidemiology of infectious myxomatosis of rabbits: VII. The virulence of strains of myxoma virus recovered from Australian wild rabbits between 1951 and 1959 » *Epidemiology and Infection*, vol. 58, n°4, p./id. 485–488.
- MARTIN S. W., PAWLOSKI L., WILLIAMS M., WEENING K., DEBOLT C., QIN X., REYNOLDS L., KENYON C., GIAMBRONE G., KUDISH K., MILLER L., SELVAGE D., LEE A., SKOFF T. H., KAMIYA H., CASSIDAY P. K., TONDELLA M. L. et CLARK T. A. (2015) : « Pertactin-Negative *Bordetella pertussis* Strains: Evidence for a Possible Selective Advantage » *Clinical Infectious Diseases*, vol. 60, n°2, p./id. 223–227.
- MAY R. M. et ANDERSON R. M. (1990) : « Parasite—host coevolution » *Parasitology*, vol. 100, n°S1, p./id. S89–S101.
- MCGEACH D. J. et COOK S. (1994) : « Molecular Phylogeny of the Alphaherpesvirinae Subfamily and a Proposed Evolutionary Timescale » *Journal of Molecular Biology*, vol. 238, n°1, p./id. 9–22.
- MCGEACH D. J., COOK S., DOLAN A., JAMIESON F. E. et TELFORD E. A. (1995) : « Molecular Phylogeny and Evolutionary Timescale for the Family of Mammalian Herpesviruses » *Journal of Molecular Biology*, vol. 247, n°3, p./id. 443–458.
- McKENDRICK A. G. (1925) : « Applications of Mathematics to Medical Problems » *Proceedings of the Edinburgh Mathematical Society*, vol. 44, p./id. 98–130.
- McLEOD D. V. et GANDON S. (2022) : « Effects of epistasis and recombination between vaccine-escape and virulence alleles on the dynamics of pathogen adaptation » *Nature Ecology & Evolution*, vol. 6, n°6, p./id. 786–793.

- MEDA N., NDOYE I., M'BOUP S., WADE A., NDIAYEE S., NIANG C., SARR F., DIOP I. et CARAËL M. (1999) : « Low and stable HIV infection rates in Senegal: natural course of the epidemic or evidence for success of prevention? » *AIDS*, vol. 13, n°11, p./id. 1397–1405.
- MELIKECHI O., YOUNG A. L., TANG T., BOWMAN T., DUNSON D. et JOHNDROW J. (2022) : « Limits of epidemic prediction using SIR models » *Journal of Mathematical Biology*, vol. 85, n°4, p./id. 36.
- MERCER A. J. (1985) : « Smallpox and Epidemiological-Demographic Change in Europe: The Role of Vaccination » *Population Studies*, vol. 39, n°2, p./id. 287–307.
- MICHEL M., SKOURTANIOTI E., PIERINI F., GUEVARA E. K., MÖTSCH A., KOCHER A., BARQUERA R., BIANCO R. A., CARLHOFF S., COPPOLA BOVE L., FREILICH S., GIFFIN K., HERMES T., HISS A., KNOLLE F., NELSON E. A., NEUMANN G. U., PAPAC L., PENSKE S., ROHRLACH A. B., SALEM N., SEMERAU L., VILLALBA-MOUCO V., ABADIE I., ALDENDERFER M., BECKETT J. F., BROWN M., CAMPUS F. G. R., CHENGHWA T., CRUZ BERROCAL M., DAMAŠEK L., DUFFETT CARLSON K. S., DURAND R., ERNÉE M., FĂNTĂNEANU C., FRENZEL H., GARCÍA ATIÉNZAR G., GUILLÉN S., HSIEH E., KARWOWSKI M., KELVIN D., KELVIN N., KHOKHLOV A., KINASTON R. L., KOROLEV A., KRETTEK K.-L., KÜSSNER M., LAI L., LOOK C., MAJANDER K., MANDL K., MAZZARELLO V., MCCORMICK M., DE MIGUEL IBÁÑEZ P., MURPHY R., NÉMETH R. E., NORDQVIST K., NOVOTNY F., OBENAU M., OLMO-ENCISO L., ONKAMO P., ORSCHIEDT J., PATRUSHEV V., PELTOLA S., ROMERO A., RUBINO S., SAJANTILA A., SALAZAR-GARCÍA D. C., SERRANO E., SHAYDULLAEV S., SIAS E., ŠLAUS M., STANČO L., SWANSTON T., TESCHLER-NICOLA M., VALENTIN F., VAN DE VIJVER K., VARNEY T. L., VIGIL-ESCALERA GUIRADO A., WATERS C. K., WEISS-KREJCI E., WINTER E., LAMNIDIS T. C., PRÜFER K., NÄGELE K., SPYROU M., SCHIFFELS S., STOCKHAMMER P. W., HAAK W., POSTH C., WARINNER C., BOS K. I., HERBIG A. et KRAUSE J. (2024) : « Ancient *Plasmodium* genomes shed light on the history of human malaria » *Nature*, vol. 631, n°8019, p./id. 125–133.
- MILNE G. J., KELSO J. K., KELLY H. A., HUBAND S. T. et McVERNON J. (2008) : « A Small Community Model for the Transmission of Infectious Diseases: Comparison of School Closure as an Intervention in Individual-Based Models of an Influenza Pandemic » *PLOS ONE*, vol. 3, n°12, p./id. e4005.
- MIYAJI Y., OTSUKA N., TOYOIZUMI-AJISAKA H., SHIBAYAMA K. et KAMACHI K. (2013) : « Genetic Analysis of *Bordetella pertussis* Isolates from the 2008–2010 Pertussis Epidemic in Japan » *PLOS ONE*, vol. 8, n°10, p./id. e77165.
- MOHAMAD K. Y., KALTENBOECK B., RAHMAN K. S., MAGNINO S., SACHSE K. et RODOLAKIS A. (2014) : « Host Adaptation of *Chlamydia pecorum* towards Low Virulence Evident in Co-Evolution of the *ompA*, *incA*, and ORF663 Loci » *PLOS ONE*, vol. 9, n°8, p./id. e103615.
- MOODY E. R. R., ÁLVAREZ-CARRETERO S., MAHENDRARAJAH T. A., CLARK J. W., BETTS H. C., DOMBROWSKI N., SZÁNTHÓ L. L., BOYLE R. A., DAINES S., CHEN X., LANE N., YANG Z., SHIELDS G. A., SZÖLLŐSI G. J., SPANG A., PISANI D., WILLIAMS T. A., LENTON T. M. et DONOGHUE P. C. J. (2024) : « The nature of the last universal common ancestor and its impact on the early Earth system » *Nature Ecology & Evolution*.
- MOOI F. R., VAN LOO I. H., VAN GENT M., HE Q., BART M. J., HEUVELMAN K. J., DE GREEFF S. C., DIAVATOPOULOS D., TEUNIS P., NAGELKERKE N. et MERTSOLA J. (2009) : « *Bordetella pertussis* Strains with Increased Toxin Production Associated with Pertussis Resurgence » *Emerging Infectious Diseases*, vol. 15, n°8, p./id. 1206–1213.
- MOOSA F., DU PLESSIS M., WEIGAND M. R., PENG Y., MOGALE D., DE GOUVEIA L., NUNES M. C., MADHI S. A., ZAR H. J., REUBENSON G., ISMAIL A., TONDELLA M. L., COHEN C., WALAZA S., VON GOTTBURG A. et WOLTER N. (2023a) : « Genomic characterization of *Bordetella pertussis* in South Africa, 2015–2019 » *Microbial Genomics*, vol. 9, n°12.

-
- MOOSA F., TEMPIA S., KLEYNHANS J., MCMORROW M., MOYES J., DU PLESSIS M., CARRIM M., TREURNICHT F. K., HELFERSCEE O., MKHENGELE T., MATHUNJWA A., MARTINSON N. A., KAHN K., LEBINA L., WAFAWANAKA F., COHEN C., VON GOTTBURG A. et WOLTER N. (2023b) : « Incidence and Transmission Dynamics of *Bordetella pertussis* Infection in Rural and Urban Communities, South Africa, 2016-2018 » *Emerging Infectious Diseases*, vol. 29, n°2, p./id. 294–303.
- MORIS D. et SCHIZAS D. (2020) : « Lockdown During COVID-19: The Greek Success » *In Vivo*, vol. 34, n°3, p./id. 1695–1699.
- MOSSONG J., HENS N., JIT M., BEUTELS P., AURANEN K., MIKOLAJCZYK R., MASSARI M., SALMASO S., SCALIA TOMBA G., WALLINGA J., HEIJNE J., SADKOWSKA-TODYS M., ROSINSKA M. et EDMUNDS W. J. (2008) : « Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases » *PLOS Medicine*, vol. 5, n°3, p./id. e74.
- MU J., JOY D. A., DUAN J., HUANG Y., CARLTON J., WALKER J., BARNWELL J., BEERLI P., CHARLESTON M. A., PYBUS O. G. et SU X.-z. (2005) : « Host Switch Leads to Emergence of *Plasmodium vivax* Malaria in Humans » *Molecular Biology and Evolution*, vol. 22, n°8, p./id. 1686–1693.
- MUSSER J. M., BEMIS D. A., ISHIKAWA H. et SELANDER R. K. (1987) : « Clonal Diversity and Host Distribution in *Bordetella bronchiseptica* » *Journal of Bacteriology*, vol. 169, n°6, p./id. 2793–2803.
- NAGARAJA T. G. (2022) : « *Corynebacterium* » *Veterinary Microbiology* 4^{ème} édition, p. 265–272, Wiley (Hoboken, États-Unis).
- NAKAYAMA T. (2013) : « Vaccine chronicle in Japan » *Journal of Infection and Chemotherapy*, vol. 19, n°5, p./id. 787–798.
- NASIR A., ROMERO-SEVERSON E. et CLAVERIE J.-M. (2020) : « Investigating the Concept and Origin of Viruses » *Trends in Microbiology*, vol. 28, n°12, p./id. 959–967.
- NGUYEN-TRUNG A. (2024), *Odysée d'Helicobacter pylori : Itinéraires, Rencontres, Contacts*, Thèse de doctorat Université Claude Bernard Lyon 1.
- NICHOLS J. E. (1951) : « RABBIT CONTROL IN AUSTRALIA: PROBLEMS AND POSSIBILITIES » *Nature*, vol. 168, n°4283, p./id. 932–934.
- NICOLLE C. et CONOR A. (1913) : « Vaccinothérapie dans la coqueluche » *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, vol. 156, n°2, p./id. 1849–1851.
- NICOSIA A. et RAPPUOLI R. (1987) : « Promoter of the Pertussis Toxin Operon and Production of Pertussis Toxin » *Journal of Bacteriology*, vol. 169, n°6, p./id. 2843–2846.
- NYAMWEYA S., HEGEDUS A., JAYE A., ROWLAND-JONES S., FLANAGAN K. L. et MACALLAN D. C. (2013) : « Comparing HIV-1 and HIV-2 infection: Lessons for viral immunopathogenesis » *Reviews in Medical Virology*, vol. 23, n°4, p./id. 221–240.
- OCTAVIA S., SINTCHENKO V., GILBERT G. L., LAWRENCE A., KEIL A. D., HOGG G. et LAN R. (2012) : « Newly Emerging Clones of *Bordetella pertussis* Carrying *prn2* and *ptxP3* Alleles Implicated in Australian Pertussis Epidemic in 2008–2010 » *The Journal of Infectious Diseases*, vol. 205, n°8, p./id. 1220–1224.
- OTSUKA N., HAN H.-J., TOYOIZUMI-AJISAKA H., NAKAMURA Y., ARAKAWA Y., SHIBAYAMA K. et KAMACHI K. (2012) : « Prevalence and Genetic Characterization of Pertactin-Deficient *Bordetella pertussis* in Japan » *PLOS ONE*, vol. 7, n°2, p./id. e31985.
- OWOLABI A. T. Y., SCHNEIDER P. et REECE S. E. (2024) : « Virulence is associated with daily rhythms in the within-host replication of the malaria parasite *Plasmodium chabaudi* » *Evolutionary Applications*, vol. 17, n°5, p./id. e13696.

- PAPPENHEIMER, Jr. A. M. et GILL D. M. (1973) : « Diphtheria: Recent studies have clarified the molecular mechanisms involved in its pathogenesis » *Science*, vol. 182, n°4110, p./id. 353–358.
- PARKHILL J., SEBAHIA M., PRESTON A., MURPHY L. D., THOMSON N., HARRIS D. E., HOLDEN M. T. G., CHURCHER C. M., BENTLEY S. D., MUNGALL K. L., CERDEÑO-TÁRRAGA A. M., TEMPLE L., JAMES K., HARRIS B., QUAIL M. A., ACHTMAN M., ATKIN R., BAKER S., BASHAM D., BASON N., CHEREVACH I., CHILLINGWORTH T., COLLINS M., CRONIN A., DAVIS P., DOGGETT J., FELTWELL T., GOBLE A., HAMLIN N., HAUSER H., HOLROYD S., JAGELS K., LEATHER S., MOULE S., NORBERCZAK H., O'NEIL S., ORMOND D., PRICE C., RABBINOWITSCH E., RUTTER S., SANDERS M., SAUNDERS D., SEEGER K., SHARP S., SIMMONDS M., SKELTON J., SQUARES R., SQUARES S., STEVENS K., UNWIN L., WHITEHEAD S., BARRELL B. G. et MASKELL D. J. (2003) : « Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica* » *Nature Genetics*, vol. 35, n°1, p./id. 32–40.
- PAVESI A. (2001) : « Origin and Evolution of GBV-C/Hepatitis G Virus and Relationships with Ancient Human Migrations » *Journal of Molecular Evolution*, vol. 53, n°2, p./id. 104–113.
- PAVESI A. (2003) : « African Origin of Polyomavirus JC and Implications for Prehistoric Human Migrations » *Journal of Molecular Evolution*, vol. 56, n°5, p./id. 564–572.
- PAVESI A. (2004) : « Detecting Traces of Prehistoric Human Migrations by Geographic Synthetic Maps of Polyomavirus JC » *Journal of Molecular Evolution*, vol. 58, n°3, p./id. 304–313.
- PAVESI A. (2005) : « Utility of JC polyomavirus in tracing the pattern of human migrations dating to prehistoric times » *Journal of General Virology*, vol. 86, n°5, p./id. 1315–1326.
- PAWLOSKI L. C., QUEENAN A. M., CASSIDAY P. K., LYNCH A. S., HARRISON M. J., SHANG W., WILLIAMS M. M., BOWDEN K. E., BURGOS-RIVERA B., QIN X., MESSONNIER N. et TONDELLA M. L. (2014) : « Prevalence and Molecular Characterization of Pertactin-Deficient *Bordetella pertussis* in the United States » *Clinical and Vaccine Immunology*, vol. 21, n°2, p./id. 119–125.
- PEETERS M., HONORÉ C., HUET T., BEDJABAGA L., OSSARI S., BUSSI P., COOPER R. W. et DELAPORTE E. (1989) : « Isolation and partial characterization of an HIV-related virus occurring naturally in chimpanzees in Gabon » *AIDS*, vol. 3, n°10, p./id. 625–630.
- PESCO P., BERGERO P., FABRICIUS G. et HOZBOR D. (2014) : « Modelling the effect of changes in vaccine effectiveness and transmission contact rates on pertussis epidemiology » *Epidemics*, vol. 7, p./id. 13–21.
- PESCO P., BERGERO P., FABRICIUS G. et HOZBOR D. (2015) : « Mathematical modeling of delayed pertussis vaccination in infants » *Vaccine*, vol. 33, n°41, p./id. 5475–5480.
- PETERSEN R. F., DALBY T., DRAGSTED D. M., MOOI F. et LAMBERTSEN L. (2012) : « Temporal Trends in *Bordetella pertussis* Populations, Denmark, 1949–2010 » *Emerging Infectious Diseases*, vol. 18, n°5, p./id. 767–774.
- PLJESA T., HE Q., DAKIC G., VIGNJEVIC-KRASTAVCEVIC M. et CIRKOVIC I. (2016) : « Polymorphism in *Bordetella Pertussis* Virulence Factors in Serbia before Switch from Whole Cell to a Cellular *Pertussis* Vaccine » *International Journal of Vaccines & Vaccination*, vol. 3, n°1, p./id. 00058.
- PLYUSNINA A., KRAJINOVIĆ L. C., MARGALETIĆ J., NIEMIMAA J., NEMIROV K., LUNDKVIST Å., MARKOTIĆ A., MILETIĆ-MEDVED M., AVŠIČ-ŽUPANC T., HENTTONEN H. et PLYUSNIN A. (2011) : « Genetic evidence for the presence of two distinct hantaviruses associated with *Apodemus* mice in Croatia and analysis of local strains » *Journal of Medical Virology*, vol. 83, n°1, p./id. 108–114.
- POINAR G. (2021) : « Fossil Record of Viruses, Parasitic Bacteria and Parasitic Protozoa » *The Evolution and Fossil Record of Parasitism: Identification and Macroevolution of Parasites* numéro 50 de Topics in Geobiology, p. 29–68, Springer (Cham, Suisse).

-
- POINAR, JR. G. et POINAR R. (2005) : « Fossil evidence of insect pathogens » *Journal of Invertebrate Pathology*, vol. 89, n°3, p./id. 243–250.
- POLAK M., ZASADA A. A., MOSIEJ E., KRYSZTOPA-GRZYBOWSKA K., WITKOWSKI L., RZECZKOWSKA M., PIEKARSKA K. et LUTYŃSKA A. (2019) : « Pertactin-deficient *Bordetella pertussis* isolates in Poland—a country with whole-cell pertussis primary vaccination » *Microbes and Infection*, vol. 21, n°3-4, p./id. 170–175.
- POOLMAN J. T. (2014) : « Shortcomings of pertussis vaccines: why we need a third generation vaccine » *Expert Review of Vaccines*, vol. 13, n°10, p./id. 1159–1162.
- POULIN R. (2014) : « Parasite biodiversity revisited: frontiers and constraints » *International Journal for Parasitology*, vol. 44, n°9, p./id. 581–589.
- POULIN R. et COMBES C. (1999) : « The Concept of Virulence: Interpretations and Implications » *Parasitology Today*, vol. 15, n°12, p./id. 474–475.
- POULIN R. et MORAND S. (2000) : « THE DIVERSITY OF PARASITES » *The Quarterly Review of Biology*, vol. 75, n°3, p./id. 277–293.
- POURRUT X., SOURIS M., TOWNER J. S., ROLLIN P. E., NICHOL S. T., GONZALEZ J.-P. et LEROY E. (2009) : « Large serological survey showing cocirculation of Ebola and Marburg viruses in Gabonese bat populations, and a high seroprevalence of both viruses in *Rousettus aegyptiacus* » *BMC Infectious Diseases*, vol. 9, p./id. 159.
- RACHAH A. et TORRES D. F. M. (2015) : « Mathematical Modelling, Simulation, and Optimal Control of the 2014 Ebola Outbreak in West Africa » *Discrete Dynamics in Nature and Society*, vol. 2015, n°1, p./id. 842792.
- RAFFEL T. R., REGISTER K. B., MARKS S. A. et TEMPLE L. (2002) : « PREVALENCE OF *BORDETELLA AVIUM* INFECTION IN SELECTED WILD AND DOMESTICATED BIRDS IN THE EASTERN USA » *Journal of Wildlife Diseases*, vol. 38, n°1, p./id. 40–46.
- RATCLIFFE F. N., MYERS K., FENNESSY B. V. et CALABY J. H. (1952) : « MYXOMATOSIS IN AUSTRALIA: A STEP TOWARDS THE BIOLOGICAL CONTROL OF THE RABBIT » *Nature*, vol. 170, n°4314, p./id. 7–11.
- REACHER M., RAMSAY M., WHITE J., DE ZOYA A., EFSTRATIOU A., MANN G., MACKAY A. et GEORGE R. C. (2000) : « Nontoxigenic *Corynebacterium diphtheriae*: An Emerging Pathogen in England and Wales? » *Emerging Infectious Diseases*, vol. 6, n°6, p./id. 640–645.
- READ A. F., BAIGENT S. J., POWERS C., KGOSANA L. B., BLACKWELL L., SMITH L. P., KENNEDY D. A., WALKDEN-BROWN S. W. et NAIR V. K. (2015) : « Imperfect Vaccination Can Enhance the Transmission of Highly Virulent Pathogens » *PLOS Biology*, vol. 13, n°7, p./id. e1002198.
- REGNERY D. C. et MARSHALL I. D. (1971) : « STUDIES IN THE EPIDEMIOLOGY OF MYXOMATOSIS IN CALIFORNIA: IV. THE SUSCEPTIBILITY OF SIX LEPORID SPECIES TO CALIFORNIAN MYXOMA VIRUS AND THE RELATIVE INFECTIVITY OF THEIR TUMORS FOR MOSQUITOES » *American Journal of Epidemiology*, vol. 94, n°5, p./id. 508–513.
- REID A. H., FANNING T. G., HULTIN J. V. et TAUBENBERGER J. K. (1999) : « Origin and evolution of the 1918 “Spanish” influenza virus hemagglutinin gene » *Proceedings of the National Academy of Sciences*, vol. 96, n°4, p./id. 1651–1656.
- REID A. H., FANNING T. G., JANCZEWSKI T. A. et TAUBENBERGER J. K. (2000) : « Characterization of the 1918 “Spanish” influenza virus neuraminidase gene » *Proceedings of the National Academy of Sciences*, vol. 97, n°12, p./id. 6785–6790.

- REID A. H., TAUBENBERGER J. K. et FANNING T. G. (2004) : « Evidence of an absence: the genetic origins of the 1918 pandemic influenza virus » *Nature Reviews Microbiology*, vol. 2, n°11, p./id. 909–914.
- RING N., DAVIES H., MORGAN J., SUNDARESAN M., TIONG A., PRESTON A. et BAGBY S. (2022) : « Comparative genomics of *Bordetella pertussis* isolates from New Zealand, a country with an uncommonly high incidence of whooping cough » *Microbial Genomics*, vol. 8, n°1.
- ROBERTS M. (2007) : « The pluses and minuses of \mathcal{R}_0 » *Journal of The Royal Society Interface*, vol. 4, n°16, p./id. 949–961.
- ROBIN M.-M. (2021) : *La Fabrique des pandémies - Préserver la biodiversité, un impératif pour la santé planétaire*, La Découverte (Paris, France).
- RODA W. C., VARUGHESE M. B., HAN D. et LI M. Y. (2020) : « Why is it difficult to accurately predict the COVID-19 epidemic? » *Infectious Disease Modelling*, vol. 5, p./id. 271–281.
- ROHANI P. et SCARPINO S. V. (2018) : « Prologue: The pertussis problem » *Pertussis: Epidemiology, Immunology and Evolution*, p. 1–5, Oxford University Press (Oxford, Royaume-Uni).
- ROHANI P., ZHONG X. et KING A. A. (2010) : « Contact Network Structure Explains the Changing Epidemiology of Pertussis » *Science*, vol. 330, n°6006, p./id. 982–985.
- ROSS R. (1908) : *Report on the Prevention of Malaria in Mauritius*, Waterlow and Sons Limited (Londres, Royaume-Uni).
- ROSS R. (1911) : *The Prevention of Malaria* 2^{ème} édition, John Murray (Londres, Royaume-Uni).
- RUMBO M. et HOZBOR D. (2014) : « Development of improved pertussis vaccine » *Human Vaccines & Immunotherapeutics*, vol. 10, n°8, p./id. 2450–2453.
- SAFARCHI A., SAEDI S., OCTAVIA S., SEDAGHATPOUR M., BOLOURCHI N., TAY C. Y., LAMICHHANE B., SHAHCHERAGHI F. et LAN R. (2021) : « Evolutionary genomics of recent clinical *Bordetella pertussis* isolates from Iran: wide circulation of multiple *ptxP3* lineages and report of the first *ptxP3* filamentous hemagglutinin-negative *B. pertussis* » *Infection, Genetics and Evolution*, vol. 93, p./id. 104970.
- SALVADORI F. et VIGNAUD L.-H. (2019) : *Antivax : La résistance aux vaccins du XVIII^e siècle à nos jours*, Vendémiaire (Paris, France).
- SANDERS K. L. et LEE M. S. Y. (2010) : « Arthropod molecular divergence times and the Cambrian origin of pentastomids » *Systematics and Biodiversity*, vol. 8, n°1, p./id. 63–74.
- SCANLON K., SKERRY C. et CARBONETTI N. (2019a) : « Association of Pertussis Toxin with Severe Pertussis Disease » *Toxins*, vol. 11, n°7, p./id. 373.
- SCANLON K., SKERRY C. et CARBONETTI N. (2019b) : « Role of Major Toxin Virulence Factors in Pertussis Infection and Disease Pathogenesis » *Pertussis Infection and Vaccines* volume 1183, p. 35–51, Springer (Cham, Suisse).
- SCHENZLE D. (1984) : « An Age-Structured Model of Pre- and Post-Vaccination Measles Transmission » *Mathematical Medicine and Biology*, vol. 1, n°2, p./id. 169–191.
- SCHMIDTKE A. J., BONEY K. O., MARTIN S. W., SKOFF T. H., TONDELLA M. L. et TATTI K. M. (2012) : « Population Diversity among *Bordetella pertussis* Isolates, United States, 1935–2009 » *Emerging Infectious Diseases*, vol. 18, n°8.
- SCHUEGGER R., SCHOERNER C., DLUGAICZYK J., LICHTENFELD I., TROUILLIER A., ZELLER-PERONNET V., BUSCH U., BERGER A., KUGLER R., HÖRMANSDORFER S. et SING A. (2009) : « Pigs as Source for Toxigenic *Corynebacterium ulcerans* » *Emerging Infectious Diseases*, vol. 15, n°8, p./id. 1314–1315.

-
- SCORPIO A., CHABOT D. J., DAY W. A., HOOVER T. A. et FRIEDLANDER A. M. (2010) : « Capsule depolymerase overexpression reduces *Bacillus anthracis* virulence » *Microbiology*, vol. 156, n°5, p./id. 1459–1467.
- SEPPÄLÄ O., LILJEROOS K., KARVONEN A. et JOKELA J. (2008) : « Host condition as a constraint for parasite reproduction » *Oikos*, vol. 117, n°5, p./id. 749–753.
- SETH A. K. et BAYNE T. (2022) : « Theories of consciousness » *Nature Reviews Neuroscience*, vol. 23, n°7, p./id. 439–452.
- SHAPIRO-SHAPIN C. G. (2010) : « Pearl Kendrick, Grace Eldering, and the Pertussis Vaccine » *Emerging Infectious Diseases*, vol. 16, n°8, p./id. 1273–1278.
- SHATTOCK A. J., JOHNSON H. C., SIM S. Y., CARTER A., LAMBACH P., HUTUBESSY R. C. W., THOMPSON K. M., BADIZADEGAN K., LAMBERT B., FERRARI M. J., JI T. M., FU H., SILAL S. P., HOUNSELL R. A., WHITE R. G., MOSSER J. F., GAYTHORPE K. A. M., TROTTER C. L., LINDSTRAND A., O'BRIEN K. L. et BAR-ZEEV N. (2024) : « Contribution of vaccination to improved survival and health: modelling 50 years of the Expanded Programme on Immunization » *The Lancet*, vol. 403, n°10441, p./id. 2307–2316.
- SHCHELKUNOV S. N. (2009) : « How long ago did smallpox virus emerge? » *Archives of Virology*, vol. 154, n°12, p./id. 1865–1871.
- SOUBEYRAND B. et PLOTKIN S. A. (2002) : « Antitoxin vaccines and pathogen virulence » *Nature*, vol. 417, n°6889, p./id. 609–610.
- SOUMANA I. H., LINZ B. et HARVILL E. T. (2017) : « Environmental Origin of the Genus *Bordetella* » *Frontiers in Microbiology*, vol. 8.
- SPENCE P. J., BRUGAT T. et LANGHORNE J. (2015) : « Mosquitoes Reset Malaria Parasites » *PLOS Pathogens*, vol. 11, n°7, p./id. e1004987.
- SPENCE P. J., JARRA W., LÉVY P., NAHRENDORF W. et LANGHORNE J. (2012) : « Mosquito transmission of the rodent malaria parasite *Plasmodium chabaudi* » *Malaria Journal*, vol. 11, n°1, p./id. 407.
- SPENCE P. J., JARRA W., LÉVY P., REID A. J., CHAPPELL L., BRUGAT T., SANDERS M., BERRIMAN M. et LANGHORNE J. (2013) : « Vector transmission regulates immune control of *Plasmodium* virulence » *Nature*, vol. 498, n°7453, p./id. 228–231.
- SPYROU M. A., BOS K. I., HERBIG A. et KRAUSE J. (2019) : « Ancient pathogen genomics as an emerging tool for infectious disease research » *Nature Reviews Genetics*, vol. 20, n°6, p./id. 323–340.
- STANGE N. (2021) : « Politics of Plague: Ancient Epidemics and Their Impact on Society » *2021 Claremont Colleges Library Undergraduate Research Award*, vol. 4.
- STEWART G. T. (1977) : « VACCINATION AGAINST WHOOPING-COUGH: Efficacy versus Risks » *The Lancet*, vol. 309, n°8005, p./id. 234–237.
- STIBITZ S. et GARLETTS T. L. (1992) : « Derivation of a physical map of the chromosome of *Bordetella pertussis* Tohama I » *Journal of Bacteriology*, vol. 174, n°23, p./id. 7770–7777.
- STING R., PÖLZELBAUER C., EISENBERG T., BONKE R., BLAZEY B., PETERS M., RISSE K., SING A., BERGER A., DANGEL A. et RAU J. (2023) : « *Corynebacterium ulcerans* Infections in Eurasian Beavers (*Castor fiber*) » *Pathogens*, vol. 12, n°8, p./id. 979.
- STRÖM J. (1960) : « IS UNIVERSAL VACCINATION AGAINST PERTUSSIS ALWAYS JUSTIFIED? » *British Medical Journal*, vol. 2, n°5207, p./id. 1184–1186.

- STRÖM J. (1967) : « Further Experience of Reactions, Especially of a Cerebral Nature, in Conjunction with Triple Vaccination: A Study Based on Vaccinations in Sweden 1959-65 » *British Medical Journal*, vol. 4, n°5575, p./id. 320–323.
- STROUHAL E. (1996) : « TRACES OF A SMALLPOX EPIDEMIC IN THE FAMILY OF RAMESSES V OF THE EGYPTIAN 20th DYNASTY » *Anthropologie*, vol. 34, n°3, p./id. 315–319.
- SUH A., BROSIUS J., SCHMITZ J. et KRIEGS J. O. (2013) : « The genome of a Mesozoic paleovirus reveals the evolution of hepatitis B viruses » *Nature Communications*, vol. 4, n°1, p./id. 1791.
- SUSAT J., LÜBKE H., IMMEL A., BRINKER U., MACĀNE A., MEADOWS J., STEER B., THOLEY A., ZAGORSKA I., GERHARDS G., SCHMÖLCKE U., KALNIŠ M., FRANKE A., PĒTERSONE-GORDINA E., TESSMAN B., TÖRV M., SCHREIBER S., ANDREE C., BĒRZIŠ V., NEBEL A. et KRAUSE-KYORA B. (2021) : « A 5,000-year-old hunter-gatherer already plagued by *Yersinia pestis* » *Cell Reports*, vol. 35, n°13, p./id. 109278.
- SUTTER R. W. et COCHI S. L. (1992) : « Pertussis Hospitalizations and Mortality in the United States, 1985-1988: Evaluation of the Completeness of National Reporting » *JAMA*, vol. 267, n°3, p./id. 386.
- SZWEJSER-ZAWISLAK E., WILK M. M., PISZCZEK P., KRAWCZYK J., WILCZYŃSKA D. et HOZBOR D. (2022) : « Evaluation of Whole-Cell and Acellular Pertussis Vaccines in the Context of Long-Term Herd Immunity » *Vaccines*, vol. 11, n°1.
- TARDY L., GIRAUDEAU M., HILL G. E., MCGRAW K. J. et BONNEAUD C. (2019) : « Contrasting evolution of virulence and replication rate in an emerging bacterial pathogen » *Proceedings of the National Academy of Sciences*, vol. 116, n°34, p./id. 16927–16932.
- TAUBENBERGER J. K., REID A. H., LOURENS R. M., WANG R., JIN G. et FANNING T. G. (2005) : « Characterization of the 1918 influenza virus polymerase genes » *Nature*, vol. 437, n°7060, p./id. 889–893.
- TAYLOR L. H., LATHAM S. M. et WOOLHOUSE M. E. J. (2001) : « Risk factors for human disease emergence » *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, vol. 356, n°1411, p./id. 983–989.
- TAYLOR M. W. (1881) : « ON THE FUNGOID ORIGIN OF DIPHTHERIA » *The British Medical Journal*, vol. 2, n°1070, p./id. 6–9.
- TIWARI T. S. P. et WHARTON M. (2012) : « Diphtheria Toxoid » *Vaccines* 6^{ème} édition, p. 153–166, Elsevier (Amsterdam, Pays-Bas).
- TORSTENSON M. et SHAW A. K. (2024) : « Pathogen evolution following spillover from a resident to a migrant host population depends on interactions between host pace of life and tolerance to infection » *Journal of Animal Ecology*, vol. 93, n°4, p./id. 1365–2656.14075.
- TOURE-KANE C., MONTAVON C., FAYE M. A., MANDOUMB P., SOW P. S., NDOYE I., GAYE-DIALLO A., DELAPORTE E., PEETERS M. et MBOUP S. (2000) : « Identification of All HIV Type 1 Group M Subtypes in Senegal, a Country with Low and Stable Seroprevalence » *AIDS Research and Human Retroviruses*, vol. 16, n°6, p./id. 603–609.
- TOWERS S., VOGT GEISSE K., TSAI C.-C., HAN Q. et FENG Z. (Thu Mar 01 00:00:00 UTC 2012) : « THE IMPACT OF SCHOOL CLOSURES ON PANDEMIC INFLUENZA: ASSESSING POTENTIAL REPERCUSSIONS USING A SEASONAL SIR MODEL » *Mathematical Biosciences & Engineering*, vol. 9, n°2, p./id. 413–430.
- TREVISANATO S. I. (2004) : « Did an epidemic of tularemia in Ancient Egypt affect the course of world history? » *Medical Hypotheses*, vol. 63, n°5, p./id. 905–910.

-
- TREVISANATO S. I. (2007) : « The ‘Hittite plague’, an epidemic of tularemia and the first record of biological warfare » *Medical Hypotheses*, vol. 69, n°6, p./id. 1371–1374.
- TRIMPERT J., GROENKE N., JENCKEL M., HE S., KUNEC D., SZPARA M. L., SPATZ S. J., OSTERRIEDER N. et MCMAHON D. P. (2017) : « A phylogenomic analysis of Marek’s disease virus reveals independent paths to virulence in Eurasia and North America » *Evolutionary Applications*, vol. 10, n°10, p./id. 1091–1101.
- TSANG R. S., SHUEL M., CRONIN K., DENG S., WHYTE K., MARCHAND-AUSTIN A., MA J., BOLOTIN S., CROWCROFT N., SCHWARTZ K., VAN DOMSELAAR G., GRAHAM M. et JAMIESON F. B. (2019) : « The evolving nature of *Bordetella pertussis* in Ontario, Canada, 2009–2017: strains with shifting genotypes and pertactin deficiency » *Canadian Journal of Microbiology*, vol. 65, n°11, p./id. 823–830.
- TSANG R. S., SHUEL M., JAMIESON F. B., DREWS S., HOANG L., HORSMAN G., LEFEBVRE B., DESAI S. et ST-LAURENT M. (2014) : « Pertactin-negative *Bordetella pertussis* strains in Canada: characterization of a dozen isolates based on a survey of 224 samples collected in different parts of the country over the last 20 years » *International Journal of Infectious Diseases*, vol. 28, p./id. 65–69.
- TUFFET R., CARVALHO G., GODEUX A.-S., MAZZAMURRO F., ROCHA E. P. C., LAABERKI M.-H., VENNER S. et CHARPENTIER X. (2024) : « Manipulation of natural transformation by AbaR-type islands promotes fixation of antibiotic resistance in *Acinetobacter baumannii* » *Proceedings of the National Academy of Sciences*, vol. 121, n°39, p./id. e2409843121.
- VAN BAALLEN M. et SABELIS M. W. (1995) : « The Dynamics of Multiple Infection and the Evolution of Virulence » *The American Naturalist*, vol. 146, n°6, p./id. 881–910.
- VAN BOVEN M., MOOI F. R., SCHELLEKENS J. F. P., DE MELKER H. E. et KRETZSCHMAR M. (2005) : « Pathogen adaptation under imperfect vaccination: implications for pertussis » *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 272, n°1572, p./id. 1617–1624.
- VAN DEN DRIESCHE P. et WATMOUGH J. (2002) : « Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission » *Mathematical Biosciences*, vol. 180, n°1-2, p./id. 29–48.
- VAN DER ZEE A., MOOI F., VAN EMBDEN J. et MUSSEY J. (1997) : « Molecular Evolution and Host Adaptation of *Bordetella* spp.: Phylogenetic Analysis Using Multilocus Enzyme Electrophoresis and Typing with Three Insertion Sequences » *Journal of Bacteriology*, vol. 179, n°21, p./id. 6609–6617.
- VAN GENT M., HEUVELMAN C. J., VAN DER HEIDE H. G., HALLANDER H. O., ADVANI A., GUISSO N., WIRSING VON KÖNIG C. H., VESTRHEIM D. F., DALBY T., FRY N. K., PIERARD D., DETEMMERMAN L., ZAVADILOVA J., FABIANOVA K., LOGAN C., HABBINGTON A., BYRNE M., LUTYŃSKA A., MOSIEJ E., PELAZ C., GRÖNDAHL-YLI-HANNUKSELA K., BARKOFF A. M., MERTSOLA J., ECONOMOPOULOU A., HE Q. et MOOI F. R. (2015) : « Analysis of *Bordetella pertussis* clinical isolates circulating in European countries during the period 1998–2012 » *European Journal of Clinical Microbiology & Infectious Diseases*, vol. 34, n°4, p./id. 821–830.
- VAN VALEN L. (1973) : « A NEW EVOLUTIONARY LAW » *Evolutionary Theory*, vol. 1, p./id. 1–30.
- VERBEKE R., LENTACKER I., DE SMEDT S. C. et DEWITTE H. (2019) : « Three decades of messenger RNA vaccine development » *Nano Today*, vol. 28, p./id. 100766.
- VINCETI M., FILIPPINI T., ROTHMAN K. J., FERRARI F., GOFFI A., MAFFEIS G. et ORSINI N. (2020) : « Lockdown timing and efficacy in controlling COVID-19 using mobile phone tracking » *eClinicalMedicine*, vol. 25, p./id. 100457.
- VODZAK J., QUEENAN A. M., SOUDER E., EVANGELISTA A. T. et LONG S. S. (2017) : « Clinical Manifestations and Molecular Characterization of Pertactin-Deficient and Pertactin-Producing *Bordetella pertussis* in Children, Philadelphia 2007–2014 » *Clinical Infectious Diseases*, vol. 64, n°1, p./id. 60–66.

- VON WINTZINGERODE F., SCHATTKA A., SIDDIQUI R. A., RÖSICK U., GÖBEL U. B. et GROSS R. (2001) : « *Bordetella petrii* sp. nov., isolated from an anaerobic bioreactor, and emended description of the genus *Bordetella*. » *International Journal of Systematic and Evolutionary Microbiology*, vol. 51, n°4, p./id. 1257–1265.
- WAKIMOTO Y., OTSUKA N., YANAGAWA Y., KOIDE K., KAMACHI K., SHIBAYAMA K., TERUYA K., KIKUCHI Y. et OKA S. (2022) : « The First Reported Case of *Bordetella pertussis* Bacteremia in a Patient With Human Immunodeficiency Virus Infection » *Open Forum Infectious Diseases*, vol. 9, n°3, p./id. ofac020.
- WALLACE B. (1989) : « CAN “STEPPING STONES” FORM STAIRWAYS? » *The American Naturalist*, vol. 133, n°4, p./id. 578–579.
- WALTER A. et LION S. (2021) : « Epidemiological and evolutionary consequences of periodicity in treatment coverage » *Proceedings of the Royal Society B: Biological Sciences*, vol. 288, n°1946, p./id. 20203007.
- WANDERSEE J. H. (1988) : « The Terminology Problem in Biology Education: A Reconnaissance » *The American Biology Teacher*, vol. 50, n°2, p./id. 97–100.
- WANG J., QIU J., ZHU Y., ZHOU H., YU L., DING Y., ZHANG L., GUO Z. et DONG C. (2017) : « Molecular evolution of hepatitis B vaccine escape variants in China, during 2000–2016 » *Vaccine*, vol. 35, n°43, p./id. 5808–5813.
- WANG R., CHEN J., GAO K. et WEI G.-W. (2021) : « Vaccine-escape and fast-growing mutations in the United Kingdom, the United States, Singapore, Spain, India, and other COVID-19-devastated countries » *Genomics*, vol. 113, n°4, p./id. 2158–2170.
- WAPLES R. S. (2022) : « What Is N_e , Anyway? » *Journal of Heredity*, vol. 113, n°4, p./id. 371–379.
- WARFEL J. M., ZIMMERMAN L. I. et MERKEL T. J. (2014) : « Acellular pertussis vaccines protect against disease but fail to prevent infection and transmission in a nonhuman primate model » *Proceedings of the National Academy of Sciences*, vol. 111, n°2, p./id. 787–792.
- WEIGAND M. R., PENG Y., LOPAREV V., BATRA D., BOWDEN K. E., BURROUGHS M., CASSIDAY P. K., DAVIS J. K., JOHNSON T., JUIENG P., KNIPE K., MATHIS M. H., PRUITT A. M., ROWE L., SHETH M., TONDELLA M. L. et WILLIAMS M. M. (2017) : « The History of *Bordetella pertussis* Genome Evolution Includes Structural Rearrangement » *Journal of Bacteriology*, vol. 199, n°8.
- WEISS R. A. et McMICHAEL A. J. (2004) : « Social and environmental risk factors in the emergence of infectious diseases » *Nature Medicine*, vol. 10, n°s12, p./id. S70–S76.
- WENDELBOE A. M., VAN RIE A., SALMASO S. et ENGLUND J. A. (2005) : « Duration of Immunity Against Pertussis After Natural Infection or Vaccination » *Pediatric Infectious Disease Journal*, vol. 24, n°5, p./id. S58–S61.
- WERTHEIM J. O., SMITH M. D., SMITH D. M., SCHEFFLER K. et KOSAKOVSKY POND S. L. (2014) : « Evolutionary Origins of Human Herpes Simplex Viruses 1 and 2 » *Molecular Biology and Evolution*, vol. 31, n°9, p./id. 2356–2364.
- WESTON R. (2012) : « Whooping Cough: A Brief History to the 19th Century » *Canadian Bulletin of Medical History*, vol. 29, n°2, p./id. 329–349.
- WILLIAMS M. M., SEN K., WEIGAND M. R., SKOFF T. H., CUNNINGHAM V. A., HALSE T. A., TONDELLA M. L. et CDC PERTUSSIS WORKING GROUP (2016) : « *Bordetella pertussis* Strain Lacking Pertactin and Pertussis Toxin » *Emerging Infectious Diseases*, vol. 22, n°2, p./id. 319–322.

-
- WILSON J. N., NOKES D. J. et CARMAN W. F. (1998) : « Current status of HBV vaccine escape variants – a mathematical model of their epidemiology » *Journal of Viral Hepatitis*, vol. 5, n°2, p./id. 25–30.
- WINDSOR D. A. (1998) : « Controversies in parasitology: Most of the species on Earth are parasites » *International Journal for Parasitology*, vol. 28, n°12, p./id. 1939–1941.
- WITTER R. L. (1997) : « Increased Virulence of Marek’s Disease Virus Field Isolates » *Avian Diseases*, vol. 41, n°1, p./id. 149.
- WOLF Y. I., KAZLAUSKAS D., IRANZO J., LUCÍA-SANZ A., KUHN J. H., KRUPOVIC M., DOLJA V. V. et KOONIN E. V. (2018) : « Origins and Evolution of the Global RNA Virome » *mBio*, vol. 9, n°6, p./id. e02329–18.
- WOOD J. et ASHBY B. (2023) : « Hyperparasitism and the evolution of parasite virulence » *Evolution*, vol. 77, n°12, p./id. 2631–2641.
- WREN M. W. D. et SHETTY N. (2005) : « Infections with *Corynebacterium diphtheriae*: six years’ experience at an inner London teaching hospital » *British Journal of Biomedical Science*, vol. 62, n°1, p./id. 1–4.
- XU Z., OCTAVIA S., LUU L. D. W., PAYNE M., TIMMS V., TAY C. Y., KEIL A. D., SINTCHENKO V., GUISSO N. et LAN R. (2019) : « Pertactin-Negative and Filamentous Hemagglutinin-Negative *Bordetella pertussis* , Australia, 2013–2017 » *Emerging Infectious Diseases*, vol. 25, n°6, p./id. 1196–1199.
- YEUNG K. H. T., DUCLOS P., NELSON E. A. S. et HUTUBESSY R. C. W. (2017) : « An update of the global burden of pertussis in children younger than 5 years: a modelling study » *The Lancet Infectious Diseases*, vol. 17, n°9, p./id. 974–980.
- ZHANG Z., STROTZ L. C., TOPPER T. P., CHEN F., CHEN Y., LIANG Y., ZHANG Z., SKOVSTED C. B. et BROCK G. A. (2020) : « An encrusting kleptoparasite-host interaction from the early Cambrian » *Nature Communications*, vol. 11, n°1, p./id. 2625.
- ZINK A., REISCHL U., WOLF H., NERLICH A. G. et MILLER R. (2001) : « *Corynebacterium* in ancient Egypt » *Medical History*, vol. 45, n°2, p./id. 267–272.
- ZOMER A., OTSUKA N., HIRAMATSU Y., KAMACHI K., NISHIMURA N., OZAKI T., POOLMAN J. et GEURTSSEN J. (2018) : « *Bordetella pertussis* population dynamics and phylogeny in Japan after adoption of acellular pertussis vaccines » *Microbial Genomics*, vol. 4, n°5, p./id. e000180.
- ZUKSWERT J. M., BARKER M. K. et McDONNELL L. (2019) : « Identifying Troublesome Jargon in Biology: Discrepancies between Student Performance and Perceived Understanding » *CBE—Life Sciences Education*, vol. 18, n°1, p./id. ar6.

CHAPITRE II

MODÉLISATION THÉORIQUE : VACCINATION ANATOXINIQUE ET ÉVOLUTION DE LA VIRULENCE

Les travaux présentés dans ce chapitre ont été partiellement publiés sous la forme d'un article paru sous licence Creative Commons (CC-BY) dans Ecological Modelling ([Lecorvaisier et al., 2024](#)).

2.1 – Introduction

2.1.1 – Retours sur l'impact des vaccins sur l'écologie et l'évolution des pathogènes

Comme présenté dans la Section 1.2, la vaccination représente l'un des plus grands progrès médicaux du XX^e siècle, période à laquelle son usage se démocratise le plus largement. Pour rappel, son but premier est de protéger les individus vaccinés contre une infection ou les symptômes d'une maladie, mais les bénéfices de la vaccination dépassent le seul individu et peuvent s'étendre à la population entière. J'ai introduit dans la Section 1.3.2 le nombre de reproduction de base d'un pathogène, communément dénoté \mathcal{R}_0 , représentant le nombre d'infections que va provoquer un unique individu infecté dans une population autrement complètement susceptible au pathogène, donc généralement en début d'épidémie. Cette valeur théorique représente un seuil clé en épidémiologie : s'il est plus bas qu'un, alors la maladie est vouée à l'extinction. À l'inverse, s'il est supérieur à un, alors la maladie peut se répandre dans la population.

Le 8 mai 1980, l'Organisation mondiale de la santé (OMS) déclare la variole éradiquée sur

tous les continents. Cette éradication a été rendue possible en partie grâce à une couverture vaccinale très haute, permettant aux populations d'atteindre l'immunité de groupe, ou *herd immunity*, faisant ainsi chuter son nombre de reproduction (Fine, 1993). Puisque la variole se transmet uniquement entre humains et que la propagation ne pouvait plus se faire au sein d'une population fortement vaccinée, la variole a fini par disparaître.

Ce type d'effet de la vaccination à l'échelle de la population peut uniquement être observé si une proportion suffisamment importante de la population est vaccinée, et si le vaccin utilisé bloque la transmission du pathogène ciblé. Ainsi, il est classiquement établi que la proportion critique d'individus à vacciner pour parvenir à éradiquer un pathogène est $1 - 1/\mathcal{R}_0$ (Fine, 1993). Dans la pratique, les choses ne sont pas si simples. Les populations sont fortement hétérogènes, et les comportements individuels vis-à-vis de la vaccination peuvent avoir des effets complexes et parfois difficiles à prédire sur le devenir d'une épidémie (Lu *et al.*, 2019 ; Zou *et al.*, 2023). Ce phénomène a été notamment décrit durant la pandémie récente de Covid-19, l'influence des groupes « antivax » (Encadré 1.2), sur les réseaux sociaux notamment, ayant possiblement contribué à augmenter la durée ainsi que les dégâts de la pandémie (*e.g.*, Pullan et Dey, 2021).

Malgré ces difficultés, et grâce à ces effets à la fois individuels et populationnels (Section 1.4), la vaccination a semblé pendant un temps être la panacée grâce à laquelle l'humanité pourrait se débarrasser de toutes les maladies infectieuses. Cependant, les vaccins peuvent également avoir des effets délétères, et certains ont la réputation d'induire de graves effets secondaires comme les vaccins à germe entier (WCV, pour *whole-cell vaccine*) utilisés contre la coqueluche (Edwards et Decker, 2012). La vaccination de masse a également eu d'autres effets inattendus, comme une augmentation de l'âge de première infection pour la rougeole (Li *et al.*, 2017) et la coqueluche (Lavine *et al.*, 2011 ; Rohani *et al.*, 2010), deux maladies touchant particulièrement les très jeunes enfants. Cela n'a rien d'anodin puisqu'il a été montré que les risques de développer des symptômes et d'être hospitalisé augmentent avec l'âge chez l'enfant et l'adolescent (mais pas chez l'adulte), à la fois pour la rougeole (Giladi *et al.*, 1987 ; Yasunaga *et al.*, 2010) et pour la coqueluche (De Serres *et al.*, 2000). La vaccination, en ciblant certaines souches de pathogènes, agit également sur leur évolution en altérant les pressions de sélection qu'ils subissent (Section 1.4). Ainsi, elle favorise les souches bactériennes ou virales capables d'échapper à l'immunité induite par les vaccins (Gandon et Day, 2008 ; McLeod et Gandon, 2022).

Un autre effet que peut avoir la vaccination de masse sur l'évolution des pathogènes est la sélection de souches mutantes ayant une virulence différente des souches sauvages. Au début des années 2000, il a été théorisé que l'utilisation de vaccins imparfaits (*leaky vaccines* ou *imperfect vaccines*), c'est-à-dire des vaccins bloquant les symptômes d'une maladie mais pas l'infection par son pathogène, devrait mener à la sélection de souches de plus forte virulence. L'argument principal de cette théorie est que de tels vaccins enlèvent le coût en termes de morbidité et de mortalité à la production de facteurs de virulence par les pathogènes sans réduire le bénéfice associé à leur production (Gandon *et al.*, 2001, 2003 ; McLeod et Gandon, 2022).

Ces travaux ont permis de mettre en évidence à la fois les risques associés aux programmes de vaccination trop simplistes, qui consisteraient à vacciner un maximum d'individus sans questionner les possibles conséquences évolutives, ainsi que le besoin d'impliquer des travaux théoriques dans l'établissement de politiques de santé (Bull et Antia, 2022 ; Walter et Lion, 2021). Néanmoins, ce travail pionnier a été critiqué pour sa simplification des mécanismes biologiques (Soubeyrand et Plotkin, 2002). Le modèle proposé par Gandon *et al.* (2001) décrit comment un vaccin qui stimulerait une immunité spécifiquement contre une toxine devrait sélectionner des souches produisant plus de toxine, et donc des souches plus virulentes. Mais ce modèle ne prend pas en considération le coût métabolique associé à la production de la toxine (Soubeyrand et Plotkin, 2002). Gandon *et al.* (2002) ont répondu à cette remarque en incorporant un coût à la production de toxine dans leur modèle et ont conclu que lorsque ce coût est élevé, la vaccination devrait sélectionner des souches produisant moins de toxine.

Cette conclusion concorde avec plusieurs observations empiriques pour la diphtérie : plusieurs pays rapportent désormais une plus grande proportion de souches non toxigènes que de souches toxigènes en circulation (*e.g.*, Badell-Ocando *et al.*, 2022 ; Benamrouche *et al.*, 2016 ; Czajka *et al.*, 2018 ; Dangel *et al.*, 2018 ; Guiso *et al.*, 2022). Le modèle reste néanmoins généraliste et fatalement limité face à la grande diversité des systèmes hôte-pathogène et des mécanismes et interactions en jeu. Autrement dit, le modèle ne peut par essence pas prendre en considération les mécanismes propres à chaque relation hôte-pathogène. Par exemple, leur modèle ne considère ni compétition entre souches, ni perte d'immunité, deux mécanismes qui pourraient influencer l'impact évolutif de la vaccination. Il existe donc un besoin de développer des modèles plus spécifiques, basés sur la biologie de systèmes particuliers, pour mieux comprendre comment les

vaccins et notamment les vaccins anti-toxine peuvent influencer l'évolution de leur pathogène cible.

2.1.2 – La diphtérie comme modèle d'étude

La diphtérie est une maladie souvent grave causée par des bacilles à Gram positif du genre *Corynebacterium*, le plus souvent par l'espèce *Corynebacterium diphtheriae*. La maladie existe sous trois formes : invasive, cutanée et respiratoire. Cette dernière forme est la plus commune pour les cas indigènes dans les climats tempérés, alors que la diphtérie cutanée est plus commune dans les climats tropicaux ou au sein de populations avec une mauvaise hygiène (Tiwari et Wharton, 2012). Sous sa forme respiratoire, la diphtérie se déclenche par l'invasion des voies respiratoires supérieures, suivie de la production et de la sécrétion de toxine diphtérique (TD), une toxine encodée par un gène de bactériophage du nom de *tox* (Pappenheimer et Gill, 1973).

La production de TD conduit à la mort des cellules du larynx de l'hôte infecté et à la formation d'une pseudomembrane, cette pseudomembrane étant l'un des symptômes les plus communs de la maladie (Dobie et Tobey, 1979). Elle provoque souvent une détresse respiratoire pouvant dans certains cas s'avérer fatale, mais la plupart des morts sont imputées à un effet secondaire de la TD : celle-ci peut être absorbée par d'autres organes et induire des complications graves comme des cardiopathies et des neuropathies (Tiwari et Wharton, 2012).

Il existe deux manières de lutter contre la maladie : la thérapie et la prophylaxie (Tiwari et Wharton, 2012). Lorsque la maladie est cliniquement confirmée chez un individu, la thérapie consiste généralement en l'utilisation d'une antitoxine pour réduire la morbidité et en un traitement antibiotique pour éliminer la bactérie. La prophylaxie consiste à se vacciner avec un vaccin composé d'une forme désactivée de TD qu'on appelle une anatoxine (Tiwari et Wharton, 2012). Une étude publiée en 2019 a montré que les vaccins anti-TD peuvent en réalité contenir plusieurs centaines d'antigènes de *C. diphtheriae*. Néanmoins, les vaccins contre la diphtérie sont conçus pour stimuler l'immunité du vacciné contre la TD en particulier et pas contre la bactérie entière spécifiquement dans le but de protéger contre les effets de la toxine en minimisant le risque d'effets secondaires graves (Scheifele et Ochnio, 2009). La TD peut donc être considérée comme le facteur de virulence de *C. diphtheriae*.

Par construction, les vaccins anatoxiques peuvent protéger contre les formes symptomatiques d'une infection (en stimulant les défenses immunitaires de l'hôte contre le(s) facteur(s) de virulence d'un pathogène) mais pas contre l'infection elle-même. De tels vaccins se retrouvent donc incapables de stopper la propagation d'une maladie (Fine, 1993). Cette particularité des vaccins anatoxique fait de la diphtérie un modèle intéressant dans l'étude de l'impact d'un vaccin imparfait au sens de Gandon *et al.* (2001). D'après les prédictions de ces auteurs, un tel vaccin devrait sélectionner des souches de plus forte virulence (produisant plus de TD), or la multiplication des souches isolées de *C. diphtheriae* ne produisant plus la toxine dans différents pays ayant une forte couverture vaccinale affaiblit cette théorie (Pappenheimer, 1984; Soubeyrand et Plotkin, 2002).

Le but de ce travail a été de développer un modèle mathématique basé sur la biologie de la diphtérie pour déchiffrer comment la pression vaccinale, couplée aux relations écologiques entre une souche virulente (*i.e.*, produisant la TD) et une souche avirulente (*i.e.*, ne produisant pas la toxine) de *C. diphtheriae* pourrait diriger l'évolution de ce pathogène. Ce modèle se base sur l'hypothèse que la prévalence relative de la souche avirulente devrait augmenter lorsque la pression vaccinale est suffisamment élevée puisque cette souche ne paye aucun coût métabolique à la production de la TD, ainsi qu'il a été prédit par Soubeyrand et Plotkin (2002) puis par Guiso *et al.* (2022). Le résultat de la compétition entre une souche avirulente et une souche virulente dans un contexte vaccinal fort devrait ainsi être intimement lié au coût à la production de la TD, et pourrait mener à l'éradication de la souche virulente.

2.2 – Description du modèle

2.2.1 – Formulation mathématique

Le modèle représente la dynamique de deux souches hypothétiques de *C. diphtheriae* dans une population d'hôtes avec différents statuts. La première souche (souche 1) est une souche mutante, *tox*-négative (*tox*-), ne produisant donc pas la TD. La seconde souche (souche 2) est une souche sauvage, *tox*-positive (*tox*+), produisant donc la toxine et étant capable d'induire des

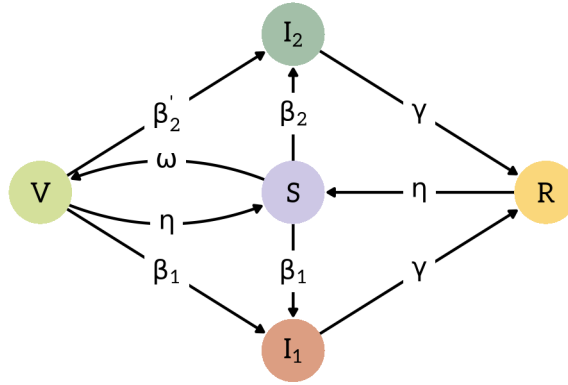


FIGURE 2.1 – **Modèle épidémiologique** décrit par les équations (2.2). Les paramètres sont détaillés dans le tableau 2.1 et dans le texte. La définition des compartiments est donnée dans le texte. Les flux démographiques (natalité et mortalité) ne sont pas représentés pour simplifier la figure.

dégâts chez les individus non immunisés¹. Pour des raisons de simplicité, j’ai considéré dans ce travail que la souche 1, ne produisant pas la toxine, est complètement avirulente.

Dans ce modèle, la maladie est transmise par contact direct avec un individu infecté au sein d’une population homogène. Le modèle est en réalité une extension du classique modèle SIR avec cinq compartiments qui classent les individus en susceptibles (S), infectés par la souche 1 (I₁), infectés par la souche 2 (I₂), vaccinés (V) et guéris (R). La Figure 2.1 représente les flux d’individus entre les cinq compartiments.

La reproduction des hôtes se fait avec un taux constant b . Une fraction p_p des individus nouveaux-nés, représentant la part de la population vaccinée à la naissance, entre dans le compartiment V tandis qu’une fraction $1 - p_p$ entre dans le compartiment S. Tous les hôtes, peu importe leur statut, meurent à un taux d .

Les individus susceptibles sont vaccinés à un taux ω , qui est le produit de la fréquence des boosters, f , et de la proportion de la population qui fait ses rappels vaccinaux, p_b . L’immunité vaccinale est perdue à un taux η , et les individus qui perdent leur immunité retournent au compartiment S.

Le modèle prend en considération trois taux de transmission : β_1 , β_2 et β_2' , représentant les

1. Cette symbologie est inversée par rapport à l’article publié et inclus dans l’Annexe A mais permet une plus grande cohérence avec les autres chapitres.

taux de transmission de la souche 1 dans n'importe quelle population, le taux de transmission de la souche 2 dans une population susceptible et le taux de transmission de la souche 2 dans une population vaccinée, respectivement.

La production de la TD confère un avantage écologique à la souche 2 dans une population susceptible en augmentant la disponibilité en fer, un élément essentiel au développement bactérien (Oram et Holmes, 2006). En population vaccinée, la production de TD devient désavantageuse pour la souche 2 à cause du coût métabolique à produire la toxine rendue inefficace par l'action du vaccin (*i.e.*, l'effet de la TD est « annulé » par le système immunitaire stimulé par le vaccin qui détruit la toxine). Cette fluctuation entre avantage et désavantage à la production de la toxine a des conséquences sur la capacité qu'a la bactérie à exploiter son hôte, donc à se multiplier, et donc à se transmettre. Pour la souche 1, la transmission est la même peu importe le statut vaccinal de l'hôte puisque cette souche ne produit pas la TD. Cette hypothèse est vraisemblable pour un pur vaccin anatoxinique, et bien que l'étude présentée plus tôt de Möller *et al.* (2019) montre que d'autres molécules immunogènes peuvent être trouvées dans les vaccins anti-diphthériques, il n'a pas été prouvé pour l'instant que ces protéines se trouvent en une concentration suffisamment élevée pour induire une immunité protégeant contre ces antigènes. J'ai donc considéré dans ces travaux que les vaccins anti-diphthériques sont des vaccins anatoxiniques purs.

Le taux de transmission β_1 peut être considéré comme le taux de transmission de base de *C. diphtheriae*, c'est-à-dire le taux de transmission sans coût ni bénéfice à la production de la TD. Pour que la production de la toxine présente un bénéfice et ait été sélectionnée jusqu'à l'ère vaccinale, elle doit conférer un gain de fitness, au moins au sein d'une population susceptible. Si la population est immunisée contre la toxine, alors le coût à sa production est supposé dépasser son bénéfice. Les taux de transmission β_1 , β_2 et β'_2 sont donc reliés par la relation

$$\beta_2 > \beta_1 > \beta'_2, \tag{2.1}$$

qui est l'inéquation fondamentale de notre modèle.

Pour des raisons de simplicité, le présent modèle ne prend en compte aucune forme de super-infection ou co-infection, c'est-à-dire qu'un individu infecté par une souche ne peut pas être simultanément infecté par une autre. Les deux souches se retrouvent donc en compétition pour

l'accès aux hôtes : la croissance (multiplication) d'une souche est limitée par la croissance de l'autre souche. De plus, il est considéré dans le modèle que l'infection ne provoque pas de surmortalité. Ainsi, la virulence n'est pas définie dans ce travail comme un gain de mortalité induit par le pathogène, on considère plus simplement que la souche 2 est virulente et que la souche 1 est avirulente.

Les individus infectés guérissent à un taux γ et entrent alors dans le compartiment R en acquérant une immunité contre toute nouvelle infection. Les individus guéris perdent leur immunité à un taux η et retournent dans le compartiment S. En l'absence de données de qualité pour la diphtérie, il est établi ici que l'immunité acquise naturellement (*i.e.*, par l'infection) et celle acquise artificiellement (*i.e.*, par la vaccination) disparaissent au même taux η . Cette simplification peut être remise en question : elle est notamment fautive pour la coqueluche, un système hôte-pathogène proche ([Wendelboe et al., 2005](#)).

Il est néanmoins attendu que l'immunité acquise par la vaccination diffère de celle acquise naturellement. La vaccination apporte une immunité spécifiquement contre la TD alors qu'une infection apporte une immunité contre une multitude d'autres antigènes de la bactérie. Cela signifie qu'un individu non vacciné infecté par la souche 2 développera une immunité contre la TD et les autres antigènes bactériens, alors qu'un individu infecté par la souche 1 développera une immunité contre ces antigènes s'il n'a pas été vacciné, ou contre ces antigènes et la TD s'il a été vacciné. Pour des raisons de simplicité, on considère que n'importe quelle forme d'infection provoque une immunité complète et protectrice contre toute nouvelle infection, qu'elle soit par la souche 1 ou la souche 2.

La dynamique épidémique du modèle est décrite avec un système de cinq équations différentielles ordinaires telles que

$$\begin{aligned} \frac{dS(t)}{dt} = & (1 - p_p)b + \eta(V(t) + R(t)) - \beta_1 \frac{S(t)}{N(t)} I_1(t) - \beta_2 \frac{S(t)}{N(t)} I_2(t) \\ & - \omega S(t) - dS(t), \end{aligned} \quad (2.2a)$$

$$\frac{dI_1(t)}{dt} = \beta_1 \frac{S(t)}{N(t)} I_1(t) + \beta_1 \frac{V(t)}{N(t)} I_1(t) - \gamma I_1(t) - dI_1(t), \quad (2.2b)$$

$$\frac{dI_2(t)}{dt} = \beta_2 \frac{S(t)}{N(t)} I_2(t) + \beta_2' \frac{V(t)}{N(t)} I_2(t) - \gamma I_2(t) - dI_2(t), \quad (2.2c)$$

$$\begin{aligned} \frac{dV(t)}{dt} = & p_p b + \omega S(t) - \beta_1 \frac{V(t)}{N(t)} I_1(t) - \beta_2' \frac{V(t)}{N(t)} I_2(t) - \eta V(t) \\ & - dV(t), \end{aligned} \quad (2.2d)$$

$$\frac{dR(t)}{dt} = \gamma (I_1(t) + I_2(t)) - \eta R(t) - dR(t), \quad (2.2e)$$

où $S(t)$, $I_1(t)$, $I_2(t)$, $V(t)$, et $R(t)$ représentent le nombre d'individus dans les compartiments S, I_1 , I_2 , V et R, respectivement, à l'instant t , et $N(t)$ représente le nombre total d'individus à l'instant t ($N(t) = S(t) + I_1(t) + I_2(t) + V(t) + T(t)$).

2.2.2 – Paramétrisation

Le tableau 2.1 résume les paramètres du modèle. Les valeurs des paramètres peuvent être trouvées dans Hammarlund *et al.* (2016) pour la durée de l'immunité, dans Truelove *et al.* (2020) pour la durée de l'infection, le nombre de reproduction de base et la réduction du taux de transmission de la souche 2 chez les individus vaccinés. Le taux de mortalité est issu du site Internet de l'Institut national de la statistique et des études économiques INSEE². Quand les données manquaient, les paramètres ont été calculés ou choisis arbitrairement comme présenté plus bas. Il est d'intérêt de noter que la littérature est contradictoire sur la définition du niveau d'anticorps assurant une protection contre une infection, le seuil variant entre auteurs ; ainsi, la durée estimée pendant laquelle perdure l'immunité est elle aussi discutée. Dans ce travail, j'ai choisi de considérer un niveau de 0.01 UI.mL⁻¹ comme protecteur et choisi une valeur de η adapté à ce seuil (Amanna *et al.*, 2007 ; Hammarlund *et al.*, 2016).

L'OMS recommande aux adultes de prendre un rappel anti-diphtérique tous les dix ans (World Health Organisation, 2017). Cette recommandation est suivie dans un grand nombre de pays – au moins dans les calendriers vaccinaux, mais plus rarement dans la pratique – mais pas en France, où les rappels sont recommandés à 25, 45 et 65 puis tous les dix ans au-delà (Ministère des Solidarités et de la Santé, 2022). Si la proportion d'enfants vaccinés dans chaque pays du monde est estimée par l'OMS, ce n'est pas le cas pour les adultes. Elle peut néanmoins être estimée à partir d'études dans plusieurs pays comme la France (Baratin *et al.*, 2012), le Mexique (Cruz-Hervert *et al.*, 2013), les États-Unis (Williams *et al.*, 2017) et l'Allemagne (Poethko-Müller

2. Disponible sur https://www.insee.fr/fr/outil-interactif/5367857/tableau/20_DEM/23_DME (dernier accès le 02/01/2023).

TABLE 2.1 – **Paramètres du modèle** décrit par les équations (2.2) et valeurs utilisées dans les simulations. j : jours.

Paramètre	Signification	Valeur [unité]	Référence
b	Natalité	$\frac{1}{3}$ [j^{-1}]	Arbitraire
d	Mortalité	$\frac{1}{83 \times 365}$ [j^{-1}]	INSEE
p_p	Couverture vaccinale primaire	[0,1] [sans unité]	Arbitraire
p_b	Couverture vaccinale secondaire	[0,1] [sans unité]	Arbitraire
f	Fréquence des rappels	$\frac{1}{10 \times 365}$ [j^{-1}]	World Health Organisation (2017)
ω	Taux de rappel	$f p_b$	Voir section 2.2.1
η	Taux de perte de l'immunité	$\frac{1}{42 \times 365}$ [j^{-1}]	Hammarlund <i>et al.</i> (2016)
γ	Taux de guérison	$\frac{1}{18.5}$ [j^{-1}]	Truelove <i>et al.</i> (2020)
\mathcal{R}_0	Nombre de reproduction de base de la souche 2	2.6 [sans unité]	Truelove <i>et al.</i> (2020)
r_1	Réduction de transmission de la souche 1	0.42 [sans unité]	Voir section 2.2.2
r_2	Réduction de transmission de la souche 2	0.60 [sans unité]	Truelove <i>et al.</i> (2020)
β_1	Taux de transmission de la souche 1	$(1 - r_1) \beta_2$ [j^{-1}]	(2.5)
β_2	Taux de transmission de la souche 2 chez les individus S	$\mathcal{R}_0 (\gamma + d)$ [j^{-1}]	(2.3)
β_2'	Taux de transmission de la souche 2 chez les individus V	$(1 - r_2) \beta_2$ [j^{-1}]	(2.4)

et Schmitz, 2013). Globalement, pour la plupart des pays et des vaccins, la couverture vaccinale des adultes est faible, notamment en Europe Kanitz *et al.* (2012).

La valeur du \mathcal{R}_0 pour une maladie est généralement calculée (ou plutôt, devrait être calculée) au début d'une épidémie, quand la population est encore naïve face au pathogène (Delamater *et al.*, 2019). Dans le cas de la diphtérie, le \mathcal{R}_0 est habituellement calculé lors d'épidémies se déclenchant dans des populations partiellement immunisées et ne devrait donc pas, en toute rigueur, être dénommé \mathcal{R}_0 (*e.g.*, Matsuyama *et al.*, 2018). Cependant, en utilisant des outils d'analyse statistique, il est possible d'inférer la valeur qu'aurait le \mathcal{R}_0 dans une population entièrement susceptible (Truelove *et al.*, 2020). Une fois le \mathcal{R}_0 obtenu, il est possible de calculer les taux de transmission de chacune des souches. Puisque seules les souches toxigènes peuvent être efficacement surveillées (donc la souche 2 dans notre modèle), tout \mathcal{R}_0 calculé à partir d'études de terrain ne peut donc décrire que la circulation de ce type de souche. Le taux de transmission de la souche 2 dans une population susceptible, β_2 , vaut

$$\beta_2 = \mathcal{R}_0 (\gamma + d). \quad (2.3)$$

Le taux de transmission de la souche 2 dans une population vaccinée, β_2' , peut quant à lui

être obtenu par la formule

$$\beta'_2 = (1 - r_2) \beta_2, \quad (2.4)$$

où r_2 représente la réduction de transmission de la souche 2 due à la vaccination. Cette valeur peut également être considérée comme l'efficacité du vaccin contre la souche 2 : une valeur de zéro indique un vaccin totalement inefficace, tandis qu'une valeur d'un indique un vaccin bloquant totalement la circulation de la souche 2.

Comme exprimé plus tôt, le \mathcal{R}_0 ne peut pas être obtenu pour une souche ne produisant pas la TD (donc la souche 1 dans notre modèle) puisque ce type de souche est difficile à surveiller en population. Il n'est donc pas possible de calculer le taux de transmission de la souche 1 à partir de l'équation (2.6). À la place, la valeur du taux de transmission de la souche 1, β_1 , est calculée comme étant une réduction du taux de transmission de la souche 2 dans une population susceptible avec la formule

$$\beta_1 = (1 - r_1) \beta_2, \quad (2.5)$$

où r_1 représente la réduction de transmission de la souche 1 due au fait que cette souche ne produit pas la TD et n'en tire donc aucun bénéfice. On pose $r_2 > r_1$ en accord avec l'équation (2.1). Comme exposé plus haut, il est difficile de suivre les souches non toxigènes. Il en résulte qu'à ma connaissance aucune valeur estimée pour r_1 n'est disponible dans la littérature, et j'ai donc choisi une valeur arbitraire comme base. Puisque la production de la toxine est essentielle pour l'acquisition de fer dans les voies respiratoires supérieures – un milieu pauvre en fer – j'ai estimé que la non production de la TD représente un fort handicap pour la souche et donc que r_1 pouvait être fixé à une valeur élevée (Murphy, 1996 ; Parveen *et al.*, 2019).

2.2.3 – Analyse d'invasion

Pour rappel, le nombre de reproduction de base (\mathcal{R}_0) est défini comme le nombre d'infections secondaires générées par un unique individu infecté au sein d'une population entièrement susceptible. Cette valeur décrit la capacité ($\mathcal{R}_0 > 1$) ou l'incapacité ($\mathcal{R}_0 < 1$) pour une souche à

envahir une population. Dans le modèle, le \mathcal{R}_0 de la souche s est obtenu par la formule

$$\mathcal{R}_0 = \frac{\beta_s}{\gamma + d}. \quad (2.6)$$

Si une proportion de la population se fait vacciner, alors la population n'est plus « entièrement susceptible » et l'équation (2.6) ne peut plus décrire la capacité d'invasion du pathogène dans ce contexte. Il faut donc définir de nouvelles équations pour obtenir le nombre de reproduction de chaque souche.

On commence par définir une population hypothétique à l'équilibre sans maladie (DFE, pour *Disease-Free Equilibrium*). Une population au DFE peut être soit mixte (composée à la fois d'individus susceptibles et d'individus vaccinés), soit complètement susceptible. Dans un cas simple où une seule souche circule, il est établi que cette souche peut envahir la population et devenir épidémique si son nombre de reproduction est supérieur à un. Soit \mathcal{R}_s le nombre de reproduction de la souche s ($s \in \{1,2\}$). On considère qu'une invasion commence par l'arrivée d'un unique individu infecté dans la population (*i.e.*, $I_s(0) = 1$ pour une invasion par la souche s).

Dans une population mixte, la souche 2 peut envahir la population au DFE si son nombre de reproduction au DFE, $\mathcal{R}_{2,DFE}$, est supérieur à un. Sa valeur est obtenue par la formule

$$\mathcal{R}_{2,DFE} = \frac{\beta_2 s_{DFE}^*}{\gamma + d} + \frac{\beta_2' v_{DFE}^*}{\gamma + d}, \quad (2.7)$$

où

$$s_{DFE}^* = \frac{(1 - p_p) d + \eta}{\omega + \eta + d} \quad (2.8)$$

est la proportion d'individus susceptibles dans la population au DFE et

$$v_{DFE}^* = \frac{p_p d + \omega}{\omega + \eta + d} \quad (2.9)$$

est la proportion d'individus vaccinés au DFE ($s_{DFE}^* + v_{DFE}^* = 1$). La valeur de $\mathcal{R}_{2,DFE}$ dépend donc de la couverture vaccinale.

Réciproquement, la souche 1 peut envahir la population au DFE si son nombre de reproduc-

tion au DFE, $\mathcal{R}_{1,DFE}$, est supérieur à un. Sa valeur est obtenue par la formule

$$\mathcal{R}_{1,DFE} = \frac{\beta_1}{\gamma + d}, \quad (2.10)$$

qui est similaire à l'équation (2.6) et implique que la valeur de $\mathcal{R}_{1,DFE}$ est indépendante de la couverture vaccinale. C'est attendu, puisque la vaccination n'interfère qu'avec la transmission de la souche toxigène.

On s'intéresse aux conditions permettant la coexistence au long terme des deux souches en utilisant un critère d'invasion mutuel. Cela revient à chercher dans quelles conditions chacune des souches peut envahir une population d'hôtes, c'est-à-dire que les deux souches ont un taux de croissance positif à l'équilibre lorsque l'autre souche circule seule dans la population (Grainger *et al.*, 2019). Techniquement, cela revient à introduire la nouvelle souche dans une population à l'équilibre mais avec la première souche qui s'y est déjà établie. On cherche alors les conditions pour lesquelles $dI_1(t)/dt > 0$ (respectivement $dI_2(t)/dt > 0$) quand la population est à l'équilibre endémique avec la souche 2 (respectivement la souche 1) seulement. On obtient

$$\frac{s_2^*}{v_2^*} < \frac{\beta_1 - \beta_2'}{\beta_2 - \beta_1} < \frac{s_1^*}{v_1^*}, \quad (2.11)$$

où s_s^* (respectivement v_s^*) est la proportion d'individus susceptibles (respectivement vaccinés) dans la population totale, à l'équilibre où la souche s est endémique. Sachant l'inégalité (2.1), on a $\beta_1 - \beta_2' > 0$ et $\beta_2 - \beta_1 > 0$. La démonstration de (2.11) est donnée dans l'Annexe B.1.

La valeur $\beta_1 - \beta_2'$ représente la différence positive de fitness entre la souche 1 et la souche 2 au sein d'une population d'hôtes vaccinés. Elle met en évidence l'avantage de la souche 1 dans une population d'hôtes vaccinés. Réciproquement, la valeur $\beta_2 - \beta_1$ représente la différence positive de fitness entre la souche 2 et la souche 1 au sein d'une population d'hôtes susceptibles, mettant en évidence l'avantage de la souche 2 dans une population d'hôtes susceptibles. Les termes de l'équation peuvent être séparés pour mieux comprendre les conditions dans lesquelles les deux souches peuvent coexister. On obtient alors d'un côté

$$\frac{\beta_1 - \beta_2'}{\beta_2 - \beta_1} < \frac{s_1^*}{v_1^*} \Leftrightarrow (\beta_2 - \beta_1) s_1^* > (\beta_1 - \beta_2') v_1^*, \quad (2.12)$$

et de l'autre

$$\frac{\beta_1 - \beta'_2}{\beta_2 - \beta_1} > \frac{s_2^*}{v_2^*} \Leftrightarrow (\beta_1 - \beta'_2) v_2^* > (\beta_2 - \beta_1) s_2^*. \quad (2.13)$$

L'inégalité (2.12) montre que la souche 2 peut envahir une population où la souche 1 est endémique si l'avantage de la souche 2 au sein de la population susceptible, relativement à la proportion d'individus susceptibles dans la population, est supérieure à l'avantage de la souche 1 au sein de la population vaccinée, relativement à la proportion d'individus vaccinés dans la population.

Réciproquement, l'inégalité (2.13) montre que la souche 1 peut envahir une population où la souche 2 est endémique si l'avantage de la souche 1 au sein de la population vaccinée, relativement à la proportion d'individus vaccinés dans la population, est supérieure à l'avantage de la souche 2 au sein de la population susceptible, relativement à la proportion d'individus susceptibles dans la population.

L'invasion implique la compétition : si une souche ne peut pas envahir une population où l'autre souche est endémique, cela signifie que cette dernière exclut compétitivement la première.

Dans le reste de ce chapitre, j'appellerai « nombre de reproduction effectif », dénoté \mathcal{R}_s où s désigne la souche 1 ou la souche 2, le nombre de reproduction de la souche s dans n'importe quelle population non complètement susceptible.

2.3 – Résultats des simulations

Le modèle mathématique présenté dans les équations (2.2) a été implémenté dans un script C. Les simulations du modèle ont ensuite été faites en utilisant le langage R v4.1.0 (R Core Team, 2022) et le package `deSolve` v1.34 (Soetaert *et al.*, 2010, 2023). Plusieurs séries de simulations ont été faites pour qualifier et quantifier l'effet de la couverture vaccinale couplé à celui de six autres paramètres clés sur la dynamique de compétition des deux souches. Chaque simulation a tourné sur un temps suffisamment long pour que la population atteigne un état proche de l'équilibre (10^6 jours, soit approximativement 2740 années³).

Pour chaque simulation, deux types de conditions initiales ont été utilisées selon que le système soit étudié avec deux souches en circulation ou une seule. Dans les simulations où les deux souches

3. Une explication du choix de cette durée est présentée dans l'Annexe B.2.

circulent, les conditions initiales sont $S(0) = b/d - 2$, $I_1(0) = 1$, $I_2(0) = 1$, $V(0) = 0$ et $R(0) = 0$. Dans les simulations où seule la souche s circule, les conditions initiales ont été $S(0) = b/d - 1$, $I_s(0) = 1$, $I_{q|q \neq s}(0) = 0$, $V(0) = 0$ et $R(0) = 0$. Ces conditions initiales assurent une taille de population constante et un début de dynamique épidémique avec uniquement un individu par souche.

2.3.1 – Prévalence de la souche 1

Le premier résultat des simulations est la prévalence de la souche 1 – donc la souche non toxigène – à l'équilibre. Cette prévalence a été calculée comme étant le nombre d'infections par la souche 2 pour 10 000 habitants. À noter que j'ai choisi de m'intéresser particulièrement aux prévalences des différents types de souches plutôt qu'à leurs fréquences relatives parce que la coexistence des souches à l'équilibre n'est en réalité possible que pour une faible gamme de valeurs de paramètres (voir Section 2.3.4 plus loin). Ainsi, pour la plupart des valeurs, la fréquence relative des souches est soit zéro, soit un, et les conclusions les plus intéressantes apparaissent en s'intéressant à la prévalence plutôt qu'à la fréquence relative des souches.

Couverture vaccinale

Lorsque la souche 2 circule dans la même population que la souche 1, la prévalence de la souche 1 augmente lorsque la couverture vaccinale augmente avec un effet seuil (Figure 2.2, tous les panels). Cela s'explique par la structure du modèle : plus la couverture vaccinale est élevée, plus le nombre d'individus dans le compartiment V est élevé, or la souche 1 est avantagée (par rapport à la souche 2) dans une population fortement vaccinée, donc sa prévalence augmente lorsque la part vaccinée de la population augmente. L'effet seuil, quant à lui, représente l'issue de la compétition entre les deux souches. D'un côté du seuil (que j'appellerai dans la suite « seuil de remplacement »), la souche 1 est exclue de la population à l'équilibre et sa prévalence est quasi-nulle. De l'autre côté du seuil de remplacement, la souche 2 est exclue de la population à l'équilibre et la prévalence de la souche 1 est soit constante, soit dépendante de la valeur du paramètre observé. Aux alentours du seuil se dessine une gamme de valeurs, plus ou moins étroite selon le paramètre observé, pour lesquelles les deux souches coexistent à l'équilibre. Ce point sera

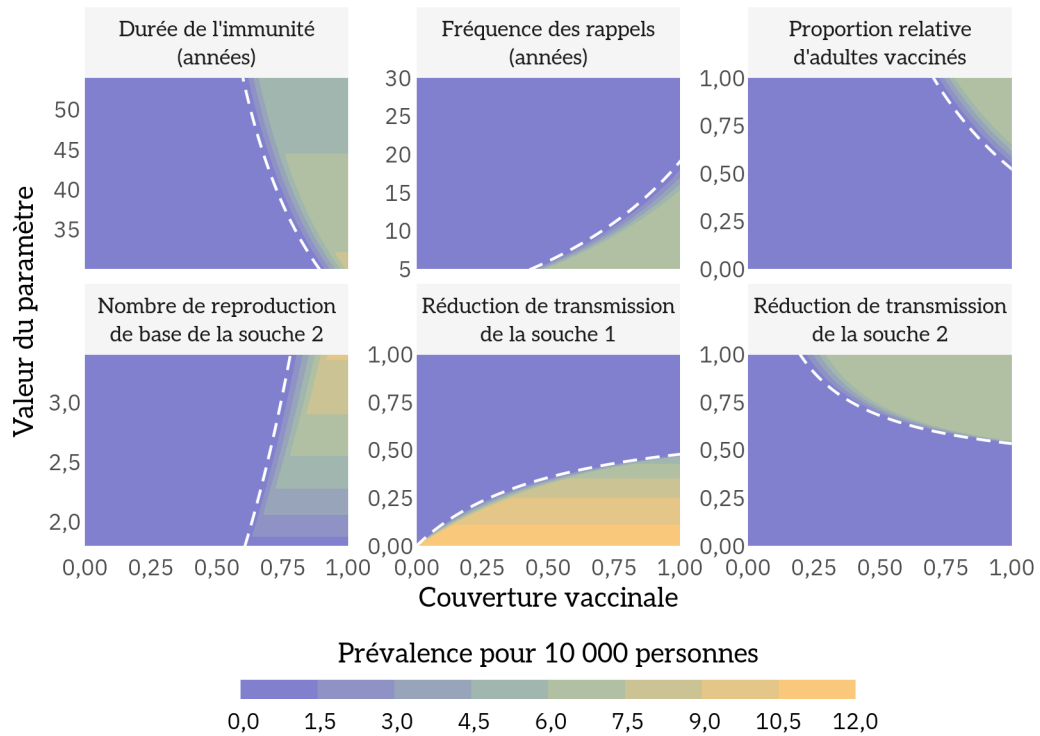


FIGURE 2.2 – **Prévalence de la souche 1** lorsque les deux souches circulent dans la population (nombre d’infectés à l’équilibre pour 10 000 habitants). Chaque panel représente un paramètre différent : durée de l’immunité $((365 \times \eta)^{-1})$, fréquence des rappels (f) , proportion relative d’adultes vaccinés (p_b/p_p) , nombre de reproduction de base de la souche toxigène (\mathcal{R}_0) , réduction de transmission de la souche 1 (r_1) et réduction de transmission de la souche 2 (r_2) . La ligne en tirets indique le seuil pour lequel la prévalence de la souche 1 à l’équilibre est non nulle.

discuté plus tard. Sur la Figure 2.3 est représentée la prévalence à l’équilibre de la souche 1 quand la souche 2 est initialement absente de la population. On observe que les résultats présentés sur cette figure correspondent à ceux montrés sur la Figure 2.2 du côté de l’équilibre où la prévalence de la souche 1 est non nulle.

Durée de l’immunité

Nous pouvons observer que le seuil de remplacement est atteint pour de plus faibles couvertures vaccinales lorsque la durée de l’immunité est élevée (Figure 2.2, panel haut-gauche). En revanche, une fois le seuil franchi, nous observons que plus la durée de l’immunité est forte, moins la prévalence à l’équilibre est élevée. Cela s’explique encore une fois par la structure du

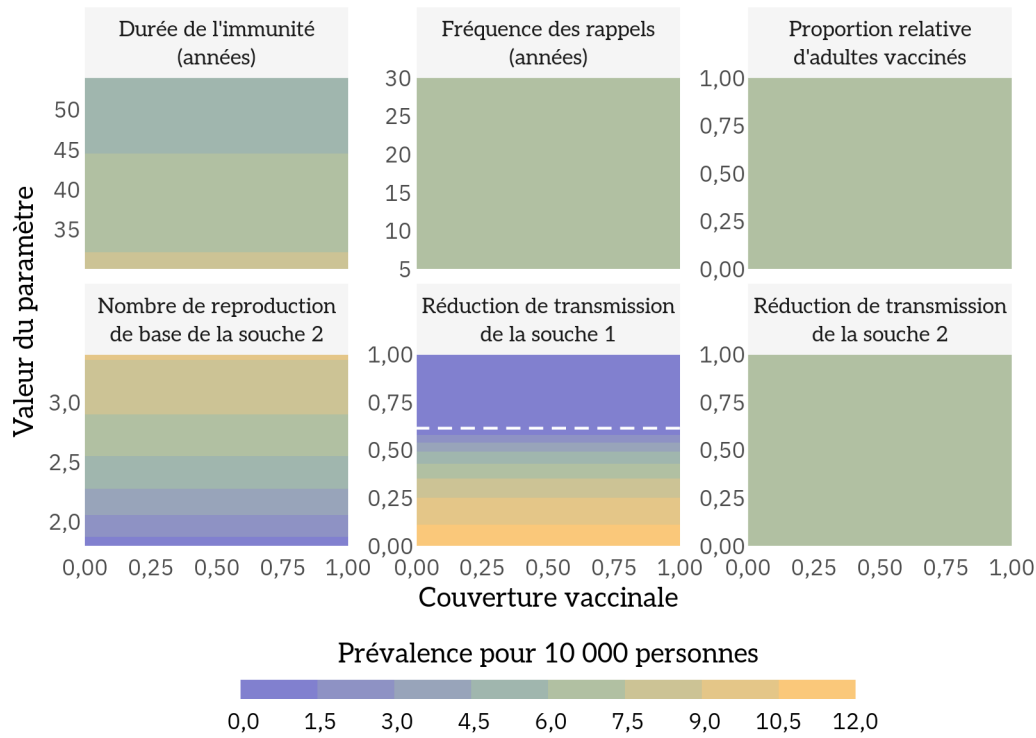


FIGURE 2.3 – **Prévalence de la souche 1** lorsque la souche 1 circule seule dans la population (nombre d’infectés à l’équilibre pour 10 000 habitants). Chaque panel représente un paramètre différent : durée de l’immunité $((365 \times \eta)^{-1})$, fréquence des rappels (f), proportion relative d’adultes vaccinés (p_b/p_p), nombre de reproduction de base de la souche toxigène (\mathcal{R}_0), réduction de transmission de la souche 1 (r_1) et réduction de transmission de la souche 2 (r_2). La ligne en tirets indique le seuil pour lequel la prévalence de la souche 1 à l’équilibre est non nulle.

modèle : plus la durée de l’immunité est élevée, plus le temps passé dans le compartiment V par les individus est grand, et donc plus la part vaccinée de la population est grande, à l’avantage de la souche 1. Au-delà de ce seuil, en revanche, cet effet s’inverse : plus la durée de l’immunité est longue, plus la prévalence est faible. La raison derrière ce phénomène est la dualité de la durée de l’immunité : celle-ci est la même que l’individu soit dans le compartiment V ou dans le compartiment R, or les individus dans le compartiment R ne peuvent pas être infectés, donc plus la durée de l’immunité est élevée, plus la part de la population immunisée contre toutes les souches est élevée, et plus la prévalence de la souche 1 est faible. Ce phénomène est explicite dans le cas où la souche 1 circule seule au début des simulations. Dans ce cas, la couverture vaccinale n’a pas d’effet sur la prévalence de la souche 1, puisque celle-ci se transmet de la même

manière entre individus susceptibles et vaccinés. Nous observons alors bien que lorsque la durée de l'immunité est plus élevée, la prévalence de la souche 1 est toujours plus faible (Figure 2.3, panel haut-gauche).

Durée entre rappels

Le seuil de remplacement de la souche 2 par la souche 1 est atteint plus rapidement lorsque la durée entre rappels est courte (Figure 2.2, panel haut-milieu). Une forte fréquence des rappels (donc une faible durée entre les rappels) assure un renouvellement important de la population vaccinée, à l'avantage de la souche 1, ce qui explique que lorsque les rappels sont suffisamment proches la prévalence de la souche 1 augmente. Contrairement à ce qui a été montré pour la durée de l'immunité, du côté du seuil où la prévalence de la souche 1 est non nulle, il n'y a pas de lien entre la valeur du paramètre et la prévalence à l'équilibre. Cela s'explique par le fait que sans compétition avec la souche 2 la prévalence de la souche 1 est indépendante de la répartition des individus entre les compartiments S et V (Figure 2.3, panel haut-milieu).

Proportion relative d'adultes vaccinés

La proportion relative d'adultes vaccinés agit de concert avec la couverture vaccinale primaire, les deux étant directement liés. On remarque que plus cette proportion est élevée, plus le seuil se décale vers de (relativement) faibles valeurs de couverture vaccinale primaire (Figure 2.2, panel haut-droite). Cela s'explique par le fait que si les adultes se font vacciner en masse, la part vaccinée de la population se maintient à un niveau élevé, à l'avantage de la souche 1. En revanche, avec les valeurs des autres paramètres utilisées dans le modèle, on remarque qu'il faut que la couverture vaccinale primaire soit déjà très élevée pour que la vaccination des adultes ait un effet observable. Comme attendu, lorsque seule la souche 1 circule dans la population, la proportion relative d'adultes vaccinés n'a pas d'impact sur la prévalence de la bactérie (Figure 2.3, panel haut-droite).

Nombre de reproduction de base

Plus le nombre de reproduction de base de la souche virulente (\mathcal{R}_0) est élevé, plus le seuil de remplacement est difficile à atteindre. En revanche, lorsque celui-ci est atteint, plus le \mathcal{R}_0 est

élevé et plus la prévalence de la souche 1 est élevée (Figure 2.2, panel bas-gauche). Cela s'explique par la relation entre le nombre de reproduction de base de la souche 2 et celui de la souche 1, ce dernier valant $(1 - r_1) \mathcal{R}_0$, donc directement proportionnel au premier. On remarque ainsi que lorsque la souche 1 circule seule dans la population, pour toute valeur de couverture vaccinale, la prévalence de la souche 1 augmente lorsque le \mathcal{R}_0 augmente (Figure 2.3, panel bas-gauche).

Réduction de transmission de la souche 1

La réduction de transmission de la souche 1 peut être interprétée comme le « coût » à ne pas produire la TD en termes de non-gain de *fitness*. Si $r_1 = 0$, alors il n'y a aucun coût et la production de la toxine n'apporte en réalité aucun avantage. Si $r_1 = 1$, alors la souche 1 perd toute capacité à se transmettre, autrement dit la production de la TD est indispensable à la multiplication du pathogène ou à sa transmission. Entre ces deux extrêmes, la souche 1 peut se transmettre plus ou moins bien par rapport à la souche 2. Plus cette réduction est élevée, plus il faut une couverture vaccinale élevée pour parvenir à éradiquer la souche 2 (Figure 2.2, panel bas-milieu). À l'inverse, on remarque que si réduction est très faible (*i.e.*, le coût à ne pas produire la TD est minime), alors une couverture vaccinale même très basse suffit à faire franchir le seuil de remplacement. À droite du seuil de remplacement, plus la réduction de transmission est faible plus la prévalence de la souche 1 est élevée, un résultat que l'on retrouve dans les simulations où seule la souche 2 est absente de la population (Figure 2.2, panel bas-milieu) et qui s'explique simplement par la définition du paramètre r_1 : il réduit le taux de transmission de la souche 1. Ainsi, même sans être en compétition avec la souche 2, il est possible que la souche 1 ne puisse pas se maintenir dans la population si ce paramètre r_1 est plus élevé qu'un certain seuil, représenté par la ligne en tirets sur la figure.

Réduction de transmission de la souche 2

La réduction de transmission de la souche 2 peut être interprétée comme le coût à produire la TD lorsque l'hôte est immunisé contre la toxine. Plus le coût à la production de la TD est élevé, plus le seuil de remplacement est bas, autrement dit moins il y a besoin d'une couverture vaccinale élevée pour réussir à remplacer la souche 2 par la souche 1. Une fois le seuil franchi, la réduction de transmission de la souche 2 n'a pas d'impact sur la prévalence de la souche 1,

comme attendu. De la même manière, lorsque seule la souche 1 circule dans la population, il n'y a pas d'effet de la réduction de transmission de la souche 2 sur la prévalence de la souche 1 (Figure 2.3, panel bas-droite).

2.3.2 – Prévalence de la souche 2

Couverture vaccinale

La prévalence de la souche 2 lorsque la souche 1 circule reflète celle de la souche 1 détaillée dans la section précédente. On retrouve l'effet seuil présenté plus tôt, mais cette fois-ci inversé : à gauche du seuil de remplacement, la souche 2 est prévalente, puis elle disparaît à droite du seuil, c'est-à-dire lorsque la couverture vaccinale est élevée (Figure 2.4, tous les panels). Les résultats lorsque la souche 1 est absente de la population montrent que, même sans compétition, la souche 2 est impactée par la couverture vaccinale de la population : plus celle-ci est élevée, plus la prévalence de la souche 2 est faible (Figure 2.5, tous les panels). Ces résultats diffèrent de ceux que j'ai présentés plus tôt pour la souche 1 circulant seule (2.3), ce qui s'explique par la biologie de la souche 2, dont le taux de transmission diffère entre les individus vaccinés et ceux susceptibles.

Durée de l'immunité

La durée de l'immunité impacte fortement la prévalence de la souche 2 : plus celle-ci est élevée et plus le seuil de remplacement par la souche 1 est décalé vers de basses valeurs de couverture vaccinale (Figure 2.4, panel haut-gauche). La durée de l'immunité a un impact double sur la prévalence de la souche 2 : en prolongeant la durée de séjour des individus dans le compartiment R, elle réduit la part de la population pouvant être infectée par la souche, et en prolongeant la durée de séjour des individus dans le compartiment V, elle réduit le taux de transmission global de la souche. Ces deux éléments combinés font que, même lorsque la souche 1 est absente de la population, la prévalence de la souche 2 diminue lorsque la durée de l'immunité et la couverture vaccinale sont élevées (Figure 2.5, panel haut-gauche).

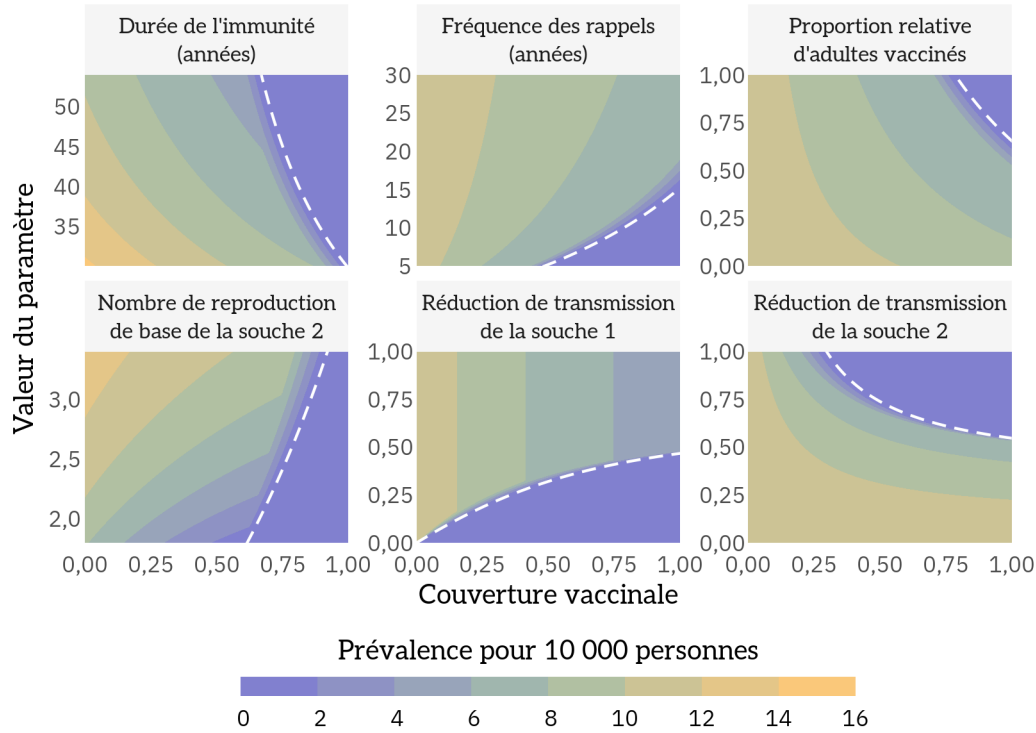


FIGURE 2.4 – **Prévalence de la souche 2** lorsque les deux souches circulent dans la population (nombre d'infectés à l'équilibre pour 10 000 habitants). Chaque panel représente un paramètre différent : durée de l'immunité $((365 \times \eta)^{-1})$, fréquence des rappels (f), proportion relative d'adultes vaccinés (p_b/p_p), nombre de reproduction de base de la souche toxigène (\mathcal{R}_0), réduction de transmission de la souche 1 (r_1) et réduction de transmission de la souche 2 (r_2). La ligne en tirets indique le seuil pour lequel la prévalence de la souche 2 à l'équilibre est non nulle.

Durée entre rappels

Comme cela a été présenté plusieurs fois dans ce chapitre, la souche 2 est avantagée lorsque la proportion vaccinée de la population est faible et désavantagée lorsque celle-ci est élevée. Lorsque la durée entre deux rappels est grande (*i.e.*, la fréquence des rappels est faible), la prévalence de la souche 2 est plus élevée, donc le seuil de remplacement par la souche 1 se déplace vers de hautes valeurs de couverture vaccinale, voire la possibilité d'un remplacement disparaît (Figure 2.4, panel haut-milieu). De même, lorsque la souche 1 est absente de la population, une fréquence élevée des rappels contribue à maintenir une fraction vaccinée de la population élevée, réduisant la prévalence de la souche virulente (Figure 2.5, panel haut-milieu).

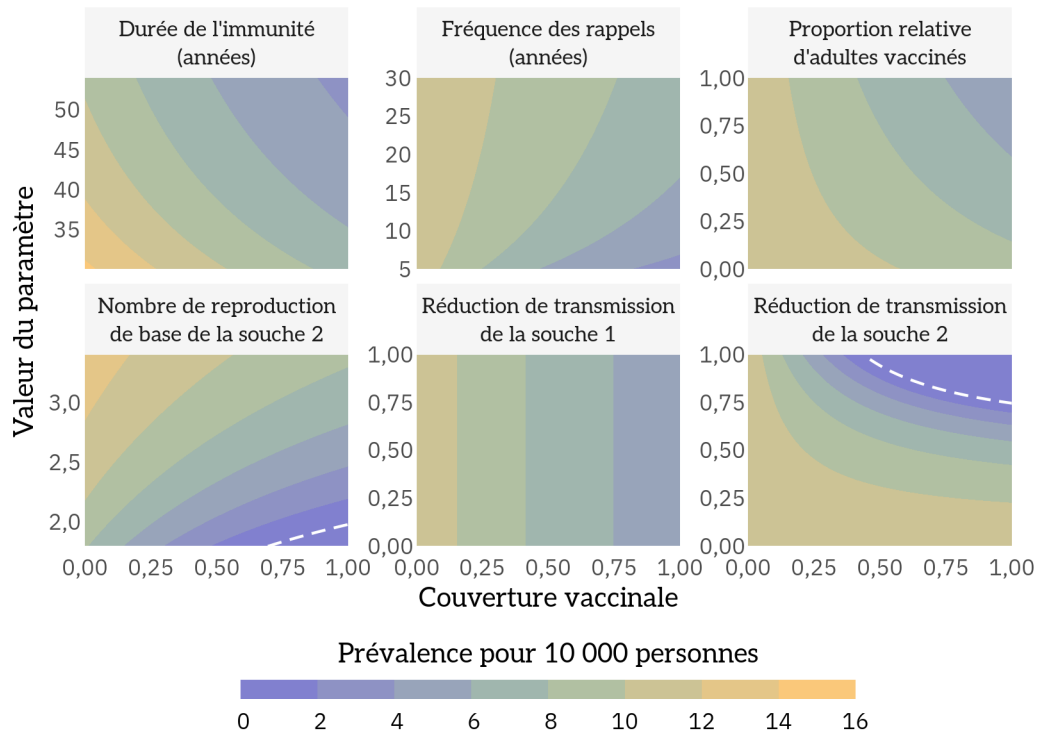


FIGURE 2.5 – **Prévalence de la souche 2** lorsque la souche 2 circule seule dans la population (nombre d’infectés à l’équilibre pour 10 000 habitants). Chaque panel représente un paramètre différent : durée de l’immunité $((365 \times \eta)^{-1})$, fréquence des rappels (f), proportion relative d’adultes vaccinés (p_b/p_p), nombre de reproduction de base de la souche toxigène (\mathcal{R}_0), réduction de transmission de la souche 1 (r_1) et réduction de transmission de la souche 2 (r_2). La ligne en tirets indique le seuil pour lequel la prévalence de la souche 2 à l’équilibre est non nulle.

Proportion relative d’adultes vaccinés

Une forte proportion relative d’adultes vaccinés contribue, comme une faible durée entre rappels et une forte couverture vaccinale, à maintenir une part de population vaccinée élevée, au détriment de la transmission de la souche 2. Ainsi, lorsque cette proportion est élevée, le seuil de remplacement par la souche avirulente se décale vers de plus faibles valeurs de couverture vaccinale (Figure 2.4, panel haut-droite). De la même manière, en l’absence de la souche compétitrice dans la population, une forte proportion relative d’adultes vaccinés permet de diminuer la prévalence de la souche 2 (Figure 2.5, panel haut-droite).

Nombre de reproduction de base

Le nombre de reproduction de base (\mathcal{R}_0) défini dans ce travail est celui de la souche 2. Il est donc logique que lorsque celui-ci augmente, la prévalence de la souche 2 augmente; et ainsi plus celui-ci est élevé, plus le seuil de remplacement est décalé vers de fortes valeurs de couverture vaccinale (Figure 2.4, panel bas-gauche). Cela signifie que plus la souche 2 se transmet à un taux élevé (les taux de guérison et de mortalité étant par ailleurs constants), plus elle est avantagée par rapport à la souche 1, alors même que le taux de transmission de la souche 1 est directement proportionnel à celui de la souche 2 (équation (2.5)). Dans le cas où la souche 2 circule seule dans la population, on remarque qu'une couverture vaccinale forte et un \mathcal{R}_0 bas conduisent à la disparition de la souche 2 (Figure 2.5, panel bas-gauche).

Cette particularité s'explique par la structure de la population dans laquelle se propage la souche 2 et de son lien avec le nombre de reproduction de la souche 2. La souche 1 peut se maintenir dans la population tant que celle-ci se renouvelle en individus susceptibles et vaccinés, alors que la souche 2 ne le peut que si la population se renouvelle suffisamment en individus susceptibles ou, si la population se renouvelle essentiellement en individus vaccinés, que la diminution de la transmission chez les individus vaccinés est assez faible. Dans la suite de ce chapitre, je discuterai plus en détails de l'impact de la structure de la population sur la capacité à envahir de la souche 2.

Réduction de transmission de la souche 1

Lorsque la réduction de transmission de la souche 1 est faible, le seuil de remplacement de la souche 2 par la souche 1 est rapidement atteint, un résultat déjà discuté plus tôt (Figure 2.4, panel bas-milieu). On remarque également que lorsque la souche 2 domine, sa prévalence est indépendante de la valeur de r_1 , un résultat attendu et qui correspond à ce que l'on observe lorsque la souche 1 est absente de la population (Figure 2.5, panel bas-milieu).

Réduction de transmission de la souche 2

Lorsque la réduction de transmission de la souche 2 est élevée, le seuil de remplacement est atteint pour de plus basses valeurs de couverture vaccinale, et même lorsque la souche 2 domine

sa prévalence est réduite (Figure 2.4, panel bas-droite). Il est particulièrement intéressant de remarquer que même lorsque la souche 1 est absente de la population il existe toujours un seuil au-delà duquel la souche 2 s'éteint (Figure 2.5, panel bas-droite). En revanche, ce seuil est resserré autour de plus hautes valeurs de couverture vaccinale et de réduction de transmission de la souche 2.

2.3.3 – Invasion par une unique souche d'une population saine

La capacité d'invasion de la souche 1 d'une population saine est indépendante de la structure vaccinale de la population, comme indiqué par l'équation (2.10). La souche 1 peut donc envahir une population saine si $r_1 < 1 - 1/\mathcal{R}_0$. Avec une valeur de $\mathcal{R}_0 = 2.6$, cela signifie que la souche 1 peut envahir tant que $r_1 < 0.615$. Cette valeur correspond à la ligne en tirets sur la Figure 2.3 (panel bas-milieu).

La capacité d'invasion de la souche 2 dépend, contrairement à celle de la souche 1, de la structure de la population saine, comme indiqué par l'équation (2.7). Lorsque la part vaccinée de la population au début de l'invasion est plus importante, le nombre de reproduction de la souche 2 diminue; de même lorsque la réduction de transmission de la souche 2 en population vaccinée est élevée (Figure 2.6). À noter que la couverture vaccinale n'est pas égale à la proportion d'individus vaccinés dans ce modèle (voir Annexe B.3 pour plus de détails à ce sujet).

2.3.4 – Invasion endémique, compétition et coexistence

Invasion par la souche 1

La souche 1 peut envahir une population où la souche 2 est à l'équilibre endémique quand l'inégalité (2.13) est respectée, donc lorsque $dI(0)/dt > 0$ avec la souche 2 à l'équilibre, c'est-à-dire lorsque le nombre de reproduction effectif de la souche 1 est supérieur à un. Les valeurs des paramètres pour lesquelles cette condition est respectée correspondent aux valeurs pour lesquelles la prévalence de la souche 1 est non nulle lorsque les deux souches circulent initialement dans la population (Figure 2.7).

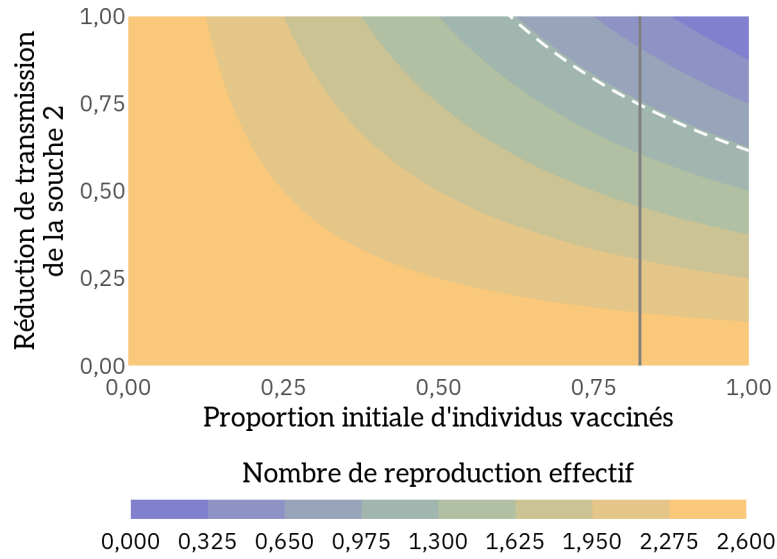


FIGURE 2.6 – **Nombre de reproduction effectif de la souche 2** en fonction de la proportion initiale d’individus vaccinés et de la réduction de transmission de la souche 2 en population vaccinée (paramètres \mathcal{R}_2 , v_{DFE} et r_2 , respectivement). La ligne blanche en tirets indique le seuil $\mathcal{R}_2 = 1$ et la ligne verticale grise indique la proportion d’individus vaccinés dans la population pour $p_p = p_b = 1$.

Invasion par la souche 2

La capacité d’invasion de la souche 2 est analogue à celle de la souche 1. Ici encore, les valeurs des paramètres pour lesquelles la souche 2 peut envahir une population où la souche 1 est à l’équilibre endémique correspondent aux valeurs pour lesquelles la souche 2 a une prévalence non nulle lorsque les deux souches circulent initialement dans la population (Figure 2.8).

Invasion mutuelle et coexistence

Les Figures 2.2, 2.4, 2.7 et 2.8 laissent entrevoir que la coexistence n’est globalement possible que pour une gamme resserrée de valeurs de paramètres. Ce phénomène apparaît de manière plus flagrante encore lorsque l’on combine les figures, que ce soit au niveau de la prévalence des souches (Figure 2.9) ou de leur nombre reproductif effectif (Figure 2.10).

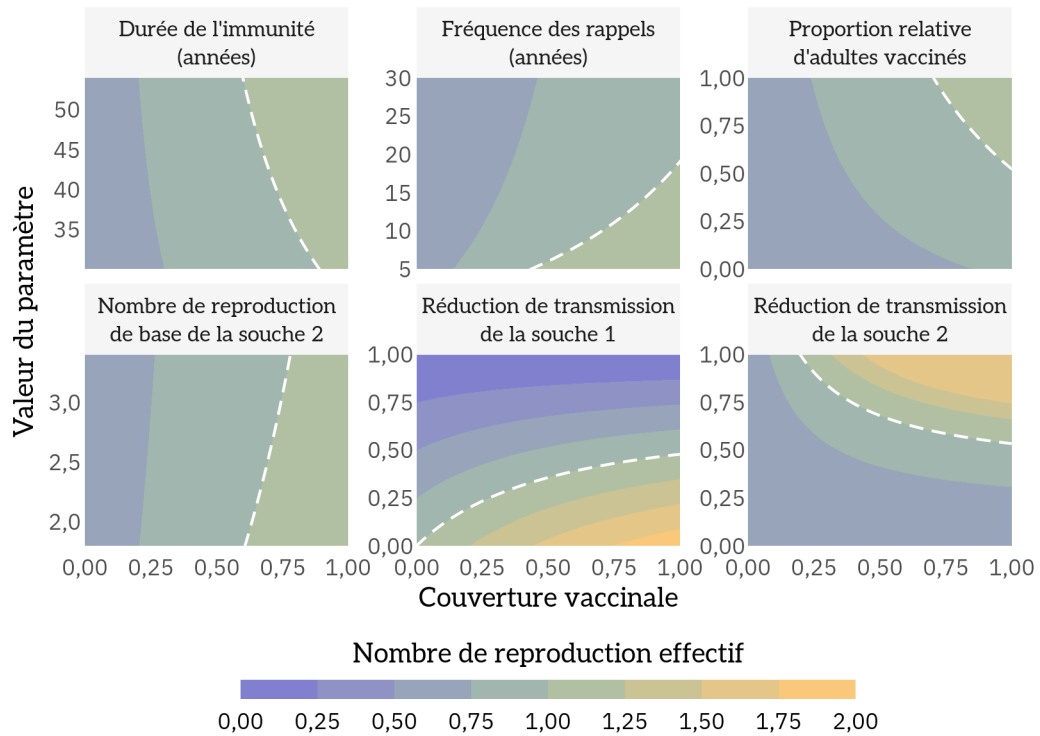


FIGURE 2.7 – **Nombre de reproduction effectif de la souche 1** lorsque la souche 2 est à l'équilibre endémique. Chaque panel représente un paramètre différent : durée de l'immunité $((365 \times \eta)^{-1})$, fréquence des rappels (f), proportion relative d'adultes vaccinés (p_b/p_p), nombre de reproduction de base de la souche toxigène (\mathcal{R}_0), réduction de transmission de la souche 1 (r_1) et réduction de transmission de la souche 2 (r_2). La ligne en tirets indique le seuil pour lequel le nombre de reproduction effectif de la souche 1 est égal à un.

Efficacité vaccinale et compétition

Comme cela a déjà été présenté, la réduction de transmission de la souche 2 (r_2) peut être interprétée comme l'efficacité du vaccin. Si $r_2 = 0$, le vaccin est totalement inefficace, et si $r_2 = 1$ alors le vaccin est parfaitement efficace et la souche 2 ne peut pas du tout se transmettre parmi des individus vaccinés. Combinées, une forte couverture vaccinale et une forte efficacité du vaccin conduisent à réduire le nombre de reproduction effectif de la souche 2 (\mathcal{R}_2). Dans une population où seule la souche 2 circulerait, les couverture et efficacité vaccinales nécessaires pour faire passer \mathcal{R}_2 sous le seuil critique d'éradication ($\mathcal{R}_2 < 1$) sont très élevées (Figure 2.11, panel droit). En revanche, lorsque la souche 1 circule également dans la population, ce seuil est décalé vers de plus basses valeurs d'efficacité et de couverture vaccinales (Figure 2.11, panel gauche). Cela signifie

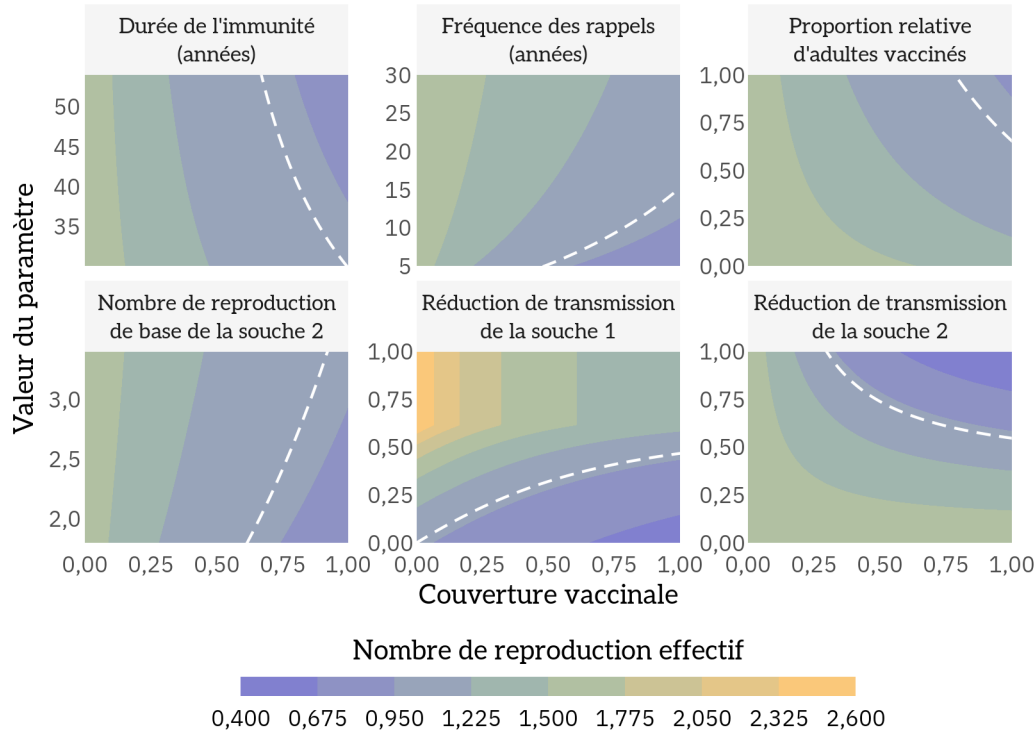


FIGURE 2.8 – **Nombre de reproduction effectif de la souche 2** lorsque la souche 1 est à l'équilibre endémique. Chaque panel représente un paramètre différent : durée de l'immunité ($(365 \times \eta)^{-1}$), fréquence des rappels (f), proportion relative d'adultes vaccinés (p_b/p_p), nombre de reproduction de base de la souche toxino-gène (\mathcal{R}_0), réduction de transmission de la souche 1 (r_1) et réduction de transmission de la souche 2 (r_2). La ligne en tirets indique le seuil pour lequel le nombre de reproduction effectif de la souche 1 est égal à un.

que lorsque les deux souches sont en compétition la pression exercée par la souche 1 facilite l'éradication de la souche 2. Cet important résultat est discuté plus en détails dans la discussion générale.

2.4 – Discussion générale

Le but de ce travail a été de développer un modèle pour comprendre l'évolution de la virulence de la bactérie responsable de la diphtérie, *C. diphtheriae*, en étudiant la dynamique de deux souches de cette bactérie dans un contexte de vaccination de masse. Le modèle produit prend en considération à la fois le coût associé à la production de la toxine diphtérique et l'effet d'un vaccin anatoxinique visant spécifiquement la TD (et donc les souches toxino-gènes). Son étude a permis

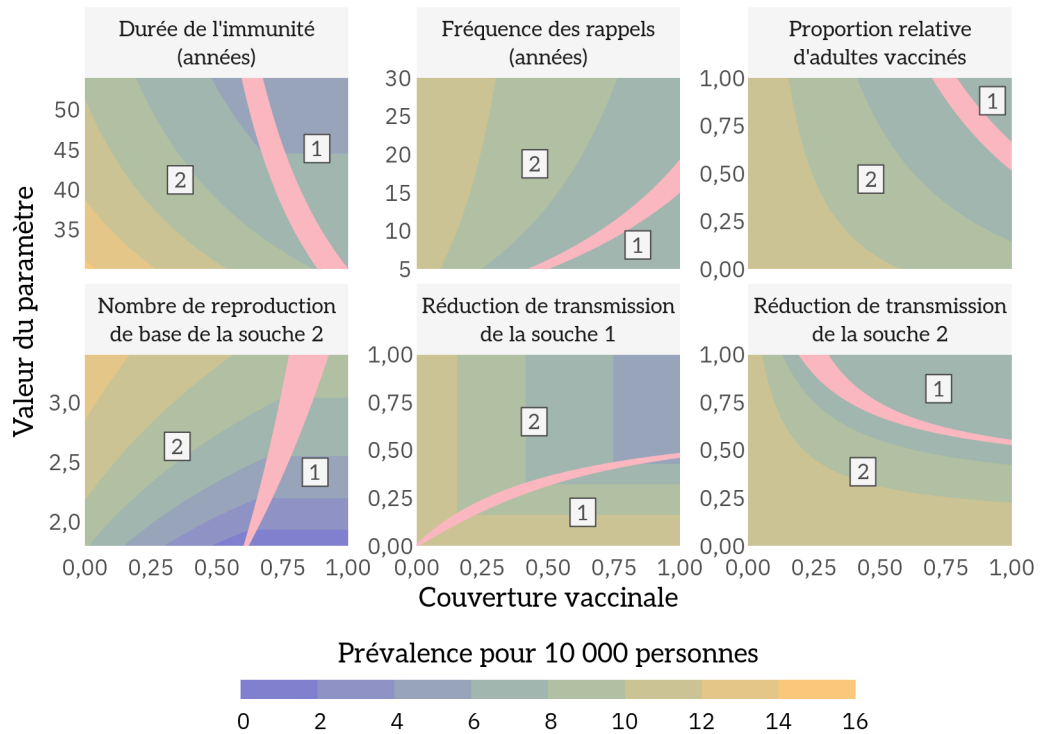


FIGURE 2.9 – **Prévalence de la souche 1 ou de la souche 2** lorsque les deux souches circulent dans la population (nombre d'infectés pour 10 000 habitants, zones labellisées « 1 » et « 2 », respectivement). La zone rose séparant les zones labellisées indique la gamme de valeurs dans laquelle les deux souches coexistent. Chaque panel représente un paramètre différent : durée de l'immunité $((365 \times \eta)^{-1})$, fréquence des rappels (f), proportion relative d'adultes vaccinés (p_b/p_p), nombre de reproduction de base de la souche toxino-gène (\mathcal{R}_0), réduction de transmission de la souche 1 (r_1) et réduction de transmission de la souche 2 (r_2).

de conclure que l'utilisation d'un tel vaccin peut mener à la sélection de souches bactériennes ne produisant pas la TD et donc, *de facto*, moins virulentes. En utilisant ce modèle, j'ai évalué le rôle de différents paramètres sur la possibilité pour chaque souche d'être sélectionnée et les conditions sous lesquelles la coexistence serait possible. J'ai pu mettre en évidence que lorsque la couverture vaccinale est forte à la fois chez les nouveaux-nés (vaccination primaire) et chez les adolescents/adultes (rappels), la souche non toxino-gène est sélectionnée. Une durée réduite entre deux rappels vaccinaux chez les adultes favorise également la sélection de la souche non toxino-gène.

Ces trois paramètres, la couverture vaccinale primaire, la couverture vaccinale des rappels, et la fréquence des rappels, varient fortement entre pays et sont ceux sur lesquels il est le plus

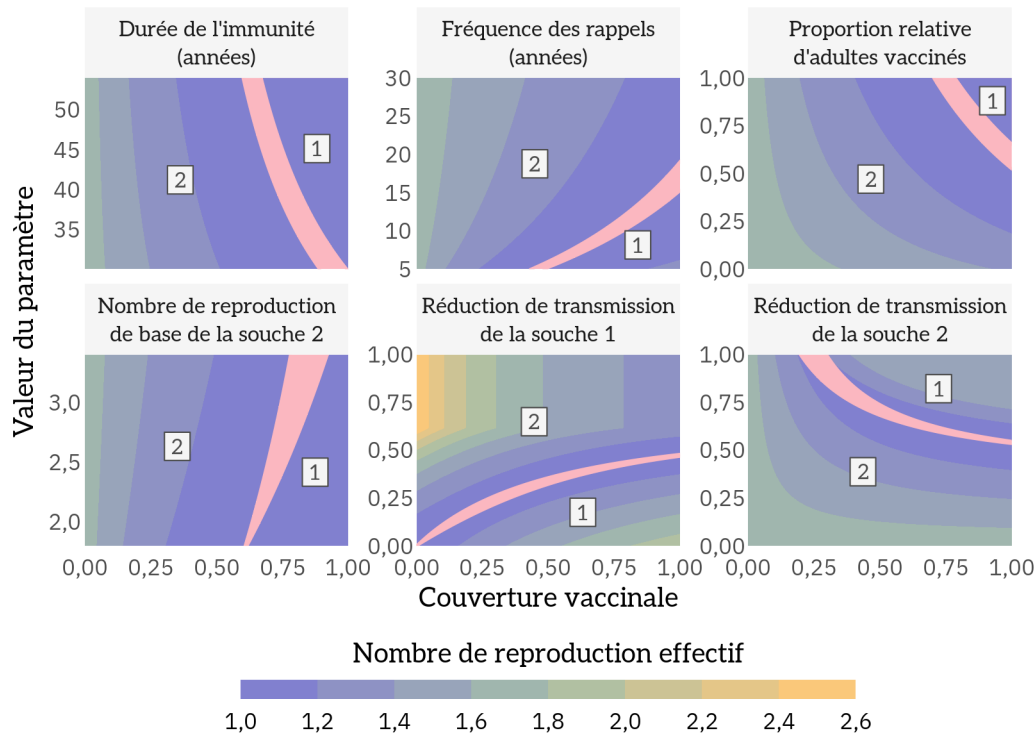


FIGURE 2.10 – Nombre de reproduction effectif de la souche 1 ou de la souche 2 lorsque la souche antagoniste est à l'équilibre endémique dans la population (zones labellisées « 1 » et « 2 », respectivement). La zone rose séparant les zones labellisées indique la gamme de valeurs dans laquelle les deux souches coexistent. Chaque panel représente un paramètre différent : durée de l'immunité ($(365 \times \eta)^{-1}$), fréquence des rappels (f), proportion relative d'adultes vaccinés (p_b/p_p), nombre de reproduction de base de la souche toxigène (\mathcal{R}_0), réduction de transmission de la souche 1 (r_1) et réduction de transmission de la souche 2 (r_2).

possible d'agir par des mesures de santé publique. Si la couverture vaccinale contre la diphtérie chez les très jeunes enfants est globalement très élevée dans la plupart des pays, d'importantes fluctuations peuvent apparaître dans certaines régions, à certaines périodes, ou encore à certains âges de la vie (Abdi *et al.*, 2021; Lee et Choi, 2017; Marchal *et al.*, 2021; Murhekar *et al.*, 2009; Sein *et al.*, 2016). Cela peut mener au sein de certaines populations à une réduction de la pression de sélection contre les souches toxigènes. Les rappels vaccinaux sont généralement recommandés sans être obligatoires. Ainsi, en France, alors qu'ils sont recommandés à 25, 45 et 65 ans, puis tous les dix ans au-delà (Ministère des Solidarités et de la Santé, 2022), le pourcentage d'adultes faisant leurs rappels est dans les faits très bas et diminue fortement avec l'âge.

Les conclusions de ce travail concordent avec d'autres travaux publiés précédemment. [Deka](#)

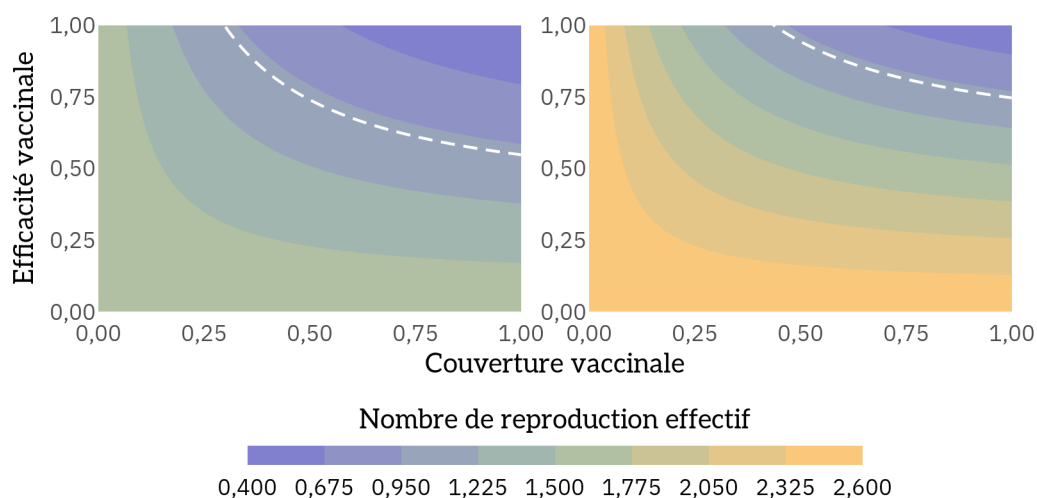


FIGURE 2.11 – **Nombre de reproduction effectif de la souche 2** en fonction de la couverture vaccinale et de l'efficacité du vaccin (paramètres \mathcal{R}_2 , $p_p = p_b$ et r_2 , respectivement) lorsque la souche 2 envahit une population où la souche 1 est endémique (panel gauche) ou une population où la population est saine (panel droit). La ligne blanche en tirets indique le seuil $\mathcal{R}_2 = 1$.

et Bhattacharyya (2022) ont montré, avec un modèle générique similaire à celui présenté ici, qu'une souche avec un \mathcal{R}_0 bas peut émerger et se propager lorsqu'un vaccin ciblant une autre souche ayant un plus grand \mathcal{R}_0 est en circulation. Ces conclusions concordent également avec les données cliniques obtenues en Algérie (Benamrouche *et al.*, 2016), en France (Badell-Ocando *et al.*, 2022), en Allemagne (Dangel *et al.*, 2018) ou encore en Pologne (Czajka *et al.*, 2018; Zasada, 2013; Zasada et Rzeckowska, 2019). Dans chacun de ces pays, l'utilisation des vaccins anatoxiques a été reliée à une augmentation de la prévalence des souches non toxigènes. Il est important de noter cependant que certains pays avec une bonne surveillance des souches de *C. diphtheriae*, comme la France, continuent de détecter des souches toxigènes dans leurs populations (Badell-Ocando *et al.*, 2022).

Comme expliqué plus tôt, la couverture vaccinale au sein d'une population peut fortement varier, provoquant un retour des souches toxigènes. Mais il est également possible que la vaccination ait été introduite dans certains pays ou territoires trop récemment pour que les souches non toxigènes aient eu le temps de remplacer complètement les souches toxigènes. Dans mes simulations, l'équilibre est long à atteindre parce que le modèle considère des paramètres démographiques et la perte de l'immunité, des phénomènes biologiques s'étendant sur une durée

largement plus élevée que les phénomènes épidémiologiques. Il aurait donc été intéressant, au-delà de l'étude des équilibres du système, de regarder la manière dont la dynamique compétitive entre souche toxigène et souche non toxigène est perturbée dans les premiers temps de la vaccination et dans les décennies suivantes; autrement dit, de s'intéresser à l'évolution transitoire de *C. diphtheriae* à l'ère vaccinale. Cette évolution transitoire est de première importance en santé publique, puisqu'elle permet de comprendre à une échelle temporelle courte à moyenne le rôle de la vaccination sur l'évolution des pathogènes.

Il reste que ce travail montre que sur le long terme, et sous réserve que les politiques de santé publique se renforcent, les vaccins anatoxiques devraient mener à la sélection de souches ne produisant plus la TD à cause du coût métabolique associé à la production de la toxine qui n'est plus compensé par un bénéfice si l'hôte est résistant à l'action de la toxine.

Les autres paramètres du modèle (*i.e.*, les paramètres non vaccinaux) sont plus contraints, dépendant de la biologie du pathogène et de l'hôte plutôt que de mesures de santé publique. Comme attendu, plus la capacité d'une souche à se reproduire est grande, plus celle-ci a de chances de se maintenir dans la population. Il en résulte que plus la souche toxigène est naturellement avantagée par rapport à la souche non toxigène, plus elle sera apte à se maintenir dans la population. Il est difficile de directement contrôler la transmission d'un pathogène, mais des mesures de santé publique peuvent contribuer à réduire son nombre de reproduction de base (\mathcal{R}_0) comme cela a été démontré au cours de la récente pandémie de Covid-19 (Talic *et al.*, 2021). En plus de mesures visant à réduire les contacts des individus infectés, il est possible de réduire la période infectieuse (*e.g.*, avec des antibiotiques) pour réduire la transmission d'un pathogène (Truelove *et al.*, 2020). Ces mesures, en ciblant des individus particuliers, pourraient réduire la transmissibilité des souches virulentes (*i.e.*, des souches toxigènes) sans impacter la circulation des souches avirulentes.

Les résultats de ce travail illustrent également comment la compétition entre différentes souches pour l'accès à une ressource (ici, les hôtes) peut faciliter l'éradication d'une souche toxigène. La coexistence des deux types de souche est ici improbable, cependant il est possible que des fluctuations dans la couverture vaccinale, en lien avec la perception du public du risque associé à l'infection, crée des contextes dans lesquels la coexistence devient possible (Deka *et Bhattacharyya*, 2022). La souche non toxigène réduit le nombre d'hôtes potentiels que la

souche toxigène pourrait utiliser pour se maintenir dans la population. Dans les simulations, j'ai montré que même en étudiant une gamme étendue de valeurs des différents paramètres, la souche toxigène pouvait difficilement être éliminée par la simple action d'une forte couverture vaccinale. Cela est dû au fait que les vaccins antidiphtériques ne bloquent pas l'infection par le pathogène ni sa transmission. Comme cela a déjà été dit, la diminution observée de la transmission due à la vaccination s'explique par la réduction du bénéfice lié à la production de la TD associé au coût métabolique de celle-ci.

Cependant, si le coût à la production de la TD est suffisamment élevé, alors la vaccination seule peut mener à l'éradication de la souche toxigène. Les données sont encore limitées sur l'impact de la vaccination sur la réduction de transmissibilité de *C. diphtheriae* (mais voir [Truelove et al., 2020](#)), mais des études sur la coqueluche montrent que l'utilisation de vaccins à anatoxine – qui contiennent en fait un à cinq antigènes – n'est pas suffisante pour stopper la propagation des souches pathogènes ([Warfel et al., 2014](#)). De la même manière, la compétition seule ne permet pas l'éradication de la souche toxigène, la souche non toxigène étant systématiquement éradiquée à cause de son taux de transmission trop bas lorsque la vaccination est inexistante. C'est donc la combinaison de la vaccination et de la compétition qui peut mener à l'éradication des souches virulentes de *C. diphtheriae*.

2.5 – Conclusion

Le but du travail présenté dans ce chapitre était de mettre en évidence l'impact d'un vaccin anatoxinique sur la sélection de souches ne produisant plus la toxine contre laquelle le vaccin stimule le système immunitaire de l'hôte. J'ai utilisé la diphtérie comme modèle biologique pour sa relative simplicité : il s'agit d'un pathogène bien connu, contre lequel on utilise depuis des décennies un vaccin ciblant spécifiquement une toxine produite par son pathogène, la toxine diphtérique. La présence d'un unique antigène dans les vaccins permet en effet de réduire la quantité d'effets que ces derniers peuvent avoir sur l'évolution de *C. diphtheriae*.

D'autres systèmes similaires à la diphtérie et de son vaccin existent, notamment celui de la coqueluche. Les vaccins anticoquelucheux modernes ciblent un ou quelques antigènes produits par *Bordetella pertussis*, la bactérie responsable de la maladie, comme les vaccins antidiphtériques

ciblent un antigène en particulier de *C. diphtheriae*. Cependant, le système coqueluche a un avantage par rapport au système diphtérie : il est plus facile, dans la pratique, de mettre en évidence l'émergence de souches ne produisant plus un ou plusieurs antigènes vaccinaux. Par exemple, la littérature foisonne de publications montrant l'émergence de souches ne produisant pas la pertactine, un antigène commun des vaccins (Section 1.4.3). Dans le prochain chapitre, j'utilise un jeu de données précédemment publié pour essayer de mettre en évidence les facteurs autres que vaccinaux influençant la vitesse d'émergence des souches dépourvues de la pertactine aux États-Unis.



BIBLIOGRAPHIE

- ABDI I., GIDDING H., LEONG R. N., MOORE H. C., SEALE H. et MENZIES R. (2021) : « Vaccine coverage in children born to migrant mothers in Australia: A population-based cohort study » *Vaccine*, vol. 39, n°6, p./id. 984–993.
- AMANNA I. J., CARLSON N. E. et SLIFKA M. K. (2007) : « Duration of Humoral Immunity to Common Viral and Vaccine Antigens » *New England Journal of Medicine*, vol. 357, n°19, p./id. 1903–1915.
- BADELL-OCANDO E., TOUBIANA J. et BRISSE S. (2022) : « Rapport annuel d'activité du Centre national de Référence des Corynébactéries du complexe *diphtheriae* – Bilan des activités 2017-2021 » Institut Pasteur (Paris, France).
- BARATIN D., DEL SIGNORE C., THIERRY J., CAULIN E. et VANHEMS P. (2012) : « Evaluation of adult dTPaP vaccination coverage in France: experience in Lyon city, 2010–2011 » *BMC Public Health*, vol. 12, p./id. 940.
- BENAMROUCHE N., HASNAOUI S., BADELL E., GUETTOU B., LAZRI M., GUIO N. et RAHAL K. (2016) : « Microbiological and molecular characterization of *Corynebacterium diphtheriae* isolated in Algeria between 1992 and 2015 » *Clinical Microbiology and Infection*, vol. 22, n°12, p./id. 1005.e1–1005.e7.
- BULL J. J. et ANTIA R. (2022) : « Which 'imperfect vaccines' encourage the evolution of higher virulence? » *Evolution, Medicine, and Public Health*, vol. 10, n°1, p./id. 202–213.
- CRUZ-HERVERT L. P., FERREIRA-GUERRERO E., DÍAZ-ORTEGA J. L., TREJO-VALDIVIA B., TÉLLEZ-ROJO M. M., MONGUA-RODRÍGUEZ N., HERNÁNDEZ-SERRATO M. I., MONTOYA-RODRÍGUEZ A. A. et GARCÍA-GARCÍA L. (2013) : « Cobertura de vacunación en adultos y adultos mayores en México » *Salud Pública de México*, vol. 55, p./id. S300–S306.
- CZAJKA U., WIATRZYK A., MOSIEJ E., FORMIŃSKA K. et ZASADA A. A. (2018) : « Changes in MLST profiles and biotypes of *Corynebacterium diphtheriae* isolates from the diphtheria outbreak period to the period of invasive infections caused by nontoxigenic strains in Poland (1950–2016) » *BMC Infectious Diseases*, vol. 18, p./id. 121.
- DANGEL A., BERGER A., KONRAD R., BISCHOFF H. et SING A. (2018) : « Geographically Diverse Clusters of Nontoxigenic *Corynebacterium diphtheriae* Infection, Germany, 2016–2017 » *Emerging Infectious Diseases*, vol. 24, n°7, p./id. 1239–1245.
- DE SERRES G., SHADMANI R., DUVAL B., BOULIANNE N., DÉRY P., DOUVILLE FRADET M., ROCHETTE L. et HALPERIN S. A. (2000) : « Morbidity of Pertussis in Adolescents and Adults » *The Journal of Infectious Diseases*, vol. 182, n°1, p./id. 174–179.
- DEKA A. et BHATTACHARYYA S. (2022) : « The effect of human vaccination behaviour on strain competition in an infectious disease: An imitation dynamic approach » *Theoretical Population Biology*, vol. 143, p./id. 62–76.
- DELAMATER P. L., STREET E. J., LESLIE T. F., YANG Y. T. et JACOBSEN K. H. (2019) : « Complexity of the Basic Reproduction Number (R_0) » *Emerging Infectious Diseases*, vol. 25, n°1, p./id. 1–4.

-
- DOBIE R. A. et TOBEY D. N. (1979) : « Clinical Features of Diphtheria in the Respiratory Tract » *JAMA*, vol. 242, n°20, p./id. 2197.
- EDWARDS K. M. et DECKER M. D. (2012) : « Pertussis vaccines » *Vaccines* 6^{ème} édition, p. 447–492, Elsevier (Amsterdam, Pays-Bas).
- FINE P. E. M. (1993) : « Herd Immunity: History, Theory, Practice » *Epidemiologic Reviews*, vol. 15, n°2, p./id. 265–302.
- GANDON S. et DAY T. (2008) : « Evidences of parasite evolution after vaccination » *Vaccine*, vol. 26, n°3, p./id. C4–C7.
- GANDON S., MACKINNON M., NEE S. et READ A. (2003) : « Imperfect vaccination: some epidemiological and evolutionary consequences » *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 270, p./id. 1129–1136.
- GANDON S., MACKINNON M. J., NEE S. et READ A. F. (2001) : « Imperfect vaccines and the evolution of pathogen virulence » *Nature*, vol. 414, n°6865, p./id. 751–756.
- GANDON S., MACKINNON M. J., NEE S. et READ A. F. (2002) : « Antitoxin vaccines and pathogen virulence » *Nature*, vol. 417, n°6889, p./id. 610.
- GILADI M., SCHULMAN A., KEDEM R. et DANON Y. L. (1987) : « Measles in adults: a prospective study of 291 consecutive cases. » *BMJ*, vol. 295, n°6609, p./id. 1314.
- GRAINGER T. N., LEVINE J. M. et GILBERT B. (2019) : « The Invasion Criterion: A Common Currency for Ecological Research » *Trends in Ecology & Evolution*, vol. 34, n°10, p./id. 925–935.
- GUISSO N., SOUBEYRAND B. et MACINA D. (2022) : « Can vaccines control bacterial virulence and pathogenicity? *Bordetella pertussis*: the advantage of fitness over virulence » *Evolution, Medicine, and Public Health*, vol. 10, n°1, p./id. 363–370.
- HAMMARLUND E., THOMAS A., POORE E. A., AMANNA I. J., RYNKO A. E., MORI M., CHEN Z. et SLIFKA M. K. (2016) : « Durability of Vaccine-Induced Immunity Against Tetanus and Diphtheria Toxins: A Cross-sectional Analysis » *Clinical Infectious Diseases*, vol. 62, n°9, p./id. 1111–1118.
- KANITZ E. E., WU L. A., GIAMBI C., STRIKAS R. A., LEVY-BRUHL D., STEFANOFF P., MERECKIENE J., APPELGREN E., D’ANCONA F., VENICE (VACCINE EUROPEAN et NEW INTEGRATED COLLABORATION EFFORT) NATIONAL GATEKEEPERS, CONTACT POINTS (2012) : « Variation in adult vaccination policies across Europe: An overview from VENICE network on vaccine recommendations, funding and coverage » *Vaccine*, vol. 30, n°35, p./id. 5222–5228.
- LAVINE J. S., KING A. A. et BJØRNSTAD O. N. (2011) : « Natural immune boosting in pertussis dynamics and the potential for long-term vaccine failure » *Proceedings of the National Academy of Sciences*, vol. 108, n°17, p./id. 7259–7264.
- LECORVAISIER F., PONTIER D., SOUBEYRAND B. et FOUCHET D. (2024) : « Using a dynamical model to study the impact of a toxoid vaccine on the evolution of a bacterium: The example of diphtheria » *Ecological Modelling*, vol. 487, p./id. 110569.
- LEE H.-J. et CHOI J.-H. (2017) : « Tetanus–diphtheria–acellular pertussis vaccination for adults: an update » *Clinical and Experimental Vaccine Research*, vol. 6, n°1, p./id. 22.
- LI S., MA C., HAO L., SU Q., AN Z., MA F., XIE S., XU A., ZHANG Y., DING Z., LI H., CAIRNS L., WANG H., LUO H., WANG N., LI L. et FERRARI M. J. (2017) : « Demographic transition and the dynamics of measles in six provinces in China: A modeling study » *PLOS Medicine*, vol. 14, n°4, p./id. e1002255.

- LU Y., GENG Y., GAN W. et SHI L. (2019) : « Impacts of conformist on vaccination campaign in complex networks » *Physica A: Statistical Mechanics and its Applications*, vol. 526, p./id. 121124.
- MARCHAL C., BELHASSEN M., GUISSO N., JACOUD F., VAN GANSE E., LE PANNERER M., COHEN R., VERDIER R. et UHART M. (2021) : « Vaccination coverage rates for Diphtheria, Tetanus, Poliomyelitis and Pertussis booster vaccination in France between 2013 and 2017: Learnings from an analysis of National Health System Real-World Data » *Vaccine*, vol. 39, n°3, p./id. 505–511.
- MATSUYAMA R., AKHMETZHANOV A. R., ENDO A., LEE H., YAMAGUCHI T., TSUZUKI S. et NISHIURA H. (2018) : « Uncertainty and sensitivity analysis of the basic reproduction number of diphtheria: a case study of a Rohingya refugee camp in Bangladesh, November–December 2017 » *PeerJ*, vol. 6, p./id. e4583.
- MCLEOD D. V. et GANDON S. (2022) : « Effects of epistasis and recombination between vaccine-escape and virulence alleles on the dynamics of pathogen adaptation » *Nature Ecology & Evolution*, vol. 6, n°6, p./id. 786–793.
- MINISTÈRE DES SOLIDARITÉS ET DE LA SANTÉ (2022), « Calendrier des vaccinations et recommandations vaccinales 2022 ».
- MÖLLER J., KRANER M., SONNEWALD U., SANGAL V., TITTLBACH H., WINKLER J., WINKLER T. H., MELNIKOV V., LANG R., SING A., MATTOS-GUARALDI A. L. et BURKOVSKI A. (2019) : « Proteomics of diphtheria toxoid vaccines reveals multiple proteins that are immunogenic and may contribute to protection of humans against *Corynebacterium diphtheriae* » *Vaccine*, vol. 37, n°23, p./id. 3061–3070.
- MURHEKAR M. V., BITRAGUNTA S., HUTIN Y., CKAARAVARTY A., SHARMA H. J. et GUPTA M. D. (2009) : « Immunization coverage and immunity to diphtheria and tetanus among children in Hyderabad, India » *Journal of Infection*, vol. 58, n°3, p./id. 191–196.
- MURPHY J. R. (1996) : « *Corynebacterium Diphtheriae* » *Medical Microbiology* 4^{ème} édition, University of Texas Medical Branch at Galveston (Galveston, États-Unis).
- ORAM D. M. et HOLMES R. K. (2006) : « Diphtheria Toxin » *The Comprehensive Sourcebook of Bacterial Protein Toxins* 3^{ème} édition, p. 245–256, Academic Press (Paris, France).
- PAPPENHEIMER, Jr. A. M. (1984) : « Diphtheria » *Bacterial Vaccines*, p. 1–36, Academic Press (Orlando, États-Unis).
- PAPPENHEIMER, Jr. A. M. et GILL D. M. (1973) : « Diphtheria: Recent studies have clarified the molecular mechanisms involved in its pathogenesis » *Science*, vol. 182, n°4110, p./id. 353–358.
- PARVEEN S., BISHAI W. R. et MURPHY J. R. (2019) : « *Corynebacterium diphtheriae* : Diphtheria Toxin, the *tox* Operon, and Its Regulation by Fe²⁺ Activation of apo-DtxR » *Microbiology Spectrum*, vol. 7, n°4, p./id. 7.4.4.
- POETHKO-MÜLLER C. et SCHMITZ R. (2013) : « Vaccination coverage in German adults: Results of the German Health Interview and Examination Survey for Adults (DEGS1) » *Bundesgesundheitsbl*, vol. 56, n°5, p./id. 845–857.
- PULLAN S. et DEY M. (2021) : « Vaccine hesitancy and anti-vaccination in the time of COVID-19: A Google Trends analysis » *Vaccine*, vol. 39, n°14, p./id. 1877–1881.
- R CORE TEAM (2022) : *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing (Vienne, Autriche).
- ROHANI P., ZHONG X. et KING A. A. (2010) : « Contact Network Structure Explains the Changing Epidemiology of Pertussis » *Science*, vol. 330, n°6006, p./id. 982–985.

-
- SCHEIFELE D. W. et OCHNIO J. J. (2009) : « Module 2: Diphtheria – Update 2009 » *The Immunological Basis for Immunization Series*, World Health Organization (Genève, Suisse).
- SEIN C., TIWARI T., MACNEIL A., WANNEMUEHLER K., SOULAPHY C., SOULIPHONE P., REYBURN R., RAMIREZ GONZALEZ A., WATKINS M. et GOODSON J. L. (2016) : « Diphtheria outbreak in Lao People’s Democratic Republic, 2012–2013 » *Vaccine*, vol. 34, n°36, p./id. 4321–4326.
- SOETAERT K., PETZOLDT T., SETZER R. W., BROWN P. N., BYRNE G. D., HAIRER E., HINDMARSH A. C., MOLER C., PETZOLD L. R., SAAD Y. et ULRICH C. W. (2023), « deSolve: Solvers for Initial Value Problems of Differential Equations ('ODE', 'DAE', 'DDE') ».
- SOETAERT K., PETZOLDT T. et WOODROW SETZER R. (2010) : « Solving Differential Equations in R: Package deSolve » *Journal of Statistical Software*, vol. 33, n°9, p./id. 1–25.
- SOUBEYRAND B. et PLOTKIN S. A. (2002) : « Antitoxin vaccines and pathogen virulence » *Nature*, vol. 417, n°6889, p./id. 609–610.
- TALIC S., SHAH S., WILD H., GASEVIC D., MAHARAJ A., ADEMI Z., LI X., XU W., MESA-EGUIAGARAY I., ROSTRON J., THEODORATOU E., ZHANG X., MOTEE A., LIEW D. et LIC D. (2021) : « Effectiveness of public health measures in reducing the incidence of covid-19, SARS-CoV-2 transmission, and covid-19 mortality: systematic review and meta-analysis » *BMJ*, p. e068302.
- TIWARI T. S. P. et WHARTON M. (2012) : « Diphtheria Toxoid » *Vaccines* 6^{ème} édition, p. 153–166, Elsevier (Amsterdam, Pays-Bas).
- TRUELOVE S. A., KEEGAN L. T., MOSS W. J., CHAISSON L. H., MACHER E., AZMAN A. S. et LESSLER J. (2020) : « Clinical and Epidemiological Aspects of Diphtheria: A Systematic Review and Pooled Analysis » *Clinical Infectious Diseases*, vol. 71, n°1, p./id. 89–97.
- WALTER A. et LION S. (2021) : « Epidemiological and evolutionary consequences of periodicity in treatment coverage » *Proceedings of the Royal Society B: Biological Sciences*, vol. 288, n°1946, p./id. 20203007.
- WARFEL J. M., ZIMMERMAN L. I. et MERKEL T. J. (2014) : « Acellular pertussis vaccines protect against disease but fail to prevent infection and transmission in a nonhuman primate model » *Proceedings of the National Academy of Sciences*, vol. 111, n°2, p./id. 787–792.
- WENDELBOE A. M., VAN RIE A., SALMASO S. et ENGLUND J. A. (2005) : « Duration of Immunity Against Pertussis After Natural Infection or Vaccination » *Pediatric Infectious Disease Journal*, vol. 24, n°5, p./id. S58–S61.
- WILLIAMS W. W., LU P.-J., O’HALLORAN A., KIM D. K., GROHSCOPF L. A., PILISHVILI T., SKOFF T. H., NELSON N. P., HARPAZ R., MARKOWITZ L. E., RODRIGUEZ-LAINZ A. et FIEBELKORN A. P. (2017) : « Surveillance of Vaccination Coverage among Adult Populations — United States, 2015 » *MMWR Surveillance Summaries*, vol. 66, n°11, p./id. 1–28.
- WORLD HEALTH ORGANISATION (2017) : « Diphtheria vaccine: WHO position paper » *Weekly Epidemiological Record*, vol. 92, n°31, p./id. 417–435.
- YASUNAGA H., SHI Y., TAKEUCHI M., HORIGUCHI H., HASHIMOTO H., MATSUDA S. et OHE K. (2010) : « Measles-related Hospitalizations and Complications in Japan, 2007-2008 » *Internal Medicine*, vol. 49, n°18, p./id. 1965–1970.
- ZASADA A. A. (2013) : « Nontoxigenic Highly Pathogenic Clone of *Corynebacterium diphtheriae*, Poland, 2004–2012 » *Emerging Infectious Diseases*, vol. 19, n°11.
- ZASADA A. A. et RZECZKOWSKA M. (2019) : « Nontoxigenic *Corynebacterium diphtheriae* Infections, Europe » *Emerging Infectious Diseases*, vol. 25, n°7, p./id. 1437–1438.
- ZOU R., DUAN X., HAN Z., LU Y. et MA K. (2023) : « What information sources can prevent the epidemic: Local information or kin information? » *Chaos, Solitons & Fractals*, vol. 168, p./id. 113104.

CHAPITRE III

MODÉLISATION STATISTIQUE : FACTEURS

ENVIRONNEMENTAUX ET PROPAGATION DE SOUCHES

MUTANTES

3.1 – Introduction

3.1.1 – De la modélisation dynamique à l’analyse de données

Dans le chapitre précédent, j’ai présenté un modèle théorique décrivant les capacités d’invasion d’une souche mutante de *Corynebacterium diphtheriae*, l’agent étiologique de la diphtérie. Cette souche mutante est dépourvue de la toxine diphtérique (TD), l’unique antigène théoriquement présent dans les vaccins antidiphtériques (mais voir [Möller *et al.*, 2019](#) sur la contamination des vaccins par d’autres agents potentiellement immunogènes). Le modèle développé a permis de montrer que la souche mutante devrait être sélectionnée lorsque la couverture vaccinale est élevée et que le coût à produire la TD est grand (conclusions publiées dans [Lecorvaisier *et al.*, 2024](#)). Enfin, j’ai montré que la compétition avec cette souche mutante facilitait l’éradication de la souche résidente.

Le remplacement de souches résidentes de pathogènes par des souches capables d’échapper à la pression vaccinale est un phénomène largement connu, comme je l’ai présenté dans la Section 1.4. Paradoxalement, bien que le phénomène soit connu et documenté, les données sur le remplacement des souches toxigènes de *C. diphtheriae* par des souches non toxigènes sont assez rares. Cette rareté s’explique partiellement par le rôle joué par la toxine diphtérique dans la

pathogénicité de *C. diphtheriae*. La TD est le principal facteur de virulence de la bactérie, ce qui signifie qu'elle est responsable de la forme clinique de la diphtérie (Guiso *et al.*, 2022; Murphy, 1996). Une infection par une souche non toxigène peut également provoquer des symptômes, mais ces symptômes ne seront habituellement pas ceux de la diphtérie (Guiso *et al.*, 2022). La détection de souches non toxigènes se fait le plus souvent lorsqu'une infection respiratoire est suspectée dans le but de déterminer le pathogène à son origine. Cependant, *C. diphtheriae* n'est pas toujours dépistée en routine, notamment en Europe où la bactérie n'est plus endémique (Wagner *et al.*, 2011). Un individu infecté par une souche non toxigène de *C. diphtheriae* aura donc moins de chances d'être diagnostiqué que s'il était infecté par une souche toxigène.

L'objectif dans ce chapitre est d'étudier les facteurs influençant la vitesse d'émergence d'une souche mutante capable d'échapper à la pression vaccinale, c'est-à-dire le différentiel dans la fréquence relative de cette souche entre pas de temps successifs. Pour cet objectif, il est nécessaire d'obtenir des données longitudinales de prévalence de souches résidentes et mutantes (*i.e.*, présentant des modifications phénotypiques facilitant l'évasion vaccinale) pour le pathogène d'étude. Cependant, comme je viens de le présenter, de telles données sont rares pour *C. diphtheriae*. Je propose donc de changer de modèle biologique et de me concentrer sur l'agent étiologique de la coqueluche, *Bordetella pertussis*. Comme présenté dans la Section 1.4.3, les vaccins anticoquelucheux sont similaires aux vaccins antidiphtériques. Ceux-ci sont composés d'un nombre réduit d'antigènes, cependant les vaccins anticoquelucheux peuvent en contenir un à cinq là où les vaccins antidiphtériques n'en contiennent qu'un. Cette multiplicité des antigènes dans les vaccins anticoquelucheux représente un avantage pour la détection de souches mutantes de *B. pertussis*, comme je l'expliquerai sans la Section 3.3.1, justifiant le choix de ce nouveau modèle biologique.

Dans ce chapitre, je présente une nouvelle méthode développée pour analyser des données longitudinales d'invasion d'une souche mutante d'un pathogène, méthode généralisable à différents pathogènes – voire, généralisable hors du champ de l'épidémiologie – et j'applique cette méthode sur un jeu de données d'invasion de souches mutantes de *B. pertussis*.

3.1.2 – Contexte et mise en situation

En écologie, une invasion peut être définie comme le déplacement d'individus d'une espèce depuis une région d'où cette espèce est native vers une nouvelle région. Une fois établie dans la nouvelle région, l'espèce invasive aura un effet délétère sur l'environnement d'accueil, et conduira souvent à l'extinction locale ou à une forte diminution d'effectifs d'espèces indigènes (Mack *et al.*, 2000). Les activités humaines telles que le transport, les déplacements et le commerce sont connues comme étant des facteurs influençant la répartition dans l'espace et le temps d'espèces non indigènes à travers le monde, et jouent ainsi un rôle de première importance dans les invasions écologiques (Mack *et al.*, 2000).

En épidémiologie, une invasion peut être définie d'une manière analogue à une invasion écologique et peut être représentée de différentes façons. Un agent pathogène peut étendre sa zone de répartition au-delà de sa région d'endémie, ou peut également passer d'une espèce hôte à une nouvelle et se propager au sein de cette nouvelle espèce (Encadré 1.1). Ce phénomène a été très bien illustré par de nombreuses maladies et pathogènes au cours des XX^e et XXI^e siècles, voire plus tôt encore (Taubenberger et Morens, 2010). Les exemples les plus parlants sont peut-être les multiples épidémies de grippe ayant entaché le XX^e siècle, telles que la grippe espagnole (1918-1920), la grippe asiatique (1957-1958) et la grippe de Hong Kong (1968-1969). On peut également citer la pandémie de SIDA ou la très récente pandémie de Covid-19.

L'invasion épidémiologique peut également être due à l'évolution d'un pathogène au sein de son hôte, et donc l'invasion peut se faire au sein d'une population où circule déjà l'agent pathogène. Dans ce cas, l'invasion peut être définie par une augmentation de fréquence du phénotype mutant avec le temps, là où l'évolution représentée est représentée par l'apparition de nouveaux mutants. Ce type d'invasion épidémiologique commençant dans une population où le pathogène est déjà établi est largement représentée dans la littérature. Un déterminant de l'évolution intra-hôte des pathogènes est le système immunitaire de l'hôte. Ce système immunitaire crée des pressions de sélection fortes contre le pathogène et peut ainsi diriger la sélection de nouveaux variants avantageés, qui finiront par se transmettre de leur hôte actuel à un nouveau (Johnson *et al.*, 2023 ; Didelot *et al.*, 2016 ; Lythgoe *et al.*, 2021). Un autre facteur favorisant l'émergence de nouvelles souches de pathogènes est l'utilisation d'antiviraux ou d'antibiotiques. Ceux-ci créent

une pression favorisant la sélection et la transmission de souches capables de résister à ces médicaments (Michael *et al.*, 2014; Plowe, 2009). Parmi les exemples les plus connus, on peut citer l'émergence de souches de tuberculose multirésistantes depuis les années 1990 (LoBue, 2009) et l'émergence de mutants résistants du virus de l'hépatite B (dos Santos *et al.*, 2017)

Enfin, un autre moyen de pression menant à la sélection de nouvelles souches est la vaccination (voir Section 1.4.1). L'utilisation massive des vaccins contre l'hépatite B a mené à la sélection de souches capables d'échapper à l'action de ces vaccins (dos Santos *et al.*, 2017; Wang *et al.*, 2017). Un autre exemple se trouve du côté de l'industrie de la volaille, où l'utilisation massive de la vaccination contre le virus de la maladie de Marek (MDV) a conduit à l'émergence de souches dites « ultra-virulentes » capables de surcharger l'immunité de l'hôte stimulée par la vaccination (Gimeno, 2008, voir également Section 1.4.1). Plus récemment, le même phénomène a été intensément décrit durant la pandémie de Covid-19 avec l'apparition de nouveaux variants capables d'échapper au moins partiellement à l'immunité induite par la vaccination (Wang *et al.*, 2021; Chakraborty *et al.*, 2022; Ciotti *et al.*, 2022). D'autres exemples du lien entre l'utilisation des vaccins et l'évolution des pathogènes sont donnés par Gandon et Day (2008).

3.1.3 – Étudier l'invasion d'une souche mutante

Comprendre l'adaptation de souches microbiennes à leur population hôte nécessite une vision écologique du processus. Cette vision écologique permet de définir la *fitness* épidémiologique du parasite. La vitesse d'invasion d'une souche ne dépend pas uniquement de l'avantage sélectif de la souche mutante sur la ou les souches résidentes. De nombreux facteurs environnementaux et épidémiologiques, et en particulier les conditions locales de transmission, peuvent varier à travers le temps et l'espace et ainsi moduler la probabilité pour une nouvelle souche d'émerger et de remplacer les souches résidentes. Par exemple, la propagation d'une espèce – ou souche – invasive peut être accélérée ou ralentie par l'absence ou la présence de compétiteurs dans l'environnement (Hall, 2011). Les conditions climatiques peuvent influencer les taux de transmission de pathogènes tels que le virus de la grippe saisonnière (Roussel *et al.*, 2016), le champignon responsable de la chytridiomycose, *Batrachochytrium dendrobatidis* (Kriger et Hero, 2007; Berger *et al.*, 2004; Yang *et al.*, 2014) ou encore le virus de la rougeole (Jia *et al.*, 2023; Peng *et al.*, 2020; Yang

et al., 2014), et peuvent ainsi influencer la vitesse de transmission de souches émergentes.

Tel qu’annoncé plus tôt, le but de ce chapitre est de présenter une nouvelle méthode pour explorer le lien entre la vitesse d’invasion d’une nouvelle souche, c’est-à-dire la manière dont fluctue la fréquence relative de cette souche mutante, et des variables environnementales et épidémiologiques. Cette méthode a été développée dans le but précis d’analyser l’émergence récente de souches mutantes de *B. pertussis*, mais elle peut être appliquée à d’autres systèmes biologiques.

Comme je l’ai exposé dans la section 1.4.3, de nombreuses mutations de *B. pertussis* ont été reliées à l’utilisation massive de la vaccination. Un caractère particulièrement intéressant à étudier est l’évolution de la production de la toxine pertussique (TP), puisqu’il s’agit de l’un des principaux facteurs de virulence de la coqueluche – au moins parmi les antigènes inclus dans les vaccins (Scanlon *et al.*, 2019a,b). Néanmoins, en s’intéressant à la TP on se retrouve avec le même problème qu’avec *C. diphtheriae* et sa TD : il existe un fort biais de détection des souches ne produisant pas un facteur de virulence puisque ces souches provoquent des infections asymptomatiques ou atypiques. Il convient donc de se concentrer sur un autre antigène vaccinal. Un de ces antigènes est la pertactine (PRN). L’émergence de souches ne produisant pas la PRN (notées PRN⁻, en opposition aux souches productrices de la PRN, notées PRN⁺) a été largement documentée et reliée à l’utilisation massive des vaccins (acellulaires), comme présenté dans la Section 1.4.3. Une étude à grande échelle, basée sur un jeu de données de 3 344 isolats collectés dans plusieurs pays utilisant différentes stratégies vaccinales, a conclu à un avantage sélectif des souches PRN⁻ lorsqu’un vaccin contenant la PRN est en circulation (Lefrancq *et al.*, 2022). Dans ce chapitre, les analyses menées ont été faites en utilisant le jeu de données publié par Lefrancq *et al.* (2022) ainsi que de multiples jeux de données de variables environnementales et démographiques.

Dans ces analyses, j’ai commencé par développer et appliquer un modèle linéaire généralisé (GLMM), puis j’ai testé la robustesse de cette approche. Les GLMM sont très classiques en écologie, les effets aléatoires étant souvent utilisés pour prendre en considération des variables spatiales ou temporelles (e.g., Essl *et al.*, 2015; Marufu *et al.*, 2020). L’utilité des GLMM dans le champ de l’écologie est exposée par Thiele et Markussen (2012). De nombreux travaux ont été publiés par le passé sur la robustesse des GLMM, notamment sur la présence valeurs aberrantes

(Yau et Kuk, 2002; Tempelman, 1998; Sinha, 2004) ou sur l'existence d'une hétéroscédasticité dans la variable réponse (Meijering et Gianola, 1985; Foulley et Gianola, 1996). Dans ce travail, la composante auto-régressive du modèle conduit à des effets à la fois non indépendants et hétéroscédastiques. Dans l'analyse de robustesse, je me suis concentré sur ces effets et j'ai pu mettre en évidence que la composante autorégressive conduit à une sous-estimation du risque d'erreur de type I lorsque l'on utilise des approches classiques de maximisation de la vraisemblance pour estimer les valeurs des paramètres du modèle. Cela a conduit à l'utilisation d'une approche basée sur la méthode Monte-Carlo par chaînes de Markov (MCMC). Ce type d'approche permet de modéliser directement les composantes autorégressives et mixtes du modèle, et ainsi de ne travailler qu'avec des effets aléatoires indépendamment et identiquement distribués. Enfin, pour conclure ce chapitre, j'ai effectué une analyse de la puissance de cette méthode bayésienne lorsque celle-ci est appliquée à un jeu de données de la taille et de la structure de celui publié par Lefrancq *et al.* (2022).

3.2 – Modèle mathématique

La base utilisée pour construire le modèle statistique employé dans la suite est le classique modèle SIR, pour *Susceptible, Infected, Recovered*. Considérant que deux souches d'un même pathogène circulent, le compartiment I est subdivisé en deux, I_1 et I_2 , représentant le nombre d'individus infectés par la souche PRN⁻ et la souche PRN⁺, respectivement. Ce modèle peut être décrit par le système d'équations différentielles ordinaires

$$\frac{dS(t)}{dt} = -\beta_1 S(t) I_1(t) - \beta_2 S(t) I_2(t), \quad (3.1a)$$

$$\frac{dI_1(t)}{dt} = \beta_1 S(t) I_1(t) - \gamma I_1(t), \quad (3.1b)$$

$$\frac{dI_2(t)}{dt} = \beta_2 S(t) I_2(t) - \gamma I_2(t), \quad (3.1c)$$

$$\frac{dR(t)}{dt} = \gamma (I_1(t) + I_2(t)). \quad (3.1d)$$

La construction du modèle statistique s'appuie sur les équations (3.1b) et (3.1c). En partant de celles-ci, nous pouvons obtenir le nombre d'individus infectés par la souche s au temps $t + 1$

sachant le nombre d'infectés au temps t ,

$$I_s(t+1) = I_s(t) e^{\int_t^{t+1} (\beta_s S(u) - \gamma) du}. \quad (3.2)$$

L'explication du passage de (3.1b) ou (3.1c) à (3.2) est donnée dans l'Annexe C.1. La proportion d'individus infectés qui le sont par la souche 1 au temps t est donnée par

$$p_1(t) = \frac{I_1(t)}{I_1(t) + I_2(t)}. \quad (3.3)$$

La formule (3.3) nous permet d'obtenir, au temps $t+1$,

$$p_1(t+1) = \frac{I_1(t+1)}{I_1(t+1) + I_2(t+1)}. \quad (3.4)$$

Maintenant, en apportant l'équation (3.2) dans l'équation (3.4) et en posant $\text{logit}(p_1(t)) = \ln\left(\frac{p_1(t)}{1-p_1(t)}\right)$, on obtient

$$\text{logit}(p_1(t+1)) = \text{logit}(p_1(t)) + \int_t^{t+1} (\beta_1 - \beta_2) S(u) du. \quad (3.5)$$

Le détail du passage de (3.4) à (3.5) est donnée dans l'Annexe C.2. La formule (3.5) est une suite arithmétique, qui peut être réécrite sous la forme

$$\text{logit}(p_1(t)) = \text{logit}(p_1(0)) + \sum_{k=0}^{t-1} \Delta\Lambda(k), \quad (3.6)$$

où $\Delta\Lambda(k) = \int_t^{t+1} (\beta_1 - \beta_2) S(u) du$ et vaut pour tout $t > 0$.

L'équation (3.6) montre que (le logit de) la proportion d'individus infectés par la souche 1 au temps t dépend de l'état initial ainsi que de la somme cumulée des valeurs prises par la variable $\Delta\Lambda(k)$ avant d'atteindre cet instant. En remontant dans le développement présenté dans cette section, on remarque que $\Delta\Lambda(k)$ est une mesure de la différence de transmissibilité entre la souche 1 et la souche 2 au temps k . D'un point de vue biologique, deux éléments peuvent expliquer un différentiel de transmissibilité entre les deux souches : une différence globale de *fitness* (coefficients de sélection différents), et un ensemble de facteurs externes, que je nommerai

dans la suite des cofacteurs. On postule de plus l'hypothèse que ces deux éléments (avantage intrinsèque aux souches mutantes capté par le temps et cofacteurs) ne suffisent pas à expliquer la variabilité dans $\Delta\Lambda(k)$, et ainsi on ajoute une variable supplémentaire $\varepsilon(k)$ qui peut être interprétée comme une forme d'erreur résiduelle. On pose

$$\Delta\Lambda(k) = a_0 + \sum_i a_i x_i(k) + \varepsilon(k), \quad (3.7)$$

où a_i est le coefficient associé avec le cofacteur i , a_0 est l'ordonnée à l'origine et représente également le coefficient de sélection de la souche 1 si les valeurs des cofacteurs sont centrées (Trucchi *et al.*, 2021), $x_i(k)$ est la valeur prise par le cofacteur i l'année k et $\varepsilon(k)$ est le montant de l'erreur résiduelle au temps k ($\varepsilon(k) \sim \mathcal{N}(0, \sigma)$).

On obtient à partir de la formule (3.7)

$$\sum_{k=0}^{t-1} \Delta\Lambda(k) = a_0 t + \sum_i a_i \sum_{k=0}^{t-1} x_i(k) + \sum_{k=0}^{t-1} \varepsilon(k). \quad (3.8)$$

Le modèle donné par la formule (3.6) peut donc être réécrit

$$\text{logit}(p_1(t)) = \text{logit}(p_1(0)) + a_0 t + \sum_i a_i \sum_{k=0}^{t-1} x_i(k) + \sum_{k=0}^{t-1} \varepsilon(k). \quad (3.9)$$

Ce modèle est la base ayant servi aux analyses que j'ai effectuées et qui sont présentées dans la suite. L'étape suivante de celles-ci est la collecte de données pour obtenir des valeurs pour $p_1(t)$ (données de prévalence relative d'une souche) et pour $x_i(k)$ (cofacteur(s) associé(s) à la prévalence relative de cette souche).

3.3 – Collection des données

3.3.1 – Données de prévalence de souches

Acquisition des données

Pour pouvoir utiliser le modèle, il a fallu trouver dans la littérature des données disponibles sur la prévalence de différentes souches de *B. pertussis*. De très nombreuses publications se sont

intéressées à la diversité des souches dans plusieurs pays et sur un plus ou moins long terme, et les auteurs ont parfois mis à disposition les données qu'ils ont utilisées dans leurs articles. Ainsi, des données provenant des États-Unis (Schmidtke *et al.*, 2012), de la Chine (Zhang *et al.*, 2010 ; Xu *et al.*, 2015), du Japon (Hiramatsu *et al.*, 2017 ; Zomer *et al.*, 2018), de la Nouvelle-Zélande (Ring *et al.*, 2022), de la Serbie (Dakic *et al.*, 2010), de la France (Bouchez *et al.*, 2021) ou encore de la Suède (Advani *et al.*, 2011) sont accessibles. Cependant, ces données ont souvent deux problèmes majeurs : un faible nombre d'isolats sont disponibles ou les données ont été collectées sur une courte période. Il y a également des cas de travaux pour lesquels un grand nombre d'isolats ont été étudiés sans que les données aient été rendues publiques au moment de la publication, comme au Brésil (Leite *et al.*, 2021). Enfin, les différentes études ne se focalisant pas exactement sur les mêmes caractéristiques des souches de *B. pertussis*, il a été difficile de trouver un jeu de données correspondant à dont que j'avais besoin pour cette analyse.

En avril 2022, un article a été publié dans *Science Translational Medicine* qui semblait résoudre ces trois problèmes (Lefrancq *et al.*, 2022). Dans cet article issu d'une collaboration internationale, les auteurs et autrices ont assemblé un jeu de données à partir de souches collectées dans de 24 pays¹ entre 1935 et 2019 (Figure 3.1). Ces données sont parfois très complètes, avec des informations détaillées sur un nombre important des 3 344 isolats inclus. Cependant, le jeu de données reste biaisé à certains égards. Il est par exemple notable que très peu de pays en développement sont représentés, notamment en Afrique (Figure 3.1A), alors que ce continent concentre une grande partie des cas de coqueluche (Yeung *et al.*, 2017). Il est bon cependant de préciser que les auteurs et autrices du travail original dont proviennent ces données avaient conscience de ce biais (Lefrancq *et al.*, 2022). Il existe également, au sein des pays représentés dans les données, un fort biais à la fois spatial et temporel. Les souches sont réparties de manière très hétérogène entre pays, les États-Unis et la France réunissant 74 % des données (1 457 et 1 011 isolats, respectivement). Au niveau temporel, bien que le travail original s'enorgueillisse de présenter des données collectées entre 1935 et 2019, dans les faits, 94 % des souches incluses dans les données ont été collectées après 1999 (Figure 3.1B).

1. Il y a bien 24 pays dans les données et pas 23 comme présenté dans l'article. Les auteurs et autrices ont visiblement compté une souche haïtienne comme étant japonaise en faisant leurs analyses.

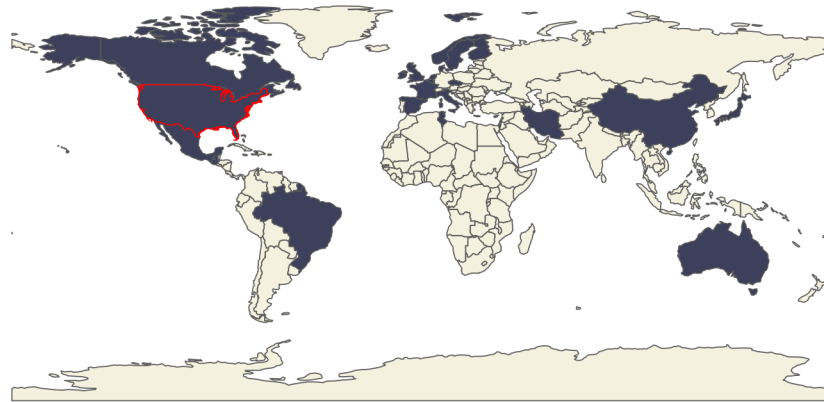
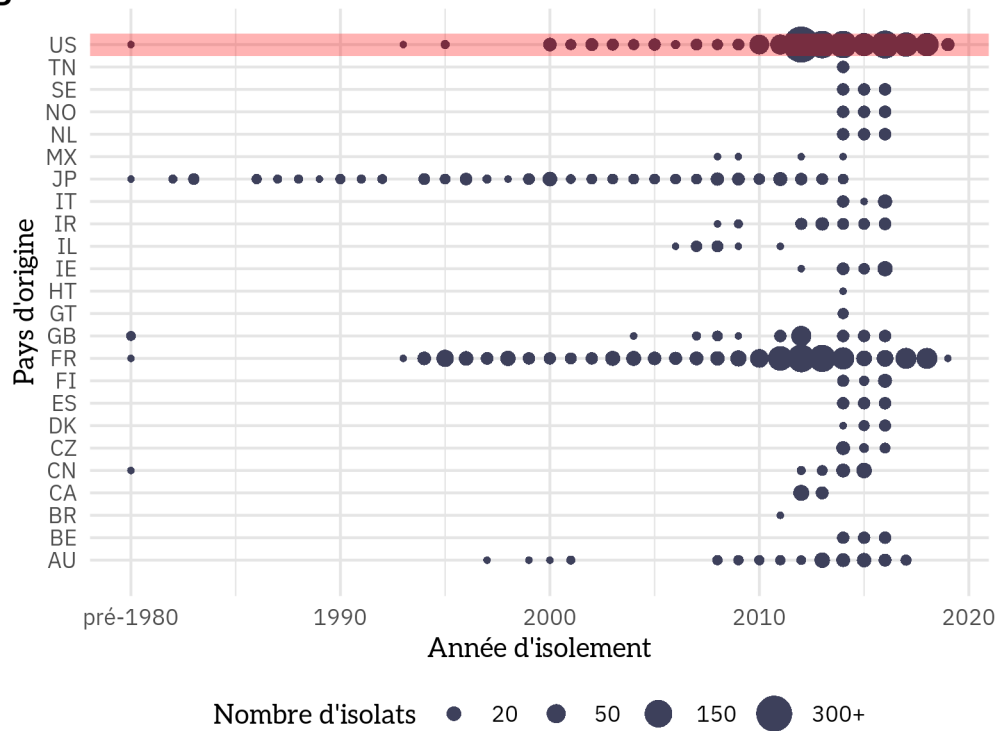
A**B**

FIGURE 3.1 – **Distribution des isolats de *B. pertussis* dans le jeu de données de Lefrancq et al. 2022.** (A) Pays inclus dans le jeu de données (en foncé). (B) Nombre d'isolats en fonction de l'année d'isolement et du pays d'origine des isolats. AU : Australie, BE : Belgique, CA : Canada, CN : Chine, CZ : République tchèque, DK : Danemark, ES : Espagne, FI : Finlande, FR : France, GT : Guatemala, HT : Haïti, IE : Irlande, IL : Italie, IR : Israël, JP : Japon, MX : Mexique, NL : Pays-Bas, NO : Norvège, SE : Suède, TN : Tunisie, GB : Royaume-Uni, US : États-Unis. Les États-Unis (contigus) sont mis en évidence sur chaque figure.

Choix des données à étudier

Comme précisé précédemment, les États-Unis (mis en évidence en rouge sur la Figure 3.1) sont le pays le plus représenté dans les données, fournissant près de 44 % des isolats. Il a donc été décidé de concentrer les analyses sur ce fragment du jeu de données, donc sur 1 457 isolats. La France et le Japon auraient également pu être des candidats intéressants pour cette étude grâce à la grande quantité de données provenant de ces pays. Cependant, pour la France, les données montrent que le phénotype PRN⁻ est encore loin d'avoir envahi la population : il atteint environ 50 % de fréquence relative² ; et pour le Japon, la fréquence relative des souches PRN⁻ a grandement fluctué dans les années récentes en lien avec des changements de vaccins sur le marché, compliquant les analyses, et la barrière linguistique complique l'acquisition des cofacteurs.

Une particularité intéressante de ce jeu de données – et mis à profit dans l'article original – est que pour de nombreux isolats des informations sont disponibles sur la distribution des isolats à une échelle infra-pays. Pour les isolats étasuniens, cette distribution est précisée à l'échelle de l'état quand elle est disponible. Pour pouvoir mener des analyses plus précises (les États-Unis étant un pays d'une très grande superficie avec une grande variété de climats, degrés d'urbanisation, etc.) il a été décidé de ne garder des données que celles pour lesquelles l'état d'origine était précisé, enlevant ainsi trois isolats du jeu de données.

Il a ensuite fallu décider sur quel type de souche focaliser les analyses. De nombreux caractères de *B. pertussis* sont « surveillés » depuis quelques années dans différents pays, et notamment l'émergence de mutations atteignant des antigènes présents dans les vaccins en circulation, comme cela a été présenté dans la section 1.4.3. Pour rappel, ces mutations peuvent être classées en trois catégories : l'émergence de nouveaux allèles, un changement de sérotype, et l'arrêt de production d'un antigène. À noter que ces mutations ne sont pas mutuellement exclusives, et qu'au contraire des changements de phénotype ou de sérotype ont été reliés à l'émergence de nouveaux génotypes (Lefrancq *et al.*, 2022).

Le changement phénotypique est certainement le plus pertinent à étudier, notamment parce que la production/non-production d'un antigène est un phénomène binaire, donc le changement de fréquence de ce phénotype peut être résumé à une dualité succès/échec dans l'étude des

2. De plus, la récente épidémie de coqueluche semble indiquer un retour en masse du phénotype PRN⁺ (Rodrigues *et al.*, 2024).

données. Dans la section 3.1.3, j'ai introduit l'idée que la production ou la non-production de la TP serait particulièrement intéressante à étudier mais que les données sont rares et parcellaires. J'ai donc proposé d'étudier à la place l'invasion de souches ne produisant plus la PRN. Les données étasuniennes de Lefrancq *et al.* (2022) permettent de visualiser de manière explicite le phénomène de remplacement des souches productrices de PRN+ par des souches PRN-. En ne conservant que les isolats pour lesquels la capacité à produire la PRN est renseignée, le jeu de données a été réduit à 1367 isolats.

Le dernier filtre appliqué aux données a été de ne conserver que les souches isolées entre 2007 et 2017, excluant ainsi celles collectées en 2018 et 2019. La raison principale est que ces deux dernières années correspondent à un retour partiel des souches PRN+ (Figure 3.2). Ce retour est possiblement dû à des fluctuations d'échantillonnage mais pourrait également être dû à des phénomènes biologiques d'intérêt. Par exemple, il est possible que la réémergence des souches PRN+ soit dû à un phénomène de sélection densité-dépendant avec un avantage aux souches rares. En effet, il est envisageable que lorsque les souches mutantes circulent au sein de la population, la population soit fortement immunisée contre cette souche mais ait perdu son immunité contre la souche sauvage à cause de sa faible circulation, favorisant ainsi une augmentation de sa fréquence.

Dans tous les cas, cette diminution de la proportion de souches PRN- pour ces deux années conduit à augmenter l'erreur résiduelle dans les données et complique les analyses. J'ai essayé de prendre en considération cette particularité dans l'analyse bayésienne qui sera présentée plus tard en ajoutant un effet de densité-dépendance avec un temps de retard. Ce temps de retard aurait été nécessaire pour expliquer pourquoi la proportion de souches PRN- augmente avant de diminuer à nouveau plutôt que simplement atteindre un plateau. Finalement, ces analyses n'ont pas été concluantes et ont été retirées du présent manuscrit, et il a été décidé de continuer les analyses en retirant les années 2018 et 2019 pour ainsi se concentrer sur la phase d'invasion du phénotype PRN-.

Finalement, le jeu de données final comportait 1195 isolats répartis dans 26 états (Figure 3.3).

Les biais spatiaux et temporels observés à l'échelle mondiale se retrouvent dans une certaine mesure dans les données à l'échelle nationale. Nous pouvons remarquer en premier lieu que de nombreux états, notamment dans la région des Rocheuses, sont absents des données (Fi-

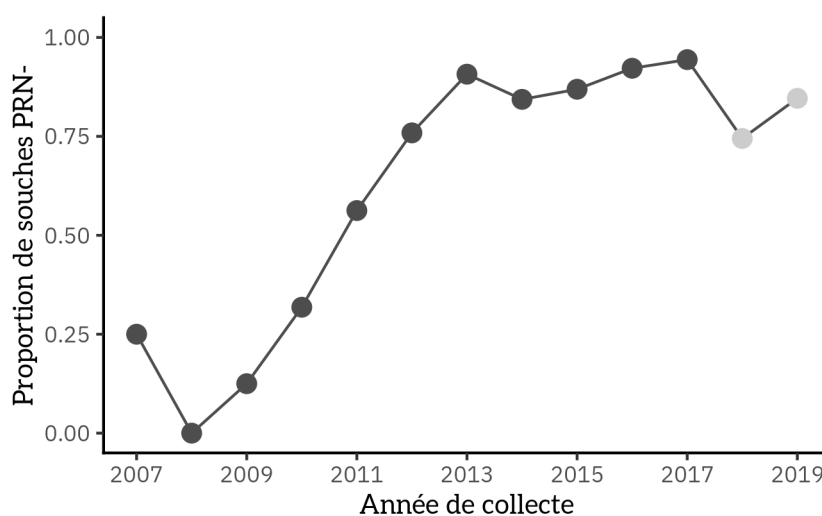


FIGURE 3.2 – **Évolution temporelle dans la proportion de souches PRN-** dans le jeu de données conservé pour les analyses, incluant les années 2018 et 2019 (en gris clair).

gure 3.3B). Ensuite, dans les états représentés, les isolats sont très inégalement répartis. Par exemple, l’Oregon, l’état de New York, l’état de Washington et le Minnesota concentrent une grande partie des données alors que d’autres contribuent très peu au jeu de données comme le Texas ou la Caroline du Sud (Figure 3.3C). Enfin, les données ne sont pas équitablement réparties dans le temps : les données pré-2010 sont rares, et l’année 2012 à elle seule concentre 40 % des données (Figure 3.3C). Cette forte hétérogénéité dans les données d’isolats va nécessairement devoir être prise en compte dans les calculs des cofacteurs employés dans le modèle, tel qu’expliqué dans la Section 3.4.

3.3.2 – Choix et acquisition des cofacteurs

Une fois les données de prévalence des souches PRN- obtenues, huit cofacteurs ont été retenus pour étudier leur impact sur la vitesse d’émergence de ces souches. Certains de ces cofacteurs ont un rôle documenté sur la dissémination d’agents pathogènes – et notamment de pathogènes à transmission aéroportée – tandis que d’autres sont plus exploratoires. Ces huit cofacteurs sont la densité de population, la température annuelle moyenne, l’âge médian de la population, la taille médiane des foyers, le revenu médian, la couverture vaccinale, la densité d’hôpitaux et la densité de médecins.

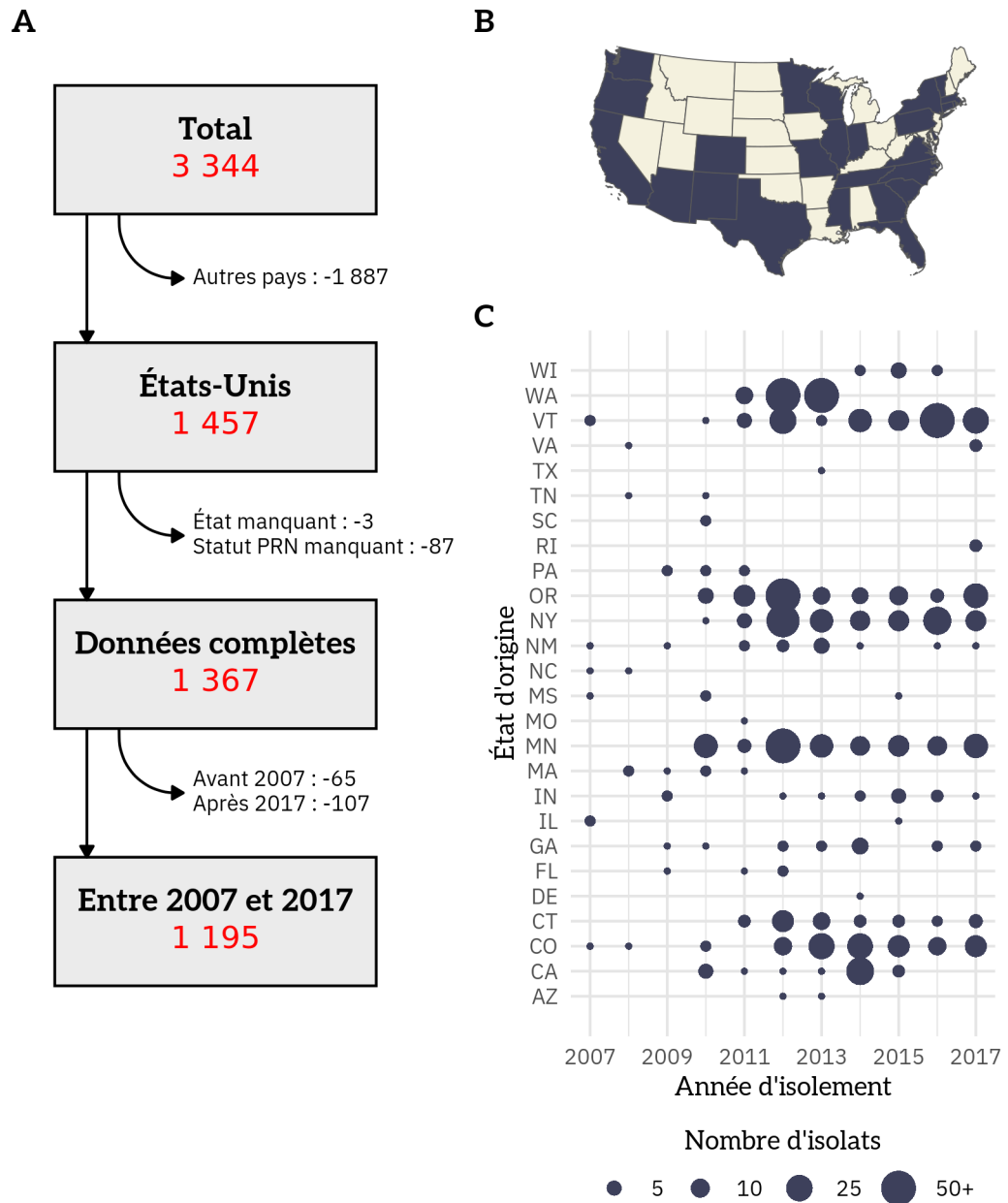


FIGURE 3.3 – Distribution des isolats de *B. pertussis* conservés dans les analyses. (A) Filtres appliqués sur les données. (B) États représentés dans les données (en foncé). (C) Nombre d'isolats en fonction de l'année d'isolement et de l'état d'origine des isolats. AZ : Arizona, CA : Californie, CO : Colorado, CT : Connecticut, DE : Delaware, FL : Floride, GA : Géorgie, IL : Illinois, IN : Indiana, MA : Massachusetts, MN : Minnesota, MO : Missouri, MS : Mississippi, NC : Caroline du Nord, NM : Nouveau-Mexique, NY : New York, OR : Oregon, PA : Pennsylvanie, RI : Rhode Island, SC : Caroline du Sud, TN : Tennessee, TX : Texas, VA : Virginie, VT : Vermont, WA : Washington, WI : Wisconsin.

À noter que sur la Figure 3.5 ainsi que dans les discussions liées, c'est le *z-score* qui est affiché et discuté. Le *z-score* représente les valeurs standardisées des variables, standardisation obtenue en soustrayant la moyenne de chaque variable et en divisant par l'écart-type de celle-ci. L'utilisation du *z-score* permet de rendre comparables des valeurs avec des unités et des étendues sinon très différentes. De plus, il permet de visualiser si la variabilité dans un cofacteur provient plutôt de sa dimension spatiale (état) ou temporelle (année). En revanche, il ne permet pas de se rendre compte de l'étendue des valeurs prises par un cofacteur.

Densité de population

La densité de population est communément reliée à la vitesse de propagation des maladies infectieuses, comme cela a pu être montré pour la rougeole, (Qin *et al.*, 2019 ; Yoshikura, 2012) ou pour le Covid-19 (Wong et Li, 2020). Le principe biologique sous-jacent est qu'une forte concentration d'individus facilite les contacts entre personnes infectées et personnes susceptibles et assure la transmission du pathogène dans la population. Ainsi, l'hypothèse est que les états avec une forte densité de population devraient plus rapidement être envahis par une souche mutante avec une plus grande valeur sélective.

La densité de population annuelle dans chaque état est obtenue à partir des recensements décennaux et de la superficie des états. Les recensements sont disponibles en ligne sur le site Internet de l'United States Census Bureau (USCB)³. Les superficies de chaque état sont également disponibles sur ce même site et correspondent aux superficies mesurées en 2010⁴.

Les rapports décennaux ne donnent, par définition, que les densités de population tous les dix ans. Pour obtenir les densités de population chaque année incluse dans le jeu de données, il a donc fallu estimer les densités de population entre les recensements. Pour ce faire, la méthode la plus simple est de considérer que la population croît de manière linéaire entre deux sondages. Les densités de population ont donc pu être estimées pour les années 2007 à 2017 à partir des recensements des années 2000, 2010 et 2020.

Il est intéressant d'observer que la densité de population varie plutôt entre états qu'entre

3. Disponible sur <https://www2.census.gov/programs-surveys/decennial/2020/data/apportionment/population-change-data-table.pdf> (dernier accès le 19/08/2024).

4. Disponible sur <https://www.census.gov/geographies/reference-files/2010/geo/state-area.html> (dernier accès le 19/08/2024).

années (Figure 3.5). Cela s’explique facilement par la courte durée de l’analyse : en une dizaine d’années, il est attendu que la densité de population fluctue peu (hors phénomène exceptionnel comme un conflit ou une catastrophe naturelle majeure).

Température annuelle moyenne

La température de l’environnement est souvent considérée comme un facteur influençant la capacité d’infection et de transmission des pathogènes. *Batrachochytrium dendrobatidis* est un pathogène fongique infectant la peau des amphibiens et responsable de la disparition de plusieurs populations et même espèces d’anoures. Plusieurs études ont mis en évidence un lien fort entre la température de l’environnement et à la fois la prévalence de la maladie et la force de l’infection (Olson *et al.*, 2021 ; Berger *et al.*, 2004 ; Kriger et Hero, 2007). Quelques travaux sur la prévalence de la rougeole en Chine semblent indiquer que la prévalence de la maladie est influencée de manière non linéaire par la température (Jia *et al.*, 2023 ; Peng *et al.*, 2020 ; Yang *et al.*, 2014), un phénomène relié à des différences dans la capacité du virus à survivre à différentes températures (de Jong, 1965). Le virus de la grippe est l’exemple typique de pathogène auquel la sagesse populaire relie une chute des températures à une augmentation de son incidence. Cela a été vérifié dans plusieurs études en population sur la propagation du virus et notamment sa saisonnalité (*e.g.*, He *et al.* (2013) ; Skog *et al.* (2014) ; Wang *et al.* (2023), voir Sooryanarain et Elankumaran (2015) pour une revue du sujet) mais également lors d’études en laboratoire (Lowen *et al.*, 2007). D’autres études encore ont montré que la propagation de certaines maladies à transmission aéroportée était influencée par le climat et notamment la température de l’environnement (références dans la Section 3.1.3). Il a donc été décidé d’intégrer cette variable dans le modèle.

Les données de température ont été obtenues à partir de la base de données GHCNd⁵. Cette base de données est accessible à travers un portail HTTPS permettant de télécharger les informations collectées par des dizaines de milliers de stations météorologiques réparties sur tout le territoire américain. L’objectif était de collecter et résumer les informations fournies par un maximum de stations réparties dans les 26 états de l’étude pour pouvoir calculer une température

5. Disponible sur <https://www.ncei.noaa.gov/products/land-based-station/global-historical-climatology-network-daily> (dernier accès le 19/08/2024).

moyenne par état et par an la plus précise possible. Pour être prise en compte dans l'étude, une station devait répondre à plusieurs critères :

- Être active durant au moins une partie de la période d'intérêt : de janvier 2007 à décembre 2019 (au moment du choix de ce critère, il n'avait pas encore été décidé de réduire le jeu de données aux années 2007 à 2017).
- Être placée dans un des états d'intérêts.
- Avoir enregistré la température minimale et maximale du jour pendant au moins une partie de sa période d'activité.

Le dernier critère est important puisque c'est à partir des températures maximales et minimales que les températures moyennes ont été calculées. En effet, certaines stations enregistrent directement la température quotidienne moyenne, mais ces stations représentent un très faible échantillon du nombre total de stations. Il y a ainsi des états pour lesquels la température quotidienne moyenne est inaccessible puisqu'aucune station ne l'enregistre. La moyenne entre température minimale et maximale a donc été utilisée comme approximation de la température quotidienne moyenne. Une fois cette moyenne proximale calculée pour chaque journée renseignée par chaque station, la température moyenne par an et par état a été calculée. Finalement, un groupe de 5 730 stations globalement bien réparties dans les états d'intérêt a été conservé (Figure 3.4). Il apparaît que la variabilité dans la température moyenne est donnée à la fois par les états, ainsi que, dans une moindre mesure, par les années (Figure 3.5).

Âge médian de la population

L'âge est un facteur clé en épidémiologie comme en clinique. Les contacts des individus au sein d'une population sont fortement structurés par âge, et cette structuration a une incidence sur la transmission et la prévalence des maladies au sein des différentes classes d'âge. Typiquement, les individus en âge d'être scolarisés vont avoir de très nombreux contacts au sein de leur classe d'âge mais également avec les autres, et concentreront la majorité des infections lors d'une épidémie de type grippe (Mossong *et al.*, 2008).

L'âge médian peut servir de proxy de la structuration en âge de la population, et donc du schéma de contacts au sein de la population. Un âge médian bas implique une population jeune, avec vraisemblablement de nombreux contacts facilitant la propagation des maladies. Un âge

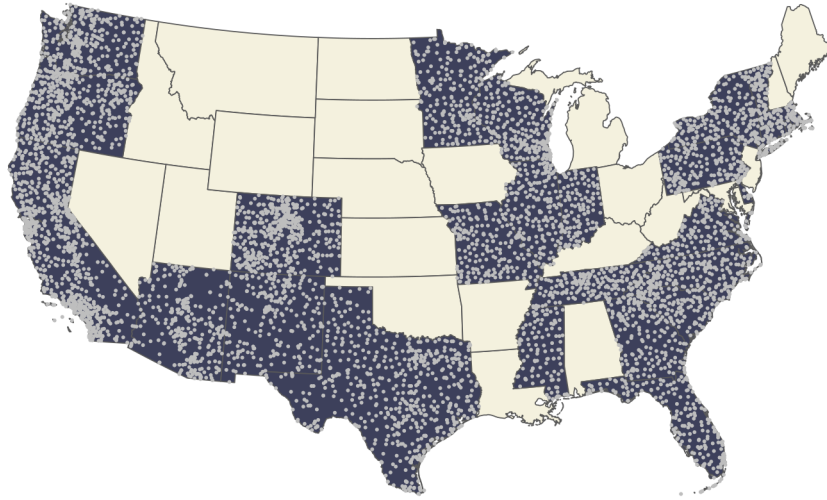


FIGURE 3.4 – **Localisation des stations météorologiques** utilisées dans le calcul des températures moyennes des états.

médian élevé implique à l'inverse une population plus vieille, et on sait que les contacts diminuent lorsque l'âge augmente, ce qui peut ralentir la propagation d'un pathogène (Mossong *et al.*, 2008). Il a été montré que la vaccination a augmenté l'âge auquel les individus sont typiquement infectés par la coqueluche, le décalant vers les classes d'âges ayant de plus forts contacts et expliquant ainsi le schéma de réémergence de la maladie observé dans de nombreux pays (Rohani *et al.*, 2010).

Pour obtenir les données sur l'âge médian au sein des différents états des États-Unis et pour les différentes années considérées, j'ai de nouveau utilisé les données de l'USCB. Ces données sont collectées annuellement dans le cadre de l'American Community Survey (ACS) depuis 2005. Ce programme a pour objectif de collecter des informations démographiques (revenus, âge, classe professionnelle, moyens de transports...) sur la population américaine en considérant à la fois les données à travers le temps (l'enquête est annuelle) et l'espace (à l'échelle de villes, de comtés ou d'états). Environ 3,5 millions d'Américains participent à ce sondage annuel, soit un peu plus d'un pourcent de la population. Cet échantillon est considéré comme étant représentatif de la population américaine.

Les données de l'ACS sont partiellement accessibles via une interface utilisateur facilitant la

visualisation et la récupération de données⁶. Néanmoins, les données antérieures à 2010 n'étant pas accessibles à partir de cette interface, j'ai dû récupérer manuellement les données dont j'avais besoin à partir du portail HTTPS de l'USCB⁷.

L'âge médian étant directement disponible pour chaque état, il n'a pas été nécessaire d'effectuer de transformation pour ce cofacteur. Comme pour la température, la variabilité de ce cofacteur semble plutôt se distribuer dans l'espace, quoique le temps y contribue partiellement (Figure 3.5).

Taille médiane des foyers

Un autre facteur impactant la dynamique des contacts et donc de la transmission d'un pathogène au sein d'une population est le nombre d'individus en cohabitation proche au sein d'un foyer. Ce foyer prend généralement la forme d'un noyau familial (parents-enfants) mais peut également représenter un couple sans enfants, une collocation... Les contacts au sein d'un foyer représentent une opportunité plus ou moins forte de transmission pour un pathogène, et on nomme ces transmissions des « cas secondaires ». Pour la tuberculose, on observe qu'environ 5 % des contacts au sein des foyers conduisent à des cas secondaires, et ce pourcentage est plus élevé chez les enfants que chez les adultes (Fox *et al.*, 2018 ; Guwatudde, 2003). Pour le SARS-CoV-2, la proportion de contacts conduisant à un cas secondaire fluctue grandement entre pays, allant de 4,6 % à Taïwan à 49,56 % en Chine – la plupart des pays affichant une proportion de cas secondaires inférieure à 20 % (Shah *et al.*, 2020).

L'augmentation de la taille des foyers conduit également à des phénomènes de *crowding*, où le nombre d'individus dans le foyer dépasse le nombre de pièces disponibles. Quoique cette conclusion soit sujette à caution⁸, il a été montré que ce *crowding* peut conduire à augmenter le risque de transmission de la tuberculose (Myers *et al.*, 2006).

Pour la coqueluche, de nombreuses études ont été menées pour comprendre comment la maladie se transmettait au sein des foyers. Habituellement, quoique la maladie soit diagnostiquée chez un jeune enfant (von König *et al.*, 1995), elle est introduite dans le foyer par un adulte (Kowalzik

6. Disponible sur <https://data.census.gov/> (dernier accès le 19/08/2024).

7. Disponible sur https://www2.census.gov/programs-surveys/acs/summary_file/ (dernier accès le 19/08/2024).

8. Dans l'étude citée, le *crowding* augmente le risque de tuberculose s'il est considéré indépendamment de toute autre variable mais pas si d'autres variables sont considérées simultanément.

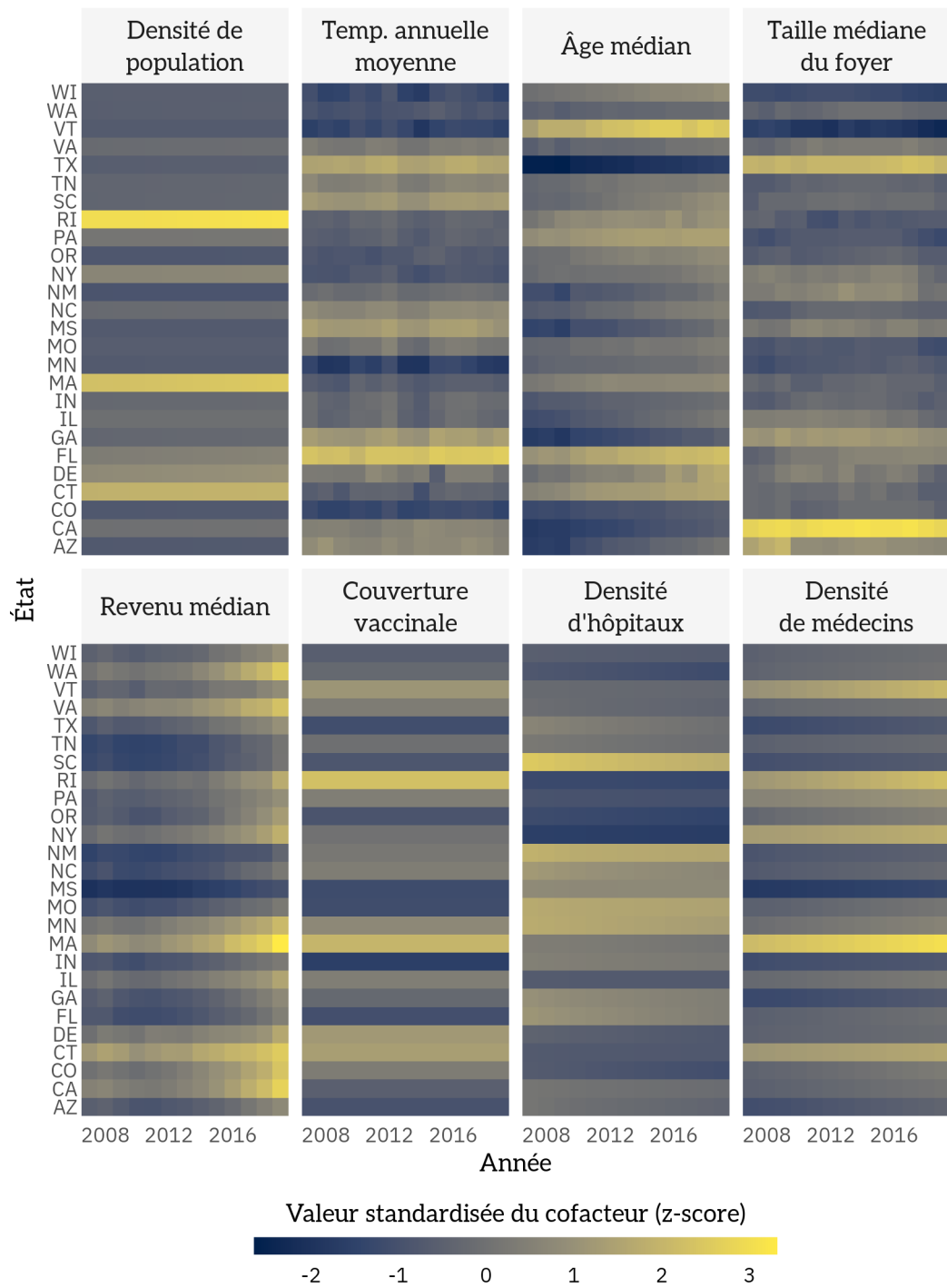


FIGURE 3.5 – **Fluctuations des z-scores des différents cofacteurs** entre états et entre années. Le z-score permet la comparaison entre des variables ayant des unités différentes. Les noms des états sont donnés dans la légende de la Figure 3.3. Temp. : Température.

et al., 2007; Deen *et al.*, 1995). La proportion de cas secondaires engendrés est généralement forte, au moins 30 % (von König *et al.*, 1995; Kowalzik *et al.*, 2007). Elle pourrait cependant être encore plus élevée, la maladie ou l'infection par *B. pertussis* étant parfois difficiles à détecter, notamment à cause d'une proportion élevée d'infections asymptomatiques (Deen *et al.*, 1995; Craig *et al.*, 2020). Dans tous les cas, il apparaît clairement que les contacts au sein d'un foyer participent fortement à la propagation de la coqueluche, et j'ai fait l'hypothèse que plus le foyer est grand, plus les contacts sont nombreux, et donc plus la souche émergente a d'opportunités de transmission.

Tout comme pour l'âge médian de la population, la taille des foyers a été obtenue à partir du portail HTTPS de l'USCB. La fluctuation interannuelle de ce cofacteur semble plutôt faible, la variabilité du cofacteur se retrouvant donc plutôt entre états. (Figure 3.5).

Revenu médian

Le revenu médian fait partie des cofacteurs très exploratoires introduits dans les analyses. En Californie, vivre dans une zone géographique avec un revenu médian bas est associé à un risque accru de tuberculose (Myers *et al.*, 2006). La tuberculose, comme la coqueluche, est une maladie respiratoire transmise directement entre individus infectés. Un revenu élevé est associé à de meilleures conditions d'hygiène, un meilleur accès aux soins, et donc un meilleur diagnostic de la maladie, pouvant freiner sa propagation. Inversement, un revenu faible entraîne des conditions favorables à la propagation d'une maladie ou d'un nouveau variant d'un pathogène.

Comme pour l'âge médian de la population ou la taille médiane des foyers, le revenu médian a été obtenu à partir du portail HTTPS de l'USCB. Le revenu pour chaque année et dans chaque état a été ajusté sur la valeur du dollar en 2021 pour prendre en considération l'inflation. Pour cela, j'ai utilisé l'indice des prix à la consommation (CPI, pour *Consumer Price Index*) fourni par l'United States Bureau of Labor Statistics⁹. Le revenu médian ajusté pour l'année X s'obtient en multipliant le revenu médian brut par le ratio du CPI de 2021 et celui de l'année X .

On remarque qu'une fois le revenu médian ajusté pour l'inflation il subsiste une hétérogénéité à la fois entre états et entre années dans le revenu médian. Globalement, pour tous les états, le revenu moyen semble diminuer autour de 2009-2011 avant d'augmenter continuellement

9. Accessible sur <https://www.bls.gov/cpi/data.htm> (dernier accès le 19/08/2024)

(Figure 3.5).

Couverture vaccinale

En 1992, l'Advisory Committee on Immunization Practices (ACIP), équivalent américain de la Haute Autorité de Santé en France, recommande l'utilisation d'un vaccin acellulaire pour les rappels vaccinaux destinés aux enfants, et en 1997 recommande de remplacer le vaccin à germe entier par un vaccin acellulaire pour la primo-vaccination (Klein, 2014). Cela signifie que, sur l'intervalle de temps compris dans mes analyses, la vaccination acellulaire a remplacé la vaccination à germe entier ; et il a été montré que c'est le changement de type de vaccin qui est le principal facteur de l'émergence des souches PRN⁻ (Ma *et al.*, 2021 ; Lefrancq *et al.*, 2022). Or, le changement de vaccin ne peut avoir une influence sur l'émergence des souches PRN⁻ que si la couverture vaccinale est suffisamment élevée. Il est donc possible – probable – que la couverture vaccinale joue un rôle de première importance dans la dynamique d'invasion des souches PRN⁻.

L'OMS propose des données en accès public sur la couverture vaccinale d'un grand nombre de pays et pour plusieurs vaccins, mis à jour annuellement ¹⁰. Malheureusement, ces données ne sont pas disponibles à une échelle géographique plus fine que le pays. Il a donc fallu utiliser une autre source de données pour obtenir la couverture vaccinale. À travers son portail ChildVaxView ¹¹, le CDC propose une variété de statistiques sur la vaccination infantile dans les différents états des États-Unis. Il a été décidé de s'intéresser en particulier à la couverture vaccinale des enfants ayant reçu les trois doses de vaccin recommandées par l'OMS. De plus, le portail permet de sélectionner une cohorte. Puisque le calendrier vaccinal américain recommande que la troisième dose de vaccin soit prise à six mois, j'ai décidé de conserver la cohorte des enfants de 13 mois.

Malheureusement, les données de ChildVaxView ne permettent pas de remonter plus tôt que l'année 2011. De plus, la couverture vaccinale ne montre pas de franche augmentation ou diminution au cours des années, ne permettant pas d'inférer la couverture vaccinale pour les années manquantes, comme j'ai pu le faire pour la densité de population. Il a donc été décidé, pour chaque état, de prendre la moyenne des couvertures vaccinales enregistrées entre 2011 et 2018 et de considérer dans les analyses cette couverture vaccinale moyenne comme étant la

10. Disponible sur <https://immunizationdata.who.int/> (dernier accès le 07/08/2024).

11. Disponible sur <https://www.cdc.gov/vaccines/imz-managers/coverage/childvaxview/> (dernier accès le 19/08/2024).

couverture de l'état pour chaque année. Pour cette raison, on peut voir sur la Figure 3.5 qu'il n'y a pas de variabilité interannuelle dans la couverture vaccinale. En revanche, on observe une certaine variabilité entre états, la couverture vaccinale fluctuant entre 86,7 et 95,8 % environ.

Densité d'hôpitaux

La densité d'hôpitaux a été ajoutée aux cofacteurs comme variable complètement exploratoire. La densité d'hôpitaux joue un rôle dans l'accessibilité aux soins ainsi que, dans une moindre mesure, au diagnostic. De plus, les hôpitaux constituent des pôles de transmission des maladies infectieuses, il est donc envisageable qu'une forte densité hospitalière conduise à la fois à une meilleure détection des souches PRN- et à une propagation plus rapide de celles-ci.

Le nombre d'hôpitaux par état est renseigné sur le site de l'United States Bureau of Labor Statistics pour l'année 2019¹². Malheureusement, je n'ai pu obtenir de donnée pour aucune autre année. Il a donc fallu considérer ce nombre comme constant dans le temps. Cependant, il est vraisemblable que le nombre d'hôpitaux n'ait pas grandement changé au cours des années précédentes. Pour obtenir la densité d'hôpitaux, il a ensuite fallu ajuster le nombre d'hôpitaux à la taille de la population. Finalement, c'est le nombre d'hôpitaux pour 100 000 habitants qui a été gardé dans les analyses.

Comme attendu, la densité d'hôpitaux varie essentiellement entre états, la fluctuation interannuelle étant plus anecdotique à cause du manque de répétitions longitudinales dans le nombre d'hôpitaux présents dans les états (Figure 3.5).

Densité de médecins

Enfin, le dernier cofacteur inclus dans les analyses est la densité de médecins. Comme la densité d'hôpitaux, il s'agit là d'un cofacteur essentiellement exploratoire, avec une incidence soupçonnée sur le diagnostic de la coqueluche et donc l'identification des souches PRN-. En revanche, la densité de médecins ne devrait pas avoir d'incidence sur la propagation réelle des souches PRN-.

Le CDC fournit une estimation du nombre de médecins en activité pour 10 000 habitants

12. Disponible sur <https://www.bls.gov/opub/ted/2020/number-of-hospitals-and-hospital-employment-in-each-state-in-2019.htm> (dernier accès le 19/08/2024).

dans chacun des états des USA ¹³ Cependant, les données ne sont pas actualisées sur une base annuelle, et seulement une poignée d'années sont disponibles (notamment 2000, 2009 et 2019). Il a donc fallu, comme pour l'estimation de la taille de population dans chaque état, estimer la densité de médecins en considérant un changement linéaire de la densité de médecins entre points temporels successifs.

Tout comme pour la densité d'hôpitaux, la variabilité dans la densité des médecins en activité est essentiellement due au facteur spatial, tandis que le facteur temporel joue un rôle mineur (Figure 3.5). Il est intéressant de noter que les états avec une forte densité de médecins ne sont généralement pas les mêmes que ceux avec une forte densité d'hôpitaux.

3.4 – Transformation des cofacteurs

Comme présenté dans la section 3.3.1, les données de prévalence des souches PRN⁻ présentent une très forte hétérogénéité spatio-temporelle, variabilité représentant une contrainte dans les analyses qui sont menées dans la suite de ce chapitre. Ainsi, certains états américains présentent peu voire pas de souches échantillonnées, tandis que d'autres sont surreprésentés. Mettre en place des analyses ouvertement spatialisées serait compliqué à cause du nombre réduit de souches par état et par pas de temps, avec certaines années ou bien états parfois absents du jeu de données. Inversement, lisser simplement les données à l'échelle nationale poserait d'autres problèmes pour les analyses. Par exemple, considérer comme cofacteur la température moyenne sur tout le territoire américain fausserait les analyses simplement par le fait que les souches ne sont pas du tout distribuées de manière homogène dans l'espace. Cela reviendrait donc à considérer des cofacteurs n'ayant pas été observés par les souches comme dirigeant la dynamique de compétition entre souches PRN⁻ et souches PRN⁺.

Il est donc nécessaire de calculer des valeurs modifiées des cofacteurs pour les incorporer dans le modèle en considérant la distribution hétérogène des souches dans l'espace (et le temps). Pour calculer ces cofacteurs modifiés, nous avons proposé la formule

$$\chi_{i,s}(k) = \frac{\sum_j n_{s,j}(k') v_{i,j}(k') + \chi_{i,s}(k-1)}{\sum_j n_{s,j}(k') + 1}, \quad (3.10)$$

13. Disponible sur <https://www.cdc.gov/nchs/data/hus/2020-2021/DocSt.pdf> (dernier accès le 19/08/2024).

où $n_{s,j}(k')$ est le nombre d'isolats de la souche s obtenus dans l'état j l'année k' , $v_{i,j}(k')$ est la valeur brute du cofacteur i dans l'état j l'année k' et $\chi_{i,s}(k-1)$ est la valeur transformée du cofacteur i pour la souche s l'année $k-1$. Ce terme est ajouté au numérateur pour servir de valeur par défaut dans le cas où il n'y aurait pas de souche s dans les données au cours d'une année, ce qui est le cas en 2008 où aucun isolat PRN- n'a été reporté. Ce terme est compensé par le terme $+1$ au dénominateur. Finalement, la valeur $\chi_{i,s}(k)$ obtenue par l'équation (3.10) représente une moyenne pondérée par le nombre de souches de la valeur du cofacteur. Elle permet de mettre en évidence les différences dans les valeurs des cofacteurs observées par les différents types de souches. On pose $\chi_{i,s}(0) = 0$.

Les données sur les souches sont collectées tout au long de l'année mais ne sont rapportées que sur une base annuelle. Pour cette raison, l'année k' dans l'équation 3.10 peut autant être interprétée comme étant l'année précédent l'année de collection d'une souche ou comme étant l'année d'isolation elle-même. Les deux interprétations de k' sont crédibles à la lumière de la discrétisation du processus continu d'émergence des souches PRN-. Soit t l'année de collection d'une souche. La première interprétation (k' est l'année avant la date de collection, donc $k' = t-1$) implique que les souches isolées pendant l'année en cours sont affectées par un phénomène mesuré en moyenne un an auparavant. La seconde interprétation (k' est l'année correspondant à l'année de collection, donc $k' = t$) implique que les souches isolées durant l'année en cours sont affectées par un phénomène qui n'a pas encore eu lieu (Figure 3.6). Les deux interprétations ont leurs avantages et désavantages et j'ai donc gardé les deux dans les analyses, permettant ainsi leur comparaison.

La formule (3.10) permet d'obtenir une valeur modifiée de cofacteur par souche et par cofacteur. À partir de ces valeurs, deux métriques ont été calculées pour être introduites dans le modèle.

La première métrique est la différence entre les valeurs calculées pour les deux souches pour un même cofacteur, $\Delta_i(k)$. Cette métrique représente la différence dans la transmission entre deux souches qui est due au cofacteur. Elle vaut

$$\Delta_i(k) = \chi_{i,1}(k) - \chi_{i,2}(k). \quad (3.11)$$

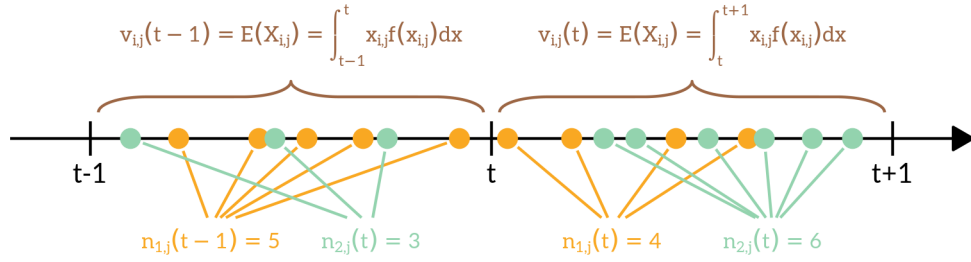


FIGURE 3.6 – **Illustration du calcul de $\chi_{i,s}(k)$.** $X_{i,j}$ est une variable aléatoire représentant la valeur brute et continue d'un cofacteur i dans l'état j (e.g., la température tout au long de l'année). Les souches annotées comme appartenant à l'année t sont collectées en continu du début de l'année (instant t) jusqu'au début de l'année suivante (instant $t+1$) et les cofacteurs sont (en théorie) calculés comme étant la moyenne des valeurs qu'ils prennent au cours de cette période. Si $k' = t$, le calcul de $\chi_{i,s}(k)$ utilise la valeur de $v_{i,j}(t)$. Or, on voit que $v_{i,j}(t)$ est obtenu en partie sur des périodes dépassant la dernière date d'acquisition des souches. Si $k' = t-1$, le calcul de $\chi_{i,s}(k)$ utilise la valeur de $v_{i,j}(t-1)$. Dans ce cas, $v_{i,j}(t-1)$ n'utilise pas de valeurs de x obtenues après la dernière acquisition de souche, mais en contrepartie la valeur du cofacteur peut largement pré-dater la période où les souches ont été acquises l'année t .

La deuxième métrique qui nous intéresse est la moyenne des valeurs d'un cofacteur, $\mu_i(k)$. L'hypothèse derrière l'utilisation de cette métrique est qu'un changement global dans les taux de transmission des deux souches pourrait affecter la vitesse à laquelle les souches PRN- remplacent les souches PRN+. Cette métrique vaut

$$\mu_i(k) = \frac{\chi_{i,1}(k) + \chi_{i,2}(k)}{2}. \quad (3.12)$$

Les valeurs obtenues avec les formules (3.11) et (3.12) sont utilisées dans toutes les analyses qui suivent en tant que cofacteurs. Ainsi, à partir des huit cofacteurs originaux présentés plus tôt, ce sont seize cofacteurs qui ont finalement été utilisés dans les analyses. À cause de leur nombre élevé, les cofacteurs ont été étudiés un à un dans les analyses plutôt que simultanément ou par groupes de cofacteurs. Leurs valeurs ont été centrées et réduites en utilisant la moyenne et l'écart-type pour toutes les années incluses dans les analyses sauf $t = 0$, pour laquelle la valeur de chaque cofacteur est 0 (i.e., les années 2008 à 2017 ont été utilisées, excluant l'année 2007). Le but de cette transformation est en partie de faciliter la convergence des modèles utilisés. De plus, centrer les données permet d'interpréter le coefficient a_0 du modèle comme étant le coefficient de sélection de la souche 1, comme présenté dans la Section 3.2 (a_0 représente le gain en logit de

la proportion de souches PRN– chaque année indépendant des cofacteurs) et les réduire permet de comparer les coefficients associés aux différents cofacteurs.

Les valeurs prises par i dans les formules (3.11) et (3.12) sont p pour la densité de population, u pour la température annuelle moyenne, a pour l'âge médian, f pour la taille médiane des foyers, r pour le revenu médian, c pour la couverture vaccinale, h pour la densité d'hôpitaux et m pour la densité de médecins.

3.5 – Approche fréquentiste

3.5.1 – Modèle linéaire généralisé

Dans une première analyse, je me suis intéressé au lien entre les différents cofacteurs et le logit de la proportion de souches PRN– à travers une approche fréquentiste. Pour cela, on pose $X_i(t) = \sum_{k=0}^t x'_i(k)$ avec $x'_i(0) = 0$ et $x'_i(k > 0) = x_i(k-1)$, et $\psi = \text{logit}(p_1(0))$. Dans ces analyses, $x'_i(k) \in (\Delta_i(k), \mu_i(k))$. En négligeant l'effet aléatoire, on peut écrire le modèle linéaire généralisé sous la forme

$$\text{logit}(p_1(t)) = \psi + a_0 t + \sum_i a_i X_i(t), \quad (3.13)$$

qui est similaire au modèle 3.9.

Les résultats des analyses menées avec $k' = t-1$ et $k' = t$ sont donnés dans les Tableaux 3.1 et 3.2, respectivement. Les modèles sont comparés à un modèle sans cofacteur (appelé modèle nul) grâce à un test du rapport de vraisemblance (LRT) à un degré de liberté.

Avec $k' = t-1$, le meilleur modèle est celui incluant le cofacteur μ_u comme variable explicative ($LRT = 21,646$, $p < 0.001$). Ce cofacteur représente la moyenne de la température observée là où les souches PRN– ont été isolées. La valeur estimée du coefficient de sélection est $a_0 = 0,529$ et celle du coefficient associé au cofacteur μ_u est $a_u = 0,479$. Cela signifie que les chances d'observer une souche PRN– plutôt qu'une souche PRN+ augmente de 69,8 % par an environ (à valeur constante du cofacteur μ_u) et de 61,5 % par écart-type du cofacteur μ_u (à valeur constante de t)¹⁴.

En fixant un seuil de rejet $\alpha = 0,05$, les autres modèles significativement meilleurs que le

14. Ces pourcentages ont été calculés par la formule $\% = (e^{a_i} - 1) \times 100$.

TABLE 3.1 – **Résultats de l’analyse par le GLM** avec les cofacteurs calculés en utilisant $k' = t - 1$. Les *p-values* sont obtenues à partir d’un test de rapport de vraisemblance (LRT) à un degré de liberté. Les modèles sont ordonnés par la valeur de leur LRT du plus grand au plus petit, avec le modèle nul en bas. NA : non applicable.

Modèle	a_0	a_i	LRT	<i>p-value</i>
μ_t	0,529	0,479	21,646	< 0,001
μ_h	0,497	0,385	15,308	< 0,001
Δ_h	0,681	-0,390	12,258	< 0,001
μ_r	0,677	-0,447	11,100	0,001
Δ_m	0,420	0,456	10,931	0,001
Δ_r	0,683	-0,472	8,237	0,004
μ_m	0,570	-0,349	6,387	0,011
Δ_t	0,621	-0,335	5,604	0,018
Δ_f	0,634	-0,327	4,556	0,033
μ_p	0,510	0,259	2,284	0,131
μ_a	0,571	-0,253	2,120	0,145
μ_f	0,585	-0,234	1,840	0,175
Δ_c	0,549	-0,177	1,366	0,242
μ_c	0,558	-0,203	1,353	0,245
Δ_a	0,426	0,147	0,982	0,322
Δ_p	0,513	-0,139	0,643	0,423
nul	0,470	NA	NA	NA

modèle nul sont ceux utilisant les cofacteurs μ_h , Δ_h , μ_r , Δ_m , Δ_r , μ_m , Δ_u et Δ_f , classés dans cet ordre (Tableau 3.1). Avec le modèle nul, le coefficient de sélection est estimé à $a_0 = 0,470$, soit une augmentation des chances d’observer une souche PRN- plutôt qu’une souche PRN+ de 60,0 % par an environ.

Avec $k' = t$, le meilleur modèle reste celui avec μ_u comme variable explicative ($LRT = 25,946$, $p < 0,001$). La valeur estimée du coefficient de sélection est $a_0 = 0,326$ et celle du coefficient associé à μ_u est $a_u = 1,262$. Cela signifie que les chances d’observer une souche PRN- plutôt qu’une souche PRN+ augmentent de 38,5 % par an environ (à valeur constante du cofacteur μ_u) et de 253,3 % par écart-type du cofacteur μ_u (à valeur constante de t). Ainsi, en calculant les cofacteurs à partir des valeurs brutes et du nombre de souches isolées de l’année en cours, le modèle estime un rythme temporel d’émergence des souches PRN- plus faible qu’en utilisant les valeurs de l’année précédente, et attribue un rôle beaucoup plus fort au cofacteur μ_u .

Les autres modèles faisant mieux que le modèle nul sont ceux utilisant les cofacteurs Δ_m , μ_m , Δ_u et μ_r , classés dans cet ordre (Tableau 3.2). Avec le modèle nul, le coefficient de sélection

TABLE 3.2 – **Résultats de l’analyse par le GLM** avec les cofacteurs calculés en utilisant $k' = t$. Les *p-values* sont obtenues à partir d’un test de rapport de vraisemblance (LRT) à un degré de liberté. Les modèles sont ordonnés par la valeur de leur LRT du plus grand au plus petit, avec le modèle nul en bas. NA : non applicable.

Modèle	a_0	a_i	LRT	<i>p-value</i>
μ_t	0,326	1,262	25,946	< 0,001
Δ_m	0,333	0,590	9,307	0,002
μ_m	0,686	-0,536	9,209	0,002
Δ_t	0,730	-0,571	8,188	0,004
μ_r	0,643	-0,414	4,772	0,029
Δ_h	0,566	-0,215	1,748	0,186
μ_a	0,577	-0,284	1,552	0,213
Δ_r	0,572	-0,258	1,382	0,240
Δ_f	0,562	-0,222	1,132	0,287
μ_p	0,462	-0,176	1,001	0,317
Δ_a	0,417	0,169	0,769	0,380
μ_c	0,519	-0,131	0,370	0,543
Δ_p	0,487	-0,065	0,098	0,755
μ_f	0,484	-0,035	0,030	0,862
Δ_c	0,474	-0,011	0,003	0,957
μ_h	0,471	-0,008	0,001	0,977
nul	0,470	NA	NA	NA

est le même que lorsque $k' = t - 1$ puisque la manière de calculer les valeurs des cofacteurs n’a pas d’incidence sur le modèle nul.

3.5.2 – Modèle linéaire généralisé mixte

Par définition, un modèle statistique comporte nécessairement une part de variations dans la variable réponse qui ne soit pas expliquée par les variables explicatives. Dans la situation que je modélise dans ce travail, cette erreur résiduelle n’est pas triviale : l’équation (3.9) nous montre que l’erreur résiduelle s’accumule avec le temps à cause de l’autocorrélation de la variable réponse. L’erreur résiduelle correspond donc à des sommes partielles distribuées normalement mais ni indépendantes ni homoscédastiques, ce qui va à l’encontre des conditions d’application des GLM. Comme je le présente plus loin dans la Section 3.6.2, le modèle linéaire généralisé n’est pas robuste à la violation de ces conditions d’application.

Une première solution simple à ce problème pourrait être de négliger ce niveau de complexité en remplaçant ces erreurs cumulées par des variables indépendantes et identiquement distribuées

suivant une loi normale. Le modèle peut ainsi s'écrire sous la forme d'un modèle linéaire généralisé mixte (GLMM) de formule

$$\text{logit}(p_1(t)) = \psi + a_0 t + \sum_i a_i X_i(t) + \mathbf{a}_r \mathbf{Z}(t) \quad (3.14)$$

où \mathbf{a}_r est le vecteur des coefficients de l'effet aléatoire et $\mathbf{Z}(t)$ est le vecteur de longueur t_m des valeurs prises par l'effet aléatoire à l'instant t . L'écart-type de l'effet aléatoire est interprété comme étant l'estimation fournie par l'approche GLMM de σ_e .

Les résultats des analyses menées avec $k' = t - 1$ et $k' = t$ sont données dans les Tableaux 3.3 et 3.4, respectivement. Comme pour les analyses avec le GLM, les modèles sont comparés au modèle nul avec un LRT à un degré de liberté.

Avec $k' = t - 1$, le meilleur modèle est une fois encore celui incluant le cofacteur μ_u comme variable explicative ($LRT = 19,161$, $p < 0,001$). La valeur estimée du coefficient de sélection est $a_0 = 0,529$ et celle du coefficient associé au cofacteur μ_u est $a_u = 0,479$. On voit que ces valeurs sont les mêmes que celles obtenues pour le GLM avec $k' = t - 1$. Ceci est dû à un problème de singularité dans ce modèle (et uniquement celui-ci). L'écart-type de l'effet aléatoire se retrouve ainsi égal à zéro et le GLMM devient un simple GLM. La présence de cette singularité est difficile à comprendre : le modèle est plutôt simple (deux covariables et un seul effet aléatoire), les données sont assez nombreuses, et la corrélation entre les covariables est faible ($R^2 = 0,120$). Quelle que soit l'origine de ce problème, il apparaît que le risque de singularité est une limite de cette approche par GLMM.

Les autres modèles meilleurs que le modèle nul sont ceux utilisant les cofacteurs μ_h et Δ_m , classés dans cet ordre (Tableau 3.3). Avec le modèle nul, le coefficient de sélection est estimé à $a_0 = 0,503$, soit une augmentation des chances d'observer une souche PRN- plutôt qu'une souche PRN+ de 65,4 % par an environ.

Avec $k' = t$, le meilleur modèle reste celui avec μ_u comme variable explicative ($LRT = 21,890$, $p < 0,001$). La valeur estimée du coefficient de sélection est $a_0 = 0,326$ et celle du coefficient associé à μ_u est $a_u = 1,262$. On voit que ces valeurs sont encore une fois les mêmes que celles obtenues pour le GLM, cette fois avec $k' = t$. Cela signifie que le problème de singularité observé lorsque l'on utilise $k' = t - 1$ se retrouve lorsque l'on utilise $k' = t$. Encore une fois, le modèle

TABLE 3.3 – Résultats de l’analyse par le GLMM avec les cofacteurs calculés en utilisant $k' = t - 1$. Les p -values sont obtenues à partir d’un test de rapport de vraisemblance (LRT) à un degré de liberté. Les modèles sont ordonnés par la valeur de leur LRT du plus grand au plus petit, avec le modèle nul en bas. NA : non applicable.

Modèle	a_0	a_i	σ_e	LRT	p -value
μ_t	0,529	0,479	0,000	19,161	< 0,001
μ_h	0,493	0,412	0,234	8,102	0,004
Δ_m	0,453	0,402	0,278	4,682	0,030
Δ_h	0,617	-0,289	0,259	3,567	0,059
μ_r	0,597	-0,276	0,283	2,568	0,109
Δ_r	0,587	-0,255	0,335	1,584	0,208
μ_m	0,550	-0,214	0,341	1,486	0,223
Δ_t	0,574	-0,204	0,355	1,295	0,255
μ_p	0,528	0,331	0,411	1,166	0,280
Δ_f	0,567	-0,166	0,371	0,833	0,362
μ_a	0,535	-0,100	0,393	0,245	0,621
Δ_p	0,530	-0,115	0,407	0,221	0,638
Δ_c	0,532	-0,085	0,401	0,197	0,657
Δ_a	0,482	0,084	0,401	0,167	0,683
μ_f	0,527	-0,063	0,401	0,099	0,753
μ_c	0,524	-0,062	0,402	0,091	0,763
nul	0,503	NA	0,410	NA	NA

avec μ_u est le seul pour lequel on trouve ce problème de singularité.

L’unique autre modèle faisant mieux que le modèle nul est celui utilisant le cofacteur Δ_m (Tableau 3.4). Avec le modèle nul, le coefficient de sélection est le même que lorsque $k' = t - 1$ puisque la manière de calculer les valeurs des cofacteurs n’a pas d’incidence sur le modèle nul.

3.6 – Analyse de robustesse

3.6.1 – Principe

Comme expliqué plus tôt, les approches par GLM et GLMM simplifient le modèle formalisé par l’équation 3.9, soit en ignorant l’accumulation de l’erreur résiduelle (GLM) soit en tentant de l’estimer par l’ajout d’un effet aléatoire (GLMM). Dans la présente section, je vérifie si ces approches sont robustes face à l’accumulation de l’erreur résiduelle, autrement dit si le risque d’erreur de type I reste constant quand l’erreur résiduelle augmente. J’étudie également la manière

TABLE 3.4 – Résultats de l’analyse par le GLMM avec les cofacteurs calculés en utilisant $k' = t$. Les p -values sont obtenues à partir d’un test de rapport de vraisemblance (LRT) à un degré de liberté. Les modèles sont ordonnés par la valeur de leur LRT du plus grand au plus petit, avec le modèle nul en bas. NA : non applicable.

Modèle	a_0	a_i	σ_e	LRT	p -value
μ_t	0,326	1,262	0,000	21,890	< 0,001
Δ_m	0,400	0,429	0,325	3,211	0,073
μ_m	0,581	-0,252	0,319	1,410	0,235
μ_h	0,451	0,441	0,443	1,080	0,299
Δ_t	0,573	-0,202	0,349	0,729	0,393
Δ_a	0,458	0,162	0,407	0,445	0,505
μ_r	0,547	-0,139	0,377	0,428	0,513
μ_f	0,474	0,094	0,415	0,149	0,700
Δ_h	0,527	-0,070	0,400	0,122	0,727
Δ_c	0,479	0,081	0,414	0,115	0,735
Δ_p	0,486	0,082	0,416	0,090	0,765
μ_c	0,490	0,045	0,414	0,033	0,856
μ_p	0,501	-0,041	0,408	0,021	0,884
μ_a	0,511	-0,026	0,407	0,011	0,916
Δ_r	0,507	-0,010	0,409	0,002	0,967
Δ_f	0,503	0,002	0,411	0,000	0,993
nul	0,503	NA	0,410	NA	NA

dont d’autres paramètres des modèles influencent le risque d’erreur de type I.

Pour cette analyse de robustesse, la première étape consiste à tirer les valeurs d’un quelconque cofacteur $x'_f(t)$. On pose $x'_f(0) = 0$ et on tire t_m valeurs de $x'_f(t > 0)$ dans une distribution normale centrée-réduite ($x'_f(t > 0) \sim \mathcal{N}(0, 1)$). Ainsi, on obtient $t_m + 1$ valeurs de x'_f où t_m est un paramètre choisi arbitrairement et que l’on peut tester. De la même manière, on pose $\varepsilon'(0) = 0$ et on tire t_m valeurs de $\varepsilon'(t > 0)$ dans une distribution normale de moyenne zéro et d’écart-type σ_e .

Les $t_m + 1$ valeurs de logit ($p_1(t)$) sont ensuite obtenues grâce à une formule rappelant les équations (3.9) et (3.13),

$$\text{logit}(p_1(t)) = \psi + a_0 t + a_f \sum_{k=0}^t x'_f(k) + \sum_{k=0}^t \varepsilon'(k), \quad (3.15)$$

où $\psi = \text{logit}(p_1(0))$, a_0 est le coefficient de sélection et a_i est le coefficient associé au cofacteur f . Les valeurs de ces paramètres sont choisies arbitrairement et peuvent être testées.

Les valeurs obtenues à partir de l'équation (3.15) sont utilisées pour calculer les valeurs de $p_1(t)$ grâce à la formule inverse du logit,

$$p_1(t) = \frac{e^{\text{logit}(p_1(t))}}{1 + e^{\text{logit}(p_1(t))}}. \quad (3.16)$$

De là, on utilise les valeurs obtenues par l'équation (3.16) pour tirer le nombre de succès (*i.e.*, le nombre de souches PRN-) pour chaque pas de temps t dans une distribution binomiale avec une probabilité de succès $p_1(t)$ et n essais ($S(t) \sim \mathcal{B}(n, p_1(t))$). Le paramètre n est choisi arbitrairement et est constant pour simplifier les analyses. Dans les approches par GLM et GLMM, j'ai commencé les analyses l'année 2007 puisque cela correspond à la première année d'isolation d'une souche PRN-. Pour mimer cela, dans les analyses de robustesse si le tirage aléatoire donne $S(0) = 0$ alors cette valeur est augmentée à un. Le nombre d'échecs (*i.e.*, le nombre de souches PRN+) pour chaque pas de temps t est quant à lui égal à $F(t) = n - S(t)$.

Enfin, on définit une nouvelle covariable, $x_d(t)$, dont les valeurs sont obtenues de la même manière que $x'_f(t)$: $x_d(0) = 0$ et $x_d(t > 0) \sim \mathcal{N}(0, 1)$. À partir de là, toutes les données nécessaires pour réaliser les analyses de robustesse sont prêtes.

3.6.2 – Modèle linéaire généralisé

L'analyse de robustesse du GLM est faite en utilisant deux modèles, le premier étant

$$\text{logit}(p_1(t)) = \psi + a_0 t + a_f X_f(k), \quad (3.17)$$

et le second

$$\text{logit}(p_1(t)) = \psi + a_0 t + a_f X_f(k) + a_d X_d(k), \quad (3.18)$$

où $X_f(t) = \sum_{k=0}^t x'_f(k)$ et $X_d(t) = \sum_{k=0}^t x_d(k)$.

Le modèle (3.17) est le modèle « correct » puisque les valeurs de $x'_f(t)$, et par extension de $X_f(t)$ sont utilisées pour produire les valeurs de $p_1(t)$. À l'inverse, le modèle (3.18) est « incorrect » parce qu'il utilise les valeurs de $x_d(t)$, et par extension de $X_d(t)$, qui n'ont pas été utilisées pour calculer les valeurs de $p_1(t)$. Pour mener l'analyse de robustesse, j'ai calculé le LRT entre les deux modèles. Sous l'hypothèse nulle, il est attendu que le LRT suive une distribution du

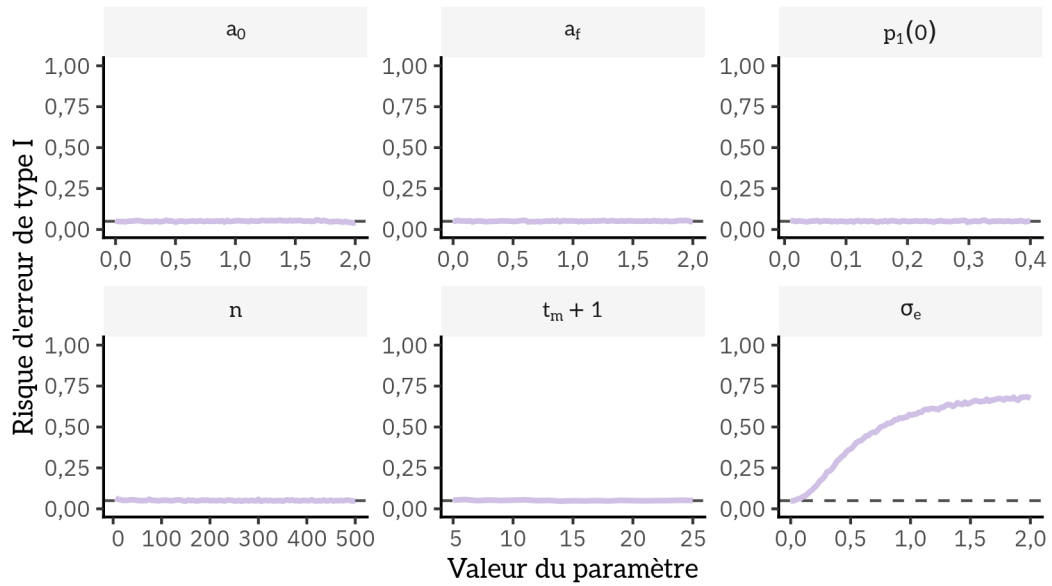


FIGURE 3.7 – **Résultats de l’analyse de robustesse du modèle linéaire généralisé** pour différentes valeurs du coefficient de sélection (a_0), du coefficient associé au cofacteur (a_f), de la proportion de souches PRN- à $t = 0$ ($p_1(0)$), du nombre de souches isolées par pas de temps (n), du nombre de pas de temps considérés ($t_m + 1$) et de l’écart-type de l’erreur résiduelle (σ_e). La ligne horizontale en tirets représente le risque attendu $\alpha = 0,05$.

χ^2 à un degré de liberté. Pour chaque combinaison de valeurs de paramètres, j’ai effectué 10 000 simulations et obtenu autant de valeurs de LRT. La mesure de la robustesse de l’approche GLM est donnée par le nombre de fois où la statistique du test a dépassé une valeur seuil $\chi_{0,05}^2 = 3,841$ représentant le 95^e quantile de la distribution du χ^2 à un degré de liberté.

Les résultats de l’analyse de robustesse sont présentés sur la Figure 3.7. Les valeurs de base des paramètres utilisées dans l’analyse sont $a_0 = a_f = 0,470$, correspondant à la valeur estimée de a_0 pour le modèle nul dans l’analyse des données par le GLM (Tableau 3.1), $p_1(0) = 0,203$, correspondant là encore à la valeur estimée dans l’analyse de données, $n = 100$, correspondant approximativement au nombre moyen de souches isolées chaque année, $t_m = 10$, correspondant au nombre de pas de temps dans l’analyse de données (en excluant $t = 0$) et $\sigma_e = 0$.

On peut voir que l’approche par GLM est robuste face à des valeurs de a_0 , a_f , $p_1(0)$, n et t_m aussi bien basses qu’élévées. Les paramètres a_0 , a_f et dans une moindre mesure $p_1(0)$ influencent la vitesse à laquelle la proportion de succès va atteindre 100 % (Figure 3.7, panels du haut). Si ces paramètres ont des valeurs élevées, la proportion de succès va rapidement augmenter, tandis que

si les valeurs sont basses alors la proportion de succès va augmenter lentement. Cela ne semble pas avoir d'impact sur la robustesse du modèle, au moins pour des valeurs « raisonnables » des paramètres. En revanche, utiliser des valeurs aberrantes (*e.g.*, $a_0 = a_f = 5$) peut conduire à des problèmes algorithmiques faussant les estimations des coefficients et réduisant artificiellement le risque d'erreur de type I.

Les paramètres n et t_m influencent la taille des échantillons et leur nombre, respectivement. On remarque ici aussi que les valeurs de ces paramètres n'ont pas d'incidence sur la robustesse du modèle (Figure 3.7, panels bas-gauche et bas-milieu). En revanche, il est vraisemblable que ces paramètres ont une incidence importante sur la puissance de l'approche par GLM.

Enfin, le dernier paramètre, et celui de première importance, est l'écart-type de l'erreur résiduelle dans les données, σ_e . Plus celui-ci est grand, plus cela implique que la vitesse de changement de la proportion de succès change indépendamment du ou des cofacteur(s) inclus dans le modèle. On remarque que l'approche par GLM n'est absolument pas robuste face à une valeur de σ_e ne serait-ce que légèrement plus élevée que zéro (Figure 3.7, panel bas-droite). Ainsi, le risque de conclure à tort qu'un modèle incluant un cofacteur est meilleur qu'un modèle excluant ce cofacteur alors même que ce cofacteur n'a pas d'impact sur la variable réponse (ici, la proportion de souches PRN-) devient excessivement élevé, rendant l'approche par GLM inutilisable en pratique.

3.6.3 – Modèle linéaire généralisé mixte

L'analyse de robustesse du GLMM est faite de manière similaire à celle du GLM. Deux modèles sont construits, le premier étant

$$\text{logit}(p_1(t)) = \psi + a_0 t + a_f X_f(k) + \mathbf{a}_r \mathbf{Z}(t), \quad (3.19)$$

et le second

$$\text{logit}(p_1(t)) = \psi + a_0 t + a_f X_f(k) + a_d X_d(k) + \mathbf{a}_r \mathbf{Z}(t). \quad (3.20)$$

Les interprétations des deux modèles sont similaires à celles présentées pour l'analyse de robustesse du GLM. L'analyse de robustesse est menée de la même manière que pour l'approche

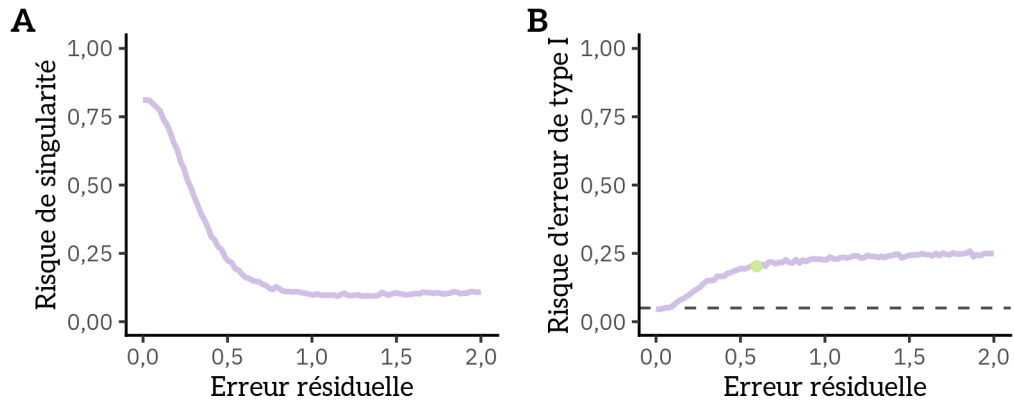


FIGURE 3.8 – Résultats de l’analyse de robustesse du modèle linéaire généralisé mixte pour différentes valeurs d’erreur résiduelle (σ_e). (A) Risque de singularité. (B) Risque d’erreur de type I. La ligne horizontale en tirets représente le risque attendu $\alpha = 0,05$. Le point indique le risque d’erreur de type I correspondant à l’écart-type de l’effet aléatoire du modèle nul, représentant lui-même une estimation de l’erreur résiduelle dans les données.

GLM, mais avec des valeurs différentes pour certains paramètres constants : $a_0 = a_f = 0,503$ et $p_1(0) = 0,159$. Ces valeurs sont celles estimées dans l’analyse des données par le GLMM nul (Tableau 3.3). Les paramètres $n = 100$ et $t_m = 10$ sont les mêmes que pour l’analyse de robustesse du GLM.

La robustesse de l’approche GLMM n’a été étudiée que pour le paramètre σ_e . La raison principale est que pour $\sigma_e = 0$, le risque de singularité est très élevé (Figure 3.8A). Cela implique qu’une faible variabilité inexpliquée par le(s) cofacteur(s) conduit le modèle à ne pas pouvoir estimer l’effet aléatoire dans le modèle et *in fine* à le considérer comme un simple GLM. Il est donc probable que l’analyse de robustesse pour les paramètres autres que σ_e donne des résultats similaires à ceux observés pour l’approche GLM.

On remarque que l’ajout d’un effet aléatoire à chaque pas de temps dans le modèle ne suffit pas à résoudre le problème de robustesse apparaissant dès que σ_e est supérieur à zéro (Figure 3.8B). Cependant, il semblerait que le risque d’erreur de type I soit légèrement moins sensible à la valeur de σ_e lorsque l’effet aléatoire est inclus dans le modèle que lorsqu’il en est exclu, suggérant que cette approche permette malgré tout de compenser au moins une partie de l’accumulation de l’erreur résiduelle.

3.7 – Approche bayésienne

3.7.1 – Principe

L'approche bayésienne consiste à modéliser la différence entre les valeurs de logit ($p_1(t)$) et logit ($p_1(t-1)$) pour chaque pas de temps $t > 0$, notée $\Delta\text{logit}(p_1(t))$. En résumé, le modèle va considérer l'erreur $\varepsilon(t)$ comme étant une variable à estimer, et utiliser cette estimation dans le calcul du $\Delta\text{logit}(p_1(t))$; puis, à partir des valeurs de $\Delta\text{logit}(p_1(t))$, le modèle va estimer les valeurs des coefficients a_0 et a_i , ce qui est donc son but premier.

Pour une explication plus détaillée, voici les différentes étapes du modèle : les valeurs des coefficients a_i sont tirées dans une distribution *a priori* normale de moyenne μ_i et d'écart-type σ_i ($a_i \sim \mathcal{N}(\mu_i, \sigma_i)$). Les valeurs de μ_i et σ_i sont quant à elles tirées dans des distributions *a priori* uniformes ($\mu_i \sim \mathcal{U}(-100, 100)$ et $\sigma_i \sim \mathcal{U}(0, 50)$). Ces distributions *a priori* et celles présentées plus loin ont été choisies naïvement et ne sont pas encore justifiées dans l'état où est arrivé ce travail. Il serait cependant intéressant de considérer différentes distributions pour étudier l'impact de celles-ci sur les capacités du modèle.

Pour le pas de temps $t = 0$, le modèle tire une valeur pour $p_1(t)$ dans une distribution *a priori* uniforme allant de zéro à un ($p_1(0) \sim \mathcal{U}(0, 1)$) et calcule la valeur de logit ($p_1(0)$). Le modèle pioche également le nombre de succès (*i.e.*, le nombre de souches PRN- isolées), $S(0)$, à partir d'une distribution *a priori* binomiale avec une probabilité de succès $p_1(0)$ et un nombre d'essais $n(0)$, et qui correspond au nombre total de souches isolées à $t = 0$ ($S(0) \sim \mathcal{B}(p_1(0), n(0))$).

Pour tous les pas de temps $t > 0$, le modèle calcule la valeur de $\Delta\text{logit}(p_1(t))$ en utilisant la formule

$$\Delta\text{logit}(p_1(t)) = a_0 + \sum_i a_i x'_i(t) + \varepsilon(t-1), \quad (3.21)$$

où $x'_i \in (\Delta_i(t), \mu_i(t))$ sont les mêmes valeurs centrées et réduites utilisées dans l'approche fréquentiste (Section 3.5).

La valeur de $\varepsilon(t)$ est modélisée explicitement dans ce modèle bayésien, ce qui est une des différences fondamentales avec l'approche fréquentiste. Cette valeur est obtenue en piochant dans une distribution *a priori* normale de moyenne zéro et d'écart-type σ_e ($\varepsilon(t) \sim \mathcal{N}(0, \sigma_e)$) avec σ_e

tiré dans une distribution *a priori* uniforme ($\sigma_e \sim \mathcal{U}(0,50)$).

Le modèle calcule ensuite les valeurs de logit ($p_1(t)$) grâce à la formule

$$\text{logit}(p_1(t)) = \text{logit}(p_1(t-1)) + \Delta \text{logit}(p_1(t)) \quad (3.22)$$

L'étape suivante pour le modèle est de calculer les valeurs de $p_1(t)$ grâce à la formule (3.16). À partir de là, les valeurs de $S(t)$ sont tirées dans une distribution *a priori* binomiale avec une probabilité de succès $p_1(t)$ et $n(t)$ essais ($S(t) \sim \mathcal{B}(n(t), p_1(t))$).

Le modèle nécessite plusieurs paramètres en entrée : le nombre de pas de temps considérés (excluant $t = 0$) t_m , le nombre de cofacteurs inclus dans l'analyse n_f , une matrice contenant les valeurs $x'_i(t)$ des cofacteurs (notée $\mathbf{X}(t)$), un vecteur contenant pour chaque année le nombre $n(t)$ de souches isolées et un vecteur contenant pour chaque année le nombre $S(t)$ de succès (*i.e.*, de souches PRN-). Le modèle a été écrit dans le langage BUGS (Thomas, 2006) et les simulations ont été conduites en utilisant le programme JAGS (Plummer, 2003). Le package `jagsUI` v1.6.2 (Kellner et Meredith, 2024) a été utilisé pour interagir avec JAGS depuis R v4.4.1 (R Core Team, 2024). Une version schématique du modèle est présentée sur la Figure 3.9.

Pour mesurer la qualité du modèle, j'ai utilisé le *Deviance Information Criterion* (DIC) (Spiegelhalter *et al.*, 2002) dans sa formulation proposée par Gelman *et al.* (2004, p. 182). On a donc

$$DIC = \overline{D(M)} + \frac{1}{2} \sigma_{D(M)}^2, \quad (3.23)$$

où $\overline{D(M)}$ is la déviance moyenne du modèle M et $\sigma_{D(M)}^2$ est la variance de la déviance du modèle M . Le principe du DIC est analogue à celui de l'AIC. On considère donc qu'un modèle avec un plus bas DIC est meilleur qu'un modèle avec un DIC plus haut.

Quoique ce modèle bayésien permette techniquement d'inclure plusieurs cofacteurs dans une même analyse, j'ai choisi d'introduire les différents cofacteurs un à un, et ce pour deux raisons. La première est que cela permet de garder une cohérence avec l'approche fréquentiste présentée plus tôt. La deuxième est que tester toutes les combinaisons de cofacteurs demanderait un temps de calcul beaucoup trop long¹⁵. Chaque simulation a été faite avec la paramétrisation qui suit :

15. Cela représenterait 65 536 combinaisons différentes, chacune tournant pendant deux minutes environ, soit plus de 90 jours de calcul.

3.7.2 – Résultats

Les résultats des analyses menées avec $k' = t - 1$ et $k' = t$ sont données dans les Tableaux 3.5 et 3.6, respectivement. Comme présenté plus tôt, la qualité des modèles est mesurée en utilisant le DIC. De plus, La « significativité » des valeurs des coefficients associés aux cofacteurs est mesurée par la proportion doublée de valeurs de a_i échantillonnées dans chaque simulation qui soient différentes de la valeur médiane de a_i . Cette mesure est notée P . Si par exemple $P = 0,038$, cela signifie que zéro n'est pas dans l'intervalle de crédibilité à 95 % de la a_f . Par construction, $0 \leq P \leq 1$ et P peut être interprétée comme un analogue bayésien de la p -value.

Avec $k' = t - 1$, le meilleur modèle est comme pour les analyses par GLM et GLMM celui incluant le cofacteur μ_u comme variable explicative ($\Delta DIC = -1,558$, $P = 0,060$). La valeur estimée du coefficient de sélection est $a_0 = 0,487$ et celle du coefficient associé au cofacteur μ_u est $a_u = 0,453$. On remarque cependant que la différence de DIC entre ce modèle et le modèle nul n'est pas particulièrement élevée ($|\Delta DIC| < 2$). De plus, l'intervalle de crédibilité à 95 % de a_u n'exclut pas zéro.

Aucun autre modèle ne présente de coefficient a_i pour lequel l'intervalle de crédibilité à 95 % n'exclut zéro (Tableau 3.5). Avec le modèle nul, le coefficient de sélection est estimé à $a_0 = 0,487$, soit une augmentation des chances d'observer une souche PRN- plutôt qu'une souche PRN+ de 62,8 % par an environ.

Avec $k' = t$, le meilleur modèle est toujours celui avec le cofacteur μ_u ($\Delta DIC = -6,318$, $P = 0,002$). On remarque cependant que la différence de qualité avec le modèle nul est beaucoup plus marquée que précédemment ($|\Delta DIC| > 2$). De plus, l'intervalle de crédibilité à 95 % exclut zéro, ce qui n'était pas le cas avec $k' = t - 1$. Ceci pourrait indiquer que la température annuelle moyenne a une influence à courte échelle temporelle sur la vitesse de propagation des souches PRN-, quoique l'analyse reste limitée pour appuyer une telle proposition. Il serait intéressant de pouvoir étudier la propagation des souches PRN- à une échelle temporelle plus réduite (intra-annuelle) pour étayer cette conclusion. Avec ce modèle, la valeur estimée du coefficient de sélection est $a_0 = 0,337$ et celle du coefficient associé au cofacteur μ_u est $a_u = 1,247$.

Il est intéressant de noter que ces valeurs sont proches de celles obtenues par la méthode du

TABLE 3.5 – **Résultats de l’analyse par le modèle bayésien** avec les cofacteurs calculés en utilisant $k' = t - 1$. Les valeurs de P représentent la proportion doublée de valeurs de a_i de signe différent de la médiane de a_i calculée pour chaque simulation. Les modèles sont ordonnés par la valeur de leur DIC du plus petit au plus grand, avec le modèle nul en bas. NA : non applicable.

Modèle	a_0	a_i	σ_e	DIC	ΔDIC	P
μ_t	0,487	0,453	0,398	58,185	-1,558	0,060
μ_c	0,431	0,245	0,680	59,511	-0,232	0,409
μ_p	0,434	0,218	0,682	59,736	-0,007	0,463
μ_a	0,436	0,226	0,685	59,795	0,052	0,448
Δ_c	0,446	0,129	0,675	59,962	0,220	0,653
μ_r	0,457	0,255	0,704	59,986	0,243	0,508
μ_f	0,455	0,111	0,698	60,126	0,383	0,708
Δ_p	0,455	0,061	0,689	60,189	0,446	0,824
Δ_t	0,456	0,045	0,687	60,330	0,587	0,870
μ_m	0,462	0,055	0,690	60,360	0,617	0,852
Δ_a	0,439	0,221	0,629	60,399	0,656	0,441
Δ_m	0,466	-0,014	0,681	60,423	0,680	0,962
Δ_r	0,458	0,020	0,673	60,469	0,726	0,950
Δ_h	0,480	-0,070	0,667	60,545	0,802	0,793
μ_h	0,457	0,340	0,521	60,810	1,067	0,166
Δ_f	0,453	0,279	0,552	62,264	2,521	0,354
nul	0,461	NA	0,602	59,743	NA	NA

GLMM/GLM¹⁶ : $a_0 = 0,326$ et $a_u = 1,262$. Ceci pourrait être dû à la faible erreur résiduelle dans ce modèle en particulier : l’approche bayésienne estime que $\sigma_e = 0,206$ pour le modèle μ_u . Celle-ci serait suffisamment faible pour ne pas perturber les estimations des valeurs des coefficients. En revanche, comme l’a montré l’analyse de robustesse, elle reste trop élevée pour permettre de tirer des conclusions fiables quant aux effets des cofacteurs sur la vitesse de propagation des souches PRN-.

Pour le modèle utilisant comme cofacteur Δ_f l’intervalle de crédibilité à 95 % exclut également zéro, cependant le DIC de ce modèle est proche de celui du modèle nul (Tableau 3.6). Pour aucun autre modèle l’intervalle de crédibilité à 95 % de a_i n’exclut zéro.

16. Pour rappel, les deux méthodes donnent le même résultat pour le modèle μ_u à cause d’un problème de singularité.

TABLE 3.6 – **Résultats de l’analyse par le modèle bayésien** avec les cofacteurs calculés en utilisant $k' = t$. Les valeurs de P représentent la proportion doublée de valeurs de a_f de signe différent de la médiane de a_i calculée pour chaque simulation. Les modèles sont ordonnés par la valeur de leur DIC du plus petit au plus grand, avec le modèle nul en bas. NA : non applicable.

Modèle	a_0	a_i	σ_e	DIC	ΔDIC	P
μ_t	0,337	1,247	0,206	53,594	-6,318	0,002
Δ_f	0,383	0,640	0,507	57,920	-1,993	0,040
Δ_a	0,395	0,504	0,613	58,311	-1,601	0,122
μ_p	0,406	0,439	0,659	58,795	-1,118	0,194
Δ_c	0,409	0,417	0,651	58,852	-1,061	0,205
μ_c	0,408	0,410	0,669	59,042	-0,871	0,232
μ_a	0,419	0,355	0,685	59,284	-0,628	0,304
Δ_m	0,423	0,320	0,684	59,398	-0,515	0,346
Δ_h	0,430	0,257	0,680	59,459	-0,454	0,386
μ_m	0,434	0,247	0,699	59,545	-0,367	0,434
Δ_p	0,425	0,305	0,682	59,557	-0,356	0,360
μ_h	0,427	0,495	0,689	59,722	-0,191	0,295
μ_f	0,448	0,166	0,702	59,999	0,087	0,611
Δ_r	0,443	0,218	0,681	60,217	0,304	0,524
Δ_t	0,462	0,062	0,680	60,465	0,553	0,854
μ_r	0,461	0,031	0,675	60,570	0,657	0,920
nul	0,463	NA	0,602	59,913	NA	NA

3.8 – Analyse de puissance

3.8.1 – Principe

J’ai mis en place une analyse de puissance du modèle bayésien pour mettre en évidence dans quelles conditions il peut être employé de manière satisfaisante. Pour ce faire, j’ai effectué un grand nombre de simulations dans lesquelles j’ai modifié les valeurs des différents paramètres du modèle. J’ai conduit deux séries de simulations.

La première série ressemble au processus présenté plus tôt pour tester la robustesse de l’approche fréquentiste (Section 3.6). On commence par tirer les valeurs d’un cofacteur hypothétique $x'_f(t)$. On pose $x'_f(0) = 0$ et on tire les t_m valeurs de $x'_f(t > 0)$ à partir d’une distribution normale centrée-réduite ($x'_f(t > 0) \sim \mathcal{N}(0, 1)$). On pose $\varepsilon'(0) = 0$ et on tire les t_m valeurs de $\varepsilon'(t > 0)$ d’une distribution normale de moyenne zéro et d’écart-type σ_e . On fixe arbitrairement les valeurs de a_0 et de p_0 . Les valeurs fixées des paramètres sont $a_0 = 0,5$, $\sigma_r^2 = 1$, $t_m = 10$,

$n = 200$ et $p_1(0) = 0,1$.

Pour obtenir la valeur de a_f , on commence par fixer la variance totale des données qui ne soit pas expliquée par le coefficient de sélection, σ_r^2 , et on détermine quelle fraction η^2 de cette variance est expliquée par le cofacteur, le reste correspondant à la variance résiduelle. On a alors $a_f = \sqrt{\eta^2} \cdot \sqrt{\sigma_r^2}$ et $\sigma_e = \sqrt{\sigma_r^2 - a_f^2}$.

On utilise ces différents paramètres pour calculer les $t_m + 1$ valeurs de $\text{logit}(p_1(t))$ à partir de la formule (3.15) puis les valeurs de $p_1(t)$ à partir de la formule (3.16). Enfin, on tire le nombre de succès pour chaque pas de temps t dans une distribution binomiale avec une probabilité de succès $p_1(t)$ et un nombre d'essais n ($S(t) \sim \mathcal{B}(n, p_1(t))$). Pour des raisons de simplification des analyses, je n'ai pas forcé $S(0)$ à être plus grand que zéro dans ces analyses.

Une fois les données prêtes, j'ai conduit des simulations en incorporant les valeurs de $S(t)$, n et $x'_f(t)$ dans le modèle bayésien présenté dans la Section 3.7. Les MCMC ont été paramétrées comme celles de la Section 3.7 à la différence que seulement 100 000 itérations ont été conservées avant l'étape de *thinning*, contre 1 000 000 dans les analyses précédentes, ceci pour économiser du temps de calcul. Pour chaque combinaison de valeurs de paramètres, 1 000 simulations ont été faites. La puissance des analyses, c'est-à-dire la capacité du modèle à détecter l'effet d'un cofacteur sur la vitesse de propagation de la souche mutante, est mesurée par le nombre de simulations pour lesquelles l'intervalle de crédibilité à 95 % de la valeur de a_f retourné par le modèle exclut zéro parmi toutes les simulations.

La deuxième série de simulations repose sur l'utilisation du jeu de données de souches PRN-utilisé tout au long de ce travail. La particularité de ces simulations est que le nombre d'isolats par année n'est pas constant, contrairement à la première série de simulations. Cette série utilise donc un vecteur $n(t)$ plutôt qu'une valeur unique n . Les valeurs de σ_r^2 , $p_1(0)$ et a_0 ne sont pas choisies arbitrairement mais estimées en faisant tourner le modèle sur les données en ne gardant que le temps comme variable explicative. Autrement, le principe de cette deuxième série de simulations est le même que celui de la première.

Ces simulations permettent de vérifier comment changer le nombre moyen et la variance dans le nombre d'individus échantillonnés par an affecte la puissance des analyses. De plus, ces simulations permettent d'étudier comment changer la fréquence d'échantillonnage (*e.g.*, en passant d'une fréquence d'un an à une fréquence de six mois) modifie la puissance des analyses.

Soient α_1 , α_2 et α_3 les facteurs permettant de modifier la variance dans la taille des échantillons, la taille moyenne des échantillons et la fréquence d'échantillonnage, respectivement. Les valeurs des paramètres a_0 , σ_r^2 et t_m doivent être modifiées pour conserver la structure des données lorsque la fréquence d'échantillonnage change : les simulations utilisent les valeurs $a'_0 = a_0\alpha_3$, $\sigma_r^{2'} = \sigma_r^2\alpha_3$ et $t'_m = t_m\alpha_3 + \alpha_3 - 1$. La taille de l'échantillon à l'instant t' devient

$$n'(t') = \left\lceil \frac{\alpha_2 \cdot \left(n \left(\left\lfloor \frac{t'}{\alpha_3} \right\rfloor \right) \cdot \sqrt{\frac{1}{\alpha_1}} + \bar{n} \cdot \left(1 - \sqrt{\frac{1}{\alpha_1}} \right) \right)}{\alpha_3} \right\rceil, \quad (3.24)$$

où $\lceil \cdot \rceil$ est la fonction partie entière supérieure, $\lfloor \cdot \rfloor$ est la fonction partie entière inférieure et \bar{n} est la taille moyenne d'échantillon entre tous les temps t .

L'équation (3.24) implique de diviser chaque année d'échantillonnage en α_3 pas de temps distincts et de distribuer les échantillons de l'année équitablement entre chaque pas de temps, en arrondissant à l'entier supérieur. Par exemple, si une année 65 souches ont été collectés et le pas de temps est divisé par quatre (*i.e.*, la fréquence d'échantillonnage multipliée par quatre), alors chacun de ces nouveaux pas de temps reçoit 17 souches.

3.8.2 – Résultats

Suite à la première série d'analyses de puissance, plusieurs observations peuvent être faites. Premièrement, il apparaît que plus la fraction de la variance expliquée par le cofacteur est élevée, plus la puissance du modèle est élevée (Figures 3.10 et 3.11). Ensuite, il semblerait que la puissance soit une fonction concave du coefficient de sélection (Figure 3.10, panel de gauche). Cela implique que si le coefficient de sélection est trop bas ou trop élevé le modèle voit sa puissance diminuer, et une valeur intermédiaire de a_0 serait ainsi optimale. Cependant, il apparaît que cette valeur optimale dépend de la part de variance résiduelle expliquée par le coefficient associé au cofacteur, η^2 .

Pour les paramètres n et t_m , les résultats sont classiques : lorsque leurs valeurs augmentent, la puissance du modèle augmente également. Cependant, on remarque que la puissance semble atteindre un plateau pour de fortes valeurs de n , et que ce plateau dépend de η^2 (Figure 3.10, panel du milieu). Ainsi, augmenter continuellement la taille des échantillons – ce qui peut re-

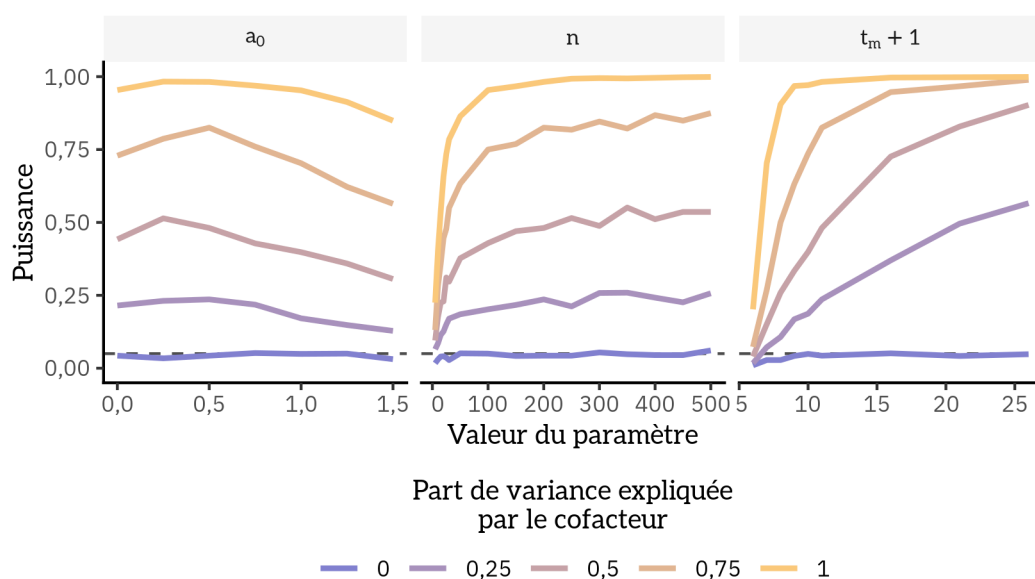


FIGURE 3.10 – **Puissance du modèle bayésien** pour la première série d’analyses en fonction du coefficient de sélection (a_0), du nombre de souches isolées par pas de temps (n) et du nombre de pas de temps considérés ($t_m + 1$). La ligne en tirets indique une puissance de 0,05.

présenter un coût important – n’est pas la meilleure mesure à mettre en place pour améliorer la puissance des analyses. En revanche, augmenter le nombre de pas de temps inclus dans les analyses semble augmenter la puissance sans effet plateau (Figure 3.10, panel de droite). Cependant, augmenter le nombre de pas de temps dans les analyses peut considérablement augmenter la durée de simulation du modèle bayésien.

Enfin, il apparaît que la puissance du modèle est partiellement conditionnée par la variance résiduelle totale (σ_r^2). Si cette dernière est trop basse (*i.e.*, si la variabilité dans les données est presque entièrement expliquée par le temps, et donc le coefficient de sélection), alors la capacité du modèle à détecter l’effet d’un cofacteur en sera réduite (Figure 3.11). Cependant, cette contrainte n’apparaît que pour des valeurs particulièrement basses de σ_r^2 .

La deuxième série d’analyses apporte à son tour plusieurs informations intéressantes. Cette série a été menée avec les valeurs de base $a_0 = 0,457$, $p_0 = 0,162$, $t_m = 10$ et $\sigma_r^2 = 0,355$. On remarque en premier lieu que diminuer la variance dans la taille des échantillons entre années augmente légèrement la puissance des analyses (Figure 3.12A). Plus précisément, on remarque que diviser la variance par deux permet d’augmenter la puissance mais que la diviser encore plus

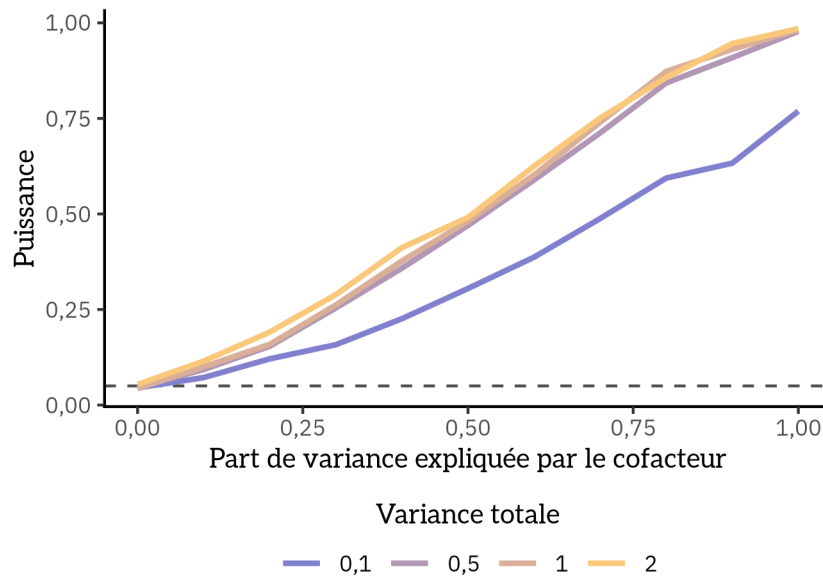


FIGURE 3.11 – **Puissance du modèle bayésien** pour la première série d’analyses en fonction de la part de variance expliquée par le cofacteur pour différentes variances résiduelles. La ligne en tirets indique une puissance de 0,05.

n’a presque plus d’effet sur la puissance des analyses. Plus qu’un besoin d’homogénéisation entre années dans l’effort d’échantillonnage, ce premier résultat montre que les années avec un très faible nombre de souches isolées réduisent la puissance des analyses plus que les années avec un fort nombre de souches ne l’augmentent.

On remarque ensuite qu’augmenter la taille moyenne des échantillons permet également d’augmenter la puissance des analyses (Figure 3.12B). Cependant, augmenter la taille des échantillons n’est pas trivial. Déjà, puisqu’il s’agit de données épidémiologiques, le nombre de souches isolées est nécessairement conditionné par la prévalence du pathogène étudié. Dans le jeu de données de [Lefrancq et al. \(2022\)](#), le nombre de souches isolées par année suit grossièrement la tendance épidémique de la coqueluche aux États-Unis. Par exemple, l’année avec le plus de souches isolées est 2012, correspondant à un pic épidémique dans le pays. De plus, la collecte, la gestion et le séquençage des souches représente un coût non négligeable qui peut représenter un frein à l’augmentation du nombre moyen de souches isolées.

Enfin, augmenter la fréquence d’échantillonnage, ou le pas de temps auquel sont annotées les données, conduit aussi à une augmentation de la puissance des analyses (Figure 3.12C). Avec

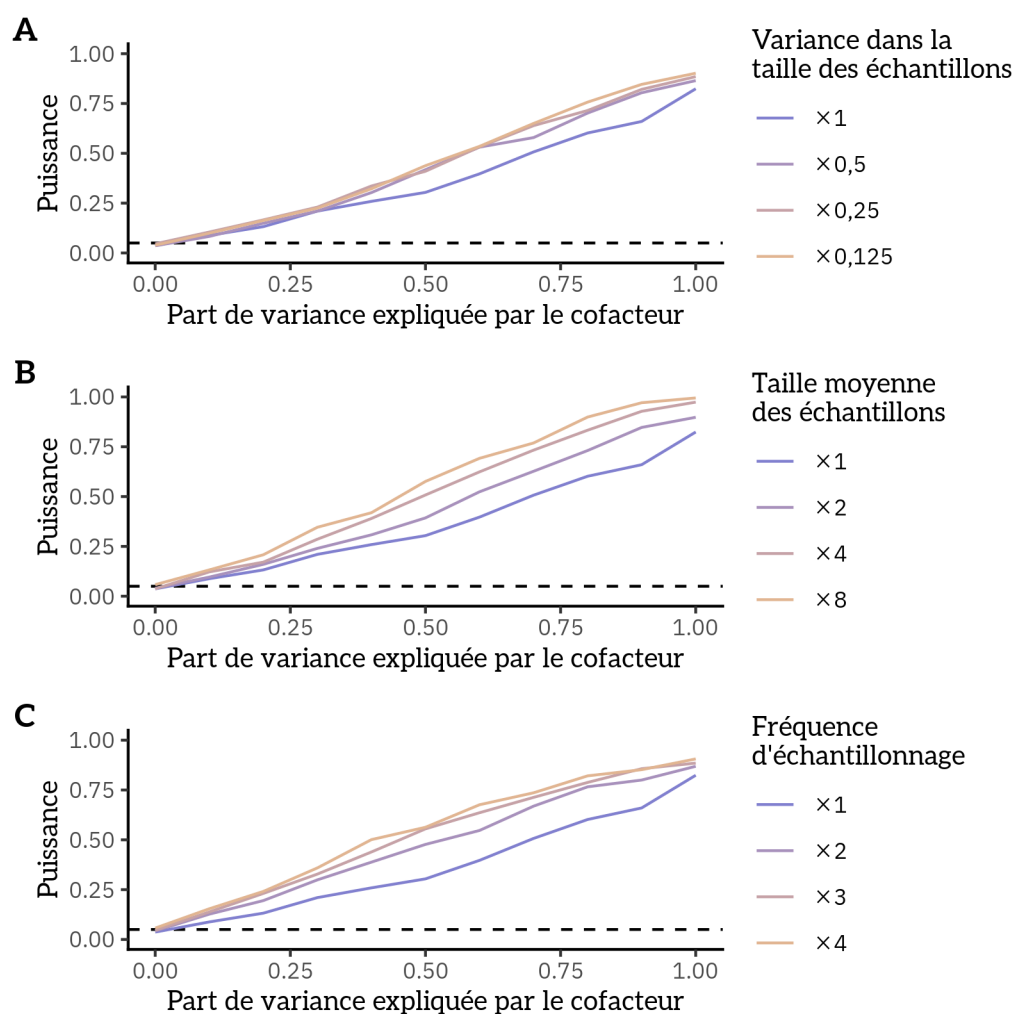


FIGURE 3.12 – **Puissance du modèle bayésien** pour la deuxième série d'analyses en fonction de la part de variance expliquée par le cofacteur. **(A)** Réduction de la variance (α_1). **(B)** Augmentation de la taille moyenne des échantillons (α_2). **(C)** Augmentation de la fréquence d'échantillonnage (α_3). La ligne en tirets indique une puissance de 0,05.

la méthode utilisée dans l'analyse de puissance pour augmenter la fréquence d'échantillonnage, doubler celle-ci conduit à diviser par deux (environ, le nombre de souches collectées par pas de temps étant arrondi à l'entier supérieur) la taille moyenne des échantillons. Or, j'ai montré juste avant que lorsque la taille moyenne des échantillons est réduite la puissance diminue. Cela signifie que le gain de puissance associé à une augmentation de la fréquence d'échantillonnage est suffisamment important pour compenser la perte de puissance associée à une diminution de

la taille moyenne des échantillons. Néanmoins, on remarque que le gain de puissance diminue rapidement lorsque l'on continue d'augmenter la fréquence d'échantillonnage, probablement à cause de l'effet de la diminution de la taille moyenne des échantillons.

3.9 – Discussion

L'évolution des pathogènes est une menace pour la santé publique. Des épidémies peuvent être déclenchées simplement par l'apparition de souches mutantes de pathogènes circulant déjà au sein de populations. Le but de ce deuxième chapitre de thèse était d'analyser un jeu de données décrivant l'émergence et la propagation de souches de *Bordetella pertussis* PRN– aux États-Unis entre 2007 et 2017 pour comprendre quels facteurs environnementaux et socio-démographiques ont pu influencer la vitesse de propagation de ces souches. Huit facteurs différents ont été inclus dans ces analyses. Certains sont réputés avoir un effet sur la propagation de pathogènes aéroportés, comme la densité de population – mais celle-ci aurait des effets contrastés selon le pathogène (Akaba *et al.*, 2022) –, la température annuelle moyenne (*e.g.*, Roussel *et al.*, 2016; Ganslmeier *et al.*, 2021; Jia *et al.*, 2023). D'autres sont beaucoup plus exploratoires, comme la densité d'hôpitaux ou de médecins.

Pour ce faire, j'ai d'abord utilisé une approche classique : un modèle linéaire généralisé (GLM)) Les GLM sont utilisés en routine pour étudier des variations de proportion en fonction d'une ou plusieurs variables explicatives. Avec cette approche simple, on conclut à un effet principalement de la température sur la vitesse de propagation des souches PRN–. D'autres cofacteurs semblent également avoir une influence, mais celle-ci est moins évidente – quoique le coefficient associé soit parfois plus élevé – et cette conclusion change selon le pas de temps considéré pour les calculs des cofacteurs (Section 3.5.1). De plus, cette approche par GLM manque en réalité cruellement de robustesse : une erreur résiduelle même faible peut mener facilement à la conclusion d'un effet d'un cofacteur quand celui-ci n'existe en réalité pas (Section 3.6.2).

Dans leur formulation usuelle, les GLM reposent sur la supposition que la variable réponse n'est pas autocorrélée. Dans le contexte d'un phénomène d'invasion étudié à travers le temps, cette hypothèse n'est pas respectée. L'autocorrélation dans la variable réponse conduit à une autocorrélation dans l'erreur résiduelle, qui prend la forme d'une erreur cumulée dans le modèle.

C'est à cause de cette erreur cumulée que l'approche par GLM n'est pas robuste à la présence d'erreur résiduelle dans les données.

Pour résoudre ce problème, la première solution que j'ai proposée a été d'inclure un effet aléatoire dans le GLM pour capturer l'excès d'erreur résiduelle gagné à chaque pas de temps. Ce faisant, le modèle est devenu un modèle linéaire généralisé mixte (GLMM). Les résultats de cette approche ont encore une fois montré que la température annuelle moyenne est le meilleur prédicteur de la vitesse de propagation des souches PRN- aux États-Unis, et ce avec les deux méthodes utilisées pour calculer les valeurs des cofacteurs (Section 3.5.2). Cependant, une analyse de robustesse a montré que cette deuxième approche était aussi insatisfaisante que la première, l'ajout d'un facteur aléatoire ne suffisant pas à contrôler l'accumulation d'erreur résiduelle (Section 3.6.3). Finalement, l'approche par GLM comme l'approche par GLMM ont le même problème : le risque d'erreur de type I augmente fortement dès que le modèle présente une erreur résiduelle même minime, et il a donc fallu développer une nouvelle méthode pour éliminer ce problème.

Cette solution s'est présentée sous la forme d'un modèle bayésien. Cette approche permet de modéliser explicitement l'erreur résiduelle du modèle à chaque pas de temps. En calculant les valeurs des cofacteurs avec $k' = t - 1$, l'analyse des données montre qu'aucun des cofacteurs n'explique particulièrement bien la manière dont fluctue la vitesse de propagation des souches PRN- aux États-Unis (Section 3.7, Tableau 3.5). Ce résultat contredit partiellement ceux obtenus à partir des approches par GLM et par GLMM, mais cela s'explique facilement par les résultats de l'analyse de robustesse : il est vraisemblable que les analyses fréquentistes ont conduit à conclure à un effet de la température annuelle moyenne – et certains autres cofacteurs – à cause de la distorsion du risque d'erreur de type I. Néanmoins, en utilisant $k' = t$ dans le calcul des cofacteurs, l'analyse bayésienne montre cette fois que le modèle incluant la moyenne des températures annuelles est bien meilleur que le modèle nul pour expliquer les données (Section 3.7, Tableau 3.6), de manière similaire à ce que les approches par GLM et GLMM montraient (Sections 3.5.1 et 3.5.2).

Il est vraisemblable que tous les facteurs inclus dans les analyses n'aient pas un effet réel sur la vitesse d'invasion des souches PRN-. Inversement, il est probable que d'autres cofacteurs non inclus dans les analyses aient pu avoir un impact sur la vitesse de propagation des souches

PRN-. Cependant, il est également possible que certains cofacteurs inclus dans les analyses aient réellement un effet sans que les méthodes utilisées parviennent à le capturer. Pour vérifier l'hypothèse que certains facteurs puissent avoir un effet sans que celui-ci soit détecté par le modèle bayésien utilisé, j'ai conduit deux séries d'analyses de puissance. Les résultats de ces analyses ont montré que le jeu de données utilisé était finalement assez limité pour répondre à la question générale de ce chapitre. Les deux problèmes majeurs sont que le nombre de pas de temps est finalement assez limité (un par an pendant onze ans) et que la taille de l'échantillon à certains pas de temps est trop faible. Ces deux facteurs expliquent en grande partie le manque de puissance statistique de l'approche bayésienne.

Une conclusion intéressante de ces analyses de puissance est qu'une manière simple d'augmenter la puissance de l'analyse bayésienne est simplement d'augmenter la fréquence d'échantillonnage dans le jeu de données. Ainsi, si le pas de temps renseigné dans le jeu de données avait été de six mois plutôt qu'un an tout en gardant le même nombre total de souches isolées par an, la puissance statistique aurait largement augmenté, alors même que le nombre de souches isolées par pas de temps est en moyenne réduit de moitié. Un autre avantage dans la réduction du pas de temps et que cela aurait permis d'étudier des fluctuations saisonnières dans la vitesse d'invasion des souches PRN-. Enfin, cela aurait permis de calculer des valeurs des cofacteurs plus précises – dans l'hypothèse où les valeurs brutes des cofacteurs sont disponibles à une échelle sub-annuelle.

La qualité des estimations des cofacteurs peut également être questionnée. Certains cofacteurs inclus dans les analyses sont estimés avec un pas de temps différent de celui des données sur les souches de *B. pertussis*. Par exemple, j'ai dû considérer la couverture vaccinale constante dans le temps par manque d'informations sur les changements temporels de celle-ci. De plus, les cofacteurs sont considérés homogènes au sein d'un état au cours d'un pas de temps, tandis qu'en réalité ils peuvent grandement fluctuer au sein d'un état (*e.g.*, la densité de population varie fortement entre les zones urbaines et rurales) et au cours de l'année (*e.g.*, la température varie fortement au cours de l'année). Autrement dit, la discrétisation spatiale et temporelle des données appauvrit la capacité des cofacteurs à expliquer la manière dont se propagent les souches PRN-. Comme noté plus haut, cet aspect aurait pu être (partiellement) mieux contrôlé avec un pas de temps plus faible. Enfin, à cause de la très forte hétérogénéité spatiale et temporelle dans l'effort d'échantillonnage des souches de *B. pertussis*, les cofacteurs transformés ont été estimés

de manière peu satisfaisante.

La méthode développée dans ce chapitre a été construite dans le but précis d'étudier le jeu de données publié par [Lefrancq et al. \(2022\)](#), et plus précisément dans le but d'étudier l'augmentation de fréquence du phénotype PRN⁻ aux États-Unis. Cependant, elle pourrait être utilisée pour comprendre d'autres émergences documentées dans le jeu de données, ce qui n'a pas pu être fait ici par manque de temps et de puissance statistique ; et plus généralement elle est applicable à des jeux de données et des contextes bien plus variés, même en dehors de l'épidémiologie. Par exemple, elle pourrait être employée pour des projets de recherche s'intéressant à la propagation d'espèces invasives remplaçant des espèces indigènes.

3.10 – Conclusion

Originellement pensé comme une « simple » analyse statistique, ce deuxième axe de thèse est finalement devenu un travail essentiellement méthodologique questionnant les approches statistiques disponibles pour étudier des phénomènes autocorrélés dans le temps, ici l'émergence d'une nouvelle souche d'un pathogène. Les résultats présentés dans ce chapitre, principalement les analyses de robustesse et de puissance, sont ainsi résumés dans un article *in prep* essentiellement méthodologique destiné à la revue *Methods in Ecology and Evolution*. Finalement, c'est là la principale force de ce deuxième axe : critiquer l'utilisation d'approches classiques de type GLM/GLMM pour modéliser des phénomènes autocorrélés et présenter une nouvelle méthode plus robuste. En effet, si la méthode développée est statistiquement adéquate pour étudier les facteurs environnementaux influençant la vitesse d'invasion d'une nouvelle souche de *B. pertussis*, j'ai montré que les résultats étaient plutôt mitigés, et qu'il était difficile de tirer des conclusions fortes notamment à cause de la structure et de la taille du jeu de données.

Dans le prochain et dernier chapitre de cette thèse, j'approfondis l'étude de l'approche bayésienne développée ici, dans un premier temps en produisant un simulateur capable de générer des données mimant la dynamique de la coqueluche aux États-Unis, et dans un deuxième temps et en analysant ces données par cette méthode bayésienne. Cette dernière étude a pour but principal de mettre en évidence les conditions dans lesquelles ladite méthode d'analyse bayésienne peut être mise en place de manière satisfaisante.



BIBLIOGRAPHIE

- ADVANI A., GUSTAFSSON L., ÅHRÉN C., MOOI F. R. et HALLANDER H. O. (2011) : « Appearance of Fim3 and *ptxP3-Bordetella pertussis* strains, in two regions of Sweden with different vaccination programs » *Vaccine*, vol. 29, n°18, p./id. 3438–3442.
- AKABA H., JINDAI K., SHOBUGAWA Y. et OSHITANI H. (2022) : « Incidence and length of outbreak period of COVID-19 and population density in comparison with seasonal influenza in Japan » *IJID Regions*, vol. 2, p./id. 154–157.
- BERGER L., SPEARE R., HINES H. B., MARANTELLI G., HYATT A., McDONALD K. R., SKERRATT L. F., OLSEN V., CLARKE J. M., GILLESPIE G., MAHONY M., SHEPPARD N., WILLIAMS C. et TYLER M. J. (2004) : « Effect of season and temperature on mortality in amphibians due to chytridiomycosis » *Australian Veterinary Journal*, vol. 82, n°7, p./id. 434–439.
- BOUCHEZ V., GUILLOT S., LANDIER A., ARMATYS N., MATCZAK S., THE FRENCH PERTUSSIS MICROBIOLOGY STUDY GROUP, TOUBIANA J. et BRISSE S. (2021) : « Evolution of *Bordetella pertussis* over a 23-year period in France, 1996 to 2018 » *Eurosurveillance*, vol. 26, n°37, p./id. 2001213.
- BROOKS S. P. et GELMAN A. (1998) : « General Methods for Monitoring Convergence of Iterative Simulations » *Journal of Computational and Graphical Statistics*, vol. 7, n°4, p./id. 434–455.
- CHAKRABORTY C., SHARMA A. R., BHATTACHARYA M. et LEE S.-S. (2022) : « A Detailed Overview of Immune Escape, Antibody Escape, Partial Vaccine Escape of SARS-CoV-2 and Their Emerging Variants With Escape Mutations » *Frontiers in Immunology*, vol. 13, p./id. 801522.
- CIOTTI M., CICOZZI M., PIERI M. et BERNARDINI S. (2022) : « The COVID-19 pandemic: viral variants and vaccine efficacy » *Critical Reviews in Clinical Laboratory Sciences*, vol. 59, n°1, p./id. 66–75.
- CRAIG R., KUNKEL E., CROWCROFT N. S., FITZPATRICK M. C., DE MELKER H., ALTHOUSE B. M., MERKEL T., SCARPINO S. V., KOELLE K., FRIEDMAN L., ARNOLD C. et BOLOTIN S. (2020) : « Asymptomatic Infection and Transmission of Pertussis in Households: A Systematic Review » *Clinical Infectious Diseases*, vol. 70, n°1, p./id. 152–161.
- DAKIC G., KALLONEN T., ELOMAA A., PLJESA T., VIGNJEVIC-KRASTAVCEVIC M. et HE Q. (2010) : « *Bordetella pertussis* vaccine strains and circulating isolates in Serbia » *Vaccine*, vol. 28, n°5, p./id. 1188–1192.
- DE JONG J. G. (1965) : « The Survival of Measles Virus in Air, in Relation to the Epidemiology of Measles » *Archiv für die gesamte Virusforschung*, vol. 16, n°1, p./id. 97–102.
- DEEN J. L., MINK C. M., CHERRY J. D., CHRISTENSON P. D., PINEDA E. F., LEWIS K., BLUMBERG D. A. et ROSS L. A. (1995) : « Household Contact Study of *Bordetella pertussis* Infections » *Clinical Infectious Diseases*, vol. 21, n°5, p./id. 1211–1219.
- DIDELOT X., WALKER A. S., PETO T. E., CROOK D. W. et WILSON D. J. (2016) : « Within-host evolution of bacterial pathogens » *Nature Reviews Microbiology*, vol. 14, n°3, p./id. 150–162.

-
- DOS SANTOS M. I. M. A., PACHECO S. R., STOCKER A., SCHINONI M. I., PARANÁ R., REIS M. G. et SILVA L. K. (2017) : « Mutations associated with drug resistance and prevalence of vaccine escape mutations in patients with chronic hepatitis B infection » *Journal of Medical Virology*, vol. 89, n°10, p./id. 1811–1816.
- ESSL F., DULLINGER S., MOSER D., STEINBAUER K. et MANG T. (2015) : « Macroecology of global bryophyte invasions at different invasion stages » *Ecography*, vol. 38, n°5, p./id. 488–498.
- FOULLEY J. L. et GIANOLA D. (1996) : « Statistical analysis of ordered categorical data via a structural heteroskedastic threshold model » *Genetics Selection Evolution*, vol. 28, n°3, p./id. 249.
- FOX G. J., NHUNG N. V., SY D. N., HOA N. L. P., ANH L. T. N., ANH N. T., HOA N. B., DUNG N. H., BUU T. N., LOI N. T., NHUNG L. T., HUNG N. V., LIEU P. T., CUONG N. K., CUONG P. D., BESTRASHNIY J., BRITTON W. J. et MARKS G. B. (2018) : « Household-Contact Investigation for Detection of Tuberculosis in Vietnam » *New England Journal of Medicine*, vol. 378, n°3, p./id. 221–229.
- GANDON S. et DAY T. (2008) : « Evidences of parasite evolution after vaccination » *Vaccine*, vol. 26, n°s3, p./id. C4–C7.
- GANSLMEIER M., FURCERI D. et OSTRY J. D. (2021) : « The impact of weather on COVID-19 pandemic » *Scientific Reports*, vol. 11, n°1, p./id. 22027.
- GELMAN A., CARLIN J. B., STERN H. S. et RUBIN D. B. (2004) : *Bayesian Data Analysis* 2^{ème} édition, Chapman & Hall/CRC (Boca Raton, États-Unis).
- GIMENO I. M. (2008) : « Marek's disease vaccines: A solution for today but a worry for tomorrow? » *Vaccine*, vol. 26, n°s3, p./id. C31–C41.
- GUISSO N., SOUBEYRAND B. et MACINA D. (2022) : « Can vaccines control bacterial virulence and pathogenicity? *Bordetella pertussis*: the advantage of fitness over virulence » *Evolution, Medicine, and Public Health*, vol. 10, n°1, p./id. 363–370.
- GUWATUDDE D. (2003) : « Tuberculosis in Household Contacts of Infectious Cases in Kampala, Uganda » *American Journal of Epidemiology*, vol. 158, n°9, p./id. 887–898.
- HALL R. J. (2011) : « Eating the competition speeds up invasions » *Biology Letters*, vol. 7, n°2, p./id. 307–311.
- HE D., DUSHOFF J., EFTIMIE R. et EARN D. J. D. (2013) : « Patterns of spread of influenza A in Canada » *Proceedings of the Royal Society B: Biological Sciences*, vol. 280, n°1770, p./id. 20131174.
- HIRAMATSU Y., MIYAJI Y., OTSUKA N., ARAKAWA Y., SHIBAYAMA K. et KAMACHI K. (2017) : « Significant Decrease in Pertactin-Deficient *Bordetella pertussis* Isolates, Japan » *Emerging Infectious Diseases*, vol. 23, n°4, p./id. 699–701.
- JIA Y., XU Q., ZHU Y., LI C., QI C., SHE K., LIU T., ZHANG Y. et LI X. (2023) : « Estimation of the relationship between meteorological factors and measles using spatiotemporal Bayesian model in Shandong Province, China » *BMC Public Health*, vol. 23, p./id. 1422.
- JOHNSON M. M., JONES C. E. et CLARK D. N. (2023) : « The Effect of Treatment-Associated Mutations on HIV Replication and Transmission Cycles » *Viruses*, vol. 15, n°1, p./id. 107.
- KELLNER K. et MEREDITH M. (2024), « jagsUI: A Wrapper Around 'rjags' to Streamline 'JAGS' Analyses ».
- KLEIN N. P. (2014) : « Licensed pertussis vaccines in the United States: History and current state » *Human Vaccines & Immunotherapeutics*, vol. 10, n°9, p./id. 2684–2690.

- KOWALZIK F., BARBOSA A. P., FERNANDES V. R., CARVALHO P. R., AVILA-AGUERO M. L., GOH D. Y. T., GOH A., DE MIGUEL M. J. G., MORAGA F., ROCA J., CAMPINS M., HUANG L. M., QUIAN J., RILEY N., BECK D. et VERSTRAETEN T. (2007) : « Prospective Multinational Study of Pertussis Infection in Hospitalized Infants and Their Household Contacts » *Pediatric Infectious Disease Journal*, vol. 26, n°3, p./id. 238–242.
- KRIGER K. M. et HERO J.-M. (2007) : « Large-scale seasonal variation in the prevalence and severity of chytridiomycosis » *Journal of Zoology*, vol. 271, n°3, p./id. 352–359.
- LECORVAISIER F., PONTIER D., SOUBEYRAND B. et FOUCHET D. (2024) : « Using a dynamical model to study the impact of a toxoid vaccine on the evolution of a bacterium: The example of diphtheria » *Ecological Modelling*, vol. 487, p./id. 110569.
- LEFRANCQ N., BOUCHEZ V., FERNANDES N., BARKOFF A.-M., BOSCH T., DALBY T., ÅKERLUND T., DARENBERG J., FABIANOVA K., VESTRHEIM D. F., FRY N. K., GONZÁLEZ-LÓPEZ J. J., GULLSBY K., HABINGTON A., HE Q., LITT D., MARTINI H., PIÉRARD D., STEFANELLI P., STEGGER M., ZAVADILOVA J., ARMATYS N., LANDIER A., GUILLOT S., HONG S. L., LEMEY P., PARKHILL J., TOUBIANA J., CAUCHEMEZ S., SALJE H. et BRISSE S. (2022) : « Global spatial dynamics and vaccine-induced fitness changes of *Bordetella pertussis* » *Science Translational Medicine*, vol. 14, n°642, p./id. eabn3253.
- LEITE D., CAMARGO C. H., KASHINO S. S., POLATTO R., MARTINS L. M., PEREIRA J. C., PAWLOSKI L., TONDELLA M. L., DE OLIVEIRA R. S. et DE ANDRADE VAZ DE LIMA L. R. (2021) : « Prevalence and characterization of pertactin deficient *Bordetella pertussis* strains in Brazil, a whole-cell vaccine country » *Vaccine: X*, vol. 8, p./id. 100103.
- LINK W. A. et EATON M. J. (2012) : « On thinning of chains in MCMC » *Methods in Ecology and Evolution*, vol. 3, n°1, p./id. 112–115.
- LOBUE P. (2009) : « Extensively drug-resistant tuberculosis: » *Current Opinion in Infectious Diseases*, vol. 22, n°2, p./id. 167–173.
- LOWEN A. C., MUBAREKA S., STEEL J. et PALESE P. (2007) : « Influenza Virus Transmission Is Dependent on Relative Humidity and Temperature » *PLOS Pathogens*, vol. 3, n°10, p./id. e151.
- LYTHGOE K. A., HALL M., FERRETTI L., DE CESARE M., MACINTYRE-COCKETT G., TREBES A., ANDERSSON M., OTECKO N., WISE E. L., MOORE N., LYNCH J., KIDD S., CORTES N., MORI M., WILLIAMS R., VERNET G., JUSTICE A., GREEN A., NICHOLLS S. M., ANSARI M. A., ABELER-DÖRNER L., MOORE C. E., PETO T. E. A., EYRE D. W., SHAW R., SIMMONDS P., BUCK D., TODD J. A., CONNOR T. R., ASHRAF S., DA SILVA FILIPE A., SHEPHERD J., THOMSON E. C., THE COVID-19 GENOMICS UK (COG-UK) CONSORTIUM, BONSALE D., FRASER C. et GOLUBCHIK T. (2021) : « SARS-CoV-2 within-host diversity and transmission » *Science*, vol. 372, n°6539, p./id. eabg0821.
- MA L., CAULFIELD A., DEWAN K. K. et HARVILL E. T. (2021) : « Pertactin-Deficient *Bordetella pertussis*, Vaccine-Driven Evolution, and Reemergence of Pertussis » *Emerging Infectious Diseases*, vol. 27, n°6, p./id. 1561–1566.
- MACK R. N., SIMBERLOFF D., MARK LONSDALE W., EVANS H., CLOUT M. et BAZZAZ F. A. (2000) : « BIOTIC INVASIONS: CAUSES, EPIDEMIOLOGY, GLOBAL CONSEQUENCES, AND CONTROL » *Ecological Applications*, vol. 10, n°3, p./id. 689–710.
- MARUFU L. T., BARSON M., PHIRI C. et NHIWATIWA T. (2020) : « Spatial and temporal distribution of an invasive crayfish (*Cherax quadricarinatus*) in Lake Kariba, Zimbabwe » *Lakes & Reservoirs: Science, Policy and Management for Sustainable Use*, vol. 25, n°4, p./id. 394–402.
- MEIJERING A. et GIANOLA D. (1985) : « Observations on Sire Evaluation with Categorical Data Using Heteroscedastic Mixed Linear Models » *Journal of Dairy Science*, vol. 68, n°5, p./id. 1226–1232.

-
- MICHAEL C. A., DOMINEY-HOWES D. et LABBATE M. (2014) : « The Antimicrobial Resistance Crisis: Causes, Consequences, and Management » *Frontiers in Public Health*, vol. 2.
- MÖLLER J., KRANER M., SONNEWALD U., SANGAL V., TITTLBACH H., WINKLER J., WINKLER T. H., MELNIKOV V., LANG R., SING A., MATTOS-GUARALDI A. L. et BURKOVSKI A. (2019) : « Proteomics of diphtheria toxoid vaccines reveals multiple proteins that are immunogenic and may contribute to protection of humans against *Corynebacterium diphtheriae* » *Vaccine*, vol. 37, n°23, p./id. 3061–3070.
- MOSSONG J., HENS N., JIT M., BEUTELS P., AURANEN K., MIKOLAJCZYK R., MASSARI M., SALMASO S., SCALIA TOMBA G., WALLINGA J., HEIJNE J., SADKOWSKA-TODYS M., ROSINSKA M. et EDMUNDS W. J. (2008) : « Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases » *PLOS Medicine*, vol. 5, n°3, p./id. e74.
- MURPHY J. R. (1996) : « *Corynebacterium Diphtheriae* » *Medical Microbiology* 4^{ème} édition, University of Texas Medical Branch at Galveston (Galveston, États-Unis).
- MYERS W. P., WESTENHOUSE J. L., FLOOD J. et RILEY L. W. (2006) : « An Ecological Study of Tuberculosis Transmission in California » *American Journal of Public Health*, vol. 96, n°4, p./id. 685–690.
- OLSON D. H., RONNENBERG K. L., GLIDDEN C. K., CHRISTIANSEN K. R. et BLAUSTEIN A. R. (2021) : « Global Patterns of the Fungal Pathogen *Batrachochytrium dendrobatidis* Support Conservation Urgency » *Frontiers in Veterinary Science*, vol. 8, p./id. 685877.
- PENG L., ZHAO X., TAO Y., MI S., HUANG J. et ZHANG Q. (2020) : « The effects of air pollution and meteorological factors on measles cases in Lanzhou, China » *Environmental Science and Pollution Research*, vol. 27, n°12, p./id. 13524–13533.
- PLOWE C. V. (2009) : « The evolution of drug-resistant malaria » *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 103, n°s1, p./id. S11–S14.
- PLUMMER M. (2003) : « JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling » *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, (Vienna, Austria).
- QIN S., DING Y., YAN R. et HE H. (2019) : « Measles in Zhejiang, China, 2004-2017: Population Density and Proportion of Floating Populations Effects on Measles Epidemic » *Health Security*, vol. 17, n°3, p./id. 193–199.
- R CORE TEAM (2024) : *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing (Vienne, Autriche).
- RING N., DAVIES H., MORGAN J., SUNDARESAN M., TIONG A., PRESTON A. et BAGBY S. (2022) : « Comparative genomics of *Bordetella pertussis* isolates from New Zealand, a country with an uncommonly high incidence of whooping cough » *Microbial Genomics*, vol. 8, n°1.
- RODRIGUES C., BOUCHEZ V., SOARES A., TROMBERT-PAOLANTONI S., AÏT EL BELGHITI F., COHEN J. F., ARMATYS N., LANDIER A., BLANCHOT T., HERVO M., REMICOQ STUDY GROUP, TOUBIANA J. et BRISSE S. (2024) : « Resurgence of *Bordetella pertussis*, including one macrolide-resistant isolate, France, 2024 » *Eurosurveillance*, vol. 29, n°31.
- ROHANI P., ZHONG X. et KING A. A. (2010) : « Contact Network Structure Explains the Changing Epidemiology of Pertussis » *Science*, vol. 330, n°6006, p./id. 982–985.
- ROUSSEL M., PONTIER D., COHEN J.-M., LINA B. et FOUCHET D. (2016) : « Quantifying the role of weather on seasonal influenza » *BMC Public Health*, vol. 16, n°1, p./id. 441.

- SCANLON K., SKERRY C. et CARBONETTI N. (2019a) : « Association of Pertussis Toxin with Severe Pertussis Disease » *Toxins*, vol. 11, n°7, p./id. 373.
- SCANLON K., SKERRY C. et CARBONETTI N. (2019b) : « Role of Major Toxin Virulence Factors in Pertussis Infection and Disease Pathogenesis » *Pertussis Infection and Vaccines* volume 1183, p. 35–51, Springer (Cham, Suisse).
- SCHMIDTKE A. J., BONEY K. O., MARTIN S. W., SKOFF T. H., TONDELLA M. L. et TATTI K. M. (2012) : « Population Diversity among *Bordetella pertussis* Isolates, United States, 1935–2009 » *Emerging Infectious Diseases*, vol. 18, n°8.
- SHAH K., SAXENA D. et MAVALANKAR D. (2020) : « Secondary attack rate of COVID-19 in household contacts: a systematic review » *QJM: An International Journal of Medicine*, vol. 113, n°12, p./id. 841–850.
- SINHA S. K. (2004) : « Robust Analysis of Generalized Linear Mixed Models » *Journal of the American Statistical Association*, vol. 99, n°466, p./id. 451–460.
- SKOG L., LINDE A., PALMGREN H., HAUSKA H. et ELGH F. (2014) : « Spatiotemporal characteristics of pandemic influenza » *BMC Infectious Diseases*, vol. 14, n°1, p./id. 378.
- SOORYANARAIN H. et ELANKUMARAN S. (2015) : « Environmental Role in Influenza Virus Outbreaks » *Annual Review of Animal Biosciences*, vol. 3, p./id. 347–373.
- SPIEGELHALTER D. J., BEST N. G., CARLIN B. P. et VAN DER LINDE A. (2002) : « Bayesian Measures of Model Complexity and Fit » *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 64, n°4, p./id. 583–639.
- TAUBENBERGER J. K. et MORENS D. M. (2010) : « Influenza: The Once and Future Pandemic » *Public Health Reports*, vol. 125, n°s3, p./id. 15–26.
- TEMPELMAN R. J. (1998) : « Generalized Linear Mixed Models in Dairy Cattle Breeding » *Journal of Dairy Science*, vol. 81, n°5, p./id. 1428–1444.
- THIELE J. et MARKUSSEN B. (2012) : « Potential of GLMM in modelling invasive spread » *CABI Reviews*, p. 1–10.
- THOMAS A. (2006) : « The BUGS Language » *R News*, vol. 6, n°1, p./id. 17–21.
- TRUCCHI E., GRATTON P., MAFESSONI F., MOTTA S., CICONARDI F., MANCIA F., BERTORELLE G., D'ANNESSA I. et DI MARINO D. (2021) : « Population Dynamics and Structural Effects at Short and Long Range Support the Hypothesis of the Selective Advantage of the G614 SARS-CoV-2 Spike Variant » *Molecular Biology and Evolution*, vol. 38, n°5, p./id. 1966–1979.
- VON KÖNIG C. H. W., POSTELS-MULTAN S., BOCK H. L. et SCHMITT H. J. (1995) : « Pertussis in adults: frequency of transmission after household exposure » *The Lancet*, vol. 346, n°8986, p./id. 1326–1329.
- WAGNER K. S., WHITE J. M., NEAL S., CROWCROFT N. S., KUPREVIČIENE N., PABERZA R., LUCENKO I., JÖKS U., AKBAŞ E., ALEXANDROU-ATHANASSOULIS H., DETCHEVA A., VUOPIO J., VON HUNOLSTEIN C., MURPHY P. G. et ANDREWS N. (2011) : « Screening for *Corynebacterium diphtheriae* and *Corynebacterium ulcerans* in patients with upper respiratory tract infections 2007–2008: a multicentre European study » *Clinical Microbiology and Infection*, vol. 17, n°4, p./id. 519–525.
- WANG D., LEI H., WANG D., SHU Y. et XIAO S. (2023) : « Association between Temperature and Influenza Activity across Different Regions of China during 2010–2017 » *Viruses*, vol. 15, n°3, p./id. 594.

-
- WANG J., QIU J., ZHU Y., ZHOU H., YU L., DING Y., ZHANG L., GUO Z. et DONG C. (2017) : « Molecular evolution of hepatitis B vaccine escape variants in China, during 2000–2016 » *Vaccine*, vol. 35, n°43, p./id. 5808–5813.
- WANG R., CHEN J., GAO K. et WEI G.-W. (2021) : « Vaccine-escape and fast-growing mutations in the United Kingdom, the United States, Singapore, Spain, India, and other COVID-19-devastated countries » *Genomics*, vol. 113, n°4, p./id. 2158–2170.
- WONG D. W. S. et LI Y. (2020) : « Spreading of COVID-19: Density matters » *PLOS ONE*, vol. 15, n°12, p./id. e0242398.
- XU Y., ZHANG L., TAN Y., WANG L., ZHANG S. et WANG J. (2015) : « Genetic diversity and population dynamics of *Bordetella pertussis* in China between 1950–2007 » *Vaccine*, vol. 33, n°46, p./id. 6327–6331.
- YANG Q., FU C., WANG N., DONG Z., HU W. et WANG M. (2014) : « The effects of weather conditions on measles incidence in Guangzhou, Southern China » *Human Vaccines & Immunotherapeutics*, vol. 10, n°4, p./id. 1104–1110.
- YAU K. K. W. et KUK A. Y. C. (2002) : « Robust Estimation in Generalized Linear Mixed Models » *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 64, n°1, p./id. 101–117.
- YEUNG K. H. T., DUCLOS P., NELSON E. A. S. et HUTUBESSY R. C. W. (2017) : « An update of the global burden of pertussis in children younger than 5 years: a modelling study » *The Lancet Infectious Diseases*, vol. 17, n°9, p./id. 974–980.
- YOSHIKURA H. (2012) : « Negative Impacts of Large Population Size and High Population Density on the Progress of Measles Elimination » *Japanese Journal of Infectious Diseases*, vol. 65, n°5, p./id. 450–454.
- ZHANG L., XU Y., ZHAO J., KALLONEN T., CUI S., XU Y., HOU Q., LI F., WANG J., HE Q. et ZHANG S. (2010) : « Effect of Vaccination on *Bordetella pertussis* Strains, China » *Emerging Infectious Diseases*, vol. 16, n°11, p./id. 1695–1701.
- ZOMER A., OTSUKA N., HIRAMATSU Y., KAMACHI K., NISHIMURA N., OZAKI T., POOLMAN J. et GEURTSEN J. (2018) : « *Bordetella pertussis* population dynamics and phylogeny in Japan after adoption of acellular pertussis vaccines » *Microbial Genomics*, vol. 4, n°5, p./id. e000180.

CHAPITRE IV

MODÉLISATION INFORMATIQUE : PROPAGATION ET ÉCHANTILLONNAGE DE SOUCHES MUTANTES

4.1 – Introduction

4.1.1 – Des données *in natura* aux données *in silico*

Dans le chapitre précédent, j’ai travaillé essentiellement sur un jeu de données issu de la collecte, du stockage, du séquençage et du traitement de centaines de souches de *Bordetella pertussis* collectées aux États-Unis entre 2007 et 2017. Ce jeu de données ne représentait qu’un fragment d’un ensemble plus grand de plus de 3 000 souches collectées entre 1935 et 2019 dans une vingtaine de pays (Lefrancq *et al.*, 2022). Ces données sont intéressantes à plusieurs égards. Elles sont issues de la collaboration d’une trentaine de chercheurs et chercheuses de nationalités très diverses et illustrent ainsi les possibilités qu’offrent les collaborations scientifiques au XXI^e siècle.

Le jeu de données était également intéressant par sa complexité : chaque souche incluse a été annotée avec son année de collecte, son pays et parfois sa région d’origine, et de multiples caractéristiques génétiques et phénotypiques. Grâce à cela, les scientifiques à l’origine de sa publication ont pu mettre en évidence le lien entre la vaccination acellulaire et l’émergence de nouveaux génotypes de *B. pertussis*, notamment l’émergence de souches déficientes en pertactine (PRN-) (Lefrancq *et al.*, 2022).

Néanmoins, le jeu de données s’est montré limité sur certains aspects. Dans le chapitre précé-

dent, j'ai montré que la discrétisation de la date de collecte en années et le nombre parfois réduit de souches collectées pouvait réduire la puissance des analyses menées sur ce jeu de données. Dans ce chapitre, j'ai développé un modèle informatique et mathématique capable de simuler une épidémie de deux souches d'un pathogène dont les caractéristiques se rapprochent de celle de *B. pertussis* au sein d'une population ressemblant à celle des États-Unis et d'échantillonner des individus infectés au sein de cette population. Plus précisément, ce modèle permet d'introduire une perturbation, sous la forme de la vaccination, dans une population au sein de laquelle le pathogène circule déjà. Finalement, ce modèle permet de produire une grande quantité de jeux de données artificiels (*in silico*) dont la structure est similaire à celui publié par [Lefrancq et al. \(2022\)](#). À partir de ces jeux de données, j'ai mené des analyses pour continuer les travaux présentés dans le chapitre précédent et étudier plus en détail les limites de la méthode d'analyse bayésienne que j'ai développée.

4.1.2 – Utilisation des données *in silico* en sciences de la vie

Le terme « *in silico* » a été forgé en référence au silicium, élément essentiel des composants des ordinateurs. Il est utilisé pour désigner les travaux utilisant la puissance de calcul des ordinateurs, et notamment en sciences de la vie, pour répondre à des questions trop complexes ou demandant trop de temps de travail par rapport aux capacités humaines. L'expression est vraisemblablement créée à la fin des années 1980 mais c'est dans les années 1990 que son usage se développe avec la démocratisation des ordinateurs.

La « préhistoire » des modèles *in silico* commence au lendemain de la Seconde Guerre mondiale. À cette époque, le célèbre mathématicien hongrois John von Neumann (1903 – 1957) se lance dans un projet d'envergure : voyant se profiler l'avènement de l'informatique, il propose d'utiliser les mathématiques pour modéliser le vivant. Sa mort prématurée mettra fin à cette ambition, mais il aura eu le mérite d'ouvrir une porte à travers laquelle d'autres chercheurs vont s'engouffrer. En 1970, le magazine de vulgarisation *Scientific American* publie un jeu proposé par le mathématicien américain John Horton Conway (1937 – 2020). Il consiste en une grille composée de cellules qui sont soit occupées, soit vides, et dont un ensemble de règles conditionne la manière dont varie l'occupation de celles-ci ([Gardner, 1970](#)). Conway appelle son jeu « life »

par analogie avec les systèmes vivants : les cellules vivent, se reproduisent, meurent. Avec la démocratisation des ordinateurs, le jeu de Conway va se populariser et se développer, permettant de créer des systèmes de plus en plus complexes, de plus en plus vivants : l'approche *in silico* est née.

À la fin du siècle dernier, les chercheurs commencent à utiliser la puissance de calcul à leur disposition pour simuler des systèmes biologiques de plus en plus complexes ou étudier des bases de données génomiques de plus en plus grandes. Cette approche permet de répondre à une variété toujours grandissante de questions : elle est par exemple utilisée dans le développement de nouveaux médicaments (Terstappen et Reggiani, 2001), pour identifier des épitopes cibles dans la recherche de nouveaux vaccins (Dash *et al.*, 2017 ; Maleki *et al.*, 2022) ou pour comprendre comment se développe le système vasculaire (Peirce *et al.*, 2004).

Les sciences de la biodiversité, écologie et biologie de l'évolution, font également un usage important des méthodes *in silico*. L'intérêt principal de celles-ci est de permettre de modéliser sur un temps court (quelques heures ou journées de calcul) des phénomènes qui peuvent prendre place sur plusieurs millions d'années. Par exemple, Stauffer *et al.* (2005) ont produit un modèle de simulation de réseau trophique montrant le lien entre prédation, compétition et évolution, qui serait difficilement observable *in natura* à cause de l'échelle temporelle à laquelle s'observent les phénomènes évolutifs. De la même manière, Würtz et Amila (2010) ont adopté une approche par simulations pour montrer le lien entre la thermodynamique et la théorie de la succession écologique. Dans le domaine de la génomique évolutive, l'utilisation d'outils informatique permet le stockage et le traitement d'énormes quantités de données améliorant nos connaissances sur la manière dont ont évolué les espèces jusqu'à aujourd'hui, que ce soit au niveau de la temporalité des événements ou des forces entrant en jeu (Rudman *et al.*, 2017).

L'épidémiologie fait usage des approches *in silico* pour répondre à une grande diversité de questions de première importance. Par exemple, certaines études prédictives ont pour but de réduire l'impact que peuvent avoir les pathogènes sur les sociétés humaines (Shatnawi *et al.*, 2013). Les approches de modélisation mathématique classiques, déjà discutées dans la Section 1.5 et dans le Chapitre 2, peuvent être utilisées pour répondre à de nombreuses questions sur le devenir des épidémies ou sur l'évolution des pathogènes, mais des approches plus empiriques de simulation peuvent également y contribuer. De plus, les approches par simulations sont plus

appropriées pour prendre en considération de fortes hétérogénéités entre les individus, notamment au niveau des taux de contacts et donc de transmission. Par exemple, [Alam *et al.* \(2009\)](#) ont modélisé un village d'un millier d'habitants environ dans lequel ils ont simulé la propagation du VIH pour comprendre la manière dont le comportement sexuel des habitants l'influçait. [Chao *et al.* \(2010\)](#) ont développé un outil en ligne¹ implémentant un modèle de transmission de la grippe permettant notamment d'évaluer comment des interventions comme l'instauration de la vaccination peuvent perturber la propagation des épidémies.

4.2 – Modèle épidémique

4.2.1 – Description du modèle

Les épidémies étudiées dans ce chapitre sont simulées grâce un modèle de Markov stochastique où les individus sont répartis en quatre compartiments : susceptibles (S), infectés par la souche mutante (I_1), infectés par la souche résidente (I_2) et guéris (R).

À chaque pas de temps, les individus peuvent ou non changer de compartiment, et les probabilités de changer de compartiment sont détaillées plus loin. Un individu susceptible peut soit être infecté par l'une des souches, soit rester susceptible, soit mourir. La mortalité dans le modèle contribue au renouvellement de la population de susceptibles, et on considère que chaque individu mort est instantanément remplacé par un nouvel individu susceptible. De cette manière, la taille de la population est constante entre chaque pas de temps.

La durée de la période infectieuse est donnée comme étant de 18,25 jours, et après chaque pas de temps les individus infectés changent obligatoirement de compartiment. Cette valeur permet de discrétiser chaque année en exactement 20 pas de temps (le modèle néglige les années bissextiles) mais est également crédible biologiquement : [Teepe *et al.* \(2015\)](#) rapportent une durée médiane de toux de 17 jours pour les patients infectés par la coqueluche, valeur réutilisée comme durée de la période infectieuse dans le modèle de [Bernhauerová \(2016\)](#). [Wearing et Rohani \(2009\)](#) utilisent une période infectieuse plus courte dans leur modèle : 15 jours ; à l'inverse, [Hethcote \(1997\)](#) utilise une période infectieuse plus longue de 21 jours dans son modèle. Finalement, la

1. Désormais inaccessible (11/09/2024).

durée de la période infectieuse est assez mal connue, d'autant plus que la période symptomatique et la période infectieuse ne sont pas totalement synchrones (Edwards et Decker, 2012), mais la durée choisie de 18,25 jours apparaît crédible.

Les individus quittent les compartiments infectés soit en mourant, et donc retournent dans le compartiment susceptible, soit en guérissant, et donc acquièrent une immunité et atteignent le compartiment guéri.

Les individus guéris peuvent soit rester dans leur compartiment, soit retourner dans le compartiment susceptible. Pour cela, ils peuvent soit perdre leur immunité, soit mourir. Dans le modèle, les deux évènements sont équivalents.

4.2.2 – Matrice de transition

L'élément central de ce modèle est la matrice de transition $\mathbf{T}(t,i)$. Celle-ci représente les probabilités pour un individu de l'état i de passer d'un compartiment à un autre à chaque pas de temps. Elle vaut

$$\mathbf{T}(t,i) = \begin{pmatrix} p_{S,S} & p_{S,I_1} & p_{S,I_2} & p_{S,R} \\ p_{I_1,S} & p_{I_1,I_1} & p_{I_1,I_2} & p_{I_1,R} \\ p_{I_2,S} & p_{I_2,I_1} & p_{I_2,I_2} & p_{I_2,R} \\ p_{R,S} & p_{R,I_1} & p_{R,I_2} & p_{R,R} \end{pmatrix} (t,i) \quad (4.1)$$

où $p_{k,l}$ indique la probabilité de passer de l'état l à l'état k . Les calculs des différentes probabilités sont détaillés plus bas et rapportées dans le Tableau 4.1.

La probabilité pour un individu de mourir d'un pas de temps au suivant est $1 - e^{-d}$, d correspondant ainsi au taux de mortalité; réciproquement sa probabilité de survivre est e^{-d} . La probabilité pour un individu susceptible d'être infecté par n'importe quelle souche est $1 - e^{-(\lambda_1(t,i) + \lambda_2(t,i))}$, $\lambda_s(t,i)$ étant la force d'infection de la souche s , détaillée plus loin. Un individu infecté par une souche s peut soit transmettre la souche qui l'a infecté, soit une souche mutante. La probabilité de transmettre la souche s est $e^{-\mu_s}$ et la probabilité de transmettre une souche mutante est $1 - e^{-\mu_s}$. Enfin, la probabilité pour un individu guéri de perdre son immunité est $1 - e^{-\eta}$, η correspondant au taux de perte de l'immunité; réciproquement sa probabilité de ne pas perdre son immunité est $e^{-\eta}$. À partir de ces valeurs de base, toutes les probabilités de transition

TABLE 4.1 – **Formules des différentes probabilités de transition.** La matrice des transitions est présentée dans la formule (4.1).

Transition	Probabilité
$S \rightarrow S$	$e^{-d} \cdot e^{-(\lambda_1(t,i)+\lambda_2(t,i))} + (1 - e^{-d})$
$S \rightarrow I_1$	$e^{-d} \cdot (1 - e^{-(\lambda_1(t,i)+\lambda_2(t,i))}) \cdot \frac{\lambda_1(t,i) \cdot e^{-\mu_1} + \lambda_2(t,i) \cdot (1 - e^{-\mu_2})}{\lambda_1(t,i) + \lambda_2(t,i)}$
$S \rightarrow I_2$	$e^{-d} \cdot (1 - e^{-(\lambda_1(t,i)+\lambda_2(t,i))}) \cdot \frac{\lambda_2(t,i) \cdot e^{-\mu_2} + \lambda_1(t,i) \cdot (1 - e^{-\mu_1})}{\lambda_1(t,i) + \lambda_2(t,i)}$
$S \rightarrow R$	0
$I_1 \rightarrow S$	$1 - e^{-d}$
$I_1 \rightarrow I_1$	0
$I_1 \rightarrow I_2$	0
$I_1 \rightarrow R$	e^{-d}
$I_2 \rightarrow S$	$1 - e^{-d}$
$I_2 \rightarrow I_1$	0
$I_2 \rightarrow I_2$	0
$I_2 \rightarrow R$	e^{-d}
$R \rightarrow S$	$e^{-d} \cdot (1 - e^{-\eta}) + (1 - e^{-d})$
$R \rightarrow I_1$	0
$R \rightarrow I_2$	0
$R \rightarrow R$	$e^{-d} \cdot e^{-\eta}$

sont construites.

Comme expliqué plus haut, les individus susceptibles le restent s'ils survivent et ne sont pas infectés, où s'ils meurent. La probabilité de rester dans l'état S est donc

$$p_{S,S} = e^{-d} \cdot e^{-(\lambda_1(t,i)+\lambda_2(t,i))} + (1 - e^{-d}). \quad (4.2)$$

Les individus susceptibles passent à un compartiment infectieux s'ils survivent et sont infectés. La probabilité d'infection globale dépend des forces d'infection des deux types de souche, et la probabilité de passer spécifiquement à l'état I_s dépend de la force d'infection relative de la souche s et des taux de mutation en jeu. Ainsi, la probabilité de passer de l'état S à l'état I_1 est

$$p_{I_1,S} = e^{-d} \cdot (1 - e^{-(\lambda_1(t,i)+\lambda_2(t,i))}) \cdot \frac{\lambda_1(t,i) \cdot e^{-\mu_1} + \lambda_2(t,i) \cdot (1 - e^{-\mu_2})}{\lambda_1(t,i) + \lambda_2(t,i)}, \quad (4.3)$$

et la probabilité de passer de l'état S à l'état I_2 est

$$p_{I_2,S} = e^{-d} \cdot (1 - e^{-(\lambda_1(t,i)+\lambda_2(t,i))}) \cdot \frac{\lambda_2(t,i) \cdot e^{-\mu_2} + \lambda_1(t,i) \cdot (1 - e^{-\mu_1})}{\lambda_1(t,i) + \lambda_2(t,i)}. \quad (4.4)$$

Les individus infectés redeviennent susceptibles s'ils meurent à la fin de leur infection, donc pour tout s on a

$$p_{S,I_s} = 1 - e^{-d}. \quad (4.5)$$

Sinon, s'ils survivent, les individus infectés par n'importe quelle souche s passent à l'état R, et on a donc

$$p_{R,I_s} = e^{-d}. \quad (4.6)$$

Les individus guéris redeviennent susceptibles s'ils survivent et perdent leur immunité, ou s'ils meurent. La probabilité de passer de l'état R à l'état S est donc

$$e^{-d} \cdot (1 - e^{-\eta}) + (1 - e^{-d}). \quad (4.7)$$

Sinon, les individus guéris peuvent le rester s'ils survivent et ne perdent par leur immunité. La probabilité de passer de rester dans l'état R est donc

$$e^{-d} \cdot e^{-\eta}. \quad (4.8)$$

Toutes les autres transitions décrites par la matrice (4.1) ont une probabilité nulle de se produire.

4.2.3 – Paramétrisation de la matrice de transition

La matrice de transition $\mathbf{T}(t,i)$ est construite à partir de sept paramètres : le taux de mortalité (d), le taux de perte de l'immunité (η), les taux de mutation de la souche 1 vers la souche 2 (μ_1) et de la souche 2 vers la souche 1 (μ_2), la matrice des contacts entre états (\mathbf{F}) et les taux de transmission des souches 1 et 2, conditionnés à l'instant t et à l'état i ($\beta_1(t,i)$ et $\beta_2(t,i)$, respectivement). Le paramètre d contrôle la probabilité de passer de n'importe quel compartiment au compartiment S, le paramètre η contrôle la probabilité de passer du compartiment R au compartiment S, et les paramètres μ_1 , μ_2 , \mathbf{F} , $\beta_1(t,i)$ et $\beta_2(t,i)$ contrôlent la probabilité de passer du compartiment S aux compartiments I_1 et I_2 .

Les valeurs de ces différents paramètres sont données dans les sections suivantes.

Taux de mortalité

L'espérance de vie (inverse du taux de mortalité d) a été fixée à 76,4 années dans le modèle, correspondant à l'espérance de vie moyenne aux États-Unis en 2021 (Arias *et al.*, 23). Les rapports du National Center for Health Statistics (NCHS) sont disponibles chaque année et apportent des informations sur l'espérance de vie pour différents groupes démographiques et géographiques. Il aurait donc été possible de faire varier le taux de mortalité en fonction du temps et de l'état, cependant cela aurait rajouté un niveau de complexité au paramétrage du modèle qui n'a pas été jugé intéressant.

Taux de perte de l'immunité

La durée de l'immunité (inverse du taux de perte de l'immunité η) acquise naturellement (*i.e.*, par infection) est aujourd'hui encore mal connue et débattue. Ainsi, Wearing et Rohani (2009) estiment qu'elle durerait plus de 30 ans tandis que Wendelboe *et al.* (2005) ont, à partir d'une revue de la littérature disponible sur ce sujet, conclu que l'immunité acquise naturellement durait entre sept et 20 ans, alors que l'immunité vaccinale, qu'elle soit induite par un WCV ou un ACV, durait entre quatre et 12 ans chez l'enfant. La durée de l'immunité induite par les ACV est souvent jugée plus basse que celle de la durée de l'immunité induite par les WCV, mais Domenech de Cellès *et al.* (2019) estiment que l'immunité induite par les ACV est toujours élevée cinq ans après une vaccination par ACV comme après une vaccination par WCV. De plus, Edelman *et al.* (2007) estiment que l'immunité apportée par les rappels vaccinaux dure également plus de cinq ans. En utilisant ces différentes sources, j'ai fixé la durée de l'immunité à 20 ans.

Taux de mutation

Dans le contexte de ce travail, une mutation désigne le changement de phénotype de la souche 1 ou de la souche 2 qui devient celui de l'autre souche. Dans le contexte de l'étude des souches productrices ou non productrices de la pertactine de *B. pertussis*, le taux de mutation μ_2 représente le taux d'acquisition d'une ou plusieurs mutations arrêtant la production de la PRN et conduisant au phénotype PRN⁻. à l'inverse, le taux de mutation μ_1 représente le taux d'acquisition d'une ou plusieurs mutations conférant la capacité à produire la PRN et conduisant

ainsi au phénotype PRN+. Ces taux s’expriment en mutation par pas de temps plutôt que par génération. Autrement dit, μ_s est le taux de passage de la souche s à l’autre souche au cours de la période infectieuse d’un individu.

Puisque le modèle considère que le pas de temps est la durée de l’infection, la mutation ne consiste pas à un passage du compartiment I_1 au compartiment I_2 , ou inversement. Ici, la mutation est considérée directement au moment de l’infection. Ainsi, une fraction des individus susceptibles deviendront infectés par la souche 1 en entrant en contact avec des individus infectés par la souche 2 et inversement.

Parvenir à donner une valeur ne serait-ce que se rapprochant de la réalité aux paramètres μ_1 et μ_2 est difficile. Pour le paramètre μ_1 , je considère que l’acquisition du phénotype PRN+ sur la durée de l’étude est improbable pour deux raisons. La première est qu’une souche avec un gène *prn* inactivé présente un fort risque d’accumulations sur ce gène de mutations qui ne sont pas écartées par la sélection naturelle. La deuxième est que l’acquisition d’un nouveau caractère est beaucoup moins probable, sur une génération, que la perte de celui-ci. Le taux de mutation permettant l’acquisition du phénotype PRN+ est ainsi fixé à $\mu_1 = 0$. Inversement, la perte du phénotype PRN+ est beaucoup plus probable, notamment à cause de la multiplicité des mutations pouvant conduire à une inactivation du gène *prn* (Otsuka *et al.*, 2012 ; Pawloski *et al.*, 2014 ; Barkoff *et al.*, 2019). Le taux de mutation de la souche 2 vers la souche 1 est fixé arbitrairement à $\mu_2 = 10^{-6}$.

Taux de contacts entre états

Les taux de contacts entre états représentent les opportunités pour les individus d’un état d’entrer en contact avec ceux d’un autre état et possiblement de leur transmettre une infection. Ces taux de contacts sont renseignés dans la matrice \mathbf{F} . Les valeurs de cette matrice sont particulièrement difficiles à estimer, notamment parce que les flux entre états sont assez mal documentés. L’USCB fournit des statistiques sur les mobilités inter-comtés journalières effectuées par les travailleurs américains², un type de déplacement que les anglophones appellent *commuting*. Ces déplacements ont l’avantage d’être de courte durée : les individus se déplacent pour travailler

2. Disponible sur <https://www.census.gov/topics/employment/commuting/data/tables.html> (dernier accès le 11/09/2024).

puis retournent à leur lieu de résidence. Ils se différencient donc, par exemple, des mouvements de migration, plus permanents. En revanche, ils négligent de nombreux déplacements possibles entre états, par exemple dans le cadre de vacances.

J'ai utilisé les tables de *commuting* de 2020 pour obtenir le nombre d'individus se déplaçant d'un état à l'autre sur une base journalière. Les mouvements de *commuting* sont très variables entre états, et très asymétriques. Par exemple, l'état de New-York reçoit de très nombreux travailleurs, tandis que le New Jersey en donne beaucoup.

Dans ce chapitre, il a été décidé de considérer que le taux de contacts entre l'état j et l'état i , f_{ij} , est proportionnel au nombre d'individus se déplaçant de l'état j à l'état i et est donné par la formule

$$f_{ij|i \neq j} = \frac{\Phi_{ij} + 1}{\max(\Phi) + 1} \varphi, \quad (4.9)$$

où Φ_{ij} est le nombre d'individus se déplaçant quotidiennement de l'état j vers l'état i et $0 \leq \varphi \leq 1$ est un facteur de calibration permettant d'influencer l'intensité des échanges entre états. Le terme $+1$ ajouté au numérateur et compensé au dénominateur est utilisé pour assurer que même si aucun déplacement de *commuting* n'est recensé entre l'état i et j , ces deux états peuvent quand même échanger des individus (autrement dit, le taux de contacts entre deux états n'est jamais nul). On fixe $f_{ii} = 1$, ce qui signifie qu'on considère la fréquence des contacts au sein d'un état comme étant la même pour tous les états et que les contacts au sein d'un état sont toujours plus importants ou égaux à ceux entre deux états puisque $0 \leq f_{ij|i \neq j} \leq 1$.

Taux de transmission des souches

Pour chacune des souches, deux formules sont proposées pour calculer le taux de transmission. La première formule implique une relation directe entre la valeur du cofacteur et la valeur des taux de transmission. Pour la souche 2 cette relation est perturbée par un facteur $(1 - r_2)$ lorsque la vaccination est introduite dans la population. On a donc, pour la souche résidente (souche 2),

$$\beta_2(t, i) = b \cdot e^{-aX(t, i)} \cdot (1 - r_2 \delta_{t \geq t_v}), \quad (4.10)$$

où a est le coefficient associé au cofacteur, b est le taux de transmission basal (*i.e.*, le taux de transmission quand $X(t, i) = 0$), r_2 est la réduction du taux de transmission de la souche 2 due au vaccin, $\delta_{t \geq t_v}$ est le delta de Kronecker (t_v est l'instant où la vaccination est introduite dans la population) et $X(t, i)$ est un cofacteur dont le calcul des valeurs est détaillé plus loin. Cette manière de calculer les valeurs de $\beta_2(t, i)$ permet d'obtenir une relation simple entre le taux de transmission de la souche 2 et la valeur du cofacteur, tout en assurant que le taux de transmission ne devienne jamais négatif.

Pour la souche mutante (souche 1), le taux de transmission est

$$\beta_1(t, i) = b \cdot e^{-aX(t, i)} \cdot (1 - r_1), \quad (4.11)$$

où r_1 est la réduction du taux de transmission de la souche 1. Le taux de transmission de la souche 1 n'est pas perturbé par la vaccination. Avec cette première formule, le rapport des taux de transmission de la souche 2 et de la souche 1 est constant entre tous les pas de temps avant la vaccination ($\frac{\beta_2(t, i)}{\beta_1(t, i)} = \frac{1}{1-r_1}$) et entre tous les pas de temps après son introduction ($\frac{\beta_2(t, i)}{\beta_1(t, i)} = \frac{1-r_2}{1-r_1}$), mais il est perturbé au moment où la vaccination débute.

La seconde formule implique que le cofacteur n'a d'effet que sur le taux de transmission de la souche 2, ce qui signifie que le ratio des taux de transmission de la souche 2 et de la souche 1 varie en fonction du temps. Typiquement, un facteur comme la couverture vaccinale devrait réduire le taux de transmission de la souche 2 mais pas de la souche 1, faisant ainsi fluctuer le rapport des taux de transmission. Avec cette deuxième approche on a, pour la souche 2,

$$\beta_2(t, i) = b \cdot e^{-aX'(t, i)}. \quad (4.12)$$

où $X'(t, i)$ est un cofacteur dont le calcul des valeurs est détaillé plus loin.

On remarque que la formule (4.12) est équivalente à la formule (4.10) pour $r_2 = 0$. Pour la souche 1, le taux de transmission devient

$$\beta_1(t, i) = b \cdot (1 - r_1), \quad (4.13)$$

et le ratio des taux de transmission de la souche 2 et de la souche 1 est $\frac{\beta_2(t, i)}{\beta_1(t, i)} = \frac{e^{-aX'(t, i)}}{1-r_1}$.

On remarque que la formule (4.13) est équivalente à la formule (4.11) pour $a = 0$.

4.2.4 – Calcul des cofacteurs

Plusieurs formules sont proposées pour les valeurs du cofacteur. Dans le cas d'un cofacteur environnemental quelconque variable dans le temps et l'espace de manière indépendante, la formule retenue pour calculer les valeurs du cofacteur $X(t, i)$ est

$$X(t, i) = X(y) + X(i) + \varepsilon(t, i), \quad (4.14)$$

où $X(y) = X(\lfloor \frac{t}{20} \rfloor)$ suit une loi normale de moyenne zéro et d'écart-type σ_y ($X(y) \sim \mathcal{N}(0, \sigma_y)$) et représente la valeur attendue du cofacteur l'année y , $X(i)$ suit une loi normale de moyenne zéro et d'écart-type σ_i ($X(i) \sim \mathcal{N}(0, \sigma_i)$) et représente la valeur attendue du cofacteur dans l'état i , et $\varepsilon(t, i)$ suit une loi normale de moyenne zéro et d'écart-type σ_ε ($\varepsilon(t, i) \sim \mathcal{N}(0, \sigma_\varepsilon)$) et représente la fluctuation de la valeur du cofacteur autour des valeurs attendues, à chaque pas de temps et dans chaque état. La fonction partie entière $\lfloor \frac{t}{20} \rfloor$ indique que la valeur de $X(y)$ change tous les 20 pas de temps, soit tous les ans. Cette variabilité annuelle dans le taux de transmission a été choisie pour faciliter les analyses statistiques qui se font sur une base annuelle. La formule (4.14) est utilisée dans la suite pour les travaux reposant sur les formules (4.10) et (4.11) pour les calculs des coefficients de transmission.

Dans le cas d'un cofacteur type couverture vaccinale, sa valeur doit être nulle avant l'arrivée de la vaccination et augmenter progressivement une fois celle-ci introduite. Le calcul des valeurs du cofacteur $X'(t, i)$ est donc plus compliqué que celui de $X(t, i)$, et on a

$$X'(t, i) = \begin{cases} 0 & \text{si } \lfloor \frac{t}{20} \rfloor < \lfloor \frac{t_v}{20} \rfloor, \\ X'_{t_v}(i) & \text{si } \lfloor \frac{t}{20} \rfloor = \lfloor \frac{t_v}{20} \rfloor, \\ \text{expit}(\text{logit}(X'(t-20, i)) + \varepsilon(\lfloor \frac{t}{20} \rfloor, i)) & \text{si } \lfloor \frac{t}{20} \rfloor > \lfloor \frac{t_v}{20} \rfloor, \end{cases} \quad (4.15)$$

où $0 \leq X'_{t_v}(i) \leq 1$ est la couverture vaccinale initiale dans l'état i , choisie arbitrairement, et $\varepsilon(\lfloor \frac{t}{20} \rfloor, i)$ représente une fluctuation aléatoire annuelle distribuée selon une loi normale de

moyenne

$$\mu_c \left(\left\lfloor \frac{t}{20} \right\rfloor, i \right) = a_c \cdot \left(1 - \frac{X'(t-20, i)}{X'_l(i)} \right), \quad (4.16)$$

où $X'_l(i)$ est la valeur maximale du cofacteur (*e.g.*, la couverture vaccinale maximale) dans l'état i et a_c est un paramètre influençant la vitesse de changement de la couverture vaccinale, et d'écart-type σ_c ($\varepsilon \left(\left\lfloor \frac{t}{20} \right\rfloor, i \right) \sim \mathcal{N} \left(\mu_c \left(\left\lfloor \frac{t}{20} \right\rfloor, i \right), \sigma_c \right)$). L'utilisation de la formule (4.16) assure la convergence de $X'(t, i)$ vers la valeur seuil $X'_l(i)$ et permet de contrôler la vitesse à laquelle cette convergence a lieu.

Puisque $X'(t, i)$ représente une couverture vaccinale et qu'on considère être dans un système type coqueluche, alors t_v représente l'instant d'introduction de l'ACV, alors qu'un WCV pouvait circuler précédemment. Une fois l'ACV introduit, les nouveaux-nés seront vaccinés avec celui-ci tandis que les classes d'âge plus âgées reçoivent des rappels. Considérer ces rappels est essentiel pour que la couverture vaccinale augmente plus rapidement que le taux de renouvellement de la population.

4.2.5 – Calcul des forces d'infection

Soit $\lambda_s(t, i)$ la force de l'infection de la souche s au temps t dans l'état i . La force de l'infection est la somme de trois éléments. Le premier, $A(t, i)$, représente les infections dans l'état i dues aux individus malades résidant dans l'état j et visitant l'état i . Il vaut

$$A_s(t, i) = \sum_{j \neq i} f_{ji} \beta_s(t, i) \frac{I_s(t, j)}{N(i)} \quad (4.17)$$

où f_{ji} est le taux de contact unidirectionnel entre l'état j et l'état i , $\beta_s(t, s)$ est le taux de transmission de la souche s au temps t dans l'état i , $I_s(t, j)$ est le nombre d'individus infectés par la souche s au temps t dans l'état j et $N(i)$ est la taille de la population dans l'état i , considérée constante. On notera que la force de l'infection est pondérée par la taille de la population de l'état i . Il aurait également été possible de pondérer par la taille de la population de l'état j ou par la moyenne des populations des états j et i , mais cela a en réalité une incidence minime sur les résultats des simulations. [Bernhauerová \(2016\)](#) utilise une pondération similaire dans son modèle de transmission de la coqueluche entre classes d'âge, où le taux de transmission est pondéré par

la taille de la classe d'âge recevant l'infection.

Le deuxième élément de $\lambda(t, i)$ est $B_s(t, i)$. Il représente les infections dues à des individus sensibles de l'état i visitant l'état j , où ils sont contaminés par des individus infectés localement.

Il vaut

$$B_s(t, i) = \sum_{j \neq i} f_{ij} \beta_s(t, j) \frac{I_s(t, j)}{N(i)}. \quad (4.18)$$

Enfin, le troisième élément de $\lambda(t, i)$ est $C_s(t, i)$. Il représente les infections acquises au sein d'un même état, c'est-à-dire par des contacts entre individus de l'état i , et vaut

$$C_s(t, i) = f_{ii} \beta_s(t, i) \frac{I_s(t, i)}{N(i)}. \quad (4.19)$$

La force d'infection pour la souche s est donc $\lambda_s(t, i) = A_s(t, i) + B_s(t, i) + C_s(t, i)$. Elle peut également être écrite sous forme matricielle. Soit $\boldsymbol{\lambda}_s(t)$ le vecteur des forces de transmissions pour l'ensemble des états au temps t . On a

$$\begin{aligned} \boldsymbol{\lambda}_s(t) &= \left(\boldsymbol{\beta}_s(t) \times \left({}^0\mathbf{F}^\top \mathbf{I}_s(t) \right) \right) \times \mathbf{N}^{-1} \\ &\quad + {}^0\mathbf{F} (\boldsymbol{\beta}_s(t) \mathbf{I}_s(t)) \times \mathbf{N}^{-1} \\ &\quad + \text{diag}(\mathbf{F}) \times (\boldsymbol{\beta}_s(t) \times \mathbf{I}_s(t) \times \mathbf{N}^{-1}), \end{aligned} \quad (4.20)$$

où $\boldsymbol{\beta}_s(t)$ est le vecteur des taux de transmission de la souche s dans les différents états au temps t , \mathbf{F} est la matrice des taux de contact entre les différents états, $\text{diag}(\mathbf{F})$ est le vecteur dont les valeurs correspondent à la diagonale de \mathbf{F} , ${}^0\mathbf{F}$ est la matrice \mathbf{F} avec les éléments de la diagonale remplacés par 0, $\mathbf{I}_s(t)$ est le vecteur du nombre d'infectés dans les différents états au temps t et \mathbf{N} est le vecteur de la taille de la population des différents états. Le symbole $^\top$ dénote la transposée et \times dénote le produit d'Hadamard ou *element-wise product*.

4.2.6 – Mise en place des simulations

Les simulations du modèle sont effectuées grâce une fonction codée en R v4.4.1 ([R Core Team, 2024](#)) prenant en entrée les éléments suivants : la structure initiale de la population (*i.e.*, le nombre d'individus dans chacun des compartiments à $t = 0$ pour chacun des états i), les paramètres utilisés pour calculer les matrices de transition $\mathbf{T}(t, i)$ et la durée de la simulation.

Chaque simulation commence avec $S(0, i) = N(i) \cdot \frac{1}{b} - 1$ individus susceptibles (arrondi à l'entier le plus proche), $I_1(0, i) = 0$ individu infecté par la souche 1, $I_2(0, i) = 1$ individu infecté par la souche 2 et $R(0, i) = N(i) \cdot (1 - \frac{1}{b})$ individus immunisés (arrondi à l'entier le plus proche). Ces conditions initiales assurent une faible probabilité que l'épidémie s'éteigne quelque temps après son commencement. Le but du modèle n'est pas d'étudier le début d'une épidémie, donc de telles conditions initiales n'entrent pas en conflit avec l'objet de ce travail.

À chaque pas de temps $t > 0$, la fonction calcule la matrice de transition dans chaque état i . À partir de là, le nombre d'individus passant d'un compartiment à un autre est obtenu à partir de distributions multinomiales (une par compartiment et par état). Soit $\mathbf{P}_j(t, i)$ le vecteur contenant le nombre d'individus déplacés vers chacun des compartiments à partir du compartiment j , à l'instant t et dans l'état i . Les valeurs de $\mathbf{P}_j(t, i)$ sont tirées dans une loi multinomiale telle que

$$\mathbf{P}_j(t, i) \sim \mathcal{M}(P_j(t-1, i), p_{S,j}, p_{I_1,j}, p_{I_2,j}, p_{R,j}) \quad (4.21)$$

où $P_j(t-1, i)$ est le nombre d'individus dans le compartiment j à l'instant $t-1$ et dans l'état i et $p_{\cdot,j}$ est la probabilité de passer du compartiment j au compartiment \cdot .

Soit $\mathbf{P}(t, i)$ le vecteur représentant le nombre d'individus dans chacun des compartiments à l'instant t dans l'état i . Autrement dit, ce vecteur représente la population à l'instant t dans l'état i et vaut

$$\mathbf{P}(t, i) = \sum_j \mathbf{P}_j(t, i) \quad (4.22)$$

Ensuite, ces étapes sont reproduites pour chaque pas de temps jusqu'à avoir atteint la limite choisie.

4.3 – Dynamique épidémiologique d'une souche unique

Le but de cette section est de montrer comment la dynamique épidémique est influencée par les différents paramètres du modèle dans un cas simple où une unique souche est en circulation. Pour des raisons de simplification, je ne considère ici ni mutation ($\mu_1 = \mu_2 = 0$) ni vaccination, ce qui implique que seule la souche 2 est en circulation. Son taux de transmission est calculé à partir de la formule (4.10). Comme nous allons le voir, la dynamique épidémiologique du modèle

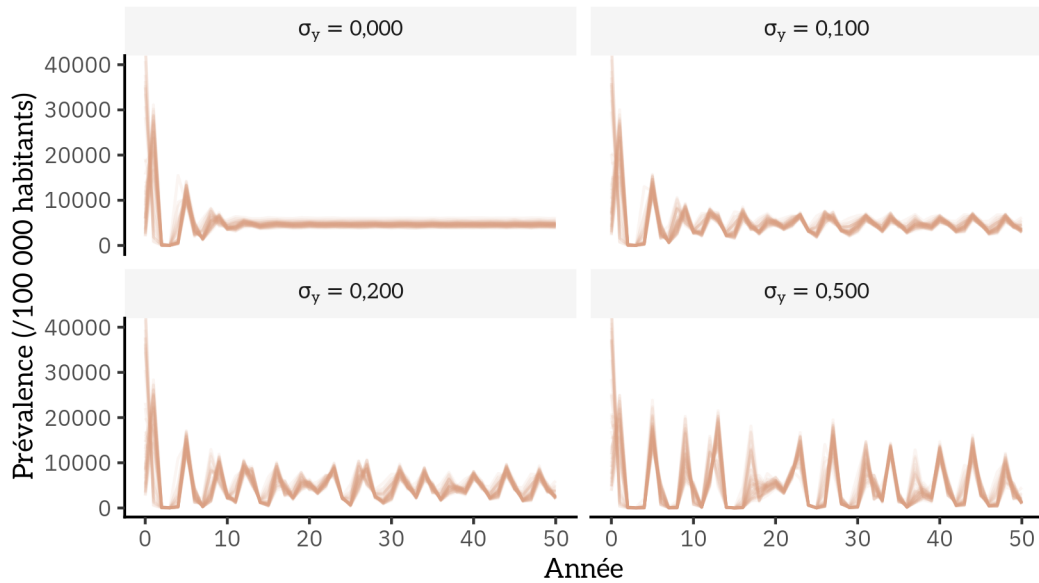


FIGURE 4.1 – **Influence de σ_y sur la prévalence de l’infection** au cours du temps pour une unique simulation. Chaque ligne indique la prévalence de la souche 2 dans un des 49 états suivis. Les autres valeurs des paramètres sont $b = 2,35$, $a = 0,2$, $\sigma_i = 0$ et $\varphi = 1$. La prévalence est donnée pour 100 000 personnes.

est dirigée par les valeurs des taux de transmission des souches, elles-mêmes contraintes par les valeurs du cofacteur $X(t,i)$.

La valeur de σ_y a une incidence sur la variabilité interannuelle dans le taux de transmission des souches. Plus σ_y est élevé, plus on remarque une importante dynamique cyclique dans la prévalence des infections (Figure 4.1). La valeur de σ_y semble avoir un impact sur l’amplitude de ces cycles mais pas sur leur périodicité. Avec les valeurs choisies pour les paramètres, les pics épidémiologiques ont lieu tous les trois ans environ, ce qui coïncide avec la périodicité des épidémies de coqueluche (Fine et Clarkson, 1982 ; Thomas, 1989). Une forte variabilité interannuelle augmente également le risque d’extinction de l’infection : une série de mauvaises années peut ralentir la circulation du pathogène jusqu’au point où celui-ci disparaît de la population avant le retour de conditions plus favorables. Le modèle ne considère pas d’échanges hors de la population qui pourraient réintroduire le pathogène après son extinction.

La valeur de σ_i a une incidence sur la variabilité spatiale dans le taux de transmission des souches. Plus σ_i est élevé, plus la prévalence moyenne des souches varie entre états (Figure 4.2).

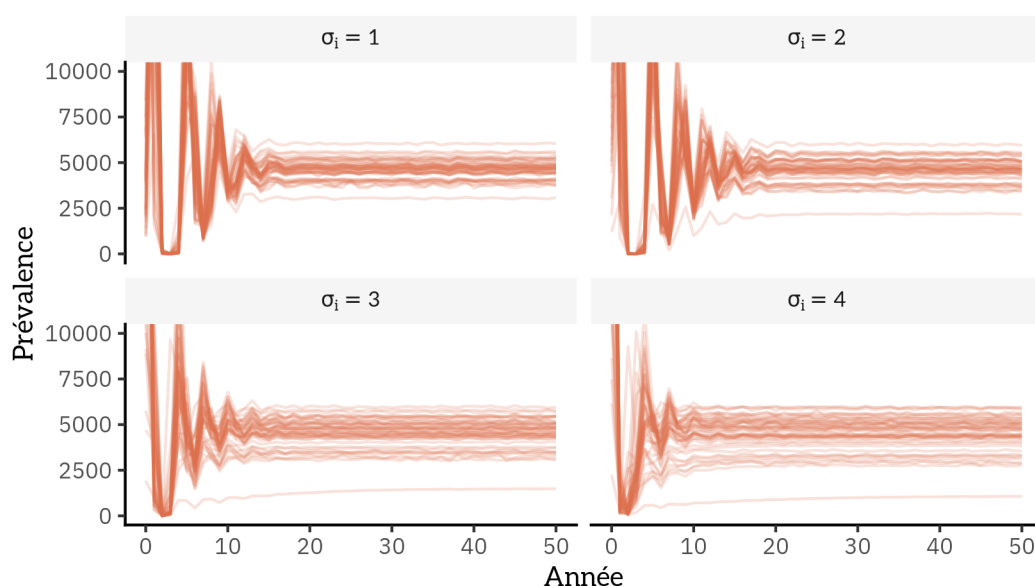


FIGURE 4.2 – **Influence de σ_i sur la prévalence de l’infection** au cours du temps pour une unique simulation. Chaque ligne indique la prévalence de la souche 2 dans un des 49 états suivis. Les autres valeurs des paramètres sont $b = 2,35$, $a = 0,2$, $\sigma_y = 0$ et $\varphi = 1$. La prévalence est donnée pour 100 000 personnes.

Avec un σ_i très élevé, il est possible que le taux de transmission dans un état soit trop bas pour que le pathogène puisse y circuler de manière endémique. Dans ce cas, les infections dans cet état seront essentiellement dues aux contacts avec d’autres états dans lesquels la circulation du pathogène est beaucoup plus efficace. Il est ainsi possible d’avoir une population très hétérogène où certains états font office de « *hubs* » centralisant la majeure partie de la dynamique épidémiologique tout en approvisionnant les autres états en infections, à la manière d’une métapopulation fonctionnant en système puits-source. Il est notable que dans la simulation présentée sur la Figure 4.2 le taux de contacts entre états est très élevé ($\varphi = 1$). Ceci explique que la différence de prévalence entre états n’est pas si marquée que ce à quoi on pourrait s’attendre (chaque état reçoit un grand nombre d’infections des autres états).

La valeur de σ_ε a une incidence à la fois sur la variabilité temporelle et spatiale dans le taux de transmission des souches. Plus σ_ε est élevé, plus le système est désordonné. De plus, ce paramètre influence les valeurs de $X(t, i)$ à l’échelle sub-annuelle. Il va donc créer du bruit statistique lorsque des analyses seront faites sur les données générées par le modèle. Pour cette

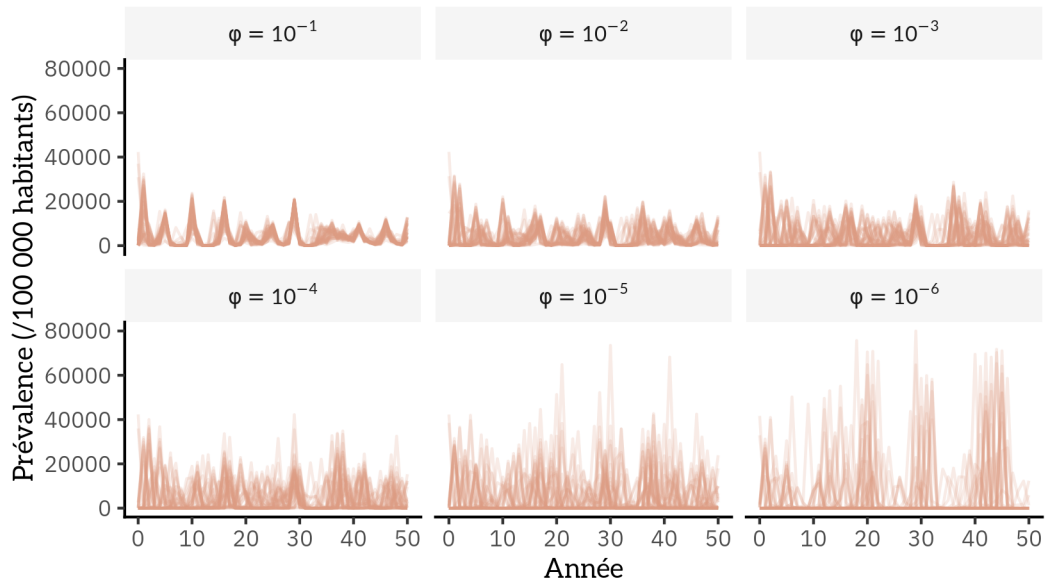


FIGURE 4.3 – **Influence de φ sur la prévalence de l’infection** au cours du temps pour une unique simulation. Chaque ligne indique la prévalence de la souche 2 dans un des 49 états suivis. Les autres valeurs des paramètres sont $b = 2,35$, $a = 0,2$, $\sigma_y = 0$ et $\sigma_i = 0$. La prévalence est donnée pour 100 000 personnes.

raison, le paramètre est fixé à $\sigma_\varepsilon = 0$ dans les simulations utilisées tout au long de ce chapitre.

La prévalence dans les différents états dépend non seulement de la valeur du paramètre σ_i mais également de la connectivité globale entre ceux-ci (paramètre φ). Ainsi, si φ est élevé alors la prévalence dans un état peut être élevée simplement à cause des contacts entre cet état et un autre où la circulation du pathogène est élevée, et les états tendent à s’homogénéiser. On se retrouve alors dans le type de système puits-source que j’évoquais plus tôt. Cette homogénéisation a comme conséquence une plus forte prédictibilité du cycle épidémique et une synchronicité entre états. Inversement, si φ est bas, alors les contacts entre états seront rares et la dynamique épidémique au sein d’un état dépendra essentiellement du taux de transmission du pathogène au sein de l’état. Dans ce cas, les cycles épidémiques deviennent plus chaotiques et se désynchronisent entre états (Figure 4.3).

Il est également intéressant de noter que plus la connectivité entre états est basse, plus le risque d’extinction locale est élevé. Ceci est illustré sur la Figure 4.4. Certaines années, la dynamique des infections dans un état peut conduire à l’extinction de l’épidémie. Cependant, si la connectivité

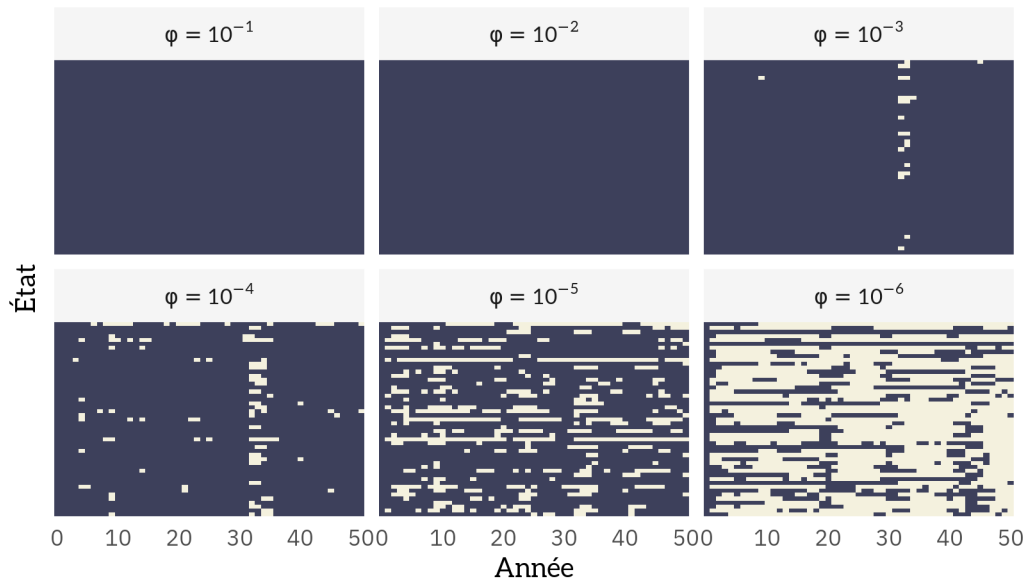


FIGURE 4.4 – **Extinction locale et persistance de l’infection.** Une case foncée indique la présence du pathogène dans l’état pour l’année donnée et une case claire indique son absence. Les données sont les mêmes que celles présentées sur la Figure 4.3.

est grande alors la probabilité que plusieurs états connaissent une extinction locale de manière simultanée est faible, et un état dans lequel le pathogène ne circulait plus pourra rapidement se faire réinfecter par un autre état. En revanche, si la connectivité entre états est faible alors on aura un effet boule de neige : un état dans lequel on observe une extinction locale pourra difficilement se faire réinfecter, et restera donc sain. D’autres états pourront à leur tour subir une extinction locale, et ne pourront pas être réinfectés par le premier état puisque celui-ci sera resté sain. Ainsi, il est possible que d’extinction locale en extinction locale le pathogène disparaisse complètement de la population, mais il est également possible que certains états soient favorables à une circulation endémique du pathogène et que l’extinction dans ces états ne survienne jamais.

Enfin, il reste possible, même avec une faible connectivité, qu’un état devenu sain soit à nouveau infecté après un temps plus ou moins long. Si ce temps est long, alors l’état aura eu le temps de renouveler au moins partiellement sa population d’individus susceptibles. Dans ce cas, la réinfection peut conduire à des épidémies d’une très grande force (Figure 4.3).

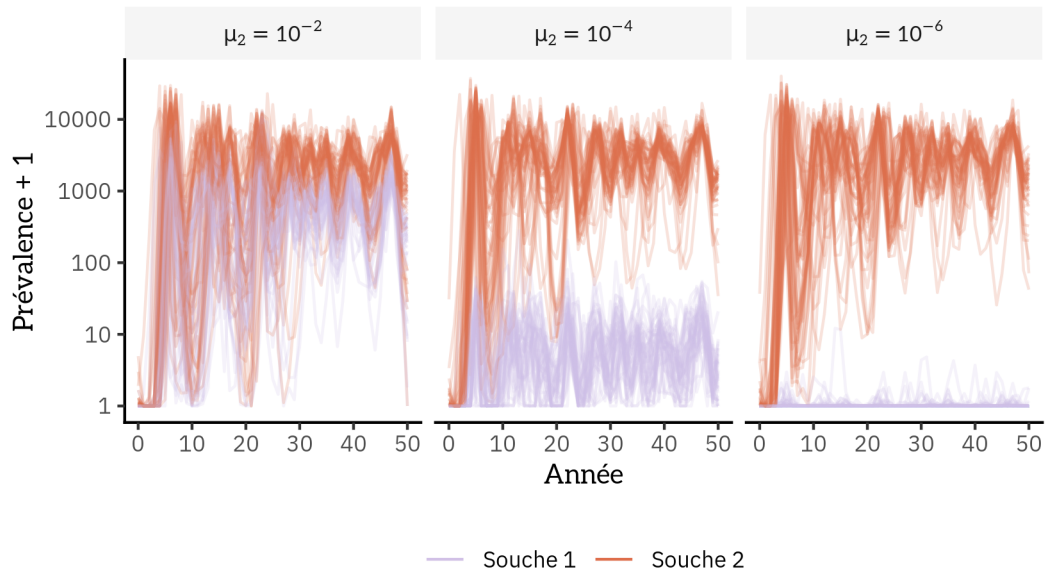


FIGURE 4.5 – **Influence de μ_2 sur la prévalence de l’infection** au cours du temps. Chaque ligne indique la prévalence de la souche 1 ou de la souche 2 dans un des 49 états suivis. La prévalence est donnée pour 100 000 personnes et est incrémentée d’une unité pour s’accorder à l’échelle logarithmique de l’axe des ordonnées.

4.4 – Compétition entre souches

Dans cette partie, les valeurs des paramètres sont fixées à $b = 2,35$, $a = 0,2$, $\sigma_y = 0,2$, $\sigma_i = 0,5$ et $\varphi = 5 \times 10^{-4}$.

L’intérêt principal du modèle est la mise en compétition de deux souches, une sauvage (nommée souche 2) et une mutante (nommée souche 1). Comme présenté dans la Section 4.2.3, la souche 1 émerge lorsque la souche 2 mute. Si aucun vaccin n’entre en circulation, la souche 1 va régulièrement émerger mais ne se maintiendra pas dans la population à cause de l’avantage compétitif de la souche 2, sauf pour des taux de mutation trop élevés pour être réalistes (Figure 4.5).

Les résultats sont différents si un vaccin entre circulation à un instant t_v de la simulation. En utilisant la formule (4.10) ou (4.12), on comprend que le taux de transmission de la souche 2 diminue à partir du moment où le vaccin est introduit, et que l’intensité de cette réduction est conditionnée par le paramètre r_2 (formule (4.10)) ou par la valeur du cofacteur (formule (4.12)). La formule (4.10) implique que lorsque la vaccination est introduite dans la population celle-ci

est instantanément et universellement adoptée. Le taux de transmission de la souche 2 va donc immédiatement devenir inférieur à celui de la souche 1 puisque $r_2 > r_1$. Quoique l'impact de cette simplification puisse être questionné, elle permet de mettre en évidence la manière dont la souche 1 remplace la souche 2 après l'introduction de la vaccination. À l'inverse, si la formule (4.12) est utilisée, alors la diminution du taux de transmission de la souche 2 sera progressive puisque celle-ci dépend de la couverture vaccinale, qui n'est pas immédiatement élevée.

La Figure 4.6 a été construite en faisant tourner 100 simulations à partir des mêmes valeurs du cofacteur calculées à partir de la formule (4.14) et des taux transmissions calculés avec les formules (4.10) et (4.11). Pour chaque pas de temps et chaque état, on regarde si la souche 1 (Figure 4.6A) ou la souche 2 (Figure 4.6B) est présente. Avant l'introduction de la vaccination, on observe que la souche 1 apparaît et disparaît régulièrement. On remarque également que dans certains états la présence de la souche 1 est très rarement observée. Cela peut s'expliquer par une circulation moindre la souche 2 dans ces états et donc à un faible nombre de mutations conduisant à l'émergence de la souche 1. De plus, cette dernière a une probabilité extrêmement faible d'être transmise d'un état à l'autre à cause de sa faible prévalence. La souche 2 quant à elle circule globalement bien dans les populations, et les extinctions locales sont rares avec les valeurs choisies pour les paramètres (les extinctions remarquées en début de simulations sont dues à la structure initiale de la population).

Toujours à partir des formules (4.10), (4.11) et (4.14), une fois la vaccination introduite la souche 1 se transmet mieux que la souche 2. On remarque que la vitesse d'invasion de la souche 1 varie entre états, ce qui peut s'expliquer par des taux de transmissions variables. Avec les paramètres choisis, on voit qu'il faut une dizaine d'années en moyenne après l'introduction de la vaccination pour que la souche s'établisse et se maintienne au sein d'un état (Figure 4.6). Pour la souche 2, une première perturbation a lieu directement après l'introduction de la vaccination, avec une vague d'extinctions locales. Celle-ci est due aux changements brutaux des taux de transmission de la souche, qui perturbent la dynamique du modèle. Une vingtaine d'années environ après l'introduction de la vaccination, on observe que la souche 2 disparaît de plusieurs états, quoique dans certains cette disparition a lieu plus précocement ou au contraire plus tardivement. Il y a donc une dizaine d'années environ au cours desquelles les deux souches vont cohabiter dans la plupart des états.

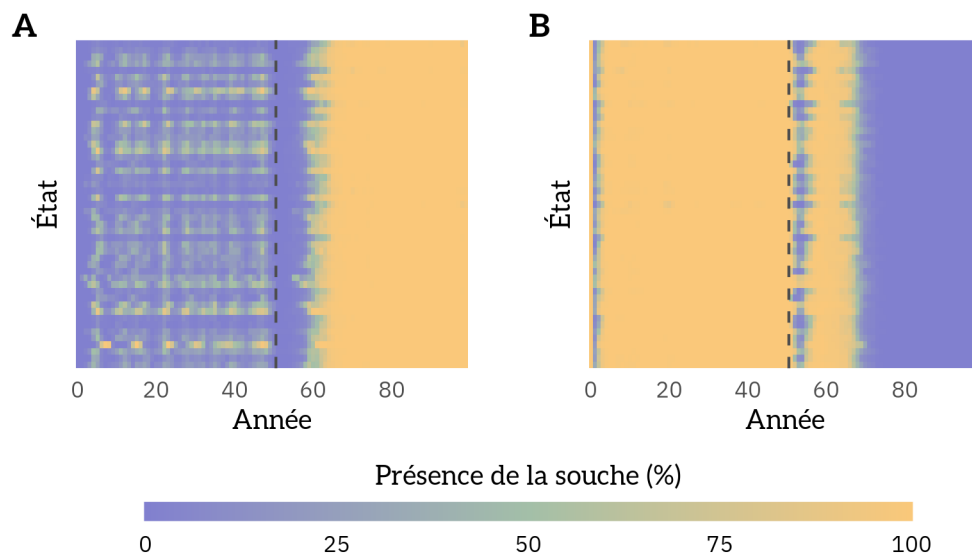


FIGURE 4.6 – **Proportion de simulations pour lesquelles les souches se maintiennent dans la population** dans chaque état et chaque année, pour la souche 1 (A) et la souche 2 (B).

4.5 – Échantillonnage, estimation et qualité des cofacteurs

4.5.1 – Simuler des données réalistes

Le modèle est capable de générer une grande variété de résultats à partir de mêmes conditions initiales et valeurs des paramètres, et cette diversité augmente encore lorsque les conditions initiales ou les valeurs des paramètres varient. Le modèle a été paramétré de manière à produire des sorties « réalistes », en se basant sur quatre critères quantitatifs.

Le premier est le nombre moyen de cas annuels. [Yeung *et al.* \(2017\)](#) estiment que 2,9 % des enfants de moins de cinq ans sur le continent américain ont été infectés par *B. pertussis* en 2014. En considérant cette proportion constante entre années, représentative de la situation des États-Unis et extensible aux autres classes d'âge de la population, cela représente un nombre d'infections moyen par an estimé à $\bar{I} = 9\,549\,000$ environ.

Le deuxième critère est le temps moyen avant qu'un état soit infecté par la souche 1 une fois la vaccination introduite dans la population américaine. Aux États-Unis, la vaccination acellulaire est introduite dans les calendriers vaccinaux en plusieurs vagues : d'abord en 1992 sous la forme

de rappels pour les enfants, puis en 1997 les vaccins acellulaires remplacent les vaccins à germe entier pour la primo-vaccination, et enfin en 2006³ un rappel vaccinal est introduit pour les adolescents (Clark *et al.*, 2012; Klein, 2014). J'utilise le jeu de données de Lefrancq *et al.* (2022) pour estimer le temps moyen avant qu'un état soit infecté par la souche mutante. Ce sont là des données issues d'un échantillonnage partiel où tous les états américains ne sont pas représentés, donc l'estimation du temps moyen avant que la souche mutante n'arrive sur un territoire est biaisée. Néanmoins, en considérant que la dynamique d'invasion dans les états échantillonnés correspond à la dynamique globale, alors le nombre moyen d'années écoulées entre l'introduction de la vaccination et l'apparition de la souche mutante dans un état est $\bar{t} = 19,542$, $\bar{t} = 14,542$ ou $\bar{t} = 5,542$ selon l'année considérée comme date de début de la vaccination (1992, 1997 ou 2006, respectivement).

Le troisième critère est l'écart-type du temps avant qu'un état soit infecté par la souche 1 une fois la vaccination introduite dans la population. En reprenant les données de Lefrancq *et al.* (2022), l'écart-type du nombre d'années avant qu'un état soit infecté est $\sigma_t = 2,553$. Cette valeur est indépendante de la date d'instauration de la vaccination.

Le quatrième critère est le nombre moyen d'années écoulées avant que 85 % des souches en circulation soient devenues des souches de type 1, noté t_{85} . Dans le jeu de données de Lefrancq *et al.* (2022), ce seuil est atteint en 2013, soit après $t_{85} = 21$, $t_{85} = 16$ ou $t_{85} = 7$ années selon celle considérée comme date de début de la vaccination acellulaire (1992, 1997 ou 2006, respectivement).

Après plusieurs essais, la combinaison de valeurs de paramètres présentée dans le Tableau 4.2 a été jugée satisfaisante pour conduire les analyses qui suivent. Obtenir des valeurs des quatre critères proches de celles attendues est compliqué par l'existence de corrélations entre les valeurs de certains d'entre eux, et notamment entre \bar{t} et σ_t ($r = 0,662$, $p < 0,001$). En conséquence, il n'a pas été possible d'obtenir des valeurs de paramètres pour lesquelles \bar{t} est élevé et σ_t est bas, comme attendu en moyenne.

On observe de plus une très forte variabilité entre simulations dans les valeurs des quatre critères (Figure 4.7). On remarque que les valeurs des paramètres choisies permettent d'obtenir un nombre moyen d'infectés proche de celui attendu (Figure 4.7A), mais que le temps moyen

3. 2005 d'après Clark *et al.* (2012)

TABLE 4.2 – Valeurs arbitraires des paramètres épidémiologiques utilisées dans le modèle. Les valeurs des paramètres biologiques sont données dans la Section 4.2.3.

Paramètre	Valeur
σ_i	0,5
σ_y	0,2
φ	10^{-4}
b	2,1
a	0,2
r_1	0,12
r_2	0,20

avant qu'un état soit infecté par la souche 1 est assez réduit (Figure 4.7B). À l'inverse, l'écart-type est quant à lui plus grand qu'attendu (Figure 4.7C). Il est cependant vraisemblable que l'absence de données dans certains états et le nombre réduit de souches dans les données aient conduit à des estimations biaisées des valeurs de \bar{t} et de σ_t . Enfin, le nombre d'années avant que la souche 1 ne représente 85 % des souches en circulation est plutôt correctement estimé, avec une médiane à $t_{85} = 19$ années, soit entre les dates estimées pour un début de vaccination entre 1992 et 1997 (Figure 4.7D). Ce dernier critère est certainement le plus important puisqu'il s'agit de celui permettant le plus de s'assurer d'une adéquation entre les simulations et les données.

4.5.2 – Estimation des cofacteurs

Une première série de simulations a été menée en utilisant les valeurs des paramètres obtenues plus tôt. Dans cette série, la formule (4.14) est utilisée pour calculer les valeurs du cofacteur et les formules (4.10) et (4.11) sont utilisées pour calculer les taux de transmissions $\beta_2(t, i)$ et $\beta_1(t, i)$, respectivement. À la fin de chaque simulation, un échantillonnage est fait avec un certain effort p_e correspondant à la probabilité pour un individu infecté d'être échantillonné. Un jeu de données est produit à partir de l'année où le premier individu infecté par la souche 1 est échantillonné et pour les dix années suivantes. Pour chacune de ces années, les valeurs des cofacteurs Δ_X et μ_X sont calculées à partir des formules (3.11) et (3.12) présentées dans le Chapitre 3, à la fois avec la méthode $k' = t - 1$ et la méthode $k' = t$, grâce aux valeurs de $X(t, i)$ moyennées sur l'année⁴ et au nombre de souches échantillonnées. Ces cofacteurs sont également calculés à partir du nombre

4. Sachant que puisque $\sigma_\varepsilon = 0$, la valeur moyennée sur l'année de $X(t, i)$ vaut $X(\lfloor \frac{t}{20} \rfloor, i)$.

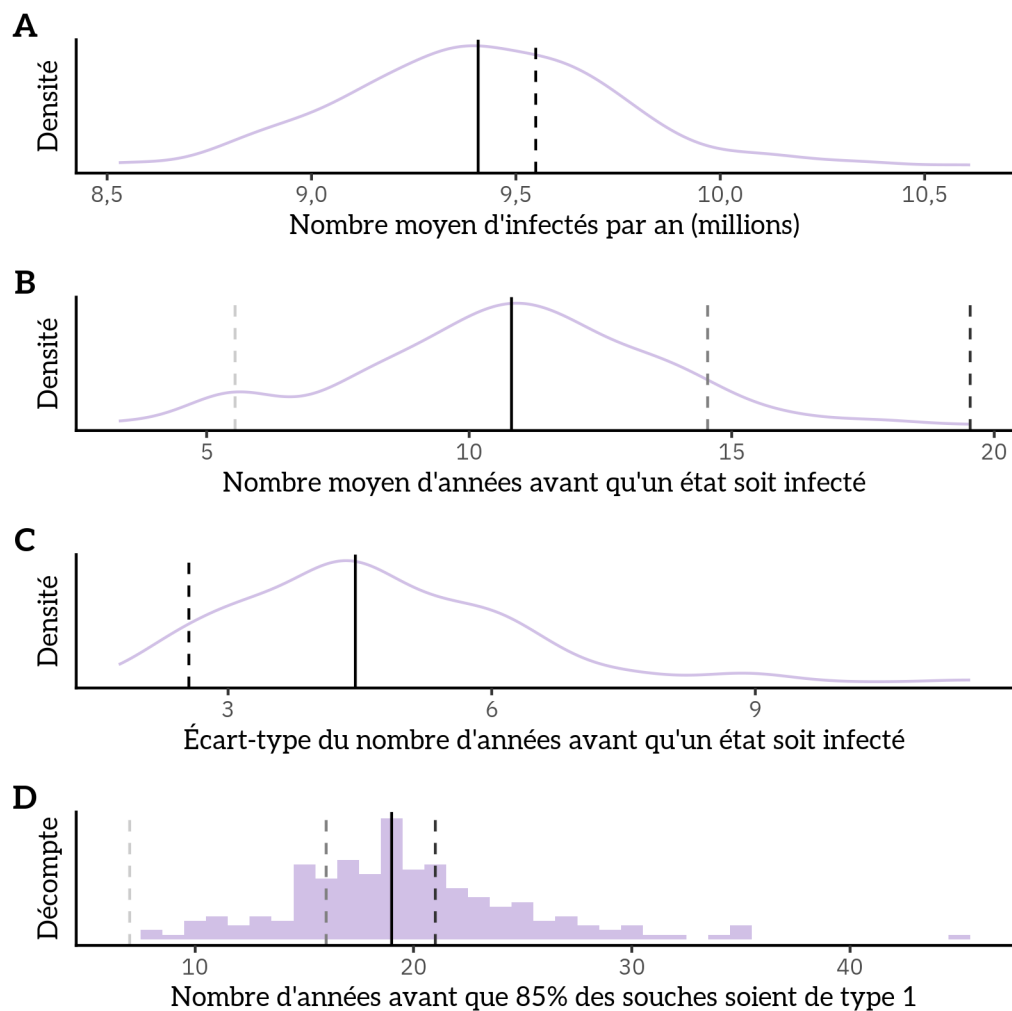


FIGURE 4.7 – Valeurs obtenues pour les différents critères d'ajustement pour 100 simulations. (A) Nombre moyen d'infectés par an (\bar{I}). (B) Nombre moyen d'années avant qu'un état soit infecté (\bar{t}). (C) Écart-type du nombre d'années avant qu'un état soit infecté (σ_t). (D) Nombre d'années avant que 85% des souches en circulation soient du type souche 1 (t_{85}). En bleu, la densité des valeurs (A-C) ou le nombre de simulations par valeur obtenue (B). La ligne verticale pleine indique la valeur médiane. Les lignes en tirets indiquent la valeur attendue pour (A et C) pour le critère ou les valeurs attendues selon la date considérée comme le début de la vaccination acellulaire (B et D, clair : 2006, intermédiaire : 1997, foncé : 1992).

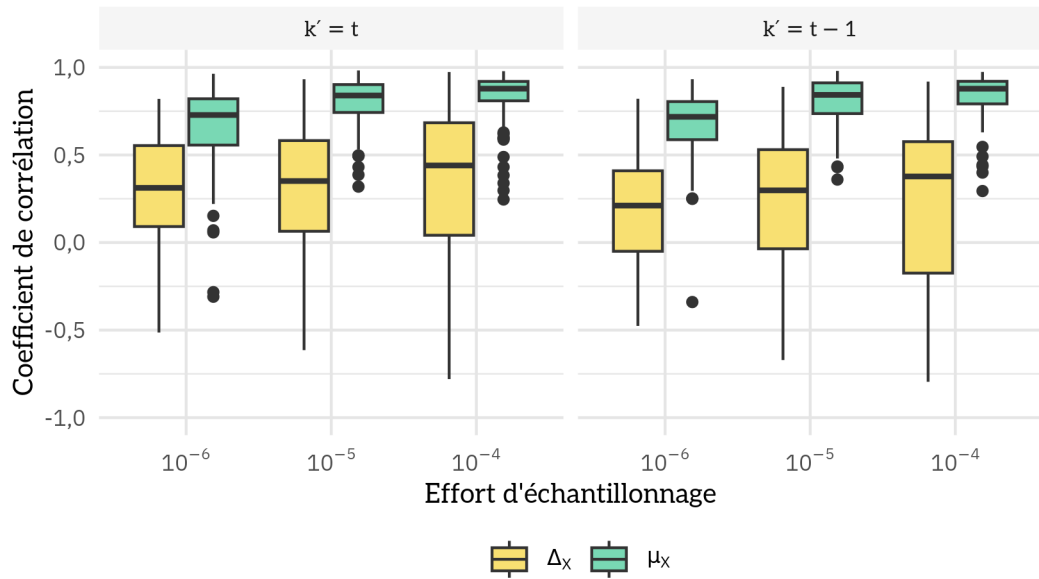


FIGURE 4.8 – **Corrélation entre la valeur réelle et la valeur estimée des cofacteurs Δ_X et μ_X** calculées à partir de $k' = t - 1$ et $k' = t$ pour différents efforts d'échantillonnage (p_e). Les cofacteurs et les taux de transmission ont été calculés avec les formules (4.14), (4.10) et (4.11).

réel d'individus infectés dans la population.

Le but de cette analyse est de vérifier si la corrélation entre les valeurs estimées et les valeurs réelles des cofacteurs est satisfaisante. Pour mesurer la corrélation entre les valeurs estimées et réelles d'un cofacteur, j'ai pris dans chaque jeu de données généré les dix valeurs estimées du cofacteur et ses dix valeurs réelles et calculé leur coefficient de corrélation.

On remarque en premier lieu qu'augmenter l'effort d'échantillonnage augmente la corrélation entre les valeurs estimées des cofacteurs et leurs valeurs réelles (Figure 4.8). Ce résultat est attendu : lorsque l'effort d'échantillonnage augmente, l'échantillon tend à se rapprocher de la population et les valeurs estimées se rapprochent donc des valeurs réelles. On remarque ensuite que les corrélations des cofacteurs Δ_X et μ_X ne sont pas équivalentes : le cofacteur μ_X montre, en moyenne, une meilleure corrélation entre sa valeur estimée et sa valeur réelle. De plus, la qualité de la corrélation pour le cofacteur Δ_X est beaucoup plus variable, les coefficients de corrélation obtenus se répartissant dans une gamme de valeurs beaucoup plus large que pour le cofacteur μ_X .

Pour la deuxième série de simulations, les valeurs des paramètres sont les mêmes que dans le Tableau 4.2, à l'exception du paramètre a dont la valeur a été fixée à 0,25. Ce changement de

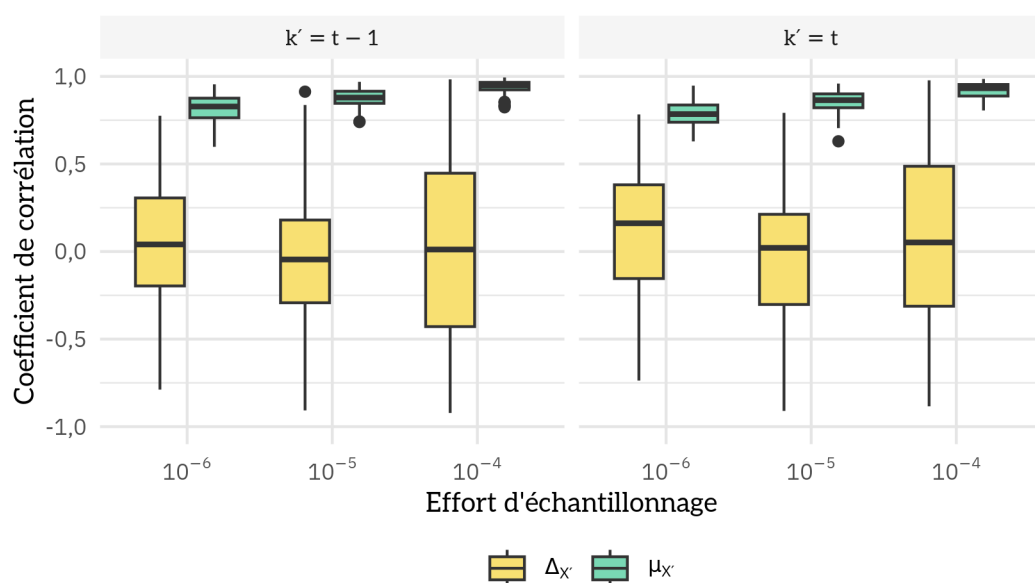


FIGURE 4.9 – **Corrélation entre la valeur réelle et la valeur estimée des cofacteurs Δ_X et μ_X** calculées à partir de $k' = t - 1$ et $k' = t$ pour différents efforts d'échantillonnage (p_e). Les cofacteurs et les taux de transmission ont été calculés avec les formules (4.15), (4.12) et (4.13).

valeur permet d'assurer que le ratio $\frac{\beta_2(t,i)}{\beta(t,i)}$ lorsque $t \rightarrow +\infty$ soit similaire à celui de la première série de simulations. Dans cette série, les valeurs du cofacteur ont été obtenues en utilisant la formule (4.15) et les valeurs des coefficients $\beta_2(t,i)$ et $\beta_1(t,i)$ ont été obtenues en utilisant les formules (4.12) et (4.13), respectivement. Les simulations fonctionnent de la même manière que pour la première série.

On retrouve des résultats similaires à ceux obtenus avec la première série de simulations : la corrélation est plus élevée pour le cofacteur $\mu_{X'}$ que pour le cofacteur $\Delta_{X'}$, et un plus grand effort d'échantillonnage semble s'accompagner d'une meilleure corrélation, quoique ce résultat ne se retrouve ici que pour $\mu_{X'}$ (Figure 4.9). De plus, pour le coefficient $\mu_{X'}$ la corrélation moyenne semble être comparable à celle obtenue avec la première série de simulations mais sa variance est réduite. En revanche, pour le cofacteur $\Delta_{X'}$, la corrélation moyenne est plus médiocre que pour la première série de simulations.

Ces différents résultats montrent que la qualité des estimations des cofacteurs dépend de la manière dont celles-ci sont exécutées. Les valeurs estimées du cofacteur μ_f (où f représente X ou X') correspondent globalement bien à celles réelles (corrélation forte), alors que pour le cofacteur

Δ_f ce n'est pas le cas. Finalement, la pertinence d'utiliser le cofacteur Δ_f dans l'étude de la vitesse de propagation d'une souche mutante est remise en question ici.

4.5.3 – Estimation des coefficients

Les jeux de données produits dans les deux séries de simulations précédentes ont été analysés en utilisant le modèle bayésien présenté dans le Chapitre 3. Seules les simulations avec $p_e = 10^{-5}$ ont été analysées. Cet effort d'échantillonnage assure une taille d'échantillon de l'ordre de la centaine de souches isolées par an, similaire à ce qui est décrit dans le jeu de données de [Lefrancq et al. \(2022\)](#).

Pour chaque série de simulations, les analyses ont été menées à la fois sur les cofacteurs calculés à partir des échantillons et sur les cofacteurs calculés à partir des populations (*i.e.*, à partir du nombre total d'individus infectés par année et par état). L'analyse bayésienne permet d'estimer, d'une part, la valeur du coefficient de sélection, a_0 , et d'autre part la valeur du coefficient associé au cofacteur, a_f . Pour la deuxième série de simulations, le modèle a également été utilisé en prenant comme variable explicative la valeur moyennée par année entre tous les états de $X'(t, i)$, que j'appelle $\tau_{X'}$. Ce nouveau cofacteur peut être interprété comme la couverture vaccinale annuelle dans la population générale sans transformation/pondération par le nombre de souches isolées dans les différents états.

Pour chaque simulation effectuée avec le modèle bayésien, le \hat{R} de [Brooks et Gelman \(1998\)](#) a été utilisé pour vérifier la convergence entre les chaînes du modèle. Si pour l'un des paramètres surveillés une simulation retournait une valeur $\hat{R} > 1,1$, alors la simulation est abandonnée et relancée, comme décrit dans la Section 3.7.1. Pour chaque jeu de données produit et chaque cofacteur étudié, si au bout de dix essais la convergence n'était toujours pas atteinte alors l'analyse était abandonnée et aucun résultat n'était produit pour ce cofacteur. La capacité du modèle à prédire la valeur d'un coefficient est donnée par une mesure que j'appellerai P et que j'ai déjà présenté dans la Section 3.7.2. Pour rappel, elle représente la proportion doublée de valeurs de la distribution postérieure des valeurs d'un coefficient qui soit du signe opposé à la valeur médiane de cette distribution. Il s'agit donc d'une valeur comprise entre zéro et un (inclus). Si $P < 0,05$, cela implique que l'intervalle de crédibilité à 95 % du paramètre étudié exclut zéro.

Les résultats de ces analyses sont encore préliminaires, mais quelques informations intéressantes peuvent en être tirées. Pour la première série de simulations, le modèle est parvenu dans la plupart des cas à estimer une valeur de a_0 dont zéro est exclu de l'intervalle de crédibilité à 95 % (Figure 4.10). Il est remarquable que les valeurs estimées de a_0 sont parfois très grandes ($a_0 > 40$ pour certaines simulations). De telles valeurs sont aberrantes mais s'expliquent facilement par la structure des jeux de données ayant conduit à ces estimations. Il s'agissait généralement de jeux de données dans lesquels la proportion initiale de souches de type 1 était élevée, donc la proportion relative de ces souches augmentait rapidement pour atteindre 100%. De telles conditions sont défavorables au bon fonctionnement du modèle, et conduisent à une mauvaise estimation des valeurs des coefficients.

Pour les valeurs de a_f , les résultats sont très différents. Globalement, et pour tous les cofacteurs, le modèle s'est retrouvé incapable d'estimer une valeur de a_f dont l'intervalle de crédibilité exclut zéro, que la valeur médiane estimée de a_f soit très éloignée ou au contraire très proche de zéro (Figure 4.11). De plus, on retrouve un grand nombre de simulations pour lesquelles la valeur estimée de a_f est aberrante, pouvant atteindre des valeurs absolues dépassant 60. Enfin, les valeurs estimées pour a_f sont distribuées de manière presque équiprobable autour de zéro, que ces estimations soient très éloignées de zéro ou au contraire très proche, accentuant la conclusion que le modèle est ici incapable d'estimer de manière robuste la valeur du coefficient a_f .

Les résultats de la deuxième série de simulations sont comparables à ceux de la première série. Pour le coefficient a_0 , le modèle s'est montré capable d'estimer pour la plupart des jeux de données des valeurs dont l'intervalle de crédibilité à 95 % exclut zéro (Figure 4.12). Une différence majeure avec la première série d'analyses est l'absence de valeurs aberrantes dans les estimations. Dans les simulations avec le cofacteur $\Delta_{X'}$ calculé avec la population, certaines estimations de a_0 sont plus grandes que 5, mais pour la majorité des simulations avec les autres cofacteurs les estimations restent en dessous de 2,5.

Pour les estimations du coefficient a_f , les estimations ont été comme pour la première série relativement mauvaises : pour une grande partie d'entre elles l'intervalle de crédibilité à 95 % comprend zéro (Figure 4.13). Cependant, on remarque une plus grande proportion de simulations pour lesquelles ce n'est pas le cas. De plus, on retrouve beaucoup moins de valeurs aberrantes dans les estimations de a_f que pour la première série de simulations. Il n'y a que pour le cofacteur

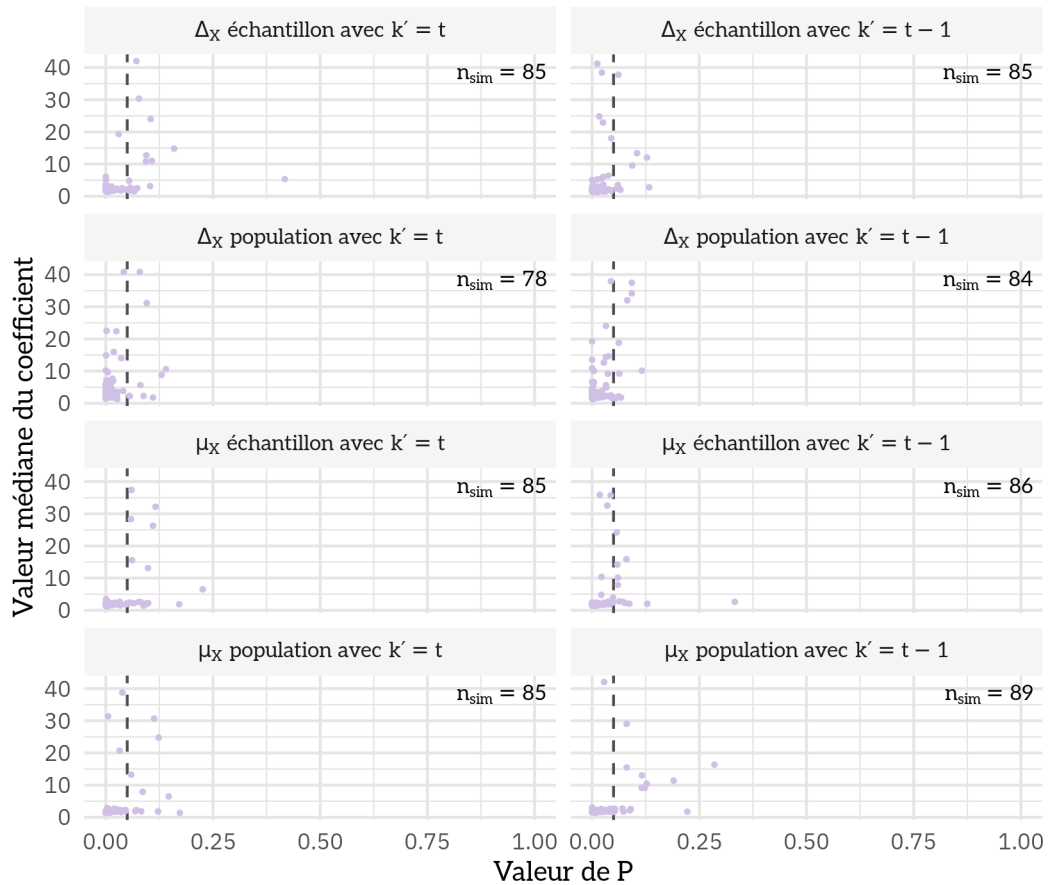


FIGURE 4.10 – Valeurs médiane de a_0 et valeur de P pour la première série de simulations avec différents cofacteurs. La valeur de n_{sim} représente le nombre de simulations pour lesquelles l’analyse bayésienne a pu être menée à bien, sur un total de 100 simulations. La ligne en tirets indique le seuil $P = 0,05$.

$\Delta_{X'}$, que des valeurs aberrantes sont apparues, comme remarqué plus tôt pour les estimations du coefficient a_0 .

La première série de simulations se basait sur l’utilisation d’un cofacteur « brut » $X(t, i)$ ayant une incidence sur le taux de transmission absolu des souches en compétition mais pas sur le ratio de ces taux de transmission. La deuxième série quant à elle se basait sur l’utilisation d’un cofacteur brut $X'(t, i)$ ayant une incidence sur le ratio des taux de transmission des deux souches. Cette différence fondamentale pourrait expliquer la – légère – différence de qualité entre les résultats de la première et de la deuxième série de simulations : si le cofacteur brut n’a pas d’incidence sur le ratio du taux de transmission des souches, alors ses fluctuations ne devraient

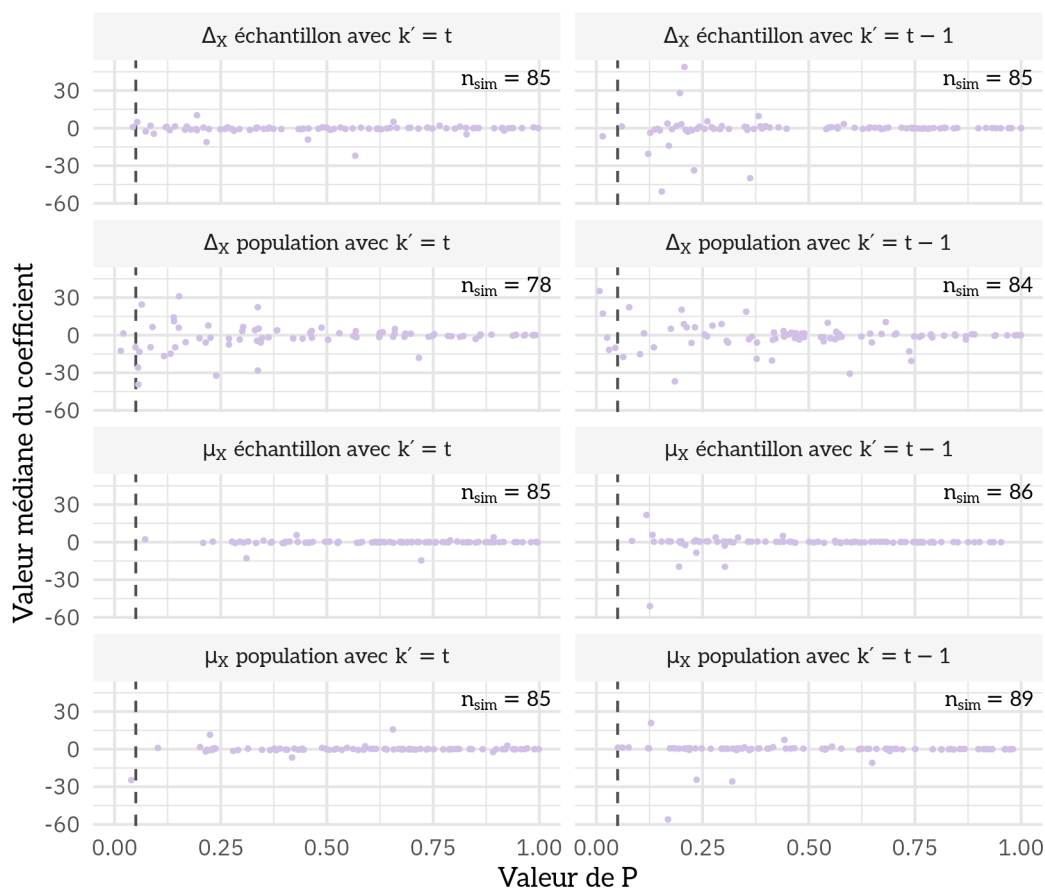


FIGURE 4.11 – Valeurs médianes de a_f et valeurs de P pour la première série de simulations avec différents cofacteurs. La valeur de n_{sim} représente le nombre de simulations pour lesquelles l’analyse bayésienne a pu être menée à bien, sur un total de 100 simulations. La ligne en tirets indique le seuil $P = 0,05$.

pas perturber la vitesse d’invasion de la souche 1, ce qui expliquerait que les estimations des coefficients des cofacteurs Δ_X et μ_X incluent zéro dans leur intervalle de crédibilité à 95 % pour la plupart des simulations.

Inversement, pour la deuxième série de simulations on s’attendrait à retrouver plus de simulations pour lesquelles l’estimation de a_f exclut zéro que ce qui est observé. Deux explications peuvent être proposées pour expliquer les résultats obtenus. Premièrement, pour de nombreuses simulations le jeu de données produit montrait une forte proportion de souches 1 dès les premiers temps de l’analyse, faussant les résultats comme expliqué plus tôt. Deuxièmement, le cofacteur $X'(t, i)$ varie en fonction du temps, et y est donc fortement corrélé. Il est donc plus que probable

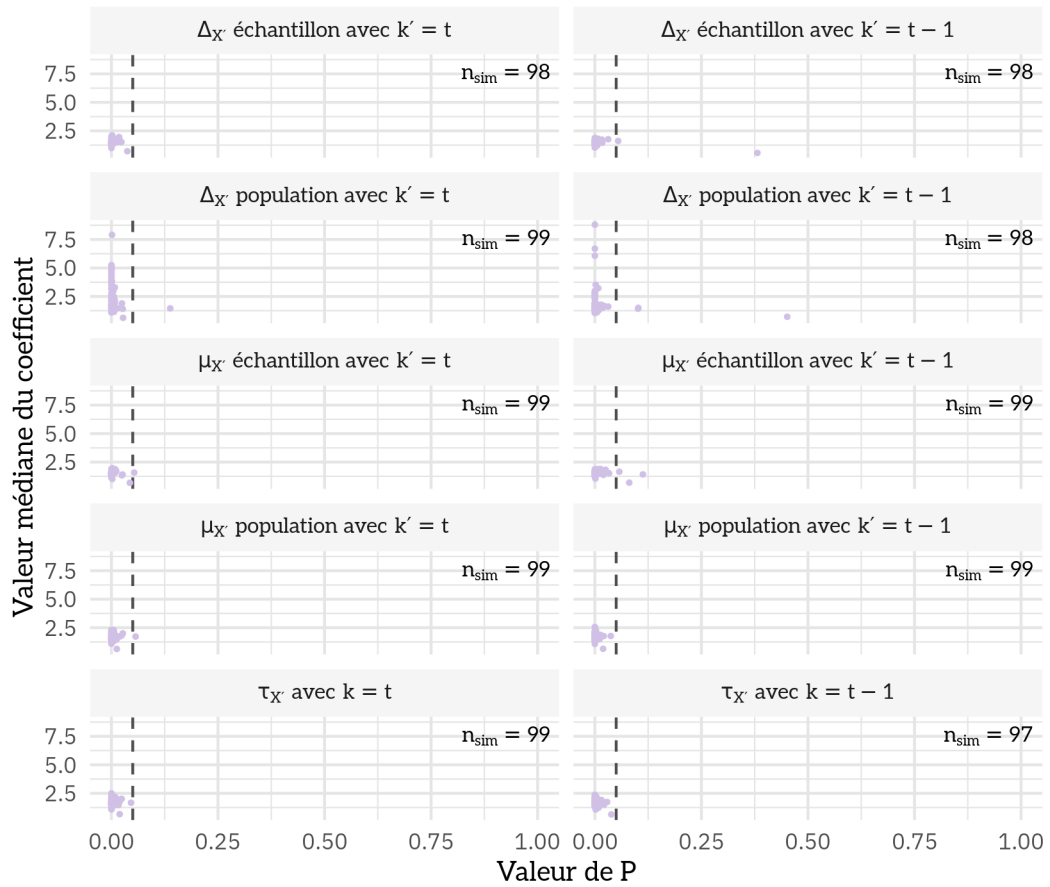


FIGURE 4.12 – Valeurs médianes de a_0 et valeurs de P pour la deuxième série de simulations avec différents cofacteurs. La valeur de n_{sim} représente le nombre de simulations pour lesquelles l’analyse bayésienne a pu être menée à bien, sur un total de 100 simulations. La ligne en tirets indique le seuil $P = 0,05$.

qu’une partie de la fluctuation de la vitesse de propagation de la souche 1 soit en fait capturée par le coefficient a_0 .

4.6 – Discussion

Dans cet ultime axe de ma thèse, j’ai développé un modèle informatique et mathématique simulant la propagation dans l’espace et le temps de deux souches d’un même pathogène. Ce modèle a été paramétré à partir d’informations sur l’épidémiologie de la coqueluche et sur certaines caractéristiques socio-démographiques des États-Unis.

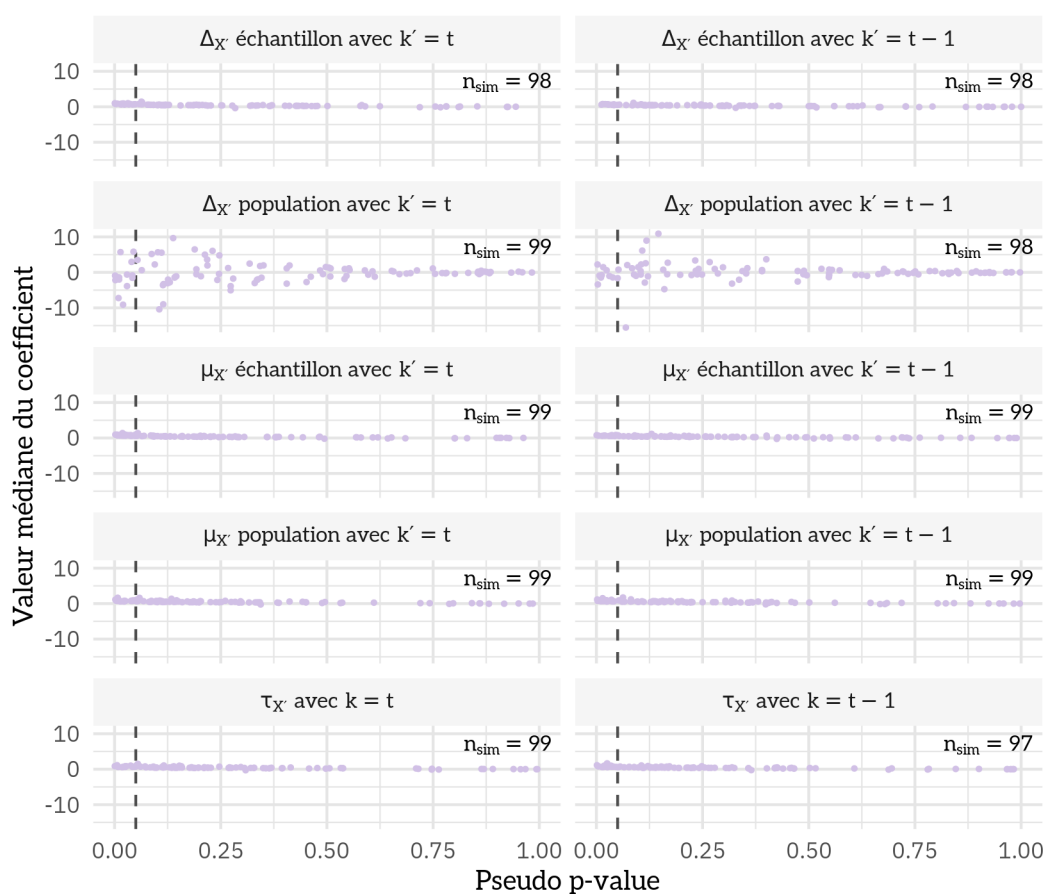


FIGURE 4.13 – Valeurs médianes de a_f et valeurs de P pour la deuxième série de simulations avec différents cofacteurs. La valeur de n_{sim} représente le nombre de simulations pour lesquelles l’analyse bayésienne a pu être menée à bien, sur un total de 100 simulations. La ligne en tirets indique le seuil $P = 0,05$.

L’un des intérêts de ce modèle est qu’il considère les taux de transmissions des souches comme étant variables dans le temps et l’espace, ce qui permet d’étudier comment les fluctuations des taux de transmissions peuvent influencer la vitesse de propagation des souches. Dans un premier temps, j’ai étudié le cas où une seule souche circule au sein d’une population sans vaccination. J’ai montré comme les fluctuations du taux de transmission et la fréquence des contacts entre entités spatiales – ici, des états américains – pouvaient perturber la dynamique épidémique de l’infection (Section 4.3). J’ai pu mettre en évidence que plus le taux de transmission fluctue dans le temps, plus la dynamique épidémique est marquée par des périodes de creux suivis de forts pics. À l’inverse, lorsque la fluctuation temporelle est faible le nombre d’individus infectés varie

peu entre deux pas de temps. J'ai également pu montrer que la fluctuation spatiale dans le taux de transmission jouait un rôle dans les différences de prévalence de l'infection entre états, quoique ces différences soient très largement influencées par le taux de contacts entre états.

Le modèle présente une structure en métapopulation : les taux de transmission diffèrent entre états, et les contacts entre ceux-ci varient d'un état à l'autre⁵. Il est courant en épidémiologie mathématique d'étudier la dynamique d'un pathogène au sein d'une métapopulation, par exemple pour comprendre comment les populations peuvent se synchroniser (Lloyd et Jansen, 2004), les conditions pour qu'une épidémie puisse s'établir (Colizza et Vespignani, 2008) ou encore la manière dont les politiques de restriction de mouvements peuvent altérer la dynamique épidémique (Tanaka *et al.*, 2014 ; Shao et Han, 2022). Certains résultats présentés dans la Section 4.3 suggèrent que lorsque la connectivité entre états est réduite le risque d'extinction locale et globale de l'épidémie augmente, mais en même temps le risque d'épidémies avec une forte incidence augmente. Ces résultats mériteraient de plus amples travaux mais ils suggèrent déjà que des interventions de santé publique à une échelle géographique large (ici, à l'échelle de l'état) pourraient contribuer à contrôler la propagation de certains pathogènes.

Par la suite, j'ai montré comment la compétition entre souches dans un contexte vaccinal conduisait à l'extinction de la souche désavantagée par l'effet du vaccin (Section 4.4). Ce résultat rejoint ce qui a déjà été discuté au Chapitre 2. Avec toutes les combinaisons de paramètres testées (celles présentées dans ce chapitre ainsi que d'autres non communiquées), la cohabitation est extrêmement rare et la souche 1, avantagée lorsqu'un vaccin circule, exclut compétitivement la souche 2 en quelques années. Pourtant, les données de Lefrancq *et al.* (2022) semblent montrer une coexistence des souches PRN+ et PRN- sur le moyen-terme. Ceci pourrait être lié au fait que le modèle néglige toute co-infection par les deux types de souche, un mécanisme qui pourrait expliquer le maintien d'un polymorphisme chez certains pathogènes (Gjini *et al.*, 2016). De plus, le modèle ne considère qu'indirectement la présence d'individus vaccinés, ce qui constitue certainement sa limite principale. Il serait important de rajouter un ou plusieurs compartiments prenant d'individus vaccinés (chez les susceptibles, mais également potentiellement chez les infectés) dans une version plus élaborée du modèle

5. Le modèle pourrait être interprété comme un modèle individu-centré, mais j'ai préféré éviter cette appellation dans mon manuscrit : l'individu serait ici l'état plutôt que l'individu biologique, ce qui crée une certaine confusion.

Enfin, et c'est là à la fois le cœur et l'objectif principal de ce chapitre, j'ai utilisé le modèle développé dans ce chapitre pour simuler des données destinées à être analysées via l'approche bayésienne développée dans l'axe précédent. J'ai d'abord montré que la qualité des estimations des cofacteurs Δ_X , $\Delta_{X'}$, μ_X et $\mu_{X'}$ dépendait du cofacteur considéré et de l'effort d'échantillonnage. Ainsi, les valeurs estimées des cofacteurs de type μ_f corrèlent globalement bien avec les valeurs réelles, tandis que pour les cofacteurs de type Δ_f la corrélation est beaucoup moins bonne. De plus, ces résultats changent selon le type de facteur environnemental envisagé : s'il s'agit d'un facteur influençant le taux de transmission de la souche sauvage uniquement, la corrélation des cofacteurs Δ_f semble encore plus mauvaise que si le facteur influence le taux de transmission des deux souches. Ces résultats questionnent la pertinence de l'utilisation de ces cofacteurs dans les analyses menées précédemment.

Ensuite, j'ai montré que le modèle bayésien est généralement capable de détecter le facteur temporel responsable du changement dans la proportion de la souche mutante mais que sa capacité à détecter l'effet d'un cofacteur est faible, qu'il s'agisse d'un cofacteur de type μ_f , de type Δ_f ou de type τ_f , c'est-à-dire un cofacteur non transformé. Pour les cofacteurs de type μ_f et Δ_f , il est possible que la méthode utilisée pour les calculer réduise la capacité du modèle à détecter l'effet du facteur environnemental sous-jacent. Pour les cofacteurs de type τ_f , deux phénomènes pourraient expliquer la non-détection d'un effet. Le premier est que le cofacteur est calculé comme étant la couverture vaccinale moyenne entre tous les états à un instant donné. Or, l'état épidémique et vaccinal des différents états peut être très variable au même instant, et cette moyenne nationale pourrait être trop « simple » pour correctement prendre en considération une dynamique se déroulant à une échelle plus fine. Le second phénomène est que le cofacteur τ est obtenu à partir d'une fonction – stochastique – du temps. Il en résulte que les deux sont très fortement corrélés : dans les jeux de données constitués pour les analyses, le coefficient de corrélation de Pearson moyen entre $\tau_{X'}$ et le temps est $r = 0,975$ pour $k' = t - 1$ et $r = 0,973$ pour $k' = t$. En conséquence, tout effet de $\tau_{X'}$ sur la vitesse de propagation de la souche mutante sera en réalité capté dans le modèle comme un effet du facteur temporel. Il serait donc pertinent d'utiliser un autre type de cofacteur dans le modèle, indépendant du temps et ayant un rôle uniquement sur le taux de transmission de la souche résidente.

Le modèle de simulation développé ici présente plusieurs limites. Comme je l'ai déjà noté, le

modèle est incapable de simuler adéquatement la coexistence des deux types de souche, ce qui entre en contradiction avec ce qui est observé dans les données de [Lefrancq *et al.* \(2022\)](#) mais aussi dans d'autres pays où les souches PRN⁻ et PRN⁺ coexistent (*e.g.*, [Hiramatsu *et al.*, 2017](#)). Il est possible que cette coexistence soit rendue possible par des fluctuations dans la composition des vaccins en circulation, un phénomène qui n'est pas pris en compte ici. Une autre limite du modèle est qu'il considère la population étudiée comme fermée, c'est-à-dire qu'aucune infection ne peut provenir de l'extérieur. Cette situation est relativement peu problématique pour une maladie endémique comme la coqueluche mais pourrait représenter un problème dans l'étude d'une maladie comme la diphtérie, pour laquelle la majorité des cas recensés sur le sol américain sont importés ([Bisgard *et al.*, 1998](#); [Golaz *et al.*, 2000](#)). Enfin, le modèle considère la vaccination simplement comme un mécanisme diminuant le taux de transmission de la souche toxigène. Cette simplification a été faite pour limiter le temps de calcul nécessaire à la réalisation des simulations ainsi que pour alléger les formules mathématiques utilisées – et ainsi limiter le risque d'erreurs. Il serait intéressant de compléter le modèle en ajoutant un compartiment d'individus vaccinés pour vérifier si les résultats observés ici sont consistants. Il serait également possible de considérer plusieurs compartiments de vaccinés pour étudier comment l'utilisation de plusieurs vaccins avec différentes compositions influence la dynamique épidémique et évolutive de la coqueluche.

4.7 – Conclusion

Cet axe final de ma thèse n'est pas complet. Il s'agit en réalité d'un travail préliminaire qui demande à être développé et approfondi. Si le modèle de simulation en lui-même a atteint un état satisfaisant – quoiqu'il puisse encore être amélioré, par exemple en ajoutant un compartiment pour les vaccinés, comme dit plus haut –, plusieurs aspects doivent impérativement être retouchés dans les analyses. Le premier est certainement la manière dont sont produites les valeurs du cofacteur, notamment lorsque leur calcul implique une fonction du temps. Comme présenté plus tôt, cette manière de calculer les valeurs contribue certainement à réduire la capacité de l'analyse bayésienne à détecter l'effet du cofacteur. Il conviendrait alors de trouver une nouvelle manière de calculer les valeurs du cofacteur qui soit à la fois indépendante du temps et qui influence le ratio des nombres de reproduction des deux souches en circulation.

Il aurait également été intéressant de faire fluctuer l'effort d'échantillonnage entre états pour étudier la manière dont celui-ci influence la qualité des données disponibles à l'analyse. Cependant, par manque de temps, de tels travaux n'ont pas pu être menés pendant la durée allouée à ma thèse.



BIBLIOGRAPHIE

- ALAM S. J., MEYER R. et NORLING E. (2009) : « A Model for HIV Spread in a South African Village » *Multi-Agent-Based Simulation IX*, p. 33–45, Springer (Berlin, Allemagne).
- ARIAS E., XU J. et KOCHANEK K. (23) : « Unites States Life Tables, 2021 » *National Vital Statistics Reports*, vol. 72, n°12, p./id. 1–64.
- BARKOFF A.-M., MERTSOLA J., PIERARD D., DALBY T., HOEGH S. V., GUILLOT S., STEFANELLI P., VAN GENT M., BERBERS G., VESTRHEIM D., GREVE-ISDAHL M., WEHLIN L., LJUNGMAN M., FRY N. K., MARKEY K. et HE Q. (2019) : « Pertactin-deficient *Bordetella pertussis* isolates: evidence of increased circulation in Europe, 1998 to 2015 » *Eurosurveillance*, vol. 24, n°7, p./id. 1700832.
- BERNHAEUEROVÁ V. (2016) : « Vaccine-driven evolution of parasite virulence and immune evasion in age-structured population: the case of pertussis » *Theoretical Ecology*, vol. 9, n°4, p./id. 431–442.
- BISGARD K. M., HARDY I. R., POPOVIC T., STREBEL P. M., WHARTON M., CHEN R. T. et HADLER S. C. (1998) : « Respiratory diphtheria in the United States, 1980 through 1995 » *American Journal of Public Health*, vol. 88, n°5, p./id. 787–791.
- BROOKS S. P. et GELMAN A. (1998) : « General Methods for Monitoring Convergence of Iterative Simulations » *Journal of Computational and Graphical Statistics*, vol. 7, n°4, p./id. 434–455.
- CHAO D. L., HALLORAN M. E., OBENCHAIN V. J. et LONGINI, Jr. I. M. (2010) : « FluTE, a Publicly Available Stochastic Influenza Epidemic Simulation Model » *PLOS Computational Biology*, vol. 6, n°1, p./id. e1000656.
- CLARK T. A., MESSONNIER N. E. et HADLER S. C. (2012) : « Pertussis control: time for something new? » *Trends in Microbiology*, vol. 20, n°5, p./id. 211–213.
- COLIZZA V. et VESPIGNANI A. (2008) : « Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations » *Journal of Theoretical Biology*, vol. 251, n°3, p./id. 450–467.
- DASH R., DAS R., JUNAID M., AKASH M. F. C., ISLAM A. et HOSEN S. Z. (2017) : « In silico-based vaccine design against Ebola virus glycoprotein » *Advances and Applications in Bioinformatics and Chemistry*, vol. 10, p./id. 11–28.
- DOMENECH DE CELLÈS M., ROHANI P. et KING A. A. (2019) : « Duration of Immunity and Effectiveness of Diphtheria-Tetanus-Acellular Pertussis Vaccines in Children » *JAMA Pediatrics*, vol. 173, n°6, p./id. 588–594.
- EDELMAN K., HE Q., MÄKINEN J., SAHLBERG A., HAANPERÄ M., SCHUERMAN L., WOLTER J. et MERTSOLA J. (2007) : « Immunity to Pertussis 5 Years after Booster Immunization during Adolescence » *Clinical Infectious Diseases*, vol. 44, n°10, p./id. 1271–1277.
- EDWARDS K. M. et DECKER M. D. (2012) : « Pertussis vaccines » *Vaccines* 6^{ème} édition, p. 447–492, Elsevier (Amsterdam, Pays-Bas).

-
- FINE P. E. M. et CLARKSON J. A. (1982) : « THE RECURRENCE OF WHOOPING COUGH: POSSIBLE IMPLICATIONS FOR ASSESSMENT OF VACCINE EFFICACY » *The Lancet*, vol. 319, n°8273, p./id. 666–669.
- GARDNER M. (1970) : « Mathematical Games: The fantastic combinations of John Conway’s new solitaire game “life” » *Scientific American*, vol. 223, n°4, p./id. 120–123.
- GJINI E., VALENTE C., SÁ-LEÃO R. et GOMES M. G. M. (2016) : « How direct competition shapes coexistence and vaccine effects in multi-strain pathogen systems » *Journal of Theoretical Biology*, vol. 388, p./id. 50–60.
- GOLAZ A., HARDY I. R., STREBEL P., BISGARD K. M., VITEK C., POPOVIC T. et WHARTON M. (2000) : « Epidemic Diphtheria in the Newly Independent States of the Former Soviet Union: Implications for Diphtheria Control in the United States » *The Journal of Infectious Diseases*, vol. 181, n°s1, p./id. S237–S243.
- HETHCOTE H. W. (1997) : « An Age-Structured Model for Pertussis Transmission » *Mathematical Biosciences*, vol. 145, n°2, p./id. 89–136.
- HIRAMATSU Y., MIYAJI Y., OTSUKA N., ARAKAWA Y., SHIBAYAMA K. et KAMACHI K. (2017) : « Significant Decrease in Pertactin-Deficient *Bordetella pertussis* Isolates, Japan » *Emerging Infectious Diseases*, vol. 23, n°4, p./id. 699–701.
- KLEIN N. P. (2014) : « Licensed pertussis vaccines in the United States: History and current state » *Human Vaccines & Immunotherapeutics*, vol. 10, n°9, p./id. 2684–2690.
- LEFRANCQ N., BOUCHEZ V., FERNANDES N., BARKOFF A.-M., BOSCH T., DALBY T., ÅKERLUND T., DARENBERG J., FABIANOVA K., VESTRHEIM D. F., FRY N. K., GONZÁLEZ-LÓPEZ J. J., GULLSBY K., HABINGTON A., HE Q., LITT D., MARTINI H., PIÉRARD D., STEFANELLI P., STEGGER M., ZAVADILOVA J., ARMATYS N., LANDIER A., GUILLOT S., HONG S. L., LEMEY P., PARKHILL J., TOUBIANA J., CAUCHEMEZ S., SALJE H. et BRISSE S. (2022) : « Global spatial dynamics and vaccine-induced fitness changes of *Bordetella pertussis* » *Science Translational Medicine*, vol. 14, n°642, p./id. eabn3253.
- LYOYD A. L. et JANSEN V. A. A. (2004) : « Spatiotemporal dynamics of epidemics: synchrony in meta-population models » *Mathematical Biosciences*, vol. 188, n°1-2, p./id. 1–16.
- MALEKI A., RUSSO G., PARASILITI PALUMBO G. A. et PAPPALARDO F. (2022) : « In silico design of recombinant multi-epitope vaccine against influenza A virus » *BMC Bioinformatics*, vol. 22, n°s14, p./id. 617.
- OTSUKA N., HAN H.-J., TOYOIZUMI-AJISAKA H., NAKAMURA Y., ARAKAWA Y., SHIBAYAMA K. et KAMACHI K. (2012) : « Prevalence and Genetic Characterization of Pertactin-Deficient *Bordetella pertussis* in Japan » *PLOS ONE*, vol. 7, n°2, p./id. e31985.
- PAWLOSKI L. C., QUEENAN A. M., CASSIDAY P. K., LYNCH A. S., HARRISON M. J., SHANG W., WILLIAMS M. M., BOWDEN K. E., BURGOS-RIVERA B., QIN X., MESSONNIER N. et TONDELLA M. L. (2014) : « Prevalence and Molecular Characterization of Pertactin-Deficient *Bordetella pertussis* in the United States » *Clinical and Vaccine Immunology*, vol. 21, n°2, p./id. 119–125.
- PEIRCE S. M., VAN GIESON E. J. et SKALAK T. C. (2004) : « Multicellular simulation predicts microvascular patterning and in silico tissue assembly » *The FASEB Journal*, vol. 18, n°6, p./id. 731–733.
- R CORE TEAM (2024) : *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing (Vienne, Autriche).
- RUDMAN S. M., BARBOUR M. A., CSILLÉRY K., GIENAPP P., GUILLAUME F., HAIRSTON, Jr. N. G., HENDRY A. P., LASKY J. R., RAFAJLOVIĆ M., RÄSÄNEN K., SCHMIDT P. S., SEEHAUSEN O., THERKILDSEN N. O., TURCOTTE M. M. et LEVINE J. M. (2017) : « What genomic data can reveal about eco-evolutionary dynamics » *Nature Ecology & Evolution*, vol. 2, n°1, p./id. 9–15.

- SHAO Q. et HAN D. (2022) : « Epidemic spreading in metapopulation networks with heterogeneous mobility rates » *Applied Mathematics and Computation*, vol. 412, p./id. 126559.
- SHATNAWI M., LAZAROVA-MOLNAR S. et ZAKI N. (2013) : « Modeling and simulation of epidemic spread: Recent advances » *2013 9th International Conference on Innovations in Information Technology (IIT)* p. 118–123, IEEE (Abu Dhabi, Émirats arabes unis).
- STAUFFER D., KUNWAR A. et CHOWDHURY D. (2005) : « Evolutionary ecology in silico: evolving food webs, migrating population and speciation » *Physica A: Statistical Mechanics and its Applications*, vol. 352, n°1, p./id. 202–215.
- TANAKA G., URABE C. et AIHARA K. (2014) : « Random and Targeted Interventions for Epidemic Control in Metapopulation Models » *Scientific Reports*, vol. 4, n°1, p./id. 5522.
- TEEPE J., BROEKHUIZEN B. D. L., IEVEN M., LOENS K., HUYGEN K., KRETZSCHMAR M., DE MELKER H., BUTLER C. C., LITTLE P., STUART B., COENEN S., GOOSSENS H. et VERHEIJ T. J. M. (2015) : « Prevalence, diagnosis, and disease course of pertussis in adults with acute cough: a prospective, observational study in primary care » *British Journal of General Practice*, vol. 65, n°639, p./id. e662–e667.
- TERSTAPPEN G. C. et REGGIANI A. (2001) : « In silico research in drug discovery » *Trends in Pharmaceutical Sciences*, vol. 22, n°1, p./id. 23–26.
- THOMAS M. G. (1989) : « Epidemiology of Pertussis » *Clinical Infectious Diseases*, vol. 11, n°2, p./id. 255–262.
- WEARING H. J. et ROHANI P. (2009) : « Estimating the Duration of Pertussis Immunity Using Epidemiological Signatures » *PLOS Pathogens*, vol. 5, n°10, p./id. e1000647.
- WENDELBOE A. M., VAN RIE A., SALMASO S. et ENGLUND J. A. (2005) : « Duration of Immunity Against Pertussis After Natural Infection or Vaccination » *Pediatric Infectious Disease Journal*, vol. 24, n°5, p./id. S58–S61.
- WÜRTZ P. et ANNILA A. (2010) : « Ecological succession as an energy dispersal process » *Biosystems*, vol. 100, n°1, p./id. 70–78.
- YEUNG K. H. T., DUCLOS P., NELSON E. A. S. et HUTUBESSY R. C. W. (2017) : « An update of the global burden of pertussis in children younger than 5 years: a modelling study » *The Lancet Infectious Diseases*, vol. 17, n°9, p./id. 974–980.

CHAPITRE V

DISCUSSION GÉNÉRALE ET PERSPECTIVES

5.1 – Retour sur les chapitres précédents

Alors que la vaccination est l'une des armes les plus efficaces de l'humanité pour se défendre contre les agents infectieux, elle est loin d'être parfaite. Si les conséquences médicales et épidémiologiques de la vaccination sont bien connues depuis des décennies, ses conséquences écologiques et évolutives le sont beaucoup moins. Il n'est plus possible aujourd'hui de considérer les différents acteurs épidémiques (hôtes, pathogènes, vaccins, antibiotiques...) d'un point de vue fixiste dans lequel le temps n'a d'incidence que sur la dynamique des populations, ni d'étudier l'épidémiologie dans un contexte isolé de l'environnement dans lequel les phénomènes étudiés prennent place.

Dans le Chapitre 2, j'ai cherché à comprendre comment l'utilisation de vaccins anatoxiques pouvait conduire à l'évolution de la virulence de la bactérie *Corynebacterium diphtheriae*. J'ai employé une méthode de modélisation théorique à base d'équations différentielles ordinaires (ODE) pour représenter un système dans lequel deux souches, une toxino-gène (virulente) et une non toxino-gène (avirulente), sont en compétition pour l'accès aux hôtes au sein d'une population où un vaccin ciblant spécifiquement la souche toxino-gène est en circulation. L'étude de ce modèle montre que c'est grâce à l'action conjuguée de la vaccination et de la compétition que l'éradication de la souche toxino-gène est rendue possible, illustrant l'importance de la prise en considération des processus écologiques pour comprendre l'impact d'une mesure de santé publique comme la vaccination sur l'évolution des pathogènes et, *in fine*, des maladies.

Dans le Chapitre 3, j'ai développé une approche statistique pour étudier la manière dont la vitesse de propagation d'une nouvelle souche de la bactérie *Bordetella pertussis* aux États-

Unis pouvait être influencée par des facteurs sociaux et environnementaux divers. Cette souche, dépourvue de pertactine (PRN), un antigène présent dans la plupart des vaccins acellulaires modernes, a émergé suite à l'utilisation massive de ces derniers. Les résultats de cette analyse indiquent que la température de l'environnement est, parmi les facteurs testés, celui qui explique le mieux la dynamique de propagation de cette nouvelle souche possiblement capable d'échapper à la pression vaccinale. Néanmoins, la structure du jeu de données utilisé limite la portée des conclusions.

Dans le Chapitre 4, j'ai développé un modèle informatique stochastique simulant la dynamique épidémique d'une infection à deux souches d'un même pathogène pour imiter les données utilisées dans le chapitre précédent. La particularité de ce modèle est de considérer le taux de transmission d'une ou des deux souches comme étant une fonction d'un facteur environnemental quelconque. Le but principal de ce dernier chapitre était de comprendre dans quelles situations la méthode développée au chapitre précédent pouvait être employée. Les résultats préliminaires présentés dans ce chapitre montrent que l'effet d'un facteur environnemental sur la vitesse de propagation d'une nouvelle souche d'un pathogène peut être difficilement identifié, notamment à cause de la dynamique de transmission du pathogène, indépendante du facteur étudié.

Finalement, dans ce manuscrit de thèse j'ai abordé l'éco-épidémiologie des maladies infectieuses à travers plusieurs approches mathématiques, allant de la modélisation déterministe théorique en temps continu à la modélisation stochastique en temps discret en passant par une approche statistique, toujours avec une vision centrée sur l'écologie et l'évolution des pathogènes. Ces différentes approches, quoiqu'utiles voire essentielles, présentent un certain nombre de limites que je me propose de présenter dans la section suivante.

5.2 – Limites des approches de modélisation

5.2.1 – De l'utilité des modèles

J'ai déjà cité la maxime de George Box « tous les modèles sont faux mais certains sont utiles »¹ (Box, 1976, 1979) dans le Chapitre 1. Quoique Box était statisticien, cette vérité va au-delà de la

1. « *all models are wrong but some are useful* »

modélisation mathématique. L’allégorie de la caverne de Platon nous renvoie déjà à cette idée : nous ne voyons jamais la réalité, seulement une projection de celle-ci. Prenons un exemple tout simple : une équipe de chercheurs travaille à développer un médicament contre les douleurs à l’estomac à destination des humains. Les essais précliniques impliquent de tester le médicament chez la souris. Mais ce n’est qu’un modèle : les chercheurs estiment que la souris donnera une réponse proche de celle d’un humain, tout en sachant qu’elle ne sera pas parfaitement identique. Ensuite, viennent les tests cliniques : on teste le médicament chez un groupe d’humain. Mais là encore, ce groupe n’est qu’un modèle : on l’estime représentatif de la population générale². Enfin, imaginons la situation invraisemblable où le médicament serait développé pour une personne en particulier, Mme X. Lors des tests sur cette personne, celle-ci ne serait à son tour qu’un modèle d’elle-même puisque les tests ne porteraient nécessairement que sur un fragment des états physiologiques et métaboliques dans laquelle Mme X pourrait se trouver.

Le but d’un modèle – qu’il soit biologique, statistique ou théorique – n’est donc pas de donner une description totale et parfaite de la réalité mais d’en donner un aperçu satisfaisant. Une partie de la tâche du modélisateur est de définir le niveau de divergence entre la réalité et le modèle qui soit acceptable. Ce niveau change nécessairement selon la précision de la question à laquelle veut répondre le chercheur : si on veut savoir si le médicament a un effet sur les mammifères, travailler sur une souris fonctionne aussi bien que travailler sur Mme X. En revanche, si on veut savoir si le médicament réduit les douleurs de Mme X les soirs de pleine lune après qu’elle a bu un verre de vin rouge, alors étudier la souris n’a aucun sens.

5.2.2 – Limites aux modèles dynamiques

Au cours de cette thèse, j’ai présenté différents modèles mathématiques plus ou moins complets et complexes et montrant, comme chaque modèle, des limites. Dans le Chapitre 2, j’ai développé un modèle dynamique déterministe construit à partir d’ODE. Comme je l’ai déjà exposé, ce type de modèle est classique en épidémiologie. Comme son nom l’indique, les résultats d’un tel modèle peuvent être prédits, ce qui facilite grandement les interprétations. Cependant, les

2. C’est par ailleurs rarement le cas : la majorité des essais cliniques sont faits sur des groupes biaisés, notamment par la surreprésentation des hommes dans les études mixtes, là où inversement les essais unisexes sont majoritairement féminins (Meinert et Gilpin, 2001 ; Prakash *et al.*, 2018 ; Buffenstein *et al.*, 2023 ; Chhaya *et al.*, 2023).

phénomènes biologiques sont en grande partie stochastiques, ce qui peut perturber l'adéquation entre les prédictions du modèle et la réalité, particulièrement lorsque la taille de la population considérée est faible (Allen et Burgin, 2000).

Le modèle que j'ai développé se base sur la biologie de la diphtérie pour tenter de comprendre les mécanismes à l'origine de la sélection de souches non toxigènes. La difficulté majeure de la construction de ce type de modèle est le niveau de détail à y inclure. Dans leur modèle illustrant le rôle de la vaccination anatoxinique sur l'évolution de la virulence des pathogènes, Gandon *et al.* (2001b) omettent de considérer le coût métabolique à la production d'une toxine, ce qui leur a été reproché (Soubeyrand et Plotkin, 2002) et qu'ils ont par la suite corrigé (Gandon *et al.*, 2002). Dans mon modèle, j'ai considéré le coût à la production de la toxine diphtérique à travers une réduction du taux de transmission de la souche produisant la toxine. Ce détail rajoute un niveau de complexité puisqu'il implique d'inclure de nouveaux paramètres dont les valeurs sont difficilement estimables. Un modèle plus simple (*i.e.*, ne considérant pas de coût métabolique) perdrait en précision biologique mais permettrait de s'affranchir d'une estimation hasardeuse des valeurs de certains paramètres, comme discuté plus bas.

D'autres limites à ce modèle peuvent être évoquées. Parmi ces limites, le modèle ne considère aucune forme de co-infection, un mécanisme qui pourrait influencer la compétition entre les deux souches (Alizon et van Baalen, 2008a). C'est particulièrement le cas dans un système impliquant *C. diphtheriae*. La TD participe à la destruction des cellules de l'hôte, et ainsi à la libération du fer qu'elles contiennent (Pappenheimer et Gill, 1973). Or, ce fer n'est pas directement absorbé par la bactérie ayant émis la toxine. Il est donc possible que des bactéries d'une souche non toxigène accèdent à ce fer, créant une dynamique particulière où la mise à disposition du fer par les souches toxigènes pourrait accélérer la multiplication des souches non toxigènes chez les hôtes co-infectés. Cette subtilité n'a pas été incluse dans le modèle mais il serait intéressant de la considérer.

De même, pour des raisons de simplification, le modèle « oublie » le statut vaccinal de l'hôte une fois celui-ci infecté, considérant que la vaccination n'a d'effet que sur le risque pour un individu d'être infecté par le pathogène et pas de le transmettre. Il s'agit là d'une hypothèse forte, et en réalité le statut vaccinal des individus infectés est une information qui pourrait avoir une importance capitale sur la dynamique de compétition entre les deux types de souches. Il serait

donc, là aussi, intéressant d’inclure cette information dans une future version plus élaborée du modèle.

Une des difficultés majeures rencontrée lors de l’étude d’un modèle dynamique tel que celui que j’ai développé est la paramétrisation, déjà rapidement évoquée plus haut. Plus le modèle est complexe, plus il comportera de paramètres auxquels il faudra attribuer des valeurs numériques. Dans certains cas, et notamment pour des travaux très théoriques, la paramétrisation peut être relativement peu problématique : les valeurs – ou gammes de valeurs – des paramètres peuvent être choisies en grande partie à la discrétion du modélisateur et n’ont souvent qu’un intérêt illustratif. Souvent, les conclusions importantes des travaux sont dérivées mathématiquement sans nécessité d’une application numérique ou sont généralisables à une large gamme de valeurs de paramètres, notamment dans le domaine de la modélisation de l’évolution de la virulence (*e.g.*, Gandon *et al.*, 2001a,b, 2003; Gandon, 2004; André et Gandon, 2006; Alizon et van Baalen, 2008a,b; Griette *et al.*, 2015; McLeod et Gandon, 2022).

À l’inverse, des modèles très système-spécifiques demandent idéalement des valeurs estimées avec un maximum de précision des paramètres, alors que celles-ci ne sont pas toujours disponibles (*e.g.*, Verguet *et al.*, 2015; Pesco *et al.*, 2015; Bernhauerová, 2016; Sornbundit *et al.*, 2017; Jayasundara *et al.*, 2020). Ainsi, pour paramétrer les modèles des Chapitres 2 et 4 j’ai dû faire une revue extensive de la littérature sur la diphtérie et la coqueluche, respectivement, en plus d’obtenir certaines informations démographiques, par exemple l’espérance de vie aux États-Unis. Pour chacun de ces modèles, un paramètre important est la durée de l’immunité, mais la valeur de ce paramètre est difficile à estimer que ce soit pour la diphtérie ou la coqueluche, et différents auteurs fournissent ainsi de très différentes estimations (*e.g.*, Wendelboe *et al.*, 2005; Amanna *et al.*, 2007; Wearing et Rohani, 2009; Hammarlund *et al.*, 2016).

La compartimentation, en discrétisant des phénomènes parfois continus, est elle-même une limite à ces modèles. Pour reprendre l’exemple de l’immunité, celle-ci correspond souvent en réalité à un gradient plutôt qu’à une présence/absence, or la plupart des modèles de type SIRS vont considérer qu’un individu est immunisé ou non et pas qu’il porte un certain degré d’immunité. Le moyen le plus simple de prendre en considération ce type de phénomène est certainement d’ajouter des compartiments de type « immunité partielle » (*e.g.*, Hethcote, 1997; Hethcote *et al.*, 1999; Mossong *et al.*, 1999; Fabricius *et al.*, 2013; Sornbundit *et al.*, 2017). Cependant,

comme discuté plus loin, multiplier le nombre de compartiments complexifie le modèle, et il est légitime de se questionner sur la nécessité d'une telle initiative.

5.2.3 – Limites aux modèles statistiques

Les approches statistiques ont également leurs limites. La principale est qu'elles sont tributaires de la quantité et de la qualité des données à disposition. Dans les Chapitres 3 et 4 de ma thèse, j'ai développé puis étudié un modèle bayésien dont la fonction est d'étudier des données longitudinales d'invasion et de quantifier l'effet de différents cofacteurs sur la vitesse de propagation d'une nouvelle souche. J'ai montré que pour être capable de détecter l'effet d'un cofacteur, le modèle a besoin d'un échantillon important réparti sur de nombreux pas de temps, comme attendu classiquement pour une analyse de puissance statistique (Cohen, 1992).

Évidemment, le rêve de tout statisticien est d'avoir à sa disposition une quantité illimitée de données de très bonne qualité, mais dans les faits les données sont limitées en quantité et leur qualité est parfois discutable. Le principal facteur limitant la quantité de données accessibles au chercheur est certainement le coût à leur collecte. Dans ce domaine, toutes les disciplines, méthodes et questions de recherche ne sont pas égales : il est beaucoup moins coûteux de collecter les tweets mentionnant les symptômes d'une maladie (*e.g.*, Achrekar *et al.*, 2011 ; Chen *et al.*, 2020) que de faire un suivi génétique sur plusieurs années de la prévalence d'un pathogène dans la population générale (*e.g.*, Olayinka *et al.*, 2016 ; Moosa *et al.*, 2023). Le temps est également un facteur limitant, particulièrement dans le cas d'études longitudinales. Parfois, aucune donnée n'est disponible pour des périodes éloignées dans le temps, réduisant le potentiel des études.

Une fois les données collectées, leur qualité n'est pas assurée. Différents facteurs peuvent influencer la qualité des données incluses dans une étude statistique, et ceux-ci fluctuent selon le type d'étude développée. Par exemple, pour les études biométriques ou morphométriques, les mesures peuvent fluctuer d'un observateur à l'autre pour un même objet d'étude, qu'on s'intéresse aux dimensions d'un crâne humain (Utermohle *et al.*, 1983) ou à celles d'un phytolithe (Out *et al.*, 2024), au nombre d'espèces végétales (Vittoz et Guisan, 2007) ou à l'abondance d'oiseaux (Cunningham *et al.*, 1999) dans une parcelle forestière. Plusieurs tests et statistiques ont été développés pour vérifier la consistance entre observateurs, illustrant l'importance de cette

problématique. Pour des données qualitatives, on peut utiliser les κ et κ_w de Cohen (1960, 1968) et le κ de Fleiss (1971) ; pour des données quantitatives, une grande variété de tests et statistiques sont aujourd’hui disponibles (Barnhart *et al.*, 2007).

Un autre facteur influençant la qualité des données est la précision ou l’efficacité des outils employés pour les récolter ou les mesurer. Cette problématique est plus complexe qu’un simple souci de précision sur une règle ou une balance. Prenons comme exemple la détection de souches de *B. pertussis* dans la population. Plusieurs méthodes existent pour détecter la bactérie : la mise en culture, la PCR, le test *direct fluorescent antibody* et le test sérologique. Chacune de ces méthodes a ses particularités en termes de sensibilité et de spécificité. La mise en culture est très spécifique mais peu sensible : elle ne réussit pas toujours à détecter une infection par *B. pertussis*, mais lorsqu’elle y parvient il est quasiment certain que la bactérie détectée est la bonne. Le test sérologique est plus sensible que la mise en culture tout en étant aussi spécifique, mais il est inefficace chez le nourrisson parce que son état sérologique dépend de celui de sa mère avant la naissance. Le test *direct fluorescent antibody* est très peu sensible mais spécifique, et la PCR est à la fois assez sensible et assez spécifique, tout en étant beaucoup plus rapide à conduire que la mise en culture (Lind-Brandberg *et al.*, 1998 ; Hallander, 1999 ; Loeffelholz *et al.*, 1999 ; Zouari *et al.*, 2012). La mise en culture est la méthode historique de détection de *B. pertussis* tandis que la PCR est plus récente, et c’est cette dernière qui est le plus souvent utilisée en routine de nos jours (Faulkner *et al.*, 2016). La succession et parfois la contemporanéité de ces différentes méthodes peut compliquer l’utilisation de méthodes statistiques, par exemple pour étudier la fluctuation de la prévalence de la coqueluche dans le temps. Ainsi, la modernisation des tests, associée à une augmentation de leur fréquence, a été proposée comme étant l’un des facteurs expliquant l’augmentation du nombre de cas rapportés de coqueluche aux États-Unis depuis les années 1980 (Marchant *et al.*, 1994 ; Cherry, 2003, 2005).

Enfin, les données peuvent parfois avoir été collectées en grande quantité et de manière très efficaces mais apparaître trop peu détaillées. Dans le Chapitre 3, j’explique que le jeu de données de Lefrancq *et al.* (2022) aurait été plus intéressant pour les analyses que j’ai menées si l’annotation temporelle des souches avait été plus fine³ (de l’ordre du semestre ou du trimestre

3. Évidemment, avoir accès à un tel jeu de données est exceptionnel et je ne peux que remercier les auteurs et autrices de l’avoir rendu public.

plutôt que de l'année). De la même manière, les données sont spatialisées à l'échelle nationale (pays d'origine des souches) et sub-nationale (pour les États-Unis, état d'origine des souches). Une échelle spatiale plus fine encore aurait pu aider à mieux identifier les facteurs influençant la vitesse de propagation des souches PRN – encore eût-il fallu que les données sur ces facteurs soient également disponibles à une échelle aussi fine, ce qui n'est souvent pas le cas. Cependant, collecter des données biomédicales à une petite échelle et en y adjoignant une grande quantité d'informations complémentaires peut représenter un risque pour la confidentialité de celles-ci, ce qui peut entrer en conflit avec les lois de protection des données personnelles et de secret médical.

Au-delà de la quantité et qualité des données, un autre facteur limitant des analyses statistiques est la compétence du statisticien. Aujourd'hui, une grande diversité de tests statistiques est à disposition des chercheurs, mais lorsque l'on fouille la littérature on se rend facilement compte que ces tests ne sont parfois pas maîtrisés : conditions d'application non respectées, pseudo-réplication, interprétations erronées, vocabulaire mal employé... sans parler de la simple malhonnêteté scientifique, où des données ou des tests sont manipulés pour changer les conclusions d'une analyse (Encadré 5.1).

ENCADRÉ 5.1 – Éthique de la recherche

En France, la loi impose aux écoles doctorales que chaque doctorant suive une formation à l'éthique de la recherche et à l'intégrité scientifique ^a, et les chercheurs doivent se conformer à cette éthique et cette intégrité ^b.

L'éthique en recherche est souvent assimilée au traitement des participants dans des études portant sur des sujets humains, et nombre de chercheurs se sont interrogés sur la manière de traiter éthiquement ces sujets, qu'il s'agisse d'enfants (Davidson et O'Brien, 2009), de personnes avec un retard mental (Karlavish, 2003), de réfugiés (Leaning, 2001) ou encore de groupes traumatisés (Seedat *et al.*, 2004).

Mais le concept d'éthique dans les pratiques de recherche va bien au-delà de la considération des sujets d'une expérience. Par exemple, une recherche éthique implique l'évitement du plagiat, défini comme étant l'appropriation en son nom propre du travail et des découvertes d'autrui sans en restituer la paternité (Helgesson et Eriksson, 2015).

Elle implique également un évitement de la fraude, c'est-à-dire la falsification consciente de données ou de résultats. La fraude peut être particulièrement difficile à détecter pour plusieurs raisons : les équipes de recherche sont peu enclines à reproduire les résultats d'autres

groupes par manque de visibilité de ces travaux, les données utilisées ne sont pas toujours disponibles, et les personnes détectant la fraude peuvent craindre des répercussions personnelles et professionnelles (Crocker et Cooper, 2011).

a. Arrêté du 26 août 2022 modifiant l'arrêté du 25 mai 2016 fixant le cadre national de la formation et les modalités conduisant à la délivrance du diplôme national de doctorat.

b. Articles L211-1 et L211-2 du Code de la recherche.

Aussi, même lorsque les données sont de qualité et disponibles en quantité et que le statisticien est parfaitement compétent, les modèles statistiques restent limités par certains aspects. Ainsi, ils sont généralement construits à partir d'hypothèses simplificatrices : par exemple, de nombreux modèles reposent sur l'hypothèse d'une relation monotone voire linéaire entre les variables testées, ce qui peut être un écart plus ou moins important face à la réalité en négligeant des relations parfois plus complexes entre les variables (interactions, relations non-monotones, effets seuils...). De même, des variables incluses dans un modèle peuvent présenter une colinéarité plus ou moins forte, réduisant la capacité du modèle à isoler l'effet de chaque variable.

Certains modèles complexes, parfois très système-spécifique, permettent dans certains cas de s'affranchir de ces limites. Cependant, cette complexité peut rendre difficile la communication des résultats et des conclusions des analyses à un public non-expert, que ce soit dans la communauté scientifique ou auprès du grand public, ce qui peut dans certains cas interroger la pertinence de l'utilisation de tels modèles. Enfin, les modèles statistiques autorisent la plupart à conclure à un effet corrélatif et non pas causal entre variables, ce qui est parfois difficile à comprendre, particulièrement pour le public général.

5.3 – La pluridisciplinarité de l'(éco-)épidémiologie

Si, au cours de ma thèse, j'ai centré mon approche de l'éco-épidémiologie autour de la modélisation – avec ses forces et ses faiblesses –, cette science pluridisciplinaire revêt de nombreuses formes, présentées dans les sections ci-après et résumées dans la Figure 5.1.

5.3.1 – Retour aux origines

Déterminer les origines d'une discipline est une tâche difficile. Certains considèrent le médecin britannique John Snow (1813 – 1858) comme le père fondateur de l'épidémiologie : une

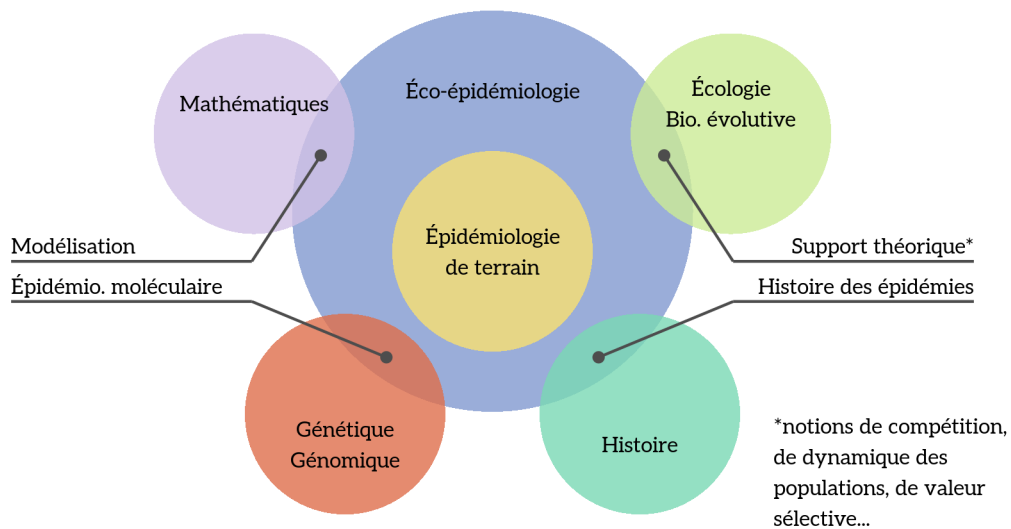


FIGURE 5.1 – **Vision personnelle de la pluridisciplinarité de l'éco-épidémiologie**, discipline à la croisée des mathématiques, de l'écologie/biologie évolutive, de la génétique/génomique et de l'histoire. Épidémiologie : épidémiologie ; bio : biologie.

plaque commémorative lui décernant ce titre est exposée au-dessus d'un *pub* portant son nom⁴. Lorsqu'une épidémie de choléra frappe le quartier londonien de Soho en 1854, il découvre que l'épidémie provenait de l'utilisation d'une pompe dont l'eau est contaminée, confortant son idée publiée quelques années plus tôt que le choléra est une maladie transmise par l'eau (Snow, 2008). En identifiant la cause de l'épidémie, il parvient à y mettre fin – simplement en retirant la poignée de la pompe.

Snow serait donc celui ayant élevé l'épidémiologie au rang de science. Cependant, bien avant ses travaux, de nombreux faits étaient connus sur les épidémies, leurs dynamiques, etc. On sait depuis longtemps que les maladies ne frappent pas au hasard et qu'il est possible de ralentir leur propagation, par exemple en instaurant des quarantaines (Gensini *et al.*, 2004). Après l'introduction de la variolisation au début du XVIII^e siècle (Section 1.2), Daniel Bernoulli montre que la mortalité due à la variole a diminué, et que le nombre de vies sauvées grâce à la variolisation excède les morts qui peuvent lui être imputées (Bernoulli et Chapelle, 2023).

On peut argumenter que l'existence de connaissances épidémiologiques qui se veulent pra-

4. Ce qui est ironique quand on sait qu'il ne buvait pas d'alcool et militait contre sa consommation.

tiques ou empiriques ne suffit pas à définir l'épidémiologie en tant que discipline scientifique. Le débat sur ce qu'est une science est complexe, et le but n'est pas ici de l'alimenter. Afin de permettre une rétrospective historique des connaissances sur les épidémies et du rapport entre les sociétés et elles, je me contenterai ici d'employer une définition large de ce qui peut être considéré comme relevant de l'épidémiologie. Ses acteurs sont multiples, allant du médecin à l'historien, du génomicien à l'écologue, du mathématicien au technicien de terrain.

5.3.2 – L'approche de terrain

La vocation du médecin est de tout temps de soigner ou guérir les malades, voire d'empêcher ses patients de tomber malade. Il est au contact des populations et peut ainsi suivre la manière dont l'infection se propage et chercher à comprendre les mécanismes expliquant comment une personne devient malade. Dès le V^e siècle avant notre ère, le médecin grec Hippocrate (*c.* 460 – *c.* 370 avant notre ère) s'intéresse à la distribution de certaines maladies et remarque que leur occurrence n'est pas aléatoire dans le temps ni dans l'espace ([Goodall, 1934](#)).

Lorsque Snow travaille sur l'épidémie de choléra qui le rendra célèbre, il le fait par une approche de terrain, au contact des populations malades. De nos jours, l'épidémiologie de terrain concentre un ensemble d'activités allant de la collecte d'échantillons biologiques à la communication avec les populations touchées. Cette approche de la discipline est de première importance, puisque c'est elle qui permet d'assurer une réponse rapide en cas de problème sanitaire ([André *et al.*, 2017](#)). Ainsi, depuis les années 1980, le CDC et l'OMS travaillent conjointement à développer des programmes pour former des épidémiologistes de terrain dans les pays à risque, notamment en Afrique ([Nsubuga *et al.*, 2008](#); [André *et al.*, 2017](#)).

5.3.3 – L'approche historique

En Europe, l'Antiquité puis le Moyen Âge fournissent une grande quantité de textes rendant compte d'épisodes épidémiques à travers les âges et les lieux. Les différents acteurs rédigeant ces textes (médecins, chroniqueurs, moines et autres lettrés) participent à cet effort de rédaction permettant de rendre compte de l'histoire épidémique du continent.

J'ai cité dans la Section [1.1.2](#) le médecin Guillaume de Baillou, qui décrit en 1578 l'épidémie

de coqueluche qui frappe Paris. Si de Baillou a été un témoin direct de l'épidémie, ce ne sont pas toujours les rédacteurs contemporains qui nous renseignent sur les épidémies d'autrefois. Le travail des historiens d'aujourd'hui nous en apprend parfois plus que les écrits anciens sur le paysage épidémique d'alors. On sait par exemple que pendant la Peste noire les écrits se taisent : personne ou presque ne parle directement des morts de l'épidémie. Les historiens reconstruisent alors, par d'autres sources (registres commerciaux, notariaux et paroissiaux notamment), le parcours de la peste.

Certains pourraient considérer que l'étude des épidémies passées à travers les écrits est du domaine de l'histoire, pas de l'épidémiologie. Je pense que les deux ne sont pas incompatibles, et que la compartimentation des disciplines ralentit les découvertes scientifiques. De nos jours, l'archéologie comme l'histoire utilisent des méthodes issues de la physique (datations par radio-isotopes...) ou de la biologie (génotypage, dendrochronologie...) pour mieux comprendre les événements et sociétés du passé. De la même manière, les outils de l'histoire et de l'archéologie peuvent permettre aux épidémiologistes de mieux connaître les épidémies du passé pour comprendre celles de l'avenir (DeWitte, 2016).

5.3.4 – L'approche moléculaire

Si les approches médicales/de terrain et historiques de l'épidémiologie ont une origine ancienne, des approches plus récentes constituent d'autres facettes de cette science. Popularisée dans les années 1990 avec l'explosion des techniques de séquençage, l'épidémiologie moléculaire a révolutionné le domaine ces dernières années. En caractérisant les pathogènes au niveau moléculaire, cette approche apporte de nombreuses réponses sur la manière dont les pathogènes évoluent. Par exemple, au cours de la récente pandémie de Covid-19, c'est l'épidémiologie moléculaire qui a permis de comprendre la succession des différents variants et la capacité de certains d'entre eux à échapper à l'immunité vaccinale (Ciotti *et al.*, 2022). Cette discipline est devenue de première importance au début du siècle, comme le montre le nombre extravagant de *reviews* publiées à cette époque sur l'épidémiologie moléculaire de parasites ou de pathogènes aussi diversifiés que le protozoaire *Giardia duodenalis* (Cacciò et Ryan, 2008), la bactérie *Mycobacterium tuberculosis* (Burgos et Pym, 2002 ; Barnes et Cave, 2003 ; Mathema *et al.*, 2006), le virus respiratoire

syncytial (Cane, 2001), l'amibe *Entamoeba histolytica* (Ali *et al.*, 2008) ou encore le champignon *Candida elegans* (McManus et Coleman, 2014).

5.3.5 – L'approche mathématique

Évidemment, les mathématiques restent un outil essentiel de l'épidémiologie. L'approche statistique permet de comprendre à des échelles spatiales et temporelles plus ou moins étendues la portée des événements épidémiques ainsi que la manière dont évoluent les pathogènes. Comme je l'ai discuté dans mes Chapitres 3 et 4, cette approche a par exemple permis de comprendre la manière dont les vaccins anticoquelucheux acellulaire ont conduit à l'émergence de souches dépourvues de la pertactine, un antigène présent dans de nombreux vaccins (Lefrancq *et al.*, 2022).

La modélisation théorique permet quant à elle de prédire le devenir d'épidémies et autres phénomènes impliquant des agents infectieux lorsque leur observation n'est pas directement possible. Le développement de modèles théoriques, généralement sous la forme d'équations mathématiques, permet par exemple d'anticiper la manière dont l'implémentation de politiques de santé publique peut perturber la dynamique d'une épidémie (*e.g.*, Fabricius *et al.*, 2013 ; Verguet *et al.*, 2015). Finalement, l'approche théorique permet à la fois de prédire *a priori* la direction dans laquelle les pathogènes évoluent et de comprendre les mécanismes à l'origine de cette évolution (*e.g.*, Kakehashi et Yoshinaga, 1992 ; Lipsitch et Nowak, 1995 ; Gandon *et al.*, 2001b ; Bernhauerová, 2016 ; Jayasundara *et al.*, 2020 ; Walter et Lion, 2021 ; Metz et Boldin, 2023 ; Torstenson et Shaw, 2024).

5.3.6 – L'approche éco-évolutive

Dans cet ultime chapitre, j'ai plusieurs fois employé le terme d'éco-épidémiologie, que j'ai généralement négligé dans les chapitres précédents. Mon opinion de néophyte dans le domaine de l'épidémiologie est qu'il est inutile aujourd'hui de distinguer « épidémiologie » et « éco-épidémiologie », une approche éco-évolutive de l'épidémiologie étant indispensable pour étudier pleinement les phénomènes infectieux. Par exemple, l'étude des goulots d'étranglement, un phénomène écologique réduisant la diversité génétique des populations suite à une diminution drastique

de population, peut nous aider à comprendre comment se structure la diversité des pathogènes, voire comment des pathogènes deviennent adaptés à l'humain (Gutiérrez *et al.*, 2012; Abel *et al.*, 2015; Weaver *et al.*, 2021). De la même manière, étudier les interactions entre différents acteurs au sein d'une communauté trophique permet de comprendre comment évoluent les pathogènes et se façonnent les dynamiques infectieuses (Parratt et Laine, 2016). Réciproquement, une approche épidémiologique apporte une contribution essentielle à l'écologie des populations : l'ubiquité des pathogènes implique un rôle important de leur part sur la dynamique des populations (Venturino, 2016).

De nombreux travaux en épidémiologie publiés ces dernières décennies incluent une approche écologique et évolutive. Une part importante d'entre eux sont des travaux de modélisation théorique. Ceux-ci incluent différents concepts écologiques ou évolutifs comme la compétition (*e.g.*, Gjini *et al.*, 2016; Lecorvaisier *et al.*, 2024), la prédation (*e.g.*, Venturino, 1994; Hilker et Malchow, 2006; Sieber et Hilker, 2011) ou encore la notion de goulot d'étranglement (*e.g.*, Geoghegan *et al.*, 2016) parmi une multitude d'autres exemples.

Des travaux plus pratiques font également usage des concepts et théories de l'écologie et de la biologie de l'évolution. Dans une expérience de laboratoire sur la malaria, de Roode *et al.* (2005) ont montré que les différents niveaux de compétition (intra-hôte et inter-hôtes) jouent un rôle dans l'évolution de la virulence du pathogène *Plasmodium chabaudi*. À partir de données environnementales collectées *in natura* (rejoignant ainsi la notion d'épidémiologie de terrain présentée plus tôt), Eby *et al.* (2023) ont montré comment les activités humaines et le climat perturbent le comportement des chauves-souris en Australie, augmentant le risque de passage de hantavirus de celles-ci vers le cheval, et donc potentiellement vers l'humain.

Finalement, l'éco-épidémiologie est une science par nature pluri-disciplinaire et au cœur des sociétés actuelles. Les chercheurs jouent un rôle de première importance dans la compréhension des maladies infectieuses, qui émergent ou réémergent avec une fréquence de plus en plus grande, en lien avec l'augmentation des activités humaines (Section 3.1.2). Mais les éco-épidémiologistes ne sont pas les seuls scientifiques à jouer un rôle important dans les sociétés.

5.4 – Quel rôle pour le chercheur et la chercheuse ?

5.4.1 – Étymologie du terme

Les termes français de chercheur et de chercheuse (ou chercheure voire, historiquement, cherchesse) ont une histoire intéressante. Ils proviennent bien évidemment du terme « chercher », qui a beaucoup évolué à travers les siècles : *cercer*, *cercier*, *cercher*, *cerchier* ou encore *sercher* sont autant de formes d’ancien ou de moyen français qui se suivent et parfois coexistent à partir du XI^e siècle. Elles proviennent du bas latin *circare*, dérivé de *circa* ou *circum* signifiant « autour » ou « faire le tour » et qui a également donné le terme « cercle »⁵. Avec un peu d’humour, on peut dire que le chercheur est « celui qui tourne en rond », mais plus exactement ce serait « celui qui fait le tour »⁶. Un chercheur ou une chercheuse est donc une personne vouant son temps et ses ressources à essayer de répondre de manière exhaustive à une ou plusieurs questions.

Le terme anglais de « *researcher* » provient lui-même du français « rechercher ». Il n’apporte donc pas beaucoup d’informations supplémentaires sur la perception du rôle du chercheur. En revanche, j’aime beaucoup la signification du terme japonais. Dans cette langue, le terme désignant un chercheur est 研究員 (Figure 5.2), romanisé en « *kenkyūin* » et prononcé « kèn-kiou-ine ». Il est composé de trois kanjis, c’est-à-dire trois idéogrammes. Le premier, 研, « *ken* », désigne l’action de polir ou d’aiguiser, ou dans un sens plus allégorique de perfectionner. Le deuxième, 究, « *kyū* », indique le fait d’étudier de manière poussée. Le dernier, 員, « *in* », désigne une personne, c’est « celui qui »⁷. Le chercheur est donc celui qui étudie pour perfectionner ses connaissances et, par extension, les connaissances humaines.

5.4.2 – Interagir avec la société

On peut considérer que les chercheur(e)s ont, particulièrement lorsque leurs recherches ou leur salaire sont financés par des fonds publics, le devoir de restituer leurs découvertes au grand nombre et notamment à travers la participation à des événements de vulgarisation scientifique.

5. *Trésor de la langue française informatisé*, ATILF, disponible sur <http://www.atilf.fr/tlfi> (dernier accès le 09/09/2024).

6. Il est amusant que le terme « quérir », venant de l’ancien français *querre*, dérivé lui-même du latin *quaerere* signifiant « chercher » ne soit aujourd’hui presque plus utilisé. La raison en est que la conjugaison du verbe chercher est plus simple que celle de quérir, ce qui a conduit à un abandon progressif de ce dernier.

7. *Kanji & Kana*, Wolfgang Hadamitzky, Pierre Durmou et Violaine Mochizuki, 2021.



FIGURE 5.2 – Calligraphie du mot 研究員 avec, de bas en haut, les caractères 研, 究, et 員. Réalisation personnelle au pinceau synthétique, encre de Chine et papier standard, avec post-traitement informatique.

Au cours de ma thèse, je n'ai pas eu beaucoup d'opportunités de participer à de tels événements. Au tout début de celle-ci (très exactement, le 9 octobre 2021), j'ai été invité à participer à la Fête de la science pour présenter mon sujet d'étude en public au cours d'un programme appelé « Ma thèse pour les nuls » dont le principe reprend celui du célèbre concours « Ma thèse en 180 secondes ». Dans mon enfance et mon adolescence – et en réalité, même dans ma vie de jeune adulte – j'ai très rarement eu l'occasion d'assister à des événements de vulgarisation. Ce n'est

que lorsque j’ai vu combien de gens se déplaçaient pour la Fête de la science que j’ai compris à quel point le public était friand de connaissances scientifiques. Finalement, quelques jours avant le rendu de ce manuscrit (le 8 octobre 2024), j’ai eu à nouveau l’occasion de participer à la Fête de la science en animant un jeu sur le microbiote à destination de collégiens.

La transmission des savoirs à travers la vulgarisation n’est en réalité pas nouvelle. Au siècle des Lumières, la science est essentiellement produite par des élites bourgeoises, assemblées en académies et laboratoires privés. Si les découvertes sont d’abord présentées à des élites intellectuelles, politiques (Louis XV était un grand amateur de ces démonstrations) et, globalement, riches, elles vont finir par atteindre un public de plus en plus largement issu du milieu populaire (Lynn, 2018).

Évidemment, il serait naïf de penser que le but final des vulgarisateurs est toujours désintéressé, qu’il ne s’agit que de transmettre au public des connaissances pour son développement et son épanouissement intellectuel. À ses débuts, la vulgarisation a une motivation essentiellement pécuniaire : mener des recherches peut coûter cher et trouver un mécène permet aux scientifiques les moins riches de poursuivre leurs recherches. Elle permet également à des acteurs extérieurs au monde de la recherche d’exercer une activité rémunératrice. Ainsi, à travers des « spectacles de science » utilisant des connaissances produites par d’autres, certains vulgarisateurs ont pu amasser fortune et célébrité dès le XVIII^e siècle (Lynn, 2018). De nos jours, la publication de magazines périodiques, très populaires en France (*Epsilon*, *Pour la Science*, *Science & Vie*, *Sciences et Avenir*...) assure la subsistance de multiples acteurs extérieurs au monde de la recherche (journalistes, éditeurs, graphistes...) tout en donnant une visibilité aux scientifiques et à leurs travaux.

Une autre manière pour les scientifiques de participer à la dissémination des connaissances est la participation à des encyclopédies collaboratives dont la plus célèbre est Wikipédia. Certes, citer son propre travail et les résultats de ses travaux de recherche sur Wikipédia peut présenter un risque de conflit d’intérêt et n’est donc pas recommandé par l’encyclopédie⁸. Néanmoins, n’importe quel chercheur peut participer à la rédaction d’un article en partageant – en sourçant – ses connaissances. Ce type d’engagement encyclopédique peut certainement être considéré

8. Plus d’informations disponibles sur https://en.wikipedia.org/wiki/Wikipedia:Conflict_of_interest#Citing_yourself (dernier accès le 19/09/2024).

comme une forme de vulgarisation, à la différence qu'il est généralement anonyme (en tout cas sur Wikipédia) là où les activités de vulgarisation plus traditionnelles apportent une certaine reconnaissance et visibilité aux chercheurs. Au cours de ma thèse, je me suis souvent investi dans l'amélioration de pages Wikipédia, profitant des connaissances que j'accumulais par mes lectures académiques. Je me suis retrouvé ainsi à écrire sur un ensemble hétéroclite de pages, parmi lesquelles se trouvent celle du kappa de Cohen, d'Edward Jenner, de l'histoire de la syphilis, de *Toxoplasma gondii*, de *Bordetella* ou encore du modèle additif. Pour cette dernière, je me suis plus particulièrement chargé de créer la page Wikipédia en français, qui n'existait pas, en traduisant la page anglaise.

Au-delà de la vulgarisation, les chercheurs doivent également être capables d'interagir avec les institutions afin d'assurer le bon déroulement de leurs travaux de recherche. La première raison est, comme pour la vulgarisation, financière. Aujourd'hui, les fonds alloués à la recherche proviennent essentiellement d'agences publiques (ANR en France, NSF aux États-Unis, UKRI au Royaume-Uni...), et les chercheurs doivent ainsi être capables de communiquer leurs projets de recherche avec ces agences pour les financer. Il est également possible – quoiqu'encore rare à ma connaissance – de solliciter le public à travers des financements participatifs. Au début de ma thèse, j'avais ainsi été sollicité pour participer à un tel financement pour un projet de recherche avec comme finalité la conservation du manul (*Otocolobus manul*) en Mongolie. Malheureusement, le projet n'a pas réuni les 3000 € nécessaires et n'a donc pas été mené à bien.

Dans une autre mesure, établir des relations de confiance avec le public et les institutions est de première importance pour développer de nombreux sujets de recherche. Ces institutions peuvent représenter des mines de données exploitables par les scientifiques, mais les intérêts divergent souvent entre le monde académique et la vie publique : les chercheurs sont encouragés à produire des recherches publiables, alors que les institutions recherchent un intérêt pratique et pouvant mener à des actions concrètes lorsque leurs données sont étudiées. Il convient donc aux chercheurs de se questionner sur la manière d'interagir avec ces institutions pour construire une relation productive entre deux mondes souvent déconnectés (Nelson *et al.*, 2015). De son côté, le public peut être sollicité pour contribuer à des projets de sciences participatives, de plus en plus répandus, dans des domaines allant des sciences de l'environnement (Fraisl *et al.*, 2022) à la santé publique (Den Broeder *et al.*, 2016). Par exemple, il peut être mis à contribution pour

identifier des espèces sur des pièges photo en amont d'une étude statistique (Pardo *et al.*, 2021 ; Nicvert *et al.*, 2024).

Finalement, les chercheur(e)s sont aujourd'hui des acteurs au cœur des sociétés, loin de l'image encore populaire du vieux savant parfois un peu fou, isolé dans son laboratoire et toute la journée en blouse blanche (Christidou, 2011). Ils et elles sont au contact des populations, à la fois pour restituer leurs découvertes, solliciter des partenariats, mais également pour prendre part aux évolutions des sociétés.

5.4.3 – Penser l'avenir

De tout temps, les scientifiques ont joué un rôle dans les grands changements qui ont dessiné l'évolution des sociétés humaines. On peut facilement penser à tous ces grands esprits d'hommes et de femmes qui apportèrent à l'humanité des technologies révolutionnèrent à travers les âges, de la poterie à la fission nucléaire, de la roue aux fusées. Mais d'autres encore ont pu changer les paradigmes, les mœurs ou les croyances en apportant des connaissances formidables : Mikolaj Kopernik (1473 – 1543) avec l'héliocentrisme, Charles Darwin (1809 – 1882) avec la théorie de l'évolution, Ada Lovelace (1815 – 1852) avec l'informatique, Albert Einstein (1879 – 1955) avec les théories de la relativité restreinte puis de la relativité générale...

Aujourd'hui encore, les scientifiques ont un rôle à jouer dans la manière dont les sociétés évoluent, notamment en apportant leurs connaissances sur des sujets de questionnement modernes, que ceux-ci touchent la manière dont sont structurées nos sociétés, à la manière dont l'humain porte atteinte aux écosystèmes ou encore à la façon dont sont dirigées les politiques de santé publique, comme je le présente ici à travers quelques exemples choisis arbitrairement – être exhaustif demanderait l'écriture d'une thèse à part entière.

Société et égalités

L'implication des scientifiques dans la vie sociale n'est pas neuve. En 1798, l'économiste britannique Thomas Malthus (1766 – 1834) publie la première édition de son *Essai sur le principe de population* dans lequel il expose sa vision des sociétés et le risque encouru par une croissance non contrôlée de la population. Par ses écrits, il militera auprès du public mais également des

politiques pour une gestion maîtrisée de la taille des populations qui conduira plus tard à la formulation du malthusianisme comme doctrine. Plus tard, ce sera au tour de Francis Galton (1822 – 1911) de s'intéresser à l'évolution des sociétés. Il initiera le mouvement eugéniste, militant pour une gestion contrôlée des naissances non pas par leur nombre mais par leur « qualité ». Les thèses de Galton auront un important succès dans la première moitié du XX^e siècle. Si on pense directement à l'Allemagne nazie, de nombreux autres gouvernements ont instauré des politiques eugénistes, que ce soit aux États-Unis, au Royaume-Uni, en Inde ou au Brésil ; et quoique fortement critiqué aujourd'hui, l'eugénisme fut pendant un temps approuvé presque unanimement par le corps scientifique comme politique (Dikötter, 1998).

Après la Seconde Guerre mondiale, de nombreux débats sociaux ont commencé à émerger, remplaçant rapidement les questions de qualité et de taille des populations : droits des femmes, des Afro-Américains, indépendance des colonies, droits des personnes LGBT, protection de la vie privée... Dans nombre de ces cas, les travaux de scientifiques, qu'ils viennent des sciences humaines et sociales ou des sciences naturelles, sont mis à contribution. Par exemple, Martin Luther King Jr. considérait que les travaux des sociologues pouvaient aider à la progression des droits civiques des Afro-Américains, notamment en étudiant et communiquant sur la condition sociale des Afro-Américains dans les états du sud des États-Unis, souvent méconnue des décideurs et des populations blanches vivant dans les états riches du Nord, et en étudiant les causes historiques comme modernes des inégalités sociales entre populations noires et blanches (King, 1968).

Un autre exemple particulièrement frappant est le rôle des scientifiques dans les questions autour du genre et du sexe des personnes. Si ces questions sont très discutées aujourd'hui, elles émergent en fait avec le féminisme post-guerre et le militantisme LGBT. L'évènement fondateur ayant ouvert la porte à ces travaux est certainement la publication de *Sex and Gender: On the Development of Masculinity and Femininity* par le psychanalyste américain Robert Stoller (1925 – 1991). Dans son livre, Stoller distingue, d'un côté, le sexe biologique (mâle ou femelle), et de l'autre le genre, qu'il définit comme étant la perception que les individus ont d'eux-mêmes comme étant mâles ou femelles (Goodwin, 1968). Ce livre ouvrira la voie à de très nombreuses études sur les questions du genre et du sexe, par exemple sur la distribution des tâches domestiques (Shaw, 1988).

De nos jours, d'autres questions émergent, notamment sur les questions de diversité des genres

et la transidentité (quoique déjà abordée par Stoller). Les travaux de sociologie permettent, par exemple, de mettre en évidence des discriminations exercées contre les personnes transgenres dans l'accès à l'emploi (Granberg *et al.*, 2020 ; Aksoy *et al.*, 2024) ou à la médecine (Winter *et al.*, 2016). Par ailleurs, la prise en considération des transidentités dans les travaux de recherche va désormais plus loin que les études portant spécifiquement sur la transidentité : dans un essai publié dans *Nature*, Ashley *et al.* (2024) invitent la communauté scientifique à repenser la manière dont sont menées les études incorporant des informations sur le genre – pas uniquement sur les transidentités – des personnes participantes dans un souci d'inclusivité, de représentativité mais également d'exactitude scientifique.

Dans le domaine médical également, les scientifiques ont un rôle à jouer sur les questions de sexe et de genre. Ils et elles permettent d'identifier l'existence de biais de genre dans les pratiques de santé. Ainsi, plusieurs études ont pu mettre en évidence un lien entre le genre du médecin ou celui des patients et le diagnostic énoncé par le médecin (Arena *et al.*, 2024). Par exemple, les médecins ont tendance à sous-diagnostiquer les problèmes cardiaques chez les femmes, notamment en attribuant les douleurs cardiaques à de l'angoisse (The EUGenMed Cardiovascular Clinical Study Group *et al.*, 2016 ; Arena *et al.*, 2024). En conséquence, le traitement est souvent différé et les femmes ont en moyenne une santé cardiaque plus mauvaise que les hommes (Azizi *et al.*, 2021). Inversement, les médecins déclarent avoir tendance à sous-diagnostiquer les problèmes de dépression ou d'anxiété chez les hommes, empêchant l'accès à un traitement chez cette partie de la population (Arena *et al.*, 2024). Ceci pourrait expliquer la différence de prévalence rapportée de ces pathologies, beaucoup plus élevée chez les femmes (Kessler *et al.*, 1994, 2003). Au-delà de simplement rapporter ces biais, les scientifiques peuvent également jouer un rôle dans leur effacement en proposant des solutions pour lutter contre eux (Geiser *et al.*, 2022 ; Arena *et al.*, 2024), et interroger ce qui relève du sexe biologique et ce qui relève du genre social dans la santé des patients (Doyal, 2003).

Environnement et biodiversité

À l'ère du réchauffement climatique et de la sixième crise de la biodiversité, les scientifiques ont un rôle d'expert à jouer pour informer les politiques et le public. En 2023, le GIEC a produit son sixième rapport d'évaluation, réparti en trois volumes représentant un total de plus de 7 400

pages et sur lesquels ont travaillé des centaines de scientifiques répartis en trois groupes de travail. Dans ces rapports, les chercheurs mais également des scientifiques extérieurs au monde académique ont rassemblé une quantité colossale d'informations scientifiques sur le réchauffement climatique pour en construire une synthèse cohérente à destination des gouvernements, pour les guider dans leurs politiques. ([Intergovernmental Panel on Climate Change, 2023a,b,c](#)).

S'il est fortement médiatisé, le travail du GIEC n'est pas le seul à intervenir dans les questions environnementales. Les rapports du GIEC traitent les questions du rôle du climat et donc du changement climatique sur les écosystèmes, ils ne s'intéressent pas à nombre de phénomènes environnementaux auxquels les scientifiques comme le public s'intéressent. Par exemple, cela fait des siècles que les voyages humains contribuent au déplacement d'espèces animales comme végétales, que ce soit par accident (rats dans les cales des bateaux, frelons asiatiques...) ou volontairement (commerce de plantes et d'animaux...). Une fois sur un nouveau territoire, les espèces introduites vont possiblement pouvoir s'y multiplier et devenir invasives (Section 3.1.2). Ces espèces introduites ont généralement mauvaise presse : elles sont accusées de provoquer des réductions d'effectif voire des extinctions chez les espèces natives ([Mollot *et al.*, 2017](#)) et leur gestion représente un coût de plusieurs milliards de dollars chaque année à l'échelle mondiale ([Diagne *et al.*, 2021](#)). En France, l'introduction d'espèces exotiques est une problématique majeure : elle est encadrée par le Code de l'Environnement (articles L 411-5 et L 411-6 notamment), et plusieurs stratégies et plans d'action ont été mis en place pour lutter contre les espèces non natives.

Les travaux des chercheurs permettent d'éclairer sur l'impact réel qu'ont les espèces introduites sur la biodiversité d'un territoire, et notamment de casser les « mythes » autour de celles-ci. Ainsi, depuis quelques années les preuves s'accumulent montrant que les espèces non natives peuvent avoir un impact positif sur l'environnement ou les populations, par exemple en représentant une nouvelle source de nourriture ou simplement en augmentant la diversité spécifique d'un territoire ([Sax *et al.*, 2022](#)). Les travaux sur les espèces introduites montrent également que celles-ci peuvent être appréciées du public, par exemple les chevaux sauvages aux États-Unis – même si ceux-ci sont responsables de dégradations des paysages –, même s'il comprend que ces espèces représentent une menace pour leur environnement ([Sharp *et al.*, 2011](#)). Or, l'intérêt du public pour une espèce, fût-elle invasive, facilite l'instauration de programmes de préservations de cette espèce qui peuvent alors rayonner et protéger d'autres espèces parfois moins spectaculaires

mais en plus grand danger d'extinction (Branton et Richardson, 2011).

Les travaux des chercheurs permettent également de vérifier la pertinence de mesures de conservation ou de restauration novatrices mais également parfois controversées comme le réensauvagement, plus connu sous son appellation anglaise de *rewilding*. Pour simplifier, le *rewilding* consiste à (ré)introduire certaines espèces sur un territoire pour qu'elles participent activement à leur restauration, en limitant généralement au maximum les interventions humaines (Perino *et al.*, 2019). L'exemple classique – quoiqu'antérieur à la théorisation du *rewilding* – de réintroduction d'une espèce ayant eu un impact majeur est celle du loup gris (*Canis lupus*) dans le parc national de Yellowstone (Smith *et al.*, 2003 ; Smith et Peterson, 2021). La réintroduction du loup, à Yellowstone comme ailleurs, est un sujet de débat public important. D'un côté, le loup représente un intérêt patrimonial et contribue à la régulation des populations d'herbivores et donc à la restauration forestière ; de l'autre côté, les éleveurs voient d'un mauvais œil ces prédateurs qui sortent des limites des parcs naturels et s'attaquent au bétail. Ainsi, les connaissances apportées par les chercheurs peuvent entrer en conflit avec les intérêts sociaux-économiques du public.

Le *rewilding* représente une autre source de débats lorsque c'est l'introduction d'une espèce nouvelle qui est considérée plutôt que la réintroduction d'une espèce historiquement présente dans un écosystème. Généralement, ce type de réintroduction est justifié par l'existence passée d'une ou plusieurs espèces jouant un rôle fort dans un écosystème et qui aurait disparu, notamment à cause des activités humaines au cours du Pléistocène ou plus récemment. On parle alors de « substitution de taxon ». Le débat porte alors sur la légitimité à transporter une espèce vers un nouvel écosystème et sur les conséquences imprévues que ces introductions pourraient avoir sur les espèces locales (Rubenstein *et al.*, 2006 ; Griffiths et Harris, 2010). Même s'ils ne sont pas nécessairement directement impliqués dans les politiques de gestion de l'environnement, les chercheurs proposant la mise en place de telles pratiques de sauvegarde des écosystèmes assument alors la responsabilité morale qu'elles échouent, voire qu'elles aggravent la dégradation des territoires dans lesquelles elles sont implémentées.

Santé publique et pandémies

S'il est une période dans laquelle la parole des scientifiques a été – parfois trop – entendue, c'est pendant la récente pandémie de Covid-19. Avides de trouver quelqu'un capable de leur expliquer

la fatalité qui leur tombait dessus et de leur trouver des solutions, les journalistes comme le public se sont concentrés autour de quelques personnalités scientifiques. On pense bien évidemment à Didier Raoult, qui a vanté les effets de la chloroquine dès la publication d'une première étude chinoise pourtant très préliminaire sur ses effets contre le virus ([Gao et al., 2020](#); [Colson et al., 2020](#)). Si d'autres chercheurs ont appelé à une tempérance ([Touret et de Lamballerie, 2020](#)), le charisme de Didier Raoult a suffi à convaincre une part importante du public de l'efficacité du médicament.

À la même période, le Président de la République française a annoncé la constitution d'un conseil scientifique composé de personnalités du monde académique et hospitalier dans le but d'aider à la prise de décision du gouvernement dans sa lutte contre la pandémie. D'autres initiatives ont émergé, comme COVISAN, portée par l'épidémiologiste Renaud Piarroux et dont le but était de casser les chaînes de transmission du virus à Paris en concentrant les efforts sur les contacts des personnes diagnostiquées positives.

Cependant, il n'y a pas eu besoin d'attendre le Covid-19 pour que les scientifiques se retrouvent impliqués dans les questions de santé publique. J'ai présenté dans la Section 1.2 l'histoire des vaccins et notamment celles des vaccins anticoquelucheux, et dans l'Encadré 1.2 j'ai traité des mouvements de contestation des vaccins. Aux origines, ce sont des médecins ou chercheurs comme Jenner et Pasteur qui développent les premiers vaccins et mettent en place des instituts pour les rendre accessibles au public, sans l'intervention des autorités publiques. Aujourd'hui encore, le développement de nouveaux vaccins est essentiel et des chercheurs issus du monde académique y contribuent. C'est par exemple le cas de la coqueluche : un très grand nombre de travaux ont été publiés ces dernières années pour proposer des alternatives aux vaccins utilisés actuellement, jugés insatisfaisants à cause de leur incapacité à la fois à induire une protection pérenne et à bloquer la transmission de la maladie ([Poolman, 2014](#); [Rumbo et Hozbor, 2014](#); [Locht, 2018](#); [Dewan et al., 2020](#)).

Le rôle des chercheurs travaillant sur les vaccins ne se limite pas à les développer : ils informent également sur les risques que ceux-ci peuvent représenter. J'ai déjà évoqué les travaux publiés dans les années 1960 et 1970 sur les risques associés à l'utilisation des vaccins anticoquelucheux (Section 1.2.3). La coqueluche n'est pas la seule maladie pour laquelle les effets secondaires des vaccins ont été dénoncés, et il est arrivé qu'une lecture précoce et unilatérale des résultats

produits par la recherche conduisent à une indignation du public face à l'obligation vaccinale.

L'exemple le plus connu est certainement le lien longtemps fantasmé entre vaccination et autisme. Tout commence lors de la publication dans *The Lancet* – la plus prestigieuse revue médicale au monde – d'un article établissant un lien entre l'utilisation du vaccin ROR et le développement de l'autisme et d'autres formes de retard mental chez les enfants vaccinés (Wakefield *et al.*, 1998). Si l'article est immédiatement critiqué par d'autres chercheurs (Chen et DeStefano, 1998), il faudra attendre presque douze ans pour qu'il soit finalement rétracté de la revue (The Editors of *The Lancet*, 2010). Entre temps, l'article a fait des dégâts : la confiance du public dans le vaccin ROR s'est effondrée (Offit et Coffin, 2003), et pendant un temps les recherches sur l'autisme se sont focalisées sur ses liens présumés avec le vaccin, faisant perdre temps et argent aux chercheurs s'intéressant à cette pathologie (The Lancet, 2002).

Ce type de scandale nous rappelle, ainsi qu'au public, que le monde de la recherche n'est pas dénué de biais, de préjugés, d'intérêts financiers et politiques. Les chercheurs et chercheuses sont humains et humaines et capables de malhonnêteté. Ils et elles doivent suivre une éthique rigoureuse (Encadré 5.1) que, malheureusement, certains et certaines ne respectent pas, parfois simplement par appât du gain ou souci de popularité, qu'elle soit académique ou publique. Les conséquences peuvent en être dramatiques, particulièrement dans le cas des sciences biomédicales où des vies peuvent être perdues à cause d'une unique publication qui, par effet boule de neige, devient responsable d'une tragédie sociétale (Chen et DeStefano, 1998 ; The Lancet, 2002).

5.5 – Questionner la recherche

Comme discuté dans la section précédente, les chercheurs et chercheuses, ou plus généralement les scientifiques, ont une place particulière dans la société. La recherche n'est pas une activité déconnectée de la réalité, une pure quête de la connaissance motivée uniquement par l'envie de déchiffrer les mystères de l'univers. La recherche est fortement politique : elle intervient sur la scène publique, que ce soit pour éduquer un public curieux, séduire une agence de financement ou prendre part aux décisions politiques et aux débats publics importants.

La recherche interagit également avec elle-même, elle se remet en question : modèles de financement, pratiques, éthique... Un sujet aussi trivial que le nom scientifique donné aux espèces

suscite des débats passionnés par publications interposées (Encadré 5.2). Alors que l'entreprise de la recherche était autrefois réservée à une élite fortunée et essentiellement européenne, la professionnalisation de la recherche scientifique a ouvert celle-ci à une gamme d'acteurs de plus en plus diversifiés. Cependant, la recherche reste fortement ancrée dans ses traditions occidentales voire capitalistes.

Depuis quelques décennies, la recherche académique est entrée dans une ère de productivisme suivant le fameux adage « *publish or perish* ». Celui-ci implique l'idée que la quantité de publications scientifiques prévôt sur leur qualité. Les structures d'accueil des scientifiques et notamment les universités se basent en grande partie sur le nombre d'articles publiés pour évaluer les chercheurs travaillant pour elles et ainsi choisir de les titulariser ou au contraire les licencier, modifier leurs salaires, autoriser ou non leurs mobilités... instituant ainsi un système compétitif où la rentabilité est le maître-mot de structures pourtant non industrielles (Miller *et al.*, 2011 ; Lee, 2014). En plus de conduire à un risque de diminution de la qualité des productions scientifiques, ce système nuit à la santé tant physique que psychologique des chercheurs et, pour les universitaires, à une diminution de la qualité de leurs enseignements, qu'ils sont obligés de sacrifier en faveur de leurs travaux de recherche (Lee, 2014).

Pour les éditeurs de revues scientifiques, en revanche, la course à la publication est une aubaine. Au 25 septembre 2024, la plateforme Web of Science de Clarivate recense plus de 33 000 journaux dans ses collections (dont certains ne sont plus publiés). L'éditeur Elsevier concentre à lui seul 2 900 journaux qui lui rapportent une quantité colossale de bénéfices : en 2023, ses bénéfices nets se sont élevés à 3,062 milliards de livres sterling, soit environ 3,660 milliards d'euros (RELX, 2024, p. 20). Si tous les éditeurs ne peuvent pas se vanter d'un bénéfice aussi conséquent, il est évident que le système d'abonnement imposé aux utilisateurs et surtout aux bibliothèques universitaires assure aux actionnaires de ces éditeurs un confortable profit. Par exemple, rien que pour l'année 2023, les abonnements (électroniques) aux revues scientifiques ont coûté un total de 1 327 167,46 € exactement aux bibliothèques universitaires de l'Université Claude Bernard Lyon 1... alors que ces abonnements excluent la majorité des revues d'Elsevier (Bibliothèques Universitaires Lyon 1, 2024, p. 32).

Une autre source majeure de bénéfice des éditeurs est la publication d'articles en accès libre, ou *open access*. De plus en plus populaires, ce format de publication est, du point de vue du

lecteur, le plus avantageux : il n'a pas besoin de souscrire à un abonnement ou de payer parfois des dizaines d'euros pour accéder à un unique PDF d'un article (qui, parfois, ne lui servira finalement pas!). En revanche, le format est plus contraignant pour l'auteur, qui doit généralement payer des frais de publication. Pour Elsevier, les journaux proposant une publication en *open access* imposent un coût variant entre 180 et 9 300 € par article⁹. Si la gratuité de lecture des articles est appréciée par la communauté scientifique, notamment dans les pays en voie de développement où les budgets alloués à la recherche sont limités, l'existence de ces frais représente dans ces mêmes pays un frein à la publication en *open access* (Papin-Ramcharan et Dawe, 2006).

En plus de représenter un coût parfois rébarbatif pour les auteurs, le modèle *open access* a une conséquence funeste : il entraîne la création de revues dites « prédatrices », c'est-à-dire des revues qui cherchent à attirer les auteurs avec des frais de publication en *open access* en général relativement peu élevés mais en produisant un travail éditorial de piètre qualité, avec une relecture minimale des articles soumis (Beall, 2013). Malheureusement, il est parfois difficile d'identifier ces revues prédatrices, d'autant plus que les articles qui y sont publiés sont pour beaucoup très souvent cités et bien référencés dans les bases de données usuellement utilisées par les chercheurs (Richtig *et al.*, 2023).

ENCADRÉ 5.2 – Taxonomie et problèmes éthiques

S'il existe un fort débat sur ce qu'est une espèce, un débat peut-être plus discret mais bien réel entoure la manière dont elles doivent être nommées. Si le débat existe depuis quelques années déjà, il s'est cristallisé autour d'un article publié par Guedes *et al.* (2023) dans *Nature Ecology & Evolution* en mars 2023^a. Dans cet article, les auteurs critiquent une tradition de la biologie : nommer les espèces d'après une personne. Ils estiment que le caractère problématique des personnes éponymes de ces espèces justifie de changer le nom de certaines d'entre elles, comme le coléoptère *Anophthalmus hitleri*, nommé d'après Adolf Hitler, et que les espèces découvertes dans le futur ne devraient plus être nommées d'après des personnes.

Ce papier a fait couler beaucoup d'encre : trois mois plus tard, *Nature Ecology & Evolution* publie une série de sept réponses aux auteurs. Selon leurs opposants, l'utilisation d'éponymes pour nommer les espèces permet à certains scientifiques de financer leurs re-

9. Disponible sur <https://www.elsevier.com/about/policies-and-standards/pricing> (dernier accès le 25/09/2024).

cherches ou des programmes de conservation et à honorer des figures locales (Jost *et al.*, 2023 ; Orr *et al.*, 2023). De plus, renommer les espèces déjà décrites créerait un certain chaos dans le monde de la biologie, notamment pour la conservation, dont les chercheurs se passeraient volontairement en ces temps de crise de la biodiversité (Orr *et al.*, 2023). Quelques mois avant la publication de l'article de Guedes *et al.* (2023) la Commission internationale de nomenclature zoologique a par ailleurs annoncé son refus de renommer des espèces jugées par certains comme étant problématiques (Ceríaco *et al.*, 2023).

Inversement, d'autres sont en faveur de changements de noms jugés offensants, et estiment que les chercheurs accordent trop d'importance à la fixité des systèmes nomenclatureaux (Raposo *et al.*, 2023). Au cours du 20^e Congrès international de botanique, 351 chercheurs contre 205 ont voté pour changer le nom de quelques centaines d'espèces de plantes afin que les épithètes basées sur la racine « ca[f][e]r », une injure raciale, soient modifiés par une racine « afra » rappelant leurs origines africaines suivant une proposition publiée quelques années plus tôt (Smith et Figueiredo, 2021 ; Callaway, 2024 ; Turland *et al.*, 2024).

Le débat reste ouvert.

a. J'ai eu l'occasion de présenter cet article au cours d'un séminaire de mon laboratoire, et je remercie tous les participants pour le débat constructif qui s'est ensuivi.

5.6 – Pour finir : quelques considérations personnelles

Dans le monde de la recherche, je ne suis qu'un nouveau-né : trois ans de thèse ne suffisent pas à en comprendre toutes les subtilités, les déboires mais également les avantages et les qualités. Mais j'ai cependant eu le temps d'observer et de vivre certaines choses, et me construire un avis sur quelques aspects de la recherche et de ce qui l'entoure.

5.6.1 – Généralités

Premièrement, j'ai pu apprécier le travail en laboratoire. Je ne peux qu'espérer que les autres laboratoires français sont des lieux aussi plaisants que le LBBE, où j'ai effectué ma thèse. Le sentiment de liberté y est puissant. Par exemple, j'ai tout au long de ma thèse été libre de mes horaires : personne ne vérifiait mes entrées et mes sorties, on me faisait confiance pour travailler mon dû. Je pouvais ainsi venir travailler très tôt – généralement, j'arrivais au laboratoire avant sept heures – et partir également tôt. Le LBBE est également un lieu de stimulation intellectuelle incroyable. Toutes les semaines, un séminaire organisé par le laboratoire nous permettait d'en

apprendre plus sur les recherches de ses membres ou d'invités extérieurs. Pour moi, écouter une personne présenter ses travaux de recherche est extrêmement motivant. En plus de ces séminaires organisés à l'échelle du laboratoire, les doctorants animent également une série de séminaires sur une base hebdomadaire à destination des non permanents. J'ai pu à plusieurs reprises faire des présentations au cours de ces séminaires : sur la manière de faire des graphiques lisibles par les personnes daltoniennes, sur la contribution à Wikipédia, sur les questions d'éthique en taxonomie (Encadré 5.2), et quelques fois pour présenter mes travaux de thèse, notamment en préparation d'une conférence.

Deuxièmement, j'ai pu être confronté à la réalité de la communication académique. Au cours de ma thèse, j'ai assisté ou participé à plusieurs conférences de différentes ampleurs. À plusieurs reprises j'ai pu y présenter les travaux de mon premier axe de thèse (je n'ai jamais eu l'occasion de présenter ceux de mes deuxième et troisième axes). La première fois, j'ai présenté mes travaux à EvoLyon, une conférence francophone biannuelle de biologie de l'évolution, en novembre 2021. C'était au tout début de ma thèse, et j'ai donc présenté un poster sur les résultats produits pendant mon stage de master 2 – les prémices de mon premier axe. En juin de l'année suivante, j'ai pu présenter à nouveau un poster sur ce premier axe de thèse au 13th International Bordetella Symposium à Vancouver au Canada. Il s'agissait cette fois d'une conférence de plus grande ampleur, sur cinq jours et en langue anglaise, réunissant quelques centaines d'experts internationaux de la coqueluche ou des bactéries du genre *Bordetella*. Enfin, j'ai pu présenter à l'oral ces mêmes travaux à deux reprises, en mars et en juillet 2023, à l'occasion d'une demi-journée *Modelling for One Health* et des rencontres du GRD Phylodynamique, respectivement.

En plus de la communication en conférences, symposiums et autres congrès, j'ai par deux fois participé au jeu de la communication des résultats de la recherche à travers la publication d'articles. Le premier article que j'ai écrit, intitulé « *Using a dynamical model to study the impact of a toxoid vaccine on the evolution of a bacterium: The example of diphtheria* », synthétise le travail de recherche présenté dans le Chapitre 2. Il a été publié en janvier 2024 après une première soumission le 4 janvier 2023 et un premier rejet le 10 mars avec invitation à resoumettre. Il a été à nouveau soumis le 6 avril et accepté le 15 novembre après révision majeure. Le deuxième article, intitulé « Impact de la vaccination sur l'évolution de *Bordetella pertussis* », est une version raccourcie et améliorée d'un mémoire de diplôme universitaire (DU) que j'ai écrit en juin 2022.

Il a été soumis pour la première fois le 23 juillet 2023, puis retourné pour révisions majeures le 27 octobre de la même année et finalement accepté le 27 novembre puis publié en février 2024.

Ces quelques conférences et publications m'ont permis d'avoir un premier aperçu de cette étape cruciale de la recherche qu'est la communication des résultats scientifiques et, plus globalement, la communication avec les pairs. J'ai pu en voir les aspects positifs mais également négatifs. Participer à des conférences est une activité plaisante, surtout lorsque la conférence correspond très bien à son domaine de compétence ; cela permet de développer ses connaissances et d'avoir des discussions enrichissantes avec les autres participants. En revanche, l'intérêt se perd lorsque la conférence est très éloignée de sa discipline. Ainsi, lors du symposium à Vancouver, je me suis rendu compte que les présentations ne correspondaient absolument pas à mon domaine de compétence. La majorité des travaux présentés traitaient d'immunologie, de biologie cellulaire, d'infectiologie ou encore de médecine, autant de disciplines sur lesquelles mes connaissances, si elles ne sont pas inexistantes, remontent au lycée ou, au mieux, à mes premières années de licence. Loin de juger de la qualité des travaux présentés, la seule présentation réellement intéressante pour moi a été celle de Noémie Lefrancq, au cours de laquelle elle a présenté ses travaux qui ont servi en partie de base à mon deuxième axe de thèse. J'ose dire qu'aller à ce symposium était une mauvaise idée, ayant coûté très cher à mon équipe et ayant eu un bilan carbone élevé pour un intérêt scientifique moindre.

Le monde de la publication n'a pas non plus été que plaisant. Travailler sur des articles, notamment en collaboration avec mes directeurs, m'a permis de grandement améliorer mes compétences de rédaction et de synthèse, et en toute honnêteté avoir son premier article publié est extrêmement plaisant – tout comme avoir sa première citation. Cependant, la lenteur du processus de publication m'est apparue très rébarbative. J'ai commencé la rédaction de mon premier article en décembre 2022, donc rapidement après le début de ma thèse et à partir des résultats préliminaires de mon stage. Cette version a été rapidement abandonnée : il manquait beaucoup trop de résultats pour qu'elle soit publiable. J'ai donc régulièrement continué à produire de nouvelles versions de l'article dans les mois suivants. Finalement, deux années après la première version de l'article, il était enfin publié. Pour mon deuxième article, les choses sont allées plus vite, notamment parce que j'étais auteur unique et ensuite parce qu'il s'agissait d'une revue bibliographique, pas d'un article de recherche. En février 2024, j'ai commencé à écrire un nouvel

article sur les résultats du deuxième axe de ma thèse. Cette fois encore, le processus est long – d’autant qu’il a été mis en pause pendant la rédaction de mon manuscrit de thèse.

L’écriture, la lecture par les co-auteurs, la réécriture, la relecture, rallongent le temps de publication des articles. C’est un fait inéluctable, surtout lorsque l’on écrit dans une optique de proposer à un éditeur la meilleure version possible de son manuscrit. Cependant, cette lenteur est frustrante, particulièrement pour un jeune chercheur. Premièrement, comme expliqué dans la Section 5.5, parce que les chercheurs sont poussés à produire un maximum d’articles pour espérer être titularisés. Deuxièmement, parce que la thèse est très cadrée temporellement, du moins en France et pour les sciences naturelles. Mon contrat doctoral couvrait trois années de salaire exactement, et l’école doctorale dont je dépendais impose de soutenir avant la fin de l’année où se termine le financement. Or, pour autoriser les doctorants à soutenir, l’école doctorale impose la publication d’au moins un article en premier auteur au cours de la thèse¹⁰. Il faut donc réussir, en trois ans, à commencer – ou dans mon cas, à continuer – un projet de recherche, le finir, rédiger un article, le corriger, l’envoyer à un éditeur, le corriger à nouveau, le renvoyer, puis qu’il soit accepté. Cette pression pousse les doctorants à vouloir à tout prix publier pendant leur thèse.

Enfin, un autre aspect que j’ai pu découvrir, parallèle à la recherche, est l’enseignement. Pendant mes trois années de doctorat, j’ai donné 208 heures de cours, principalement en statistiques et dans une moindre mesure en modélisation dynamique, pour des étudiants allant de la licence 1 au master 2. De mon avis personnel, l’enseignement est une activité plaisante mais épuisante. Préparer des cours, les restituer, répondre aux questions des étudiants, tout cela donne me donne fortement l’impression de servir à quelque chose, là où mes activités de recherche m’ont plus souvent fait me demander : « mais pourquoi est-ce que je suis là ? ». Il m’est arrivé, en effet, de me questionner sur l’intérêt de mes travaux, sur mes compétences, et sur la pertinence des approches de modélisation que j’ai mises en place au cours de ma thèse, tandis que l’enseignement m’a toujours donné un sentiment d’accomplissement.

10. Ou de trouver un troisième rapporteur de nationalité étrangère, basé dans une structure de recherche à l’étranger et ayant un statut assimilé à celui français de professeur des universités.

5.6.2 – La modélisation face la réalité du terrain

La modélisation est une approche plaisante de la biologie. Elle permet de mettre de l'ordre dans des systèmes parfois trop chaotiques, de comprendre les causes de phénomènes complexes et d'identifier les facteurs les influençant, que ce soit à travers l'utilisation de modèles dynamiques ou statistiques. Au cours de ma formation universitaire, j'ai été confronté à une multitude de modèles de différentes natures (ODE, modèles linéaires, modèles bayésiens, chaînes de Markov...) et employés dans des champs disciplinaires variés (écologie, épidémiologie, écotoxicologie... et même un peu de chimie et de physique en première année de licence). Finalement, au cours de ma thèse, j'ai conservé une certaine diversité dans les modèles employés mais en réduisant les champs disciplinaires.

Malgré mes doutes ponctuels sur mon propre travail, je reste persuadé que la modélisation est essentielle dans la recherche aujourd'hui, et notamment en (éco-)épidémiologie. Comme je l'ai discuté dans l'introduction de ce manuscrit, cette approche permet de comprendre comment se propagent les maladies, à quelle vitesse, selon quelles modalités, mais également comment les pathogènes évoluent (Section 1.5). Cependant, j'ai l'impression que le milieu de la modélisation en biologie et disciplines assimilées (dans lesquelles j'inclus l'épidémiologie) est parfois déconnecté de la réalité. J'ai ouvert ce manuscrit en citant Paul Valéry qui écrivait dans *Mauvaises pensées et autres* que « ce qui est simple est toujours faux, ce qui ne l'est pas est inutilisable ». Ces quelques mots résument bien ma vision de la modélisation. En construisant des modèles trop complexes, on s'éloigne de la biologie pour se rapprocher des mathématiques pures. Les modèles ne deviennent plus des outils au service d'une meilleure compréhension des mécanismes biologiques mais de simples jouets pour mathématiciens. Si la quête d'une connaissance mathématique pure est louable, je pense qu'il convient de rester humble dans la production de modèles pour que les résultats des études restent communicables et interprétables. En plus de cela, créer des modèles d'une trop grande complexité rend leur paramétrisation difficile, une limite que j'ai déjà évoquée plus tôt (Section 5.2.2). Si ce n'est pas nécessairement un problème pour des modèles purement théoriques, cela le devient lorsque le but du modèle est de décrire le plus fidèlement possible la réalité et de faire des prédictions.

C'est par exemple le cas lorsque l'on s'intéresse à l'effet que pourrait avoir une campagne de

vaccination sur la circulation d'une maladie. Pour citer [Schenzle \(1984\)](#), « les modélisateurs ne joueront un rôle légitime dans la mise en place de programmes vaccinaux que s'ils sont capables de montrer que leurs calculs sont quantitativement fiables »¹¹. Autrement dit, les chercheurs développant des modèles dans un objectif appliqué doivent pouvoir montrer que ceux-ci sont au moins crédibles, notamment en les paramétrant correctement à partir des données à leur disposition, s'ils veulent espérer que leurs résultats soient utilisables hors de la communauté scientifique.

Il est admis qu'un modèle ne représentera jamais totalement la réalité. Il peut tenter de s'en rapprocher au maximum mais il n'en donnera qu'un aperçu théorique et fragmentaire. C'est normal : c'est la définition même d'un modèle (Section 5.2.1). Dans certains cas, le modèle sera utile, dans d'autres non. Parfois, il sera simplement trop déconnecté des faits, du terrain. Je voudrais ici donner un exemple tiré du livre *La Vague : L'épidémie vue du terrain* de Renaud Piarroux. Dans un article publié dans *The Lancet Global Health*, [Lee et al. \(2020a\)](#) proposent une série de quatre modèles traitant de l'épidémiologie du choléra à Haïti, où la maladie sévit depuis 2010. Les conclusions du modèle sont que la probabilité d'éradiquer l'épidémie en ne considérant que les méthodes de contrôle appliquées en 2019, ces méthodes excluant l'utilisation d'un vaccin, sont très faibles, et que l'addition d'une politique vaccinale pourrait permettre de mettre un terme à l'épidémie. Sauf qu'au moment où l'étude est publiée (en août 2020), aucun cas n'a été détecté sur le territoire haïtien depuis des mois : le dernier date de janvier 2019 ([Henry et al., 2020](#) ; [Rebaudet et al., 2020](#)). Les auteurs ont par la suite répondu à ces remarques en expliquant que leur travail montre qu'une éradication de la maladie sans campagne de vaccination est théoriquement possible, quoiqu'improbable, et que l'absence de cas déclarés ne signifie pas que la maladie ne circule plus sur le territoire ([Lee et al., 2020b](#)). Cet exemple illustre la possible opposition entre les modélisateurs (et leurs modèles) et les travailleurs du terrain (et leurs données). Finalement, après plus de trois ans sans cas, le choléra est réapparu à Haïti en septembre 2022 et le profil génétique de la souche en circulation la relie à la souche responsable de l'épidémie de 2010 ([Rubin et al., 2022](#)). Il est donc possible que le pathogène ait persisté à bas bruit dans la population ou dans l'environnement, comme le proposaient [Lee et al. \(2020b\)](#).

11. « *modellers will only play a legitimate role in the planning of vaccination programs if they are able to demonstrate their calculations to be quantitatively reliable* »

Néanmoins, les prédictions de leur modèle sont restées largement déconnectées de la réalité du terrain ¹².

Finalement, quoique nécessairement complémentaires, les approches de modélisation et de terrain (ou expérimentales) sont radicalement différentes par nature. Au cours de ma thèse, je n'ai pas participé au moindre travail de terrain ni à la moindre expérience sur paillasse. Tous les ans, un appel à volontaires est fait dans le laboratoire dans lequel j'ai travaillé pour participer à un suivi sur le long cours d'une population de chevreuils dans l'ouest de la France. Chaque année, cette étude a eu lieu sur une période où j'étais occupé par les cours que je donnais, donc je n'ai jamais pu y participer. J'ai donc finalement passé les trois années de ma thèse enfermée dans mon bureau, face à mon ordinateur.

5.6.3 – Être (dé)connecté de son sujet de recherche

Travailler sur un modèle d'étude sans jamais y avoir été confronté est une expérience particulière. Pendant mon stage de master 2 et au début de ma thèse, j'ai travaillé sur la diphtérie. Si j'ai beaucoup lu sur cette maladie pour développer mon modèle, je n'y ai jamais fait face, je n'ai jamais réellement vu ni connu quelqu'un ayant souffert de la diphtérie. Pour moi, jeune chercheur vivant et travaillant dans un pays riche où l'hygiène et l'accès aux soins ont atteint un très bon niveau, la diphtérie est une maladie du passé, ou d'ailleurs : les cas observés depuis plusieurs années, quoiqu'en augmentation, sont tous importés et la maladie ne circule pas en métropole ¹³ (Badell-Ocando *et al.*, 2023).

Pour la coqueluche, que j'étudie dans mes deuxième et troisième axes, les choses sont légèrement différentes. La maladie reste présente en France et fait de temps en temps parler d'elle, comme lors de l'épidémie qui frappe la France depuis le début de l'année 2024 mais qui semble se calmer au moment où j'écris ces mots (septembre 2024). Au début de cette même année, et comme je l'ai mentionné plus tôt, j'ai publié une revue de la littérature faisant état du rôle de la vaccination sur l'évolution du pathogène de la coqueluche, *Bordetella pertussis* (Lecorvaisier, 2024). Suite à cette publication, j'ai été contacté par une mère qui m'a parlé de son enfant ma-

12. Piarroux accuse de plus les auteurs d'avoir un conflit d'intérêt non déclaré : plusieurs d'entre eux auraient reçu des fonds de la fondation Bill & Miranda Gates, qui finance des programmes de vaccination et la production d'un vaccin contre le choléra.

13. Elle est cependant endémique en Guyane, où la couverture vaccinale faible favorise son maintien.

lade de la coqueluche. Elle-même, son enfant et le père de ce dernier étaient vaccinés, et elle ne comprenait pas comment son enfant avait pu être infecté. Dans mon article, j'explique que la vaccination de masse a conduit à la sélection de souches produisant des antigènes différents de ceux contenus dans les vaccins, rendant potentiellement ces souches capables d'échapper à l'immunité induite par les vaccins. Cette mère voulait mon avis sur la possibilité que son enfant ait ainsi été infecté par une souche mutante. Par la suite, j'ai été contacté à plusieurs reprises pour mon « expertise » sur la coqueluche, notamment par l'Association du Collège des Enseignants en Immunologie et par le journal *Vaccines*. L'association me proposait d'écrire un article pour leur site Internet sur la vaccination anti-coquelucheuse, tandis que le journal me proposait de soumettre un article pour leur numéro spécial publié à l'occasion des 50 ans du Programme élargi de vaccination. J'ai refusé ces invitations pour deux raisons. La première est simplement qu'elles me sont parvenues alors que tout mon temps était monopolisé par la rédaction de mon manuscrit de thèse. La deuxième était que je ne me sentais pas légitime à y répondre positivement.

En toute franchise, être contacté pour ses travaux flatte l'ego. Néanmoins, j'ai été mal à l'aise chaque fois que c'est arrivé. Cela a été particulièrement le cas lorsque la mère de l'enfant malade m'a envoyé son mail. Je ne suis pas un expert de la coqueluche, ni médecin. Voici rapidement la genèse de l'article qui m'a valu ces différents mails : au cours de ma première année de thèse, j'ai passé un diplôme universitaire (DU) intitulé Biologie de l'Évolution et Médecine. Dans le cadre de ce DU, j'ai dû rédiger un mémoire, dont le sujet était « L'évolution de *Bordetella pertussis* dans un contexte de vaccination de masse ». Pour rédiger ce mémoire, mais également pour les besoins de ma thèse, j'ai accumulé un ensemble de connaissances sur la coqueluche, ses vaccins, et la manière dont ceux-ci ont influencé l'évolution de *B. pertussis*. Par la suite, ma directrice de thèse m'a proposé de retravailler mon mémoire pour en faire un article publiable, et ainsi est né l'article dont je discutais plus tôt. Mais je reste finalement assez déconnecté de mon sujet d'étude : je n'ai jamais été témoin des symptômes des maladies que j'ai étudiées (diphthérie comme coqueluche), et je ne me suis jamais mis en danger en étudiant ces pathologies responsables ensemble de plus d'une centaine de milliers de morts par an. J'ai simplement beaucoup lu sur ce sujet.

De manière générale, l'incidence des maladies infectieuses a beaucoup reculé dans les pays occidentaux au cours du siècle passé. Par exemple, aux États-Unis, les maladies infectieuses ont été responsables de 797 morts pour 100 000 personnes en 1900, puis ce nombre a diminué continuel-

lement pendant les huit décennies suivantes jusqu'à atteindre 36 morts pour 100 000 personnes en 1980 ([Armstrong, 1999](#)). En moins d'un siècle, la mortalité due aux maladies infectieuses a été divisé par vingt. Si elle a ensuite réaugmenté à la fin du siècle, les maladies infectieuses sont aujourd'hui loin d'être la cause majoritaire de la mortalité aux États-Unis comme dans de nombreux pays « développés ».

À l'inverse, ce n'est pas le cas dans de nombreuses autres parties du monde. Dans plusieurs pays d'Afrique de l'Ouest, plus de 10 % des enfants meurent aujourd'hui d'une infection avant d'avoir atteint l'âge de cinq ans ([Sanyang, 2019](#)). Est-ce que mes travaux permettront un jour de réduire ces chiffres ? Je l'ignore. Cependant, je ne peux qu'espérer contribuer une pierre à la fois à bâtir le mur de la connaissance qui nous permet, jour après jour, de mieux comprendre les pathogènes qui nous entourent, nous rendent malades, et parfois, nous tuent.

BIBLIOGRAPHIE

- ABEL S., ABEL ZUR WIESCH P., DAVIS B. M. et WALDOR M. K. (2015) : « Analysis of Bottlenecks in Experimental Models of Infection » *PLOS Pathogens*, vol. 11, n°6, p./id. e1004823.
- ACHREKAR H., GANDHE A., LAZARUS R., Ssu-Hsin YU et LIU B. (2011) : « Predicting flu trends using twitter data » *2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, IEEE (Shanghai, Chine).
- AKSOY B., CARPENTER C. S. et SANSONE D. (2024) : « Understanding Labor Market Discrimination Against Transgender People: Evidence from a Double List Experiment and a Survey » *Management Science*, vol. 0, n°0, p./id. mns.2023.02567.
- ALI I. K. M., CLARK C. G. et PETRI W. A. (2008) : « Molecular epidemiology of amebiasis » *Infection, Genetics and Evolution*, vol. 8, n°5, p./id. 698–707.
- ALIZON S. et VAN BAALEN M. (2008a) : « Multiple Infections, Immune Dynamics, and the Evolution of Virulence » *The American Naturalist*, vol. 172, n°4, p./id. E150–E168.
- ALIZON S. et VAN BAALEN M. (2008b) : « Transmission–virulence trade-offs in vector-borne diseases » *Theoretical Population Biology*, vol. 74, n°1, p./id. 6–15.
- ALLEN L. J. et BURGIN A. M. (2000) : « Comparison of deterministic and stochastic SIS and SIR models in discrete time » *Mathematical Biosciences*, vol. 163, n°1, p./id. 1–33.
- AMANNA I. J., CARLSON N. E. et SLIFKA M. K. (2007) : « Duration of Humoral Immunity to Common Viral and Vaccine Antigens » *New England Journal of Medicine*, vol. 357, n°19, p./id. 1903–1915.
- ANDRÉ A. M., LOPEZ A., PERKINS S., LAMBERT S., CHACE L., NOUDEKE N., FALL A. et PEDALINO B. (2017) : « Frontline Field Epidemiology Training Programs as a Strategy to Improve Disease Surveillance and Response » *Emerging Infectious Diseases*, vol. 23, n°13.
- ANDRÉ J.-B. et GANDON S. (2006) : « VACCINATION, WITHIN-HOST DYNAMICS, AND VIRULENCE EVOLUTION » *Evolution*, vol. 60, n°1, p./id. 13–23.
- ARENA F., GEISER E., AUER S., CLAIR C. et SCHWARZ J. (2024) : « Reflexivity and positionality applied to medical practice: a study on implicit gender bias with medical students in a Swiss university » *International Journal for Equity in Health*, vol. 23, n°1, p./id. 132.
- ARMSTRONG G. L. (1999) : « Trends in Infectious Disease Mortality in the United States During the 20th Century » *JAMA*, vol. 281, n°1, p./id. 61.
- ASHLEY F., BRIGHTLY-BROWN S. et RIDER G. N. (2024) : « Beyond the trans/cis binary: introducing new terms will enrich gender research » *Nature*, vol. 630, n°8016, p./id. 293–295.
- AZIZI Z., GISINGER T., BENDER U., DEISCHINGER C., RAPARELLI V., NORRIS C. M., KUBLICKIENE K., HERRERO M. T., EL EMAM K., KAUTZKY-WILLER A. et PILOTE L. (2021) : « Sex, Gender, and Cardiovascular Health in Canadian and Austrian Populations » *Canadian Journal of Cardiology*, vol. 37, n°8, p./id. 1240–1247.

-
- BADELL-OCANDO E., TOUBIANA J. et BRISSE S. (2023) : « Rapport d'Activité du Centre national de Référence des Corynébactéries du complexe diphtheriae » Institut Pasteur (Paris, France).
- BARNES P. F. et CAVE M. D. (2003) : « Molecular Epidemiology of Tuberculosis » *New England Journal of Medicine*, vol. 349, n°12, p./id. 1149–1156.
- BARNHART H. X., HABER M. J. et LIN L. I. (2007) : « AN OVERVIEW ON ASSESSING AGREEMENT WITH CONTINUOUS MEASUREMENTS » *Journal of Biopharmaceutical Statistics*, vol. 17, n°4, p./id. 529–569.
- BEALL J. (2013) : « Predatory publishing is just one of the consequences of gold open access » *Learned Publishing*, vol. 26, n°2, p./id. 79–84.
- BERNHAEUEROVÁ V. (2016) : « Vaccine-driven evolution of parasite virulence and immune evasion in age-structured population: the case of pertussis » *Theoretical Ecology*, vol. 9, n°4, p./id. 431–442.
- BERNOULLI D. et CHAPPELLE D. (2023), « Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir ».
- BIBLIOTHÈQUES UNIVERSITAIRES LYON 1 (2024) : « Rapport d'activité 2023 » Bibliothèques Universitaires Lyon 1 (Villeurbanne, France).
- BOX G. E. P. (1976) : « Science and Statistics » *Journal of the American Statistical Association*, vol. 71, n°356, p./id. 791–799.
- BOX G. E. P. (1979) : « Robustness in the Strategy of Scientific Model Building » *Robustness in Statistics*, p. 201–236, Academic Press (New York, États-Unis).
- BRANTON M. et RICHARDSON J. S. (2011) : « Assessing the Value of the Umbrella-Species Concept for Conservation Planning with Meta-Analysis: Value of the Umbrella-Species Concept » *Conservation Biology*, vol. 25, n°1, p./id. 9–20.
- BUFFENSTEIN I., KANEAKUA B., TAYLOR E., MATSUNAGA M., CHOI S. Y., CARRAZANA E., VIERECK J., LIOW K. K. et GHAFARI-RAFI A. (2023) : « Demographic recruitment bias of adults in United States randomized clinical trials by disease categories between 2008 to 2019: a systematic review and meta-analysis » *Scientific Reports*, vol. 13, p./id. 42.
- BURGOS M. V. et PYM A. S. (2002) : « Molecular epidemiology of tuberculosis » *European Respiratory Journal*, vol. 20, n°s36, p./id. 54s–65s.
- CACCIÒ S. M. et RYAN U. (2008) : « Molecular epidemiology of giardiasis » *Molecular and Biochemical Parasitology*, vol. 160, n°2, p./id. 75–80.
- CALLAWAY E. (2024) : « Hundreds of racist plant names will change after historic vote by botanists » *Nature*.
- CANE P. A. (2001) : « Molecular epidemiology of respiratory syncytial virus » *Reviews in Medical Virology*, vol. 11, n°2, p./id. 103–116.
- CERÍACO L. M. P., AESCHT E., AHYONG S. T., BALLERIO A., BOUCHARD P., BOURGOIN T., DMITRIEV D., EVENHUIS N., GRYGIER M. J., HARVEY M. S., KOTTELAT M., KLUGE N., KRELL F.-T., KOJIMA J.-I., KULLANDER S. O., LUCINDA P., LYAL C. H. C., PYLE R. L., RHEINDT F. E., LUISA SCIOSCIA C., WELTER-SCHULTES F., WHITMORE D., YANEGA D., ZHANG Z.-Q., ZHOU H.-Z. et PAPE T. (2023) : « Renaming taxa on ethical grounds threatens nomenclatural stability and scientific communication: Communication from the International Commission on Zoological Nomenclature » *Zoological Journal of the Linnean Society*, vol. 197, n°2, p./id. 283–286.

- CHEN E., LERMAN K. et FERRARA E. (2020) : « Tracking Social Media Discourse About the COVID-19 Pandemic: Development of a Public Coronavirus Twitter Data Set » *JMIR Public Health and Surveillance*, vol. 6, n°2, p./id. e19273.
- CHEN R. T. et DESTEFANO F. (1998) : « Vaccine adverse events: causal or coincidental? » *The Lancet*, vol. 351, n°9103, p./id. 611–612.
- CHERRY J. D. (2003) : « The Science and Fiction of the “Resurgence” of Pertussis » *Pediatrics*, vol. 112, n°2, p./id. 405–406.
- CHERRY J. D. (2005) : « The Epidemiology of Pertussis: A Comparison of the Epidemiology of the Disease Pertussis With the Epidemiology of *Bordetella pertussis* Infection » *Pediatrics*, vol. 115, n°5, p./id. 1422–1427.
- CHHAYA V. Y., BINION C. C., MULLES S. M., TANNHAUSER P. A., AZIZ D. Z., GREENWOOD J. D., BARLEK M. H., ROUAN J. R., WYATT T. G. et KIBBE M. R. (2023) : « Gender Bias in Clinical Trial Enrollment: Female Authorship Matters » *Annals of Vascular Surgery*, vol. 95, p./id. 233–243.
- CHRISTIDOU V. (2011) : « Interest, attitudes and images related to science: Combining students’ voices with the voices of school Science, teachers, and popular science » *International Journal of Environmental & Science Education*, vol. 6, n°2, p./id. 141–159.
- CIOTTI M., CICOZZI M., PIERI M. et BERNARDINI S. (2022) : « The COVID-19 pandemic: viral variants and vaccine efficacy » *Critical Reviews in Clinical Laboratory Sciences*, vol. 59, n°1, p./id. 66–75.
- COHEN J. (1960) : « A COEFFICIENT OF AGREEMENT FOR NOMINAL SCALE » *Educational and Psychological Measurement*, vol. 20, n°1, p./id. 37–46.
- COHEN J. (1968) : « WEIGHTED KAPPA: NOMINAL SCALE AGREEMENT WITH PROVISION FOR SCALED DISAGREEMENT OR PARTIAL CREDIT » *Psychological Bulletin*, vol. 70, n°4, p./id. 213–220.
- COHEN J. (1992) : « Statistical Power Analysis » *Current Directions in Psychological Science*, vol. 1, n°3, p./id. 98–101.
- COLSON P., ROLAIN J.-M. et RAOULT D. (2020) : « Chloroquine for the 2019 novel coronavirus SARS-CoV-2 » *International Journal of Antimicrobial Agents*, vol. 55, n°3, p./id. 105923.
- CROCKER J. et COOPER M. L. (2011) : « Addressing Scientific Fraud » *Science*, vol. 334, n°6060, p./id. 1182.
- CUNNINGHAM R. B., LINDENMAYER D. B., NIX H. A. et LINDENMAYER B. D. (1999) : « Quantifying observer heterogeneity in bird counts » *Australian Journal of Ecology*, vol. 24, n°3, p./id. 270–277.
- DAVIDSON A. J. et O’BRIEN M. (2009) : « Ethics and medical research in children » *Pediatric Anesthesia*, vol. 19, n°10, p./id. 994–1004.
- DE ROODE J. C., PANSINI R., CHEESMAN S. J., HELINSKI M. E. H., HUIJBEN S., WARGO A. R., BELL A. S., CHAN B. H. K., WALLIKER D. et READ A. F. (2005) : « Virulence and competitive ability in genetically diverse malaria infections » *Proceedings of the National Academy of Sciences*, vol. 102, n°21, p./id. 7624–7628.
- DEN BROEDER L., DEVILEE J., VAN OERS H., SCHUIT A. J. et WAGEMAKERS A. (2016) : « Citizen Science for public health » *Health Promotion International*, vol. 33, n°3, p./id. 505–514.
- DEWAN K. K., LINZ B., DEROCCO S. E. et HARVILL E. T. (2020) : « Acellular Pertussis Vaccine Components: Today and Tomorrow » *Vaccines*, vol. 8, n°2, p./id. 217.

-
- DEWITTE S. N. (2016) : « Archaeological Evidence of Epidemics Can Inform Future Epidemics » *Annual Review of Anthropology*, vol. 45, p./id. 63–77.
- DIAGNE C., LEROY B., VAISSIÈRE A.-C., GOZLAN R. E., ROIZ D., JARIĆ I., SALLES J.-M., BRADSHAW C. J. A. et COURCHAMP F. (2021) : « High and rising economic costs of biological invasions worldwide » *Nature*, vol. 592, n°7855, p./id. 571–576.
- DIKÖTTER F. (1998) : « Race Culture: Recent Perspectives on the History of Eugenics » *The American Historical Review*, vol. 103, n°2, p./id. 467–478.
- DOYAL L. (2003) : « SEX AND GENDER: THE CHALLENGES FOR EPIDEMIOLOGISTS » *International Journal of Health Services*, vol. 33, n°3, p./id. 569–579.
- EBY P., PEEL A. J., HOEGH A., MADDEN W., GILES J. R., HUDSON P. J. et PLOWRIGHT R. K. (2023) : « Pathogen spillover driven by rapid changes in bat ecology » *Nature*, vol. 613, n°7943, p./id. 340–344.
- FABRICIUS G., BERGERO P. E., ORMAZABAL M. E., MALTZ A. L. et HOZBOR D. F. (2013) : « Modelling pertussis transmission to evaluate the effectiveness of an adolescent booster in Argentina » *Epidemiology and Infection*, vol. 141, n°4, p./id. 718–734.
- FAULKNER A. E., SKOFF T. H., TONDELLA M. L., COHN A., CLARK T. A. et MARTIN S. W. (2016) : « Trends in Pertussis Diagnostic Testing in the United States, 1990 to 2012 » *Pediatric Infectious Disease Journal*, vol. 35, n°1, p./id. 39–44.
- FLEISS J. L. (1971) : « MEASURING NOMINAL SCALE AGREEMENT AMONG MANY RATERS » *Psychological Bulletin*, vol. 76, n°5, p./id. 378–382.
- FRAISL D., HAGER G., BEDESSEM B., GOLD M., HSING P.-Y., DANIELSEN F., HITCHCOCK C. B., HULBERT J. M., PIERA J., SPIERS H., THIEL M. et HAKLAY M. (2022) : « Citizen science in environmental and ecological sciences » *Nature Reviews Methods Primers*, vol. 2, n°1, p./id. 64.
- GANDON S. (2004) : « EVOLUTION OF MULTIHOST PARASITES » *Evolution*, vol. 58, n°3, p./id. 455–469.
- GANDON S., JANSEN V. A. A. et VAN BAALEN M. (2001a) : « HOST LIFE HISTORY AND THE EVOLUTION OF PARASITE VIRULENCE » *Evolution*, vol. 55, n°5, p./id. 1056.
- GANDON S., MACKINNON M., NEE S. et READ A. (2003) : « Imperfect vaccination: some epidemiological and evolutionary consequences » *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 270, p./id. 1129–1136.
- GANDON S., MACKINNON M. J., NEE S. et READ A. F. (2001b) : « Imperfect vaccines and the evolution of pathogen virulence » *Nature*, vol. 414, n°6865, p./id. 751–756.
- GANDON S., MACKINNON M. J., NEE S. et READ A. F. (2002) : « Antitoxin vaccines and pathogen virulence » *Nature*, vol. 417, n°6889, p./id. 610.
- GAO J., TIAN Z. et YANG X. (2020) : « Breakthrough: Chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies » *BioScience Trends*, vol. 14, n°1, p./id. 72–73.
- GEISER E., SCHILTER L. V., CARRIER J.-M., CLAIR C. et SCHWARZ J. (2022) : « Reflexivity as a tool for medical students to identify and address gender bias in clinical practice: A qualitative study » *Patient Education and Counseling*, vol. 105, n°12, p./id. 3521–3528.
- GENSINI G. F., YACCOUB M. H. et CONTI A. A. (2004) : « The concept of quarantine in history: from plague to SARS » *Journal of Infection*, vol. 49, n°4, p./id. 257–261.

- GEOGHEGAN J. L., SENIOR A. M. et HOLMES E. C. (2016) : « Pathogen population bottlenecks and adaptive landscapes: overcoming the barriers to disease emergence » *Proceedings of the Royal Society B: Biological Sciences*, vol. 283, n°1837, p./id. 20160727.
- GJINI E., VALENTE C., SÁ-LEÃO R. et GOMES M. G. M. (2016) : « How direct competition shapes coexistence and vaccine effects in multi-strain pathogen systems » *Journal of Theoretical Biology*, vol. 388, p./id. 50–60.
- GOODALL E. W. (1934) : « On Infectious Diseases and Epidemiology in the Hippocratic Collection » *Proceedings of the Royal Society of Medicine*, vol. 27, n°5, p./id. 525–534.
- GOODWIN D. W. (1968) : « Sex and Gender: On the Development of Masculinity and Femininity » *JAMA*, vol. 206, n°6, p./id. 1310.
- GRANBERG M., ANDERSSON P. A. et AHMED A. (2020) : « Hiring Discrimination Against Transgender People: Evidence from a Field Experiment » *Labour Economics*, vol. 65, p./id. 101860.
- GRIETTE Q., RAOUL G. et GANDON S. (2015) : « Virulence evolution at the front line of spreading epidemics » *Evolution*, vol. 69, n°11, p./id. 2810–2819.
- GRIFFITHS C. J. et HARRIS S. (2010) : « Prevention of Secondary Extinctions through Taxon Substitution » *Conservation Biology*, vol. 24, n°3, p./id. 645–646.
- GUEDES P., ALVES-MARTINS F., ARRIBAS J. M., CHATTERJEE S., SANTOS A. M. C., LEWIN A., BAKO L., WEBALA P. W., CORREIA R. A., ROCHA R. et LADLE R. J. (2023) : « Eponyms have no place in 21st-century biological nomenclature » *Nature Ecology & Evolution*, vol. 7, n°8, p./id. 1157–1160.
- GUTIÉRREZ S., MICHALAKIS Y. et BLANC S. (2012) : « Virus population bottlenecks during within-host progression and host-to-host transmission » *Current Opinion in Virology*, vol. 2, n°5, p./id. 546–555.
- HALLANDER H. O. (1999) : « Microbiological and Serological Diagnosis of Pertussis » *Clinical Infectious Diseases*, vol. 28, n°s2, p./id. S99–S106.
- HAMMARLUND E., THOMAS A., POORE E. A., AMANNA I. J., RYNKO A. E., MORI M., CHEN Z. et SLIFKA M. K. (2016) : « Durability of Vaccine-Induced Immunity Against Tetanus and Diphtheria Toxins: A Cross-sectional Analysis » *Clinical Infectious Diseases*, vol. 62, n°9, p./id. 1111–1118.
- HELGESSON G. et ERIKSSON S. (2015) : « Plagiarism in research » *Medicine, Health Care and Philosophy*, vol. 18, n°1, p./id. 91–101.
- HENRYS J. H., LEREBOURS G., ACHILLE M. A., MOISE K. et RACCURT C. (2020) : « Cholera in Haiti » *The Lancet Global Health*, vol. 8, n°12, p./id. e1469.
- HETHCOTE H. W. (1997) : « An Age-Structured Model for Pertussis Transmission » *Mathematical Biosciences*, vol. 145, n°2, p./id. 89–136.
- HETHCOTE H. W., YI L. et ZHUJUN J. (1999) : « Hopf Bifurcation in Models for Pertussis Epidemiology » *Mathematical and Computer Modelling*, vol. 30, n°11-12, p./id. 29–45.
- HILKER F. M. et MALCHOW H. (2006) : « Strange Periodic Attractors in a Prey-Predator System with Infected Prey » *Mathematical Population Studies*, vol. 13, n°3, p./id. 119–134.
- INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (2023a) : *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press (Cambridge, Royaume-Uni).
- INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (2023b) : *Climate Change 2022 – Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press (Cambridge, Royaume-Uni).

-
- INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (2023c) : *Climate Change 2022 – Mitigation of Climate Change: Working Group III Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press (Cambridge, Royaume-Uni).
- JAYASUNDARA D., LEE E., OCTAVIA S., LAN R., TANAKA M. M. et WOOD J. G. (2020) : « Emergence of pertactin-deficient pertussis strains in Australia can be explained by models of vaccine escape » *Epidemics*, vol. 31, p./id. 100388.
- JOST L., YANEZ-MUÑOZ M. H., BRITO J., REYES-PUIG C., REYES-PUIG J. P., GUAYASAMÍN J. M., RON S. R., QUINTANA C., ITURRALDE G., BAQUERO L., MONTEROS M., FREIRE-FIERRO A., FERNÁNDEZ D., MENDIETA-LEIVA G., MORALES J. F., KARREMANS A. P., VÁZQUEZ-GARCÍA J. A., SALAZAR G. A., HÁGSATER E., SOLANO R., FERNÁNDEZ-CONCHA G. C. et ARANA M. (2023) : « Eponyms are important tools for biologists in the Global South » *Nature Ecology & Evolution*, vol. 7, n°8, p./id. 1164–1165.
- KAKEHASHI M. et YOSHINAGA F. (1992) : « Evolution of airborne infectious diseases according to changes in characteristics of the host population » *Ecological Research*, vol. 7, n°3, p./id. 235–243.
- KARLAWISH J. H. (2003) : « Research Involving Cognitively Impaired Adults » *New England Journal of Medicine*, vol. 348, n°14, p./id. 1389–1392.
- KESSLER R. C., BERGLUND P., DEMLER O., JIN R., KORETZ D., MERIKANGAS K. R., RUSH A. J., WALTERS E. E. et WANG P. S. (2003) : « The Epidemiology of Major Depressive Disorder: Results From the National Comorbidity Survey Replication (NCS-R) » *JAMA*, vol. 289, n°23, p./id. 3095.
- KESSLER R. C., MCGONAGLE K. A., NELSON C. B., HUGHES M., SWARTZ M. et BLAZER D. G. (1994) : « Sex and depression in the national comorbidity survey. II: Cohort effects » *Journal of Affective Disorders*, vol. 30, n°1, p./id. 15–26.
- KING M. L. (1968) : « THE ROLE OF THE BEHAVIORAL SCIENTIST IN THE CIVIL RIGHTS MOVEMENT » *American Psychologist*, vol. 23, n°3, p./id. 180–186.
- LEANING J. (2001) : « Ethics of research in refugee populations » *The Lancet*, vol. 357, n°9266, p./id. 1432–1433.
- LECORVAISIER F. (2024) : « Impact de la vaccination sur l'évolution de *Bordetella pertussis* » *médecine/sciences*, vol. 40, n°2, p./id. 161–166.
- LECORVAISIER F., PONTIER D., SOUBEYRAND B. et FOUCHET D. (2024) : « Using a dynamical model to study the impact of a toxoid vaccine on the evolution of a bacterium: The example of diphtheria » *Ecological Modelling*, vol. 487, p./id. 110569.
- LEE E. C., CHAO D. L., LEMAITRE J. C., MATRAJT L., PASETTO D., PEREZ-SAEZ J., FINGER F., RINALDO A., SUGIMOTO J. D., HALLORAN M. E., LONGINI I. M., TERNIER R., VISSIERES K., AZMAN A. S., LESSLER J. et IVERS L. C. (2020a) : « Achieving coordinated national immunity and cholera elimination in Haiti through vaccination: a modelling study » *The Lancet Global Health*, vol. 8, n°8, p./id. e1081–e1089.
- LEE E. C., TERNIER R., LESSLER J., AZMAN A. S. et IVERS L. C. (2020b) : « Cholera in Haiti – Authors' reply » *The Lancet Global Health*, vol. 8, n°12, p./id. e1470–e1471.
- LEE I. (2014) : « Publish or perish: The myth and reality of academic publishing » *Language Teaching*, vol. 47, n°2, p./id. 250–261.
- LEFRANCQ N., BOUCHEZ V., FERNANDES N., BARKOFF A.-M., BOSCH T., DALBY T., ÅKERLUND T., DARENBERG J., FABIANOVA K., VESTRHEIM D. F., FRY N. K., GONZÁLEZ-LÓPEZ J. J., GULLSBY K., HABINGTON A., HE Q., LITT D., MARTINI H., PIÉRARD D., STEFANELLI P., STEGGER M., ZAVADILOVA J., ARMATYS N., LANDIER A., GUILLOT S., HONG S. L., LEMEY P., PARKHILL J., TOUBIANA J.,

- CAUCHEMEZ S., SALJE H. et BRISSE S. (2022) : « Global spatial dynamics and vaccine-induced fitness changes of *Bordetella pertussis* » *Science Translational Medicine*, vol. 14, n°642, p./id. eabn3253.
- LIND-BRANDBERG L., WELINDER-OLSSON C., LAGERGÅRD T., TARANGER J., TROLLFORS B. et ZACKRISSON G. (1998) : « Evaluation of PCR for Diagnosis of *Bordetella pertussis* and *Bordetella parapertussis* Infections » *Journal of Clinical Microbiology*, vol. 36, n°3, p./id. 679–683.
- LIPSITCH M. et NOWAK M. A. (1995) : « The Evolution of Virulence in Sexually Transmitted HIV/AIDS » *Journal of Theoretical Biology*, vol. 174, n°4, p./id. 427–440.
- LOCHT C. (2018) : « Will we have new pertussis vaccines? » *Vaccine*, vol. 36, n°36, p./id. 5460–5469.
- LOEFFELHOLZ M. J., THOMPSON C. J., LONG K. S. et GILCHRIST M. J. R. (1999) : « Comparison of PCR, Culture, and Direct Fluorescent-Antibody Testing for Detection of *Bordetella pertussis* » *Journal of Clinical Microbiology*, vol. 37, n°9, p./id. 2872–2876.
- LYNN M. R. (2018) : *Popular Science and Public Opinion in Eighteenth-Century France*, Manchester University Press (Manchester, Royaume-Uni).
- MARCHANT C. D., LOUGHLIN A. M., LETT S. M., TODD C. W., WETTERLOW L. H., BICCHIERI R., HIGHAM S., ETKIND P., SILVA E. et SIBER G. R. (1994) : « Pertussis in Massachusetts, 1981-1991: Incidence, Serologic Diagnosis, and Vaccine Effectiveness » *Journal of Infectious Diseases*, vol. 169, n°6, p./id. 1297–1305.
- MATHEMA B., KUREPINA N. E., BIFANI P. J. et KREISWIRTH B. N. (2006) : « Molecular Epidemiology of Tuberculosis: Current Insights » *Clinical Microbiology Reviews*, vol. 19, n°4, p./id. 658–685.
- MCLEOD D. V. et GANDON S. (2022) : « Effects of epistasis and recombination between vaccine-escape and virulence alleles on the dynamics of pathogen adaptation » *Nature Ecology & Evolution*, vol. 6, n°6, p./id. 786–793.
- MCMANUS B. A. et COLEMAN D. C. (2014) : « Molecular Epidemiology, Phylogeny and Evolution of *Candida albicans* » *Infection, Genetics and Evolution*, vol. 21, p./id. 166–178.
- MEINERT C. L. et GILPIN A. K. (2001) : « Estimation of gender bias in clinical trials » *Statistics in Medicine*, vol. 20, n°8, p./id. 1153–1164.
- METZ J. A. J. et BOLDIN B. (2023) : « Immunity-driven evolution of virulence and diversity in respiratory diseases » *Evolution*, vol. 77, n°11, p./id. 2392–2408.
- MILLER A. N., TAYLOR S. G. et BEDEIAN A. G. (2011) : « Publish or perish: academic life as management faculty live it » *Career Development International*, vol. 16, n°5, p./id. 422–445.
- MOLLOT G., PANTEL J. H. et ROMANUK T. N. (2017) : « The Effects of Invasive Species on the Decline in Species Richness: A Global Meta-Analysis » *Networks of Invasion: A Synthesis of Concepts* numéro 56 de *Advances in Ecological Research*, p. 61–83, Academic Press (Londres, Royaume-Uni).
- MOOSA F., TEMPIA S., KLEYNHANS J., MCMORROW M., MOYES J., DU PLESSIS M., CARRIM M., TREURNICHT F. K., HELFERSCEE O., MKHENCELE T., MATHUNJWA A., MARTINSON N. A., KAHN K., LEBINA L., WAFAWANAKA F., COHEN C., VON GOTTEBERG A. et WOLTER N. (2023) : « Incidence and Transmission Dynamics of *Bordetella pertussis* Infection in Rural and Urban Communities, South Africa, 2016-2018 » *Emerging Infectious Diseases*, vol. 29, n°2, p./id. 294–303.
- MOSSONG J., NOKES D. J., EDMUNDS W. J., COX M. J., RATNAM S. et MULLER C. P. (1999) : « Modeling the Impact of Subclinical Measles Transmission in Vaccinated Populations with Waning Immunity » *American Journal of Epidemiology*, vol. 150, n°11, p./id. 1238–1249.
- NELSON I. A., LONDON R. A. et STROBEL K. R. (2015) : « Reinventing the Role of the University Researcher » *Educational Researcher*, vol. 44, n°1, p./id. 17–26.

-
- NICVERT L., DONNET S., KEITH M., PEEL M., SOMERS M. J., SWANEPOEL L. H., VENTER J., FRITZ H. et DRAY S. (2024) : « Using the multivariate Hawkes process to study interactions between multiple species from camera trap data » *Ecology*, vol. 105, n°4, p./id. e4237.
- NSUBUGA P., WHITE M., FONTAINE R. et SIMONE P. (2008) : « Training programmes for field epidemiology » *The Lancet*, vol. 371, n°9613, p./id. 630–631.
- OFFIT P. A. et COFFIN S. E. (2003) : « Communicating science to the public: MMR vaccine and autism » *Vaccine*, vol. 22, n°1, p./id. 1–6.
- OLAYINKA A. T., OYEMAKINDE A., BALOGUN M. S., AJUDUA A., NGUKU P., ADERINOLA M., EGWUENU-OLADEJO A., AJISEGIRI S. W., SHA'AIBU S., MUSA B. O. P., GIDADO S. et NASIDI A. (2016) : « Seroprevalence of Hepatitis B Infection in Nigeria: A National Survey » *The American Society of Tropical Medicine and Hygiene*, vol. 95, n°4, p./id. 902–907.
- ORR M. C., HUGHES A. C., CARVAJAL O. T., FERRARI R. R., LUO A., RAJAEI H., RON S. R., WARRIT N., ZAMANI A., ZHANG Y. M. et ZHU C.-D. (2023) : « Inclusive and productive ways forward needed for species-naming conventions » *Nature Ecology & Evolution*, vol. 7, n°8, p./id. 1168–1169.
- OUT W. A., EVETT R., HOŠKOVÁ K., POWER R. C., RUIZ-PÉREZ J., TROMP M., VRYDAGHS L., WADE K. et HASLER M. (2024) : « Inter- and intra-observer variation in phytolith morphometry » *Annals of Botany*, p. mcae116.
- PAPIN-RAMCHARAN J. et DAWE R. A. (2006) : « The Other Side of the Coin for Open Access Publishing – A Developing Country View » *Libri*, vol. 56, n°1.
- PAPPENHEIMER, Jr. A. M. et GILL D. M. (1973) : « Diphtheria: Recent studies have clarified the molecular mechanisms involved in its pathogenesis » *Science*, vol. 182, n°4110, p./id. 353–358.
- PARDO L. E., BOMBACI S. P., HUEBNER S., SOMERS M. J., FRITZ H., DOWNS C., GUTHMANN A., HETEM R. S., KEITH M., LE ROUX A., MGQATSA N., PACKER C., PALMER M. S., PARKER D. M., PEEL M., SLOTOW R., STRAUSS W. M., SWANEPOEL L., TAMBLING C., TSIE N., VERMEULEN M., WILLI M., JACHOWSKI D. S. et VENTER J. A. (2021) : « Snapshot Safari: A large-scale collaborative to monitor Africa's remarkable biodiversity » *South African Journal of Science*, vol. 117, n°1-2.
- PARRATT S. R. et LAINE A.-L. (2016) : « The role of hyperparasitism in microbial pathogen ecology and evolution » *The ISME Journal*, vol. 10, n°8, p./id. 1815–1822.
- PERINO A., PEREIRA H. M., NAVARRO L. M., FERNÁNDEZ N., BULLOCK J. M., CEAUȘU S., CORTÉS-AVIZANDA A., VAN KLINK R., KUEMMERLE T., LOMBA A., PE'ER G., PLIENINGER T., REY BENAYAS J. M., SANDOM C. J., SVENNING J.-C. et WHEELER H. C. (2019) : « Rewilding complex ecosystems » *Science*, vol. 364, n°6438, p./id. eaav5570.
- PESCO P., BERGERO P., FABRICIUS G. et HOZBOR D. (2015) : « Mathematical modeling of delayed pertussis vaccination in infants » *Vaccine*, vol. 33, n°41, p./id. 5475–5480.
- POOLMAN J. T. (2014) : « Shortcomings of pertussis vaccines: why we need a third generation vaccine » *Expert Review of Vaccines*, vol. 13, n°10, p./id. 1159–1162.
- PRAKASH V. S., MANSUKHANI N. A., HELENOWSKI I. B., WOODRUFF T. K. et KIBBE M. R. (2018) : « Sex Bias in Interventional Clinical Trials » *Journal of Women's Health*, vol. 27, n°11, p./id. 1342–1348.
- RAPOSO M. A., DA SILVA H. R., FRANCISCO B. C. S., VIEIRA O., DA FONSECA O. V., DE ASSIS C. P., KIRWAN G. M., CAPDEVILE T. G., LOURENÇO A. C. C., SOBRAL G., BOCKMANN F. A., SELVATTI A. P., SOUZA DIAS P. G. B., ZACCA T., SOUZA-GUDINHO F., DE CARVALHO M. R., TARGINO M. et FRAZÃO A. (2023) : « Is stability too revered in zoological nomenclature? » *Zoological Journal of the Linnean Society*, vol. 199, n°1, p./id. 7–9.

- REBAUDET S., GAUDART J. et PIARROUX R. (2020) : « Cholera in Haiti » *The Lancet Global Health*, vol. 8, n°12, p./id. e1468.
- RELX (2024) : « 2023 Annual Report including Financial Statements and Corporate Responsibility Report » RELX (Amsterdam, Pays-Bas).
- RICHTIG G., BERGER M., KOELLER M., RICHTIG M., RICHTIG E., SCHEFFEL J., MAURER M. et SIEBENHAAR F. (2023) : « Predatory journals: Perception, impact and use of Beall's list by the scientific community—A bibliometric big data study » *PLOS ONE*, vol. 18, n°7, p./id. e0287547.
- RUBENSTEIN D. R., RUBENSTEIN D. I., SHERMAN P. W. et GAVIN T. A. (2006) : « Pleistocene Park: Does re-wilding North America represent sound conservation for the 21st century? » *Biological Conservation*, vol. 132, n°2, p./id. 232–238.
- RUBIN D. H. F., ZINGL F. G., LEITNER D. R., TERNIER R., COMPERE V., MARSEILLE S., SLATER D., HARRIS J. B., CHOWDHURY F., QADRI F., BONCY J., IVERS L. C. et WALDOR M. K. (2022) : « Reemergence of Cholera in Haiti » *New England Journal of Medicine*, vol. 387, n°25, p./id. 2387–2389.
- RUMBO M. et HOZBOR D. (2014) : « Development of improved pertussis vaccine » *Human Vaccines & Immunotherapeutics*, vol. 10, n°8, p./id. 2450–2453.
- SANYANG Y. (2019) : « Prevalence of under-five years of age mortality by infectious diseases in West African region » *International Journal of Africa Nursing Sciences*, vol. 11, p./id. 100175.
- SAX D. F., SCHLAEPFER M. A. et OLDEN J. D. (2022) : « Valuing the contributions of non-native species to people and nature » *Trends in Ecology & Evolution*, vol. 37, n°12, p./id. 1058–1066.
- SCHENZLE D. (1984) : « An Age-Structured Model of Pre- and Post-Vaccination Measles Transmission » *Mathematical Medicine and Biology*, vol. 1, n°2, p./id. 169–191.
- SEEDAT S., PIENAAR W. P., WILLIAMS D. et STEIN D. J. (2004) : « Ethics of research on survivors of trauma » *Current Psychiatry Reports*, vol. 6, n°4, p./id. 262–267.
- SHARP R. L., LARSON L. R. et GREEN G. T. (2011) : « Factors influencing public preferences for invasive alien species management » *Biological Conservation*, vol. 144, n°8, p./id. 2097–2104.
- SHAW S. M. (1988) : « Gender Differences in the Definition and Perception of Household Labor » *Family Relations*, vol. 37, n°3, p./id. 333.
- SIEBER M. et HILKER F. M. (2011) : « Prey, predators, parasites: intraguild predation or simpler community modules in disguise? » *Journal of Animal Ecology*, vol. 80, n°2, p./id. 414–421.
- SMITH D. W. et PETERSON R. O. (2021) : « Intended and unintended consequences of wolf restoration to Yellowstone and Isle Royale National Parks » *Conservation Science and Practice*, vol. 3, n°4, p./id. e413.
- SMITH D. W., PETERSON R. O. et HOUSTON D. B. (2003) : « Yellowstone after Wolves » *BioScience*, vol. 53, n°4, p./id. 330–340.
- SMITH G. F. et FIGUEIREDO E. (2021) : « (126) Proposal to add a new Article 61.6 to permanently and retroactively eliminate epithets with the root *caff[e]r-* or *caff[e]r-* from the nomenclature of algae, fungi and plants » *TAXON*, vol. 70, n°6, p./id. 1395–1396.
- SNOW S. J. (2008) : « John Snow: the making of a hero? » *The Lancet*, vol. 372, n°9632, p./id. 22–23.
- SORNBUNDIT K., TRIAMPO W. et MODCHANG C. (2017) : « Mathematical modeling of diphtheria transmission in Thailand » *Computers in Biology and Medicine*, vol. 87, p./id. 162–168.

-
- SOUBEYRAND B. et PLOTKIN S. A. (2002) : « Antitoxin vaccines and pathogen virulence » *Nature*, vol. 417, n°6889, p./id. 609–610.
- THE EDITORS OF THE LANCET (2010) : « Retraction—Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children » *The Lancet*, vol. 375, n°9713, p./id. 445.
- THE EUGENMED CARDIOVASCULAR CLINICAL STUDY GROUP, REGITZ-ZAGROSEK V., OERTELT-PRIGIONE S., PRESCOTT E., FRANCONI F., GERDTS E., FORYST-LUDWIG A., MAAS A. H. E. M., KAUTZKY-WILLER A., KNAPPE-WEGNER D., KINTSCHER U., LUDWIG K. H., SCHENCK-GUSTAFSSON K. et STANGL V. (2016) : « Gender in cardiovascular diseases: impact on clinical manifestations, management, and outcomes » *European Heart Journal*, vol. 37, n°1, p./id. 24–34.
- THE LANCET (2002) : « Time to look beyond MMR in autism research » *The Lancet*, vol. 359, n°9307, p./id. 637.
- TORSTENSON M. et SHAW A. K. (2024) : « Pathogen evolution following spillover from a resident to a migrant host population depends on interactions between host pace of life and tolerance to infection » *Journal of Animal Ecology*, vol. 93, n°4, p./id. 1365–2656.14075.
- TOURET F. et DE LAMBALLERIE X. (2020) : « Of chloroquine and COVID-19 » *Antiviral Research*, vol. 177, p./id. 104762.
- TURLAND N. J., ÁLVAREZ I., KNAPP S., MONRO A. M. et WIERSEMA J. H. (2024) : « XX International Botanical Congress, Madrid 2024: Report of Congress action on nomenclature proposals » *TAXON*.
- UTERMOHLE C. J., ZEGURA S. L. et HEATHCOTE G. M. (1983) : « Multiple Observers, Humidity, and Choice of Precision Statistics: Factors Influencing Craniometric Data Quality » *American Journal of Physical Anthropology*, vol. 61, n°1, p./id. 85–95.
- VENTURINO E. (1994) : « THE INFLUENCE OF DISEASES ON LOTKA-VOLTERRA SYSTEMS » *Rocky Mountain Journal of Mathematics*, vol. 24, n°1, p./id. 381–402.
- VENTURINO E. (2016) : « Ecoepidemiology: a More Comprehensive View of Population Interactions » *Mathematical Modelling of Natural Phenomena*, vol. 11, n°1, p./id. 49–90.
- VERGUET S., JOHRI M., MORRIS S. K., GAUVREAU C. L., JHA P. et JIT M. (2015) : « Controlling measles using supplemental immunization activities: A mathematical model to inform optimal policy » *Vaccine*, vol. 33, n°10, p./id. 1291–1296.
- VITTOZ P. et GUISAN A. (2007) : « How reliable is the monitoring of permanent vegetation plots? A test with multiple observers » *Journal of Vegetation Science*, vol. 18, n°3, p./id. 413–422.
- WAKEFIELD A. J., MURCH S. H., ANTHONY A., LINNELL J., CASSON D. M., MALIK M., BERELOWITZ M., DHILLON A. P., THOMSON M. A., HARVEY P., VALENTINE A., DAVIES S. E. et WALKER-SMITH J. A. (1998) : « Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children » *The Lancet*, vol. 351, n°9103, p./id. 637–641.
- WALTER A. et LION S. (2021) : « Epidemiological and evolutionary consequences of periodicity in treatment coverage » *Proceedings of the Royal Society B: Biological Sciences*, vol. 288, n°1946, p./id. 20203007.
- WEARING H. J. et ROHANI P. (2009) : « Estimating the Duration of Pertussis Immunity Using Epidemiological Signatures » *PLOS Pathogens*, vol. 5, n°10, p./id. e1000647.
- WEAVER S. C., FORRESTER N. L., LIU J. et VASILAKIS N. (2021) : « Population bottlenecks and founder effects: implications for mosquito-borne arboviral emergence » *Nature Reviews Microbiology*, vol. 19, n°3, p./id. 184–195.

- WENDELBOE A. M., VAN RIE A., SALMASO S. et ENGLUND J. A. (2005) : « Duration of Immunity Against Pertussis After Natural Infection or Vaccination » *Pediatric Infectious Disease Journal*, vol. 24, n°5, p./id. S58–S61.
- WINTER S., DIAMOND M., GREEN J., KARASIC D., REED T., WHITTLE S. et WYLIE K. (2016) : « Transgender people: health at the margins of society » *The Lancet*, vol. 388, n°10042, p./id. 390–400.
- ZOUARI A., SMAOUI H. et KECHRID A. (2012) : « The diagnosis of pertussis: which method to choose? » *Critical Reviews in Microbiology*, vol. 38, n°2, p./id. 111–121.

ACRONYMES

- ACIP** Advisory Committee on Immunization Practices. 162
- ACS** American Community Survey. 158
- ACV** Vaccin acellulaire (*Acellular vaccine* en anglais). 40, 41, 62–68, 206, 211
- AIC** Critère d'information d'Akaike (*Akaike Information Criterion* en anglais). 178
- BUGS** Bayesian inference Using Gibbs Sampling. 178
- CDC** Centers for Diseases Control and Prevention. 162, 163, 251
- Covid-19** Maladie à coronavirus 2019. 5, 7, 13, 30, 35, 38, 53, 58, 71, 79, 104, 133, 143, 144, 155, 252, 263, 264
- CPI** Indice des prix à la consommation (*Consumer Price Index* en anglais). 161
- DFE** Équilibre sans maladie (*Disease-Free Equilibrium* en anglais). 114, 115
- DIC** Critère d'information de déviance (*Deviance Information Criterion* en anglais). 178, 180–182
- DU** Diplôme universitaire. 269, 275
- FHA** Hémagglutinine filamenteuse. 40, 62, 68
- Fim** Fimbriae. 40, 62
- GHCND** Global Historical Climate Network daily. 156
- GIEC** Groupe d'experts intergouvernemental sur l'évolution du climat. 261, 262
- GLM** Modèle linéaire généralisé (*Generalized Linear Model* en anglais). 21, 22, 168–171, 173–176, 180, 181, 188, 189, 191
- GLMM** Modèle linéaire généralisé mixte (*Generalized Linear Mixed Model* en anglais). 22, 145, 170–173, 175, 176, 180, 181, 189, 191
- INSEE** Institut national de la statistique et des études économiques. 111
- JAGS** Just Another Gibbs Sampler. 178, 179
- LGBT** lesbiennes, gays, bisexuels et transgenres. 260
- LRT** Test du rapport des vraisemblances (*Likelihood-Ratio Test* en anglais). 167–174
- LUCA** Dernier ancêtre commun universel (*Last Universal Common Ancestor* en anglais). 25, 26
- MCMC** Monte-Carlo par chaînes de Markov (*Markov chain Monte Carlo* en anglais). 146
- MDV** Virus de la maladie de Marek (*Marek's disease virus* en anglais). 58, 59, 144
- MERS-CoV** Coronavirus du syndrome respiratoire du Moyen-Orient (*Middle East respiratory syndrome-related coronavirus* en anglais). 31

MYXV Virus de la myxomatose (*Myxoma virus* en anglais). 54-56, 74

NCHS National Center for Health Statistics. 206

ODE Équation différentielle ordinaire (*Ordinary Differential Equation* en anglais). 69, 241, 243, 272

OMS Organisation mondiale de la santé. 35, 39, 57, 79, 103, 111, 162, 251

PCR Réaction en chaîne par polymérase (*Polymerase Chain Reaction* en anglais). 247

PEV Programme élargi de vaccination. 57

PRN Pertactine. 5-8, 20, 40, 61-63, 66-68, 145, 146, 151-153, 162-168, 170, 173-175, 177, 178, 180, 181, 183, 188-191, 199, 206, 207, 232, 234, 242, 248

ROR Rougeole, oreillons, rubéole. 265

RPV Virus de la peste bovine (*Rinderpest virus* en anglais). 30

SARS-CoV Coronavirus du syndrome respiratoire aigu sévère (*Severe acute respiratory syndrome coronavirus* en anglais). 31

SIDA Syndrome d'immunodéficience acquise. 30, 50

SIR Susceptible - Infecté - Rétabli (ou Guéri). 5, 7, 70, 71, 73, 108

TD Toxine diphtérique. 32, 61, 65, 106, 107, 109, 110, 113, 121, 129, 130, 133, 134, 141, 142, 145, 244

TP Toxine pertussique. 40, 61, 62, 64-66, 68, 69, 145, 152

USCB United States Census Bureau. 155, 158, 159, 161, 207

VIH Virus de l'immunodéficience humaine. 30, 49, 50, 52, 75, 202

WAIFW Qui acquiert l'infection de qui (*Who Acquires Infection From Whom* en anglais). 72

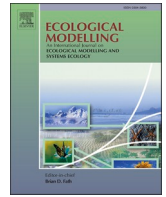
WCV Vaccin à germe entier (*whole-cell vaccine* en anglais). 39, 40, 62-64, 66, 67, 104, 206, 211

ANNEXES

A – Articles publiés

Dans cette section se trouvent les versions éditées des deux articles que j'ai publiés pendant la durée de ma thèse. Les deux articles ont été publiés en accès ouvert sous licence Creative Commons (CC-BY)¹⁴. Cette licence est différente de celle du corps principal du présent manuscrit (CC-BY-NC) et ces articles peuvent donc être réutilisés dans des conditions différentes du reste du manuscrit de la thèse.

14. Les spécificités de la licence sont précisées sur <https://creativecommons.org/licenses/by/4.0/> (dernier accès le 18/09/2024).



Using a dynamical model to study the impact of a toxoid vaccine on the evolution of a bacterium: The example of diphtheria

Florian Lecorvaisier^{a,*}, Dominique Pontier^a, Benoît Soubeyrand^b, David Fouchet^a

^a UMR CNRS 5558 Laboratoire de Biométrie et Biologie Evolutive (LBBE), Université Claude Bernard Lyon 1, Université de Lyon, 69622, Villeurbanne, France

^b Blossom Vaccinology, 13 rue d'Algérie, 69001, Lyon, France

ARTICLE INFO

Keywords:

Virulence evolution
Diphtheria
Imperfect vaccination
Competition

ABSTRACT

Vaccines are one of humankind's greatest weapons against infectious diseases. However, their usefulness is often questioned and the public tends to distrust vaccines. A mathematical model published in the early 2000s predicts the selection of more virulent strains of pathogens when populations are protected with imperfect vaccines, i.e., vaccines which reduce but do not entirely block pathogen transmission, such as toxoid vaccines. In this study, we built a disease-specific competition model to analyze the evolution of diphtheria's virulence under the pressure of a toxoid vaccine. Our results show that i) high vaccine coverage favors the emergence and increase prevalence of avirulent (or less virulent) strains of *Corynebacterium diphtheriae* (the etiologic agent of diphtheria) and ii) that competition between strains is crucial in the eradication of toxigenic strains when toxoid vaccines are used. We conclude that the use of toxoid vaccines could lead to disease eradication if the interaction between strains is taken into account. Our results could extend to biologically similar systems such as pertussis.

1. Introduction

Vaccination, one of the great achievements of the 20th century, was initially thought of as a means of protection against the disease or its severe symptoms – for vaccinated individuals. Nevertheless, its action goes beyond the individual to extend the population. When a pathogenic agent is introduced within a population, we define the basic reproduction number of the infection, denoted R_0 , as the number of secondary cases induced by one infectious individual within this population. Theoretically, knowing this epidemic threshold parameter is an important step to estimate the vaccine coverage of an infectious disease to reach for this R_0 to fall below one (Fine, 1993), as an epidemic likely fades out when R_0 becomes less than unity.

On May 8, 1980, the World Health Organization (WHO) declared the eradication of smallpox worldwide. Eradication became possible thanks to a high vaccine coverage leading to what we call herd immunity. This term encapsulates the capacity of a population to protect itself from a pathogen thanks to a collective immunity level high enough for the pathogen to be unable to spread among the population (Fine, 1993). As variola virus could only infect humans and transmission could no longer be sustained in the vaccinated populations, smallpox could only die out.

This kind of population-level effect of vaccination can only be seen if

a high enough proportion of the population is vaccinated with a vaccine blocking the transmission of the pathogen it targets. It is classically accepted that the critical proportion of vaccine coverage to reach for a disease to go extinct is given with the formula $p_c = 1 - 1/R_0$ (Fine, 1993). In practice, this is not this simple, as populations are highly heterogeneous, and the behavior of individuals toward vaccination may have complex and sometimes hardly predictable effects on the outcome of an epidemic (Lu et al., 2019; Zou et al., 2023). This phenomenon has been exemplified in the recent COVID-19 pandemic, where the influence of anti-vaccination groups, notably on social networks, may have increased the pandemic length and burden (e.g., Pullan and Dey, 2021).

Nevertheless, with both individual- and population-level benefits, vaccination seemed to be the panacea to every infectious disease. Yet, vaccines can also have deleterious effects. Some vaccines are known to have severe side effects, such as for example whole-cell pertussis vaccines (Edwards and Decker, 2012). Mass vaccination also had unexpected consequences, as for example it has been linked to an increase in the mean age of first infection for measles (Li et al., 2017) and pertussis (Lavine et al., 2011; Rohani et al., 2010), two infections that disproportionately affect children. It has been shown, among adolescents and adults, who are now more frequently infected by childhood diseases like measles and pertussis, that age increases the risk of severe symptoms and

* Corresponding author.

E-mail address: florian.lecorvaisier@univ-lyon1.fr (F. Lecorvaisier).

<https://doi.org/10.1016/j.ecolmodel.2023.110569>

Received 6 April 2023; Received in revised form 15 November 2023; Accepted 15 November 2023

Available online 26 November 2023

0304-3800/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

hospitalization (e.g., De Serres et al., 2000 for pertussis; Giladi et al., 1987; Yasunaga et al., 2010 for measles). Vaccination, by targeting specific strains in a population, also shapes the evolution of pathogens by altering selection pressures on them, thus favoring bacterial or viral strains evading the vaccine-induced immunity (Gandon and Day, 2008; McLeod and Gandon, 2022).

Another possible effect of mass vaccination on the evolution of pathogens is the selection of strains of different virulence. In the early 2000s, Gandon and coauthors, using a mathematical model, have theorized that vaccines which block the onset of symptoms but not the transmission nor the growth of pathogens – so called “imperfect” or “leaky” vaccines – could lead to the selection of more virulent strains because they remove the cost of producing virulence factors in terms of host mortality without removing the benefit or producing them (Gandon et al., 2003, 2001; but see also McLeod and Gandon, 2022).

The pioneer work by Gandon et al. (2001) brought to light the possible problems linked to vaccination programs intending to vaccinate a maximum of individuals without questioning their possible long-term (i.e., evolutionary) consequences and the need for theoretical works in the establishment of vaccination policies (Bull and Antia, 2022; Walter and Lion, 2021). Nevertheless, it has been criticized for its oversimplification of biological mechanisms (Soubeyrand and Plotkin, 2002). Notably, while the model proposed by Gandon et al. (2001) describes how vaccines inducing antitoxin immunity should select for pathogen strains with higher toxin production (i.e., higher virulence), it fails to account for the metabolic cost associated with such production (Soubeyrand and Plotkin, 2002). Gandon and coauthors responded to Soubeyrand and Plotkin by incorporating a cost to toxin production in their model and concluded that when the cost of toxin production is high a high vaccine efficacy should select for lower toxin production (Gandon et al., 2002). While this conclusion matches empirical observations for diphtheria, with countries now reporting a high proportion of non-toxicogenic strains over toxicogenic strains (e.g., Badell-Ocando et al., 2022; Benamrouche et al., 2016; Czajka et al., 2018; Dangel et al., 2018), the model is still general, meaning it cannot fully apprehend the mechanisms specific to each host-pathogen interaction. More specific models, which would be built based on the specificity of exemplary diseases, are needed to better characterize the mechanisms through which antitoxin vaccines may affect the evolution of their targeted bacteria.

Thus, the model of Gandon et al. (2001) is fatally limited in face of the great diversity in host-pathogen interactions and mechanisms at play. For example, it does not take into account phenomena such as competition between strains or waning immunity, which could have an impact on the evolutionary outcome of the use of vaccines. Diphtheria and pertussis are two human diseases for which the prevalence of non-toxicogenic strains has increased in countries with a high toxoid vaccine coverage (Guiso et al., 2022; Soubeyrand and Plotkin, 2002). Diphtheria is an acute and serious disease caused by strains of *Corynebacterium* species, the most common being *Corynebacterium diphtheriae*, a gram-positive bacillus. The disease exists in three forms: invasive, cutaneous and respiratory, the latter being the most common form of the disease for indigenous cases in temperate climates (Tiwari and Wharton, 2012). Respiratory diphtheria is caused by the invasion of the upper respiratory tract by the bacterium, followed by the production and emission of the diphtheria toxin (DT), a toxin encoded by a bacteriophage gene called *tox* (Pappenheimer and Gill, 1973).

The production of DT leads to the death of the host’s pharynx cells and thus to the formation of a characteristic pseudomembrane. This pseudomembrane on the pharynx is one of the most common symptoms of the disease (Dobie, 1979). It can provoke a respiratory distress that may be fatal, but most of the deaths are imputed to a secondary effect of the toxin: it can be absorbed by other organs and induce severe complications such as cardiopathy and neuropathy (Tiwari and Wharton, 2012).

There are two ways to fight against the disease: therapy and

prophylaxis (Tiwari and Wharton, 2012). Once the disease has been clinically confirmed in a patient, therapy can consist in the use of an antitoxin to reduce morbidity and antibiotics to kill the bacteria. Prophylactic vaccines against diphtheria are composed of a deactivated form of the DT, called a toxoid (Tiwari and Wharton, 2012). While a recent study showed that currently used vaccines may contain up to hundreds of *C. diphtheriae* proteins (Möller et al., 2019), these vaccines are designed to stimulate the immunity against this toxin particularly and not against the whole bacterium, to protect against the deleterious effects of DT (Scheifele and Ochnio, 2009). DT can thus be considered the virulence factor of *C. diphtheriae*. Toxoid vaccines can protect against the symptomatic forms of the infection (by stimulating the host’s defenses against the virulence factor) but not the infection itself. Such vaccines are thus unable to block the transmission (Fine, 1993). This particularity is what makes diphtheria an interesting model in the study of the impact of an imperfect (i.e., non-transmission blocking) vaccine on the evolution of a pathogen, as it has been predicted that such vaccine should select for strains of higher virulence (Gandon et al., 2001), but empirical evidence for diphtheria, in the form of a multiplication of isolated non-toxicogenic strains in different countries with a high vaccine coverage, weakens this theory (Pappenheimer, 1984; Soubeyrand and Plotkin, 2002).

In the present study, we build a mathematical model based on the biology of diphtheria to understand how the vaccine pressure as well as ecological relationships between a virulent (i.e., toxin-producing) and an avirulent (i.e., toxin-deficient) strain of *C. diphtheriae* could drive the evolution of this bacterium. Our model relies on the hypothesis that the incidence of the avirulent strain should increase when the proportion of the vaccinated population increases because it does not pay the cost associated with the production of a toxin, as it has been theorized by Soubeyrand and Plotkin (2002) and later by Guiso and coauthors (2022). Thus, the outcome of the competition between avirulent and virulent strains in highly vaccinated populations is closely dependent on the cost of toxin production and may lead to the disappearance of virulent strains.

2. Methods

2.1. Model description

We propose to model the dynamical relationship between two hypothetical strains of *C. diphtheriae* in a host population with different states of immunity. The first strain is a *tox* positive strain, i.e., able to produce the DT and therefore it induces damages in non-immunized individuals. The second strain is a *tox* negative strain, i.e., unable to produce the toxin. For the sake of simplicity, we consider in this work that a strain that does not produce the toxin is avirulent (i.e., non-pathogenic). Here, *tox* positive and *tox* negative strains are referred to as Cd_1 and Cd_2 , respectively.

The model represents the evolution of the disease within a closed homogeneous and uniformly mixed population. This disease is transmitted by direct contact with an infected individual. More specifically, the model is an extension of the standard SIR one, with five compartments that classifies individuals as susceptible (S), infected with the Cd_1 strain (I_1), infected with the Cd_2 strain (I_2), vaccinated (V) and recovered (R). Fig. 1 represents the flows of individuals among the five epidemiological compartments.

Host reproduction occurs at a constant rate b . A p_1 fraction of offspring, representing the part of the population vaccinated at birth, enters the V compartment and a $1 - p_1$ fraction enters the S compartment. All hosts die at a rate d representing the natural mortality rate of the population.

Susceptible individuals get vaccinated at a rate ω , which is the product of the frequency of the boosters f and the proportion of the population taking the boosters p_2 . We allow for vaccinal immunity to wane at a rate η , moving individuals to the susceptible compartment.

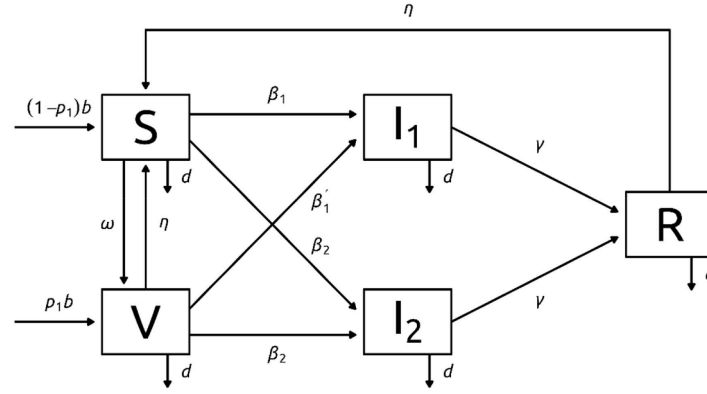


Fig. 1. Schematic diagram of the model described in equations (2). Definition of the parameters is given in Table 1 and in the main text. Definition of the compartments is given in the main text. Compartment-to-compartment arrows represent epidemiological transitions between states while void-to-compartment and compartment-to-void arrows represent demographic events (births and deaths, respectively).

We consider three transmission rates: β_1 , β_1' and β_2 , representing the transmission rate of Cd_1 to a susceptible individual, the transmission rate of Cd_1 to a vaccinated individual, and the transmission rate of Cd_2 to any kind of host, respectively.

Production of DT confers an ecological advantage to Cd_1 in susceptible hosts by increasing the availability of iron that is essential to the bacterial development (Oram and Holmes, 2006). But in vaccinated hosts, DT production becomes disadvantageous for Cd_1 because of the cost associated with producing a “useless” toxin (i.e., deactivated by the host’s immune system). This advantage/disadvantage has consequences on the pathogen’s capacity to exploit its host, thus to multiply, and therefore to transmit to a susceptible host. For Cd_2 , transmission is the same whether the host is susceptible or vaccinated, since this strain does not produce the toxin and thus is not affected by the vaccination status of its host. This is true for a pure toxoid vaccine, and while a recent study showed that other immunogenic proteins can be found in diphtheria vaccines it has not yet been demonstrated that such proteins are present in a high enough concentration to induce protective immunity against them (Möller et al., 2019). We thus consider in this work that diphtheria vaccine is a pure toxoid vaccine.

We define β_2 as the “basal” transmission rate of *C. diphtheriae*, i.e., the transmission rate without any cost nor benefits associated with the production of the toxin. For toxin production to be advantageous and thus selected, it must confer a fitness gain, at least in a susceptible population. If the population is immunized against the toxin, then the cost of its production is likely to exceed its benefit. We thus obtain the following relationship between β_1 , β_1' and β_2 :

$$\beta_1 > \beta_2 > \beta_1' \tag{1}$$

which is the fundamental inequation of our model.

For the sake of simplicity, we consider no superinfection, i.e., an individual infected with one strain cannot be infected with the other strain at the same time. This means that the two strains are in competition for access to hosts: their growth (multiplication) is limited by the growth of the other strain. Moreover, we consider no excess mortality induced by the infection. Virulence is therefore not defined as the mortality induced by the pathogen; we simply consider that Cd_1 is virulent and Cd_2 is avirulent.

Infected individuals recover after a time period γ and enter the R compartment with a newly acquired immunity. Recovered individuals can lose their immunity at a rate η and return to the susceptible state (S compartment). In absence of data for diphtheria we assume that both naturally and artificially induced immunity (vaccination) wane at the same rate η . This may represent an oversimplification since it is presumably wrong for pertussis, a close host-vaccine-pathogen system (Wendelboe et al., 2005).

Nevertheless, immunity acquired through vaccination is expected to differ from that acquired naturally. Vaccination provides immunity against DT whereas natural infection provides immunity against the “whole pathogen” (DT and other bacterial antigens). This means that a non-vaccinated individual infected with Cd_1 develops an immunity against DT and other antigens whereas an individual infected with Cd_2 develops an immunity only against these other antigens if it has not been vaccinated, or against these antigens and DT if it has been vaccinated. For the sake of simplicity, we consider that a recovered individual cannot be reinfected because we assume that the immunity developed with exposure to these other antigens protects against any new *C. diphtheriae* infection.

With these assumptions, the epidemiological dynamics can be described by the following ordinary differential equations system:

$$\frac{dS(t)}{dt} = (1 - p_1)b + \eta(V(t) + R(t)) - \beta_1 \frac{S(t)}{N(t)} I_1(t) - \beta_2 \frac{S(t)}{N(t)} I_2(t) - \omega S(t) - dS(t) \tag{2a}$$

$$\frac{dI_1(t)}{dt} = \beta_1 \frac{S(t)}{N(t)} I_1(t) + \beta_1' \frac{V(t)}{N(t)} I_1(t) - \gamma I_1(t) - dI_1(t) \tag{2b}$$

$$\frac{dI_2(t)}{dt} = \beta_2 \frac{S(t)}{N(t)} I_2(t) + \beta_2 \frac{V(t)}{N(t)} I_2(t) - \gamma I_2(t) - dI_2(t) \tag{2c}$$

$$\frac{dV(t)}{dt} = p_1 b + \omega S(t) - \beta_1' \frac{V(t)}{N(t)} I_1(t) - \beta_2 \frac{V(t)}{N(t)} I_2(t) - \eta V(t) - dV(t) \tag{2d}$$

$$\frac{dR(t)}{dt} = \gamma(I_1(t) + I_2(t)) - \eta R(t) - dR(t) \tag{2e}$$

where $S(t)$, $I_1(t)$, $I_2(t)$, $V(t)$ and $R(t)$ represent the number of individuals in each compartment of S, I₁, I₂, V and R, respectively, at time t .

2.2. Invasion analysis

The basic reproduction number (R_0) is defined as the number of secondary cases generated by a single infected individual in an entirely susceptible population at equilibrium. In our model, when focusing on Cd_1 , we have:

$$R_0 = \frac{\beta_1}{\gamma + d} \tag{3}$$

If individuals in the population begin to be vaccinated, the population is no longer “entirely susceptible”, and Eq. (3) cannot describe the invasive capacity of a pathogen in such a context. We thus define new

equations for the reproduction number of each strain.

We consider a hypothetical population at the disease-free equilibrium (DFE). We define here a population at the DFE as either a mixed population (i.e., composed of both susceptible and vaccinated individuals) or as a fully susceptible population. In the simplest case where only one strain spreads, this strain can invade the population and start an epidemic if its reproduction number is superior to one. Let $R_{0,x}$ be the R_0 (i.e., the reproduction number) of the Cd_x strain ($x \in (1, 2)$). We consider that an invasion starts with the arrival of a unique infected individual in the population (i.e., $I_x(0) = 1$ for an invasion by Cd_x).

In a mixed population, Cd_1 can invade if $R_{0,1} > 1$ with:

$$R_{0,1} = \frac{\beta_1 s_{DFE}^* + \beta_1' v_{DFE}^*}{\gamma + d} \quad (4)$$

where s_{DFE}^* is the proportion of susceptible individuals at the DFE and v_{DFE}^* is the proportion of vaccinated individuals at the DFE. The value of $R_{0,1}$ thus depends on the vaccine coverage.

Reciprocally, Cd_2 can invade a population at the DFE if $R_{0,2} > 1$ with:

$$R_{0,2} = \frac{\beta_2}{\gamma + d} \quad (5)$$

which means the value of $R_{0,2}$ is independent of the vaccine coverage. This is expected since vaccination only acts on the transmission of pathogenic strains.

We investigated the conditions promoting the long-term coexistence of the two strains based on the criterion of mutual invasibility, which means two strains can coexist if each can invade a host population (i.e., have a positive growth rate; for a review see Grainger et al., 2019) at the equilibrium in which the other strain would be present alone. Technically, after equilibrium has been established with the first strain, the second strain is introduced in the population with only one infected individual at time $t = 0$. We then search for the conditions under which $dI_1/dt > 0$ (respectively $dI_2/dt > 0$) when the population is at the endemic equilibrium with Cd_2 (respectively Cd_1) alone. We obtain:

$$\frac{s_1^*}{v_1^*} < \frac{\epsilon'}{\epsilon} < \frac{s_2^*}{v_2^*} \quad (6)$$

with $\epsilon' = \beta_2 - \beta_1'$, $\epsilon = \beta_1 - \beta_2$ and s_x^* (respectively v_x^*) the proportion of susceptible (respectively vaccinated) individuals at the equilibrium where the Cd_x strain is endemic ($x \in (1, 2)$). Knowing (1), $\epsilon' > 0$ and $\epsilon > 0$.

2.3. Parametrization

Table 1 summarizes the model parameters. The description of the assigned values can be found in Hammarlund et al. (2016) for duration of immunity, in Truelove et al. (2020) for duration of infection, basic reproduction number, and reduction of Cd_1 transmission to vaccinated individuals. The host mortality rate is taken from the *Institut National de la Statistique et des Etudes Economiques* (INSEE) website (https://www.insee.fr/fr/outil-interactif/5367857/tableau/20_DEM/23_DME, last accessed on January 2nd, 2023). When relevant data are lacking, parameters (see below) have been calculated or chosen arbitrarily. Note that literature is contradictory in the definition of an immunizing antibody level, with threshold varying to authors. In this work, we consider an antibody level of 0.01 UI.mL⁻¹ as protective, as proposed by Amanna et al. (2007) and Hammarlund et al. (2016), and choose a value of η adapted to this threshold.

The WHO recommends that adults take booster doses of diphtheria toxoid every ten years (World Health Organisation, 2017). This recommendation is followed in most countries with one exception being France where boosters are recommended at ages 25, 45, 65 and every 10 years beyond that age (Ministère des Solidarités et de la Santé, 2022). The proportion of adults taking the booster is not given by the WHO but

Table 1

Parameters and their values used in the simulations of the model described in equations (2). d: days.

Parameter	Signification	Value	Reference
b	Birth rate	$\frac{1}{3}d^{-1}$	Arbitrary
d	Death rate	$\frac{1}{83 \times 365}d^{-1}$	INSEE
p_1	Children vaccine coverage	Varies	None
p_2	Adults vaccine coverage	Varies	See Figs. 2 and 3
f	Booster frequency	$\frac{1}{10 \times 365}d^{-1}$	World Health Organization, 2017
ω	Booster rate	$f\hat{p}_2$	See text
η	Immunity loss	$\frac{1}{42 \times 365}d^{-1}$	Hammarlund et al., 2016
γ	Recovering rate	$\frac{1}{18.5}d^{-1}$	Truelove et al., 2020
R_0	Basic reproduction number	2.6	Truelove et al., 2020
r_1	Transmission reduction of Cd_1	0.60	Truelove et al., 2020
r_2	Transmission reduction of Cd_2	0.42	See Section 2.3
β_1	Transmission rate of Cd_1 to S individuals	$R_0(\gamma + d)d^{-1}$	(7)
β_1'	Transmission rate of Cd_1 to V individuals	$(1 - r_1)\beta_1 d^{-1}$	(8)
β_2	Transmission rate of Cd_2	$(1 - r_2)\beta_1 d^{-1}$	(9)

it can be estimated from multiple studies in different countries: France (Baratin et al., 2012), Mexico (Cruz-Hervert et al., 2013), the United States (Williams et al., 2017) and Germany (Poethko-Müller and Schmitz, 2013). For a review of the status of adult vaccination in Europe, see Kanitz et al. (2012).

Usually, the value of the R_0 for a disease is calculated (or rather, should be calculated) at the beginning of its spread in a completely susceptible population (Delamater et al., 2019). In the case of diphtheria, R_0 is usually calculated during diphtheria outbreaks in a partially immune population, and thus cannot be rigorously called R_0 (e.g., Matsuyama et al., 2018). Nevertheless, using statistical analysis tools, it is possible to infer the value of R_0 as it would be in a susceptible population (Truelove et al., 2020). Once R_0 is obtained, it is possible to calculate the value of the transmission rate of the strains. Since only toxigenic strains can be monitored (i.e., Cd_1 strains in our model), any R_0 from field studies can only describe the transmission of such a strain. The transmission rate of Cd_1 in a susceptible population, β_1 , can therefore be obtained from Eq. (3):

$$\beta_1 = \frac{R_0}{\gamma + d} \quad (7)$$

Then the transmission rate of Cd_1 in a vaccinated population, β_1' , can be obtained using the following equation:

$$\beta_1' = (1 - r_1)\beta_1 \quad (8)$$

where r_1 represents the decrease in the transmission of Cd_1 due to vaccination (i.e., the vaccine efficacy).

As said earlier, R_0 cannot be obtained for a nontoxigenic strain, i.e., a Cd_2 strain in our model, as such strain is not monitored. It is therefore not possible to calculate the transmission rate of Cd_2 from Eq. (3). Instead, we calculate the value of the transmission rate of Cd_2 , β_2 , as a reduction of the transmission rate of Cd_1 in a susceptible population:

$$\beta_2 = (1 - r_2)\beta_1 \quad (9)$$

where r_2 represents the decrease in the transmission of Cd_2 due to the lack of toxin production in this strain. We set $r_1 > r_2$ in agreement with (1). Because to our knowledge no information about the transmission

dynamics of non-toxicogenic strains is available yet, we assigned an arbitrary value for r_2 . Since the production of the toxin is essential for the iron uptake in an iron-depleted growth medium such as the respiratory tract, we estimate that non-production of the toxin represents a high handicap for the strain and thus that its reduction in transmission can be fixed at a high value (Murphy, 1996; Parveen et al., 2019).

The numerical values of the parameters and their sources are summarized in Table 1.

3. Results

The mathematical model presented in (2) was implemented in a C script. Simulations from the model were run using the R language (R Core Team, 2022) and the *deSolve* package (Soetaert et al., 2010). Simulations were run long enough to reach the near-equilibrium state (10^6 days, so roughly 2740 years).

3.1. Prevalence of Cd_1

We first investigate the propagation of the pathogenic strain Cd_1 in the host population, since it is an important goal from a public health point of view.

We run simulations to assess the effect of vaccine coverage and of changes in seven key parameters values on the prevalence of Cd_1 in the population, with Cd_1 prevalence estimated as the number of Cd_1 -infected individuals per 10,000 individuals at the equilibrium (Fig. 2). We started with a population at the equilibrium, either with Cd_2 (Cd_2 -endemic equilibrium or CEE, dashed lines in Fig. 2) or with no bacterium in the population (disease-free equilibrium or DFE, solid lines in Fig. 2). Note that for low values of vaccine coverage the dashed and solid lines are superimposed.

Vaccine coverage values ranged from 0 to 1. The other six parameters could be parameters directly taken from the model equations (e.g., r_2 , Fig. 2E), or relevant novel parameters adapted from model parameters (e.g., the proportion of individuals vaccinated at birth that would later take the adult booster, p_2/p_1 , Fig. 2F). For each parameter, we consider between three and five possible values. For example, we explore four booster frequencies to analyze how public health decisions could impact disease spread (Fig. 2B). When the value of one parameter is changed, the values of the other parameters stay the same (see Table 1).

For any parameter value, the higher the vaccine coverage, the lower the prevalence of Cd_1 at equilibrium (Fig. 2A–F). We also observe that the best values (from a public health perspective) for immunity-related parameters, i.e., immunity duration, booster frequency, and proportion of individuals vaccinated at birth taking the adult boosters, lead to a lower Cd_1 burden (Fig. 2A, B, F). At the same time, when Cd_1 has a better transmission capacity, i.e., a higher R_0 , its equilibrium prevalence is higher (Fig. 2D). Obviously, the effect of the transmission reduction of Cd_2 , r_2 , on the prevalence of Cd_1 can only be seen if Cd_2 is spreading in the population. In this case, we observe that the lower r_2 is, the lower the vaccine coverage needed for Cd_1 prevalence to decrease (Fig. 2E).

We can also see from Fig. 2 that when Cd_2 circulates in the population, the prevalence of Cd_1 at equilibrium drops stiffly above certain values of vaccine coverage, and this drop can occur for a more or less important vaccine coverage depending on the parameters values.

Following this observation of a sharp drop in the “absolute” prevalence of Cd_1 , we wanted to investigate how the relative prevalence of Cd_1 changes under the same conditions as those described above (Fig. 3). By relative prevalence, we mean the proportion of infected individuals carrying Cd_1 , or in other words the $I_1/(I_1+I_2)$ value at equilibrium. Since it would be useless to observe the relative prevalence of Cd_1 if there was no Cd_2 in the population, simulations were run starting with a

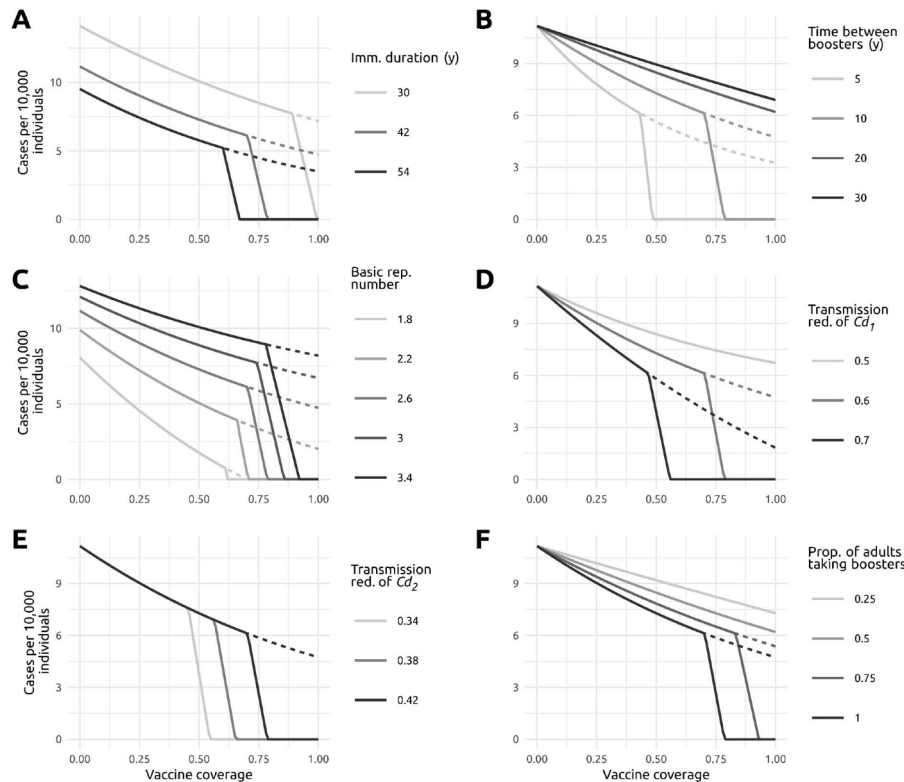


Fig. 2. Cd_1 prevalence at the system equilibrium as vaccine coverage is moved, for different values of (A) immunity duration ($\frac{1}{\gamma \times 365}$ years), (B) time between boosters ($\frac{1}{\beta \times 365}$ years), (C) basic reproduction number (R_0), (D) transmission reduction of Cd_1 (r_1), (E) transmission reduction of Cd_2 (r_2), and (F) proportion of individuals vaccinated at birth who take the vaccine booster ($\frac{p_2}{p_1}$). For plots (A)–(E) we considered that p_2 changed synchronously with p_1 (i.e., $\frac{p_2}{p_1} = 1$). Solid lines indicate simulations where Cd_2 is endemic while dashed lines indicate simulations where Cd_2 is absent from the population.

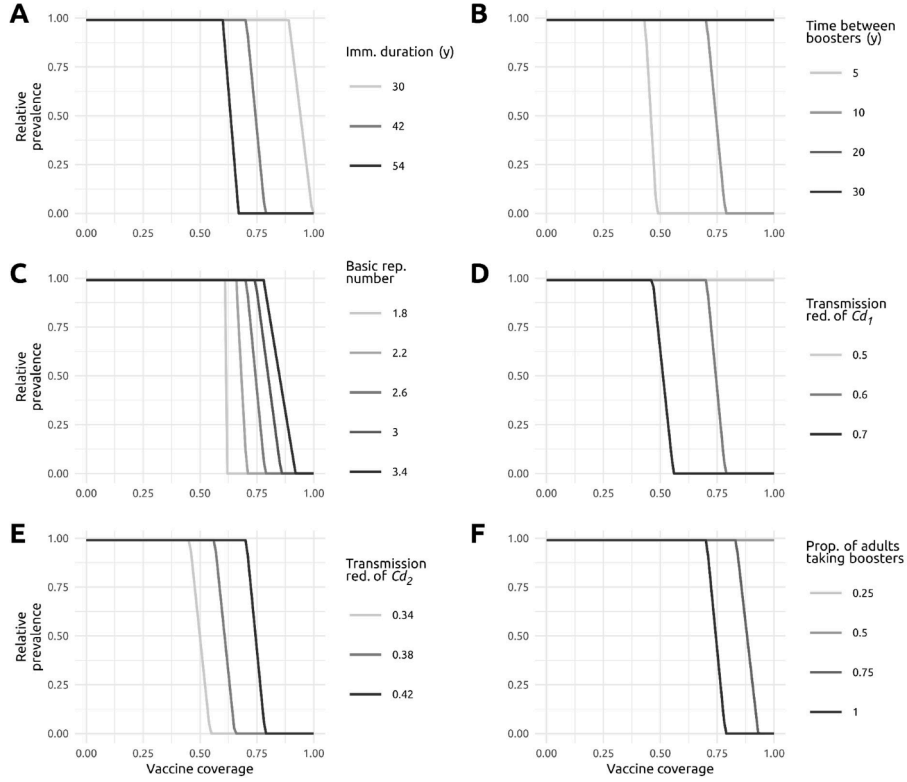


Fig. 3. Cd_1 relative prevalence at the endemic equilibrium as vaccine coverage is moved, for different values of (A) immunity duration ($\frac{1}{\eta \times 365}$ years), (B) time between boosters ($\frac{1}{\gamma \times 365}$ years), (C) basic reproduction number (R_0), (D) transmission reduction of Cd_1 (r_1), (E) transmission reduction of Cd_2 (r_2), and (F) proportion of individuals vaccinated at birth who take the vaccine booster ($\frac{p_2}{p_1}$). For plots (A)–(E) we considered that p_2 changed synchronously with p_1 (i.e., $\frac{p_2}{p_1} = 1$).

population at the CEE.

We find similar effects of parameters to those observed previously. This means that if a parameter value makes the absolute prevalence fall faster (respectively slower) along the vaccine coverage gradient, it also makes the relative prevalence fall faster (respectively slower). What is of particular interest is that for most vaccine coverage values, the infected population is either infected with Cd_1 or Cd_2 , and coexistence of both strains appears to occur only for a short range of vaccine coverage values (Fig. 3A–F).

3.2. Mutual invasion criterion

For both strains to coexist, they must be able to mutually invade each other. The criterion for such co-invasion to occur (and reciprocally, for exclusion to occur) is given by Eq. (6).

Parameter ϵ' represents the difference in fitness between Cd_2 and Cd_1 within a population of vaccinated hosts. Based on inequality (1), ϵ' is positive, showing the advantage of Cd_2 among vaccinated hosts. In the same way, parameter ϵ represents the difference of fitness between Cd_1 and Cd_2 within a susceptible population. Again, based on (1), ϵ is positive and shows the advantage of Cd_1 among susceptible hosts. We can separate the terms of the inequation to better understand the conditions under which both strains can coexist. We obtain from one side:

$$\frac{\epsilon'}{\epsilon} < \frac{s_2^*}{v_2^*} \Leftrightarrow \epsilon s_2^* > \epsilon' v_2^* \quad (10)$$

and from the other side:

$$\frac{\epsilon'}{\epsilon} > \frac{s_1^*}{v_1^*} \Leftrightarrow \epsilon' v_1^* > \epsilon s_1^* \quad (11)$$

Inequation (10) implies that the strain Cd_1 can invade a population

where Cd_2 is endemic if the advantage of Cd_1 among susceptible individuals relatively to the proportion of susceptible individuals in the population is higher than the advantage of Cd_2 among vaccinated individuals relatively to the proportion of vaccinated individuals in the population.

Reciprocally, inequation (11) implies that strain Cd_2 can invade a population where Cd_1 is endemic if the advantage of Cd_2 among vaccinated individuals relatively to the proportion of vaccinated individuals in the population is greater than the advantage of Cd_1 among susceptible individuals relatively to the proportion of susceptible individuals in the population.

Invasion here implies competition: if a strain cannot invade a population where the other one is endemic, it means that the latter strain is competitively excluding the former one.

3.3. Vaccine efficacy

Since r_1 represents the loss of transmission for Cd_1 when the targeted host is vaccinated, it can be interpreted as the vaccine efficacy against transmission of the bacterium. A value of $r_1 = 0$ means that the vaccine fails to prevent the transmission of the pathogen, while a value of $r_1 = 1$ means the vaccine perfectly blocks the transmission (a perfect vaccine).

We ran simulations to investigate the simultaneous effect of both vaccine coverage and vaccine efficacy on the reproduction number of Cd_1 (Fig. 4). The simulations started with two types of population: a population at DFE (Fig. 4A) and a population at CEE (Fig. 4B). A single individual infected with Cd_1 was then included in the population and simulations were re-run until a near-equilibrium state was reached again. In both cases, there are values of vaccine efficacy and vaccine coverage for which $R_0 < 1$, which is the condition under which the pathogen does not transmit sufficiently and goes extinct (herd immunity). This threshold is reached for lower values of vaccine coverage and

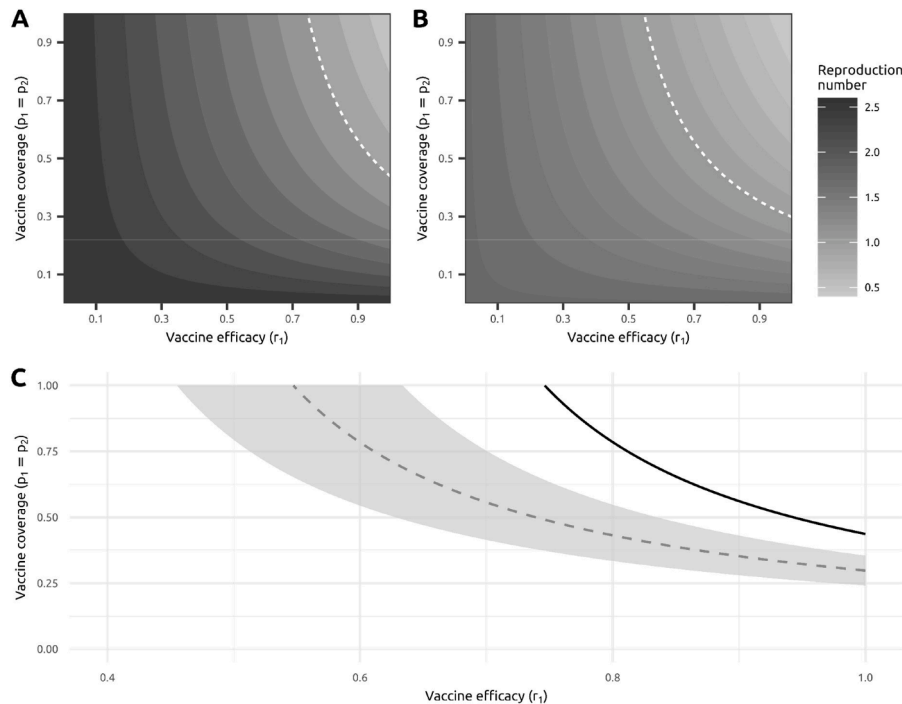


Fig. 4. Effect of the vaccine coverage and vaccine efficacy on the reproduction number of Cd_1 (A) in a population without Cd_2 or (B) with Cd_2 at the endemic equilibrium. For (A) and (B) the white dashed line indicates $R_0 = 1$. (C) Combination of vaccine coverage and vaccine efficacy for which $R_0 = 1$ in a population without Cd_2 (solid line) or in a population with Cd_2 at the endemic equilibrium (dashed line, $r_2 = 0.42$). The ribbon surrounding the dashed line indicates possible positions of the dashed line for r_2 values moving between $r_2 = 0.34$ (lower boundary) and $r_2 = 0.50$ (upper boundary).

vaccine efficacy when Cd_2 is circulating in the population than when there is no competitive strain (Fig. 4C).

We ran more simulations with a range of values to investigate how the transmission capacity of Cd_2 could affect this threshold. When the transmission rate of Cd_2 was high (i.e., the value of r_2 was low), lower vaccine coverage and efficacy were needed to reach the extinction threshold (Fig. 4C).

4. Discussion

In this study, we developed a disease-specific model to investigate the evolutionary direction followed by *C. diphtheriae* strains in a context of mass vaccination. We considered the cost associated with toxin production as well as the effect of a toxoid vaccine against toxin-producing strains and concluded that the use of such vaccine should select for bacteria strains of lower virulence. Using a two-strains model, we evaluated the role of different parameters on the selection of each strain (one toxin-deficient, one toxin-producing) and the conditions under which coexistence is possible. We found that when vaccine coverage is high both for children (primary vaccination) and adults (booster vaccination) the toxin-deficient strain is selected. A shorter booster interval in adults also contributes to the selection of toxin-deficient strains. These three parameters, primary coverage, booster coverage, and booster intervals, vary widely among countries, and are among the easiest to influence. While vaccination against diphtheria at an early age is mandatory in most countries, vaccination coverage may be low in some regions, during some periods or at certain ages (Abdi et al., 2021; Lee and Choi, 2017; Marchal et al., 2021; Murhekar et al., 2009; Sein et al., 2016), reducing selective pressure against toxigenic strains. Booster vaccination is often recommended but not mandatory, and the frequency of boosters may differ between countries. In France, it is recommended that adults receive boosters at 25, 45, and 65 years of age, and then every 10 years (Ministère des Solidarités et de la Santé, 2022), but the percentage of individuals receiving the booster is low and decreases sharply with age (Marchal et al., 2021).

Our findings match clinical data for diphtheria for example in Algeria (Benamrouche et al., 2016), France (Badell-Ocando et al., 2022), Germany (Dangel et al., 2018) and Poland (Czajka et al., 2018; Zasada, 2013; Zasada and Rzeczowska, 2019), where the use of toxoid vaccines has been associated with an increase in the prevalence of *tox-* strains. Countries with a good monitoring system on the incidence of *C. diphtheriae*, like France, still report finding *tox+* strains circulating in the population at low prevalence (Badell-Ocando et al., 2022). In addition to the phenomenon described above that vaccination can drop, these countries probably did not reach their equilibrium state since vaccination was introduced less than a century ago. In our simulations, equilibrium was long to reach because of the inclusion of demographic parameters and immunity decay (data not shown), two biological phenomena acting on a time scale much longer than the epidemiological scale. This could explain the persistence of *tox+* strains in these countries. Nevertheless, we argue that in the end toxoid vaccines should select for toxin-deficient strains, because toxin production represents a potentially high metabolic cost, which is not compensated by any benefit if the host is resistant to the toxin's action.

Other parameters may be more difficult to act upon, as they depend on the biology of both the pathogen and the host. As expected, the higher the reproductive capacity of a strain, the better it persists in the population. Thus, the greater the relative advantage of a toxigenic strain over a non-toxigenic strain, the more difficult it will be for the latter to invade the population. The transmission of a pathogen cannot be directly controlled, but public health measures can help reduce its R_0 , as was demonstrated during the recent COVID-19 pandemic (Talic et al., 2021). Combined with contact restriction, reducing the infectious period (e.g., with antibiotics) could also reduce the transmission of pathogenic strains (Truelove et al., 2020). These measures, if individual-based, could reduce the transmission of symptomatic strains (i.e., pathogenic strains) while having little to no effect on the transmission of non-pathogenic strains.

Our results also illustrate how competition between strains for access to hosts can contribute to the elimination of pathogenic strains, as non-

toxigenic strains reduce the pool of the susceptible population that could be used by toxigenic strains to maintain in the population. In our simulations, we found that, over a wide range of model parameters, the toxigenic strain could only be eliminated if the nontoxigenic strain was also present in the population. This is because the vaccine does not block either the infection by the toxigenic bacterium or its transmission. Instead, it reduces its transmission by removing the benefit of producing the toxin, leaving the bacterium to produce a costly toxin without any benefit, which impacts the transmission rate of the bacterium. In a scenario where the cost of producing the toxin is high enough that the transmission rate decreases sufficiently and the bacterium can no longer spread in the population, vaccination could result in the extinction of the toxigenic strain. Data are currently lacking on the extent to which vaccine reduces transmission of diphtheria, but studies on the similar pertussis system show that toxoid vaccines (which also contains non-toxin antigens) are not sufficient to stop transmission of the pathogen in a baboon model (Warfel et al., 2014).

This study focuses on one pathogen, *C. diphtheriae*, the etiologic agent of diphtheria, but our results may explain the occurrence of non-toxigenic strains in other host-pathogen systems such as pertussis. Most developed countries abandoned conventional whole-cell vaccines (WCVs) in favor of acellular vaccines (ACVs) in the last three decades. These new generation vaccines are similar to diphtheria vaccines in that they are not composed of the whole pathogen, but of certain antigens selected for their role in triggering disease. One of these antigens, present in all ACVs, is pertussis toxin, which is thought to be responsible for most of pertussis symptoms (Carbonetti, 2010; Gregg and Merkel, 2019; Scanlon et al., 2019a, 2019b). Interestingly, at least two countries using ACVs have reported the detection of toxin-deficient strains in recent years (Bouchez et al., 2021; Williams et al., 2016). It is possible that these toxin-deficient strains, rather than being rare, are frequent but difficult to detect, due to the lack of symptoms in the host they infect. Thus, the use of ACVs (i.e., vaccines with a toxoid component) could potentially result in a similar mechanism of competition between toxigenic and non-toxigenic strains, a hypothesis that has not been substantiated so far. Note that for pertussis the massive use of vaccines has been linked to the emergence of strains producing more toxin *in vitro*, but this finding has never been confirmed clinically (Mooi et al., 2009).

One general and important conclusion of our work is that there may be some advantages in having avirulent strains (or at least strains of reduced virulence) of a bacterium circulating in a host population. Our results may support a new and ecological way of vaccine use. Instead of targeting the eradication of pathogens, the aim of some immunization programs could be to reduce the disease to nil by favoring the selection and spread out within the human population of non-pathogenic (i.e., commensal) strains.

This new way of considering the role of vaccination can also be illustrated in rabbits, where the impact of RHDV (Rabbit Haemorrhagic Disease Virus), a highly virulent virus that has decimated rabbit populations worldwide, is reduced in populations with a high level of pre-existing circulation of non-pathogenic strains of the virus (Forrester et al., 2007; Strive et al., 2013; White et al., 2002). The same question arises for COVID-19. Indeed, new variants of importance are described every few months (Ciotti et al., 2022) with characteristics (in terms of transmission and pathogenicity) that are not fully characterized until they invade many countries. This leaves our populations vulnerable to undesirable biological events that could make a new variant both more transmissible and more virulent. Efficient circulation of the pre-existing variant, combined with high vaccination coverage of the most vulnerable part of the population, could be an interesting strategy to reduce the impact of such an unpredictable event.

Finally, it is important to keep in mind that our model has some limitations. As with many SIR-type models, we focus on the biology of the pathogen but not of the host. Diphtheria, like pertussis, can be spread by asymptomatic individuals carrying the pathogen without even knowing they are infected (Tiwari and Wharton, 2012). Our model does

not distinguish susceptible individuals based on their risk of developing a more or less severe form of the disease, but it would be interesting to investigate how asymptomatic individuals might shape the evolutionary dynamics of the disease. Symptomatic individuals are more likely to receive therapeutic treatments (antitoxin and antibiotics) than asymptomatic individuals, thus creating another pressure against virulent strains, which are more likely to be toxigenic. We also do not distinguish individuals based on their “willingness to be vaccinated”. In our model, individuals receive booster vaccination randomly, whereas one would expect the choice to receive a booster to be determined by i) the vaccination status during childhood and ii) whether or not they received a booster in previous years. A recent work also showed that despite being theoretically a toxoid vaccine, diphtheria vaccines also contain other proteins of *C. diphtheriae*, some of which showing immunogenic properties (Möller et al., 2019). This aspect could be incorporated into the model. Finally, our model could be completed by taking into account migratory flux, with new arrivals of *tox+* individuals in the population, participating in the dynamical of maintenance of toxigenic strains in countries with high vaccine coverage as it has already been described in Europe (European Centre for Disease Prevention and Control, 2022).

CRediT authorship contribution statement

Florian Lecorvaisier: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Visualization. **Dominique Pontier:** Conceptualization, Supervision, Writing – review & editing. **Benoît Soubeyrand:** Conceptualization, Writing – review & editing. **David Fouchet:** Methodology, Conceptualization, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Benoit Soubeyrand reports a relationship with Blossom Vaccinology that includes: board membership. Co-author BS previously worked for Sanofi.

Data availability

No data was used for the research described in the article.

Acknowledgments

We thank the Laboratoire d'Excellence (LabEx) ECOFECT 'Evolutionary Dynamics of Infectious Diseases' (ANR-11-LABX-0048) of Université de Lyon. We thank the Laboratoire de Biométrie et Biologie Evolutive (LBBE). This work was performed using the computing facilities of the CC LBBE/PRABI. We also thank Nicole Guiso for her expertise on the biology of diphtheria and for her advices in the revision of the manuscript. We also thank the three anonymous reviewers for their helpful comments that led to the improvement of the quality of this manuscript.

References

- Abdi, I., Gidding, H., Leong, R.N., Moore, H.C., Seale, H., Menzies, R., 2021. Vaccine coverage in children born to migrant mothers in Australia: a population-based cohort study. *Vaccine* 39, 984–993. <https://doi.org/10.1016/j.vaccine.2020.12.058>.
- Amanna, I.J., Carlson, N.E., Slifka, M.K., 2007. Duration of humoral immunity to common viral and vaccine antigens. *N. Engl. J. Med.* 357, 1903–1915. <https://doi.org/10.1056/NEJMoa066092>.
- Badell-Ocando, E., Toubiana, J., Brisse, S., 2022. Rapport Annuel D'activité Du Centre national De Référence des Corynébactéries du Complexe *Diphtheriae* – Bilan des Activités 2017-2021. Institut Pasteur, Paris, France.

- Baratin, D., Del Signore, C., Thierry, J., Caulin, E., Vanhems, P., 2012. Evaluation of adult dTPaP vaccination coverage in France: experience in Lyon city, 2010–2011. *BMC Public Health* 12, 940. <https://doi.org/10.1186/1471-2458-12-940>.
- Benamrouche, N., Hasnaoui, S., Badell, E., Guettou, B., Lazri, M., Guiso, N., Rahal, K., 2016. Microbiological and molecular characterization of *Corynebacterium diphtheriae* isolated in Algeria between 1992 and 2015. *Clin. Microbiol. Infect.* 22 <https://doi.org/10.1016/j.cmi.2016.08.013>, 1005e1–1005.e7.
- Bouchez, V., Guillot, S., Landier, A., Armatys, N., Matczak, S., , the French pertussis microbiology study group, Toubiana, J., Brisse, S., 2021. Evolution of *Bordetella pertussis* over a 23-year period in France, 1996 to 2018. *Eurosurveillance* 26, 2001213. <https://doi.org/10.2807/1560-7917.ES.2021.26.37.2001213>.
- Bull, J.J., Antia, R., 2022. Which ‘imperfect vaccines’ encourage the evolution of higher virulence? *Evol. Med. Public Health* 10, 202–213. <https://doi.org/10.1093/emph/eoac015>.
- Carbonetti, N.H., 2010. Pertussis toxin and adenylate cyclase toxin: key virulence factors of *Bordetella pertussis* and cell biology tools. *Future Microbiol.* 5, 455–469. <https://doi.org/10.2217/fmb.09.133>.
- Ciotti, M., Ciccocci, M., Pieri, M., Bernardini, S., 2022. The COVID-19 pandemic: viral variants and vaccine efficacy. *Crit. Rev. Clin. Lab. Sci.* 59, 66–75. <https://doi.org/10.1080/10408363.2021.1979462>.
- Cruz-Hervert, L.P., Ferreira-Guerrero, E., Díaz-Ortega, J.L., Trejo-Valdivia, B., Téllez-Rojo, M.M., Mongua-Rodríguez, N., Hernández-Serrato, M.I., Montoya-Rodríguez, A. A., García-García, L., 2013. Cobertura de vacunación en adultos y adultos mayores en México. *Salud Pública México* 55, S300–S306. <https://doi.org/10.21149/spm.v55s2.5128>.
- Czajka, U., Wiatrzyk, A., Mosiej, E., Formińska, K., Zasada, A.A., 2018. Changes in MLST profiles and biotypes of *Corynebacterium diphtheriae* isolates from the diphtheria outbreak period to the period of invasive infections caused by nontoxicogenic strains in Poland (1950–2016). *BMC Infect. Dis.* 18, 121. <https://doi.org/10.1186/s12879-018-3020-1>.
- Dangel, A., Berger, A., Konrad, R., Bischoff, H., Sing, A., 2018. Geographically diverse clusters of nontoxicogenic *Corynebacterium diphtheriae* infection, Germany, 2016–2017. *Emerg. Infect. Dis.* 24, 1239–1245. <https://doi.org/10.3201/eid2407.172026>.
- Delamater, P.L., Street, E.J., Leslie, T.F., Yang, Y.T., Jacobsen, K.H., 2019. Complexity of the basic reproduction number (R₀). *Emerg. Infect. Dis.* 25, 1–4. <https://doi.org/10.3201/eid2501.171901>.
- De Serres, G., Shadmani, R., Duval, B., Boulianne, N., Déry, P., Douville Fradet, M., Rochette, L., Halperin, S.A., 2000. Morbidity of pertussis in adolescents and adults. *J. Infect. Dis.* 182, 174–179. <https://doi.org/10.1086/315648>.
- Dobie, R.A., 1979. Clinical features of diphtheria in the respiratory tract. *JAMA* 242, 2197. <https://doi.org/10.1001/jama.1979.03300200027017>.
- Edwards, K.M., Decker, M.D., 2012. Pertussis vaccines. *Vaccines*. Elsevier, Amsterdam, pp. 447–492.
- European Centre for Disease Prevention and Control, 2022. Increase of reported diphtheria cases due to *Corynebacterium diphtheriae* among migrants in Europe.
- Fine, P.E.M., 1993. Herd immunity: history, theory, practice. *Epidemiol. Rev.* 15, 265–302. <https://doi.org/10.1093/oxfordjournals.epirev.a036121>.
- Forrester, N.L., Trout, R.C., Gould, E.A., 2007. Benign circulation of rabbit haemorrhagic disease virus on Lambay Island, Eire. *Virology* 358, 18–22. <https://doi.org/10.1016/j.virol.2006.09.011>.
- Gandon, S., Day, T., 2008. Evidences of parasite evolution after vaccination. *Vaccine* 26, C4–C7. <https://doi.org/10.1016/j.vaccine.2008.02.007>.
- Gandon, S., Mackinnon, M., Nee, S., Read, A., 2003. Imperfect vaccination: some epidemiological and evolutionary consequences. *Proc. R. Soc. London* 270, 1129–1136. <https://doi.org/10.1098/rspb.2003.2370>.
- Gandon, S., Mackinnon, M.J., Nee, S., Read, A.F., 2002. Antitoxin vaccines and pathogen virulence. *Nature* 417. <https://doi.org/10.1038/417610a>, 610–610.
- Gandon, S., Mackinnon, M.J., Nee, S., Read, A.F., 2001. Imperfect vaccines and the evolution of pathogen virulence. *Nature* 414, 751–756. <https://doi.org/10.1038/414751a>.
- Giladi, M., Schulman, A., Kedem, R., Danon, Y.L., 1987. Measles in adults: a prospective study of 291 consecutive cases. *BMJ* 295. <https://doi.org/10.1136/bmj.295.6609.1314>, 1314–1314.
- Grainger, T.N., Levine, J.M., Gilbert, B., 2019. The invasion criterion: a common currency for ecological research. *Trends Ecol. Evol. (Amst.)* 34, 925–935. <https://doi.org/10.1016/j.tree.2019.05.007>.
- Gregg, K.A., Merkel, T.J., 2019. Pertussis toxin: a key component in pertussis vaccines? *Toxins (Basel)* 11, 557. <https://doi.org/10.3390/toxins11100557>.
- Guiso, N., Soubeyrand, B., Macina, D., 2022. Can vaccines control bacterial virulence and pathogenicity? *Bordetella pertussis*: the advantage of fitness over virulence. *Evol. Med. Public Health* 10, 363–370. <https://doi.org/10.1093/emph/eoac028>.
- Hammarlund, E., Thomas, A., Poore, E.A., Amanna, L.J., Rynko, A.E., Mori, M., Chen, Z., Slifka, M.K., 2016. Durability of vaccine-induced immunity against tetanus and diphtheria toxins: a cross-sectional analysis. *Clin. Infect. Dis.* 62, 1111–1118. <https://doi.org/10.1093/cid/ciw066>.
- Kanitz, E.E., Wu, L.A., Giambi, C., Strikas, R.A., Levy-Bruhl, D., Stefanoff, P., Mereckiene, J., Appelgren, E., D’Ancona, F., 2012. Variation in adult vaccination policies across Europe: an overview from VENICE network on vaccine recommendations, funding and coverage. *Vaccine* 30, 5222–5228. <https://doi.org/10.1016/j.vaccine.2012.06.012>.
- Lavine, J.S., King, A.A., Bjornstad, O.N., 2011. Natural immune boosting in pertussis dynamics and the potential for long-term vaccine failure. *Proc. Natl. Acad. Sci.* 108, 7259–7264. <https://doi.org/10.1073/pnas.1014394108>.
- Lee, H.-J., Choi, J.-H., 2017. Tetanus–diphtheria–acellular pertussis vaccination for adults: an update. *Clin. Exp. Vaccine Res.* 6, 22. <https://doi.org/10.7774/cevr.2017.6.1.22>.
- Li, S., Ma, C., Hao, L., Su, Q., An, Z., Ma, F., Xie, S., Xu, A., Zhang, Y., Ding, Z., Li, H., Cairns, L., Wang, H., Luo, H., Wang, N., Li, L., Ferrari, M.J., 2017. Demographic transition and the dynamics of measles in six provinces in China: a modeling study. *PLoS Med.* 14, e1002255. <https://doi.org/10.1371/journal.pmed.1002255>.
- Lu, Y., Geng, Y., Gan, W., Shi, L., 2019. Impacts of conformist on vaccination campaign in complex networks. *Physica A: Stat. Mech. Appl.* 526, 121124. <https://doi.org/10.1016/j.physa.2019.121124>.
- Marchal, C., Belhassen, M., Guiso, N., Jacoud, F., Van Ganse, E., Le Paner, M., Cohen, R., Verdier, R., Uhart, M., 2021. Vaccination coverage rates for diphtheria, tetanus, poliomyelitis and pertussis booster vaccination in France between 2013 and 2017: learnings from an analysis of National Health System Real-World Data. *Vaccine* 39, 505–511. <https://doi.org/10.1016/j.vaccine.2020.12.021>.
- Matsuyama, R., Akhmetzhanov, A.R., Endo, A., Lee, H., Yamaguchi, T., Tsuzuki, S., Nishiura, H., 2018. Uncertainty and sensitivity analysis of the basic reproduction number of diphtheria: a case study of a Rohingya refugee camp in Bangladesh. November–December 2017 *PeerJ* 6, e4583. <https://doi.org/10.7717/peerj.4583>.
- McLeod, D.V., Gandon, S., 2022. Effects of epistasis and recombination between vaccine-escape and virulence alleles on the dynamics of pathogen adaptation. *Nat. Ecol. Evol.* 6, 786–793. <https://doi.org/10.1038/s41559-022-01709-y>.
- Ministère des Solidarités et de la Santé, 2022. Calendrier Des Vaccinations Et Recommandations Vaccinales 2022.
- Möller, J., Kraner, M., Sonnewald, U., Sangal, V., Tittlbach, H., Winkler, J., Winkler, T. H., Melnikov, V., Lang, R., Sing, A., Mattos-Guaraldi, A.L., Burkovski, A., 2019. Proteomics of diphtheria toxoid vaccines reveals multiple proteins that are immunogenic and may contribute to protection of humans against *Corynebacterium diphtheriae*. *Vaccine* 37, 3061–3070. <https://doi.org/10.1016/j.vaccine.2019.04.059>.
- Mooi, F.R., van Loo, I.H.M., van Gent, M., He, Q., Bart, M.J., Heuvelman, K.J., de Greeff, S.C., Diavatopoulos, D., Teunis, P., Nagelkerke, N., Mertsola, J., 2009. *Bordetella pertussis* strains with increased toxin production associated with pertussis resurgence. *Emerg. Infect. Dis.* 15, 1206–1213. <https://doi.org/10.3201/eid1508.081511>.
- Murhekar, M.V., Bitragunta, S., Hutin, Y., Ckavravarty, A., Sharma, H.J., Gupte, M.D., 2009. Immunization coverage and immunity to diphtheria and tetanus among children in Hyderabad, India. *J. Infect.* 58, 191–196. <https://doi.org/10.1016/j.jinf.2008.12.010>.
- Murphy, J.R., 1996. *Corynebacterium diphtheriae*. In: Baron, S. (Ed.), *Medical Microbiology*. University of Texas Medical Branch at Galveston, Galveston, Texas.
- Oram, D.M., Holmes, R.K., 2006. Diphtheria toxin. *The Comprehensive Sourcebook of Bacterial Protein Toxins*. Academic Press, Paris, pp. 245–256.
- Pappenheimer, A.M., 1984. Diphtheria. In: Germanier, R. (Ed.), *Bacterial Vaccines*. Academic Press, Orlando, pp. 1–36.
- Pappenheimer, A.M., Gill, D.M., 1973. Diphtheria: recent studies have clarified the molecular mechanisms involved in its pathogenesis. *Science* 182, 353–358. <https://doi.org/10.1126/science.182.4110.353>.
- Parveen, S., Bishai, W.R., Murphy, J.R., 2019. *Corynebacterium diphtheriae*: diphtheria toxin, the *tox* operon, and its regulation by Fe²⁺ activation of apo-DtxR. *Microbiol. Spectr.* 7. <https://doi.org/10.1128/microbiolspec.GPP3-0063-2019>, 7.4.4.
- Poethko-Müller, C., Schmitz, R., 2013. Vaccination coverage in German adults. *Bundesgesundheitsbl* 56, 845–857. <https://doi.org/10.1007/s00103-013-1693-6>.
- Pullan, S., Dey, M., 2021. Vaccine hesitancy and anti-vaccination in the time of COVID-19: a Google trends analysis. *Vaccine* 39, 1877–1881. <https://doi.org/10.1016/j.vaccine.2021.03.019>.
- R Core Team, 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rohani, P., Zhong, X., King, A.A., 2010. Contact network structure explains the changing epidemiology of pertussis. *Science* 330, 982–985. <https://doi.org/10.1126/science.1194134>.
- Scanlon, K., Skerry, C., Carbonetti, N., 2019a. Role of major toxin virulence factors in pertussis infection and disease pathogenesis. In: Fedele, G., Ausiello, C.M. (Eds.), *Pertussis Infection and Vaccines*, Advances in Experimental Medicine and Biology. Springer International Publishing, Cham, pp. 35–51. https://doi.org/10.1007/5584_2019_403.
- Scanlon, K., Skerry, C., Carbonetti, N., 2019b. Association of pertussis toxin with severe pertussis disease. *Toxins (Basel)* 11, 373. <https://doi.org/10.3390/toxins11070373>.
- Scheifele, D.W., Ochnio, J.J., 2009. Module 2: diphtheria. *The Immunological Basis for Immunization Series*. Genève.
- Sein, C., Tiwari, T., Macneil, A., Wannemuehler, K., Soulayon, C., Soulyphone, P., Reyburn, R., Ramirez Gonzalez, A., Watkins, M., Goodson, J.L., 2016. Diphtheria outbreak in Lao People’s Democratic Republic, 2012–2013. *Vaccine* 34, 4321–4326. <https://doi.org/10.1016/j.vaccine.2016.06.074>.
- Soetaert, K., Petzoldt, T., Woodrow Setzer, R., 2010. Solving differential equations in R: package deSolve. *J. Stat. Softw.* 33, 1–25. <https://doi.org/10.18637/jss.v033.i09>.
- Soubeyrand, B., Plotkin, S.A., 2002. Antitoxin vaccines and pathogen virulence. *Nature* 417, 609–610. <https://doi.org/10.1038/417609b>.
- Strive, T., Elsworth, P., Liu, J., Wright, J.D., Kovaliski, J., Capucci, L., 2013. The non-pathogenic Australian rabbit calicivirus RCV-A1 provides temporal and partial cross protection to lethal rabbit haemorrhagic disease virus infection which is not dependent on antibody titres. *Vet. Res.* 44, 51. <https://doi.org/10.1186/1297-9716-44-51>.
- Talic, S., Shah, S., Wild, H., Gasevic, D., Maharaj, A., Ademi, Z., Li, X., Xu, W., Mesa-Eguiaagaray, I., Rostron, J., Theodoratou, E., Zhang, X., Motee, A., Liew, D., Ilic, D., 2021. Effectiveness of public health measures in reducing the incidence of covid-19, SARS-CoV-2 transmission, and covid-19 mortality: systematic review and meta-analysis. *BMJ*, e068302. <https://doi.org/10.1136/bmj-2021-068302>.

- Tiwari, T.S.P., Wharton, M., 2012. Diphtheria toxoid. *Vaccines*. Elsevier, Amsterdam, pp. 153–166.
- Truelove, S.A., Keegan, L.T., Moss, W.J., Chaisson, L.H., Macher, E., Azman, A.S., Lessler, J., 2020. Clinical and epidemiological aspects of diphtheria: a systematic review and pooled analysis. *Clin. Infect. Dis.* 71, 89–97. <https://doi.org/10.1093/cid/ciz808>.
- Walter, A., Lion, S., 2021. Epidemiological and evolutionary consequences of periodicity in treatment coverage. *Proc. R. Soc. B* 288, 20203007. <https://doi.org/10.1098/rspb.2020.3007>.
- Warfel, J.M., Zimmerman, L.I., Merkel, T.J., 2014. Acellular pertussis vaccines protect against disease but fail to prevent infection and transmission in a nonhuman primate model. *Proc. Natl. Acad. Sci.* 111, 787–792. <https://doi.org/10.1073/pnas.1314688110>.
- Wendelboe, A.M., Van Rie, A., Salmaso, S., Englund, J.A., 2005. Duration of immunity against pertussis after natural infection or vaccination. *Pediatr. Infect. Dis. J.* 24, S58–S61. <https://doi.org/10.1097/01.inf.0000160914.59160.41>.
- White, P.J., Norman, R.A., Hudson, P.J., 2002. Epidemiological consequences of a pathogen having both virulent and avirulent modes of transmission: the case of rabbit haemorrhagic disease virus. *Epidemiol. Infect.* 129, 665–677. <https://doi.org/10.1017/s095026880200777x>.
- Williams, M.M., Sen, K., Weigand, M.R., Skoff, T.H., Cunningham, V.A., Halse, T.A., Tondella, M.L., CDC Pertussis Working Group, 2016. *Bordetella pertussis* strain lacking pertactin and pertussis toxin. *Emerg. Infect. Dis.* 22, 319–322. <https://doi.org/10.3201/eid2202.151332>.
- Williams, W.W., Lu, P.-J., O'Halloran, A., Kim, D.K., Grohskopf, L.A., Pilishvili, T., Skoff, T.H., Nelson, N.P., Harpaz, R., Markowitz, L.E., Rodriguez-Lainz, A., Fiebelkorn, A.P., 2017. Surveillance of vaccination coverage among adult populations — United States, 2015. *MMWR Surveill. Summ.* 66, 1–28. <https://doi.org/10.15585/mmwr.ss6611a1>.
- World Health Organisation, 2017. Diphtheria vaccine: WHO position paper. *Wkly. Epidemiol. Rec.* 92, 417–435.
- Yasunaga, H., Shi, Y., Takeuchi, M., Horiguchi, H., Hashimoto, H., Matsuda, S., Ohe, K., 2010. Measles-related hospitalizations and complications in Japan, 2007–2008. *Intern. Med.* 49, 1965–1970. <https://doi.org/10.2169/internalmedicine.49.3843>.
- Zasada, A.A., 2013. Nontoxigenic highly pathogenic clone of *Corynebacterium diphtheriae*, Poland, 2004–2012. *Emerg. Infect. Dis.* 19 <https://doi.org/10.3201/eid1911.130297>.
- Zasada, A.A., Rzeczowska, M., 2019. Nontoxigenic *Corynebacterium diphtheriae* Infections. *Europe. Emerg. Infect. Dis.* 25, 1437–1438. <https://doi.org/10.3201/eid2507.180995>.
- Zou, R., Duan, X., Han, Z., Lu, Y., Ma, K., 2023. What information sources can prevent the epidemic: local information or kin information? *Chaos, Solit. Fractals* 168, 113104. <https://doi.org/10.1016/j.chaos.2023.113104>.

> Les vaccins contre la coqueluche sont commercialisés et utilisés dans la plupart des pays du monde depuis déjà plusieurs décennies. L'histoire des vaccins anticoquelucheux est particulière, les traditionnels vaccins cellulaires ayant été remplacés par des vaccins acellulaires dans de nombreux pays développés au cours des deux décennies passées. Selon leur composition, ces vaccins ont un effet important sur l'évolution de l'agent étiologique de la coqueluche, la bactérie *Bordetella pertussis*. Cette évolution se traduit par des changements génétiques et phénotypiques chez la bactérie, et a eu comme conséquences la résurgence de la coqueluche dans de nombreux pays ainsi que l'apparition de nouvelles souches échappant aux vaccins. L'émergence de nouvelles souches par pression vaccinale illustre ainsi l'importance de la prise en compte de la biologie de l'évolution de la bactérie dans l'élaboration de nouveaux vaccins. <

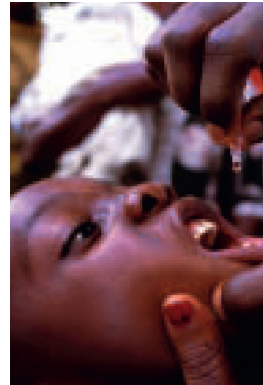
La coqueluche est une maladie infectieuse des voies respiratoires strictement humaine et causée par la bactérie *Bordetella pertussis*. Comme de nombreuses maladies humaines, l'origine de la coqueluche serait zoonotique. Les espèces du genre *Bordetella* infectent différentes espèces animales, principalement des mammifères mais certaines peuvent également infecter des oiseaux (comme *B. avium*). La lignée de *Bordetella pertussis* aurait divergé pour devenir spécialiste de l'hôte *Homo sapiens*, provoquant aujourd'hui la maladie appelée coqueluche [1]. Longtemps un fléau principalement chez les jeunes enfants, la coqueluche est combattue depuis maintenant presque un siècle par la vaccination.

L'effet de la vaccination sur la dynamique épidémiologique de la coqueluche a déjà fait l'objet de questionnements. Autrefois maladie extrêmement prévalente,

Vignette (© OMS).

Impact de la vaccination sur l'évolution de *Bordetella pertussis*

Florian Lecorvaisier



Université Claude Bernard Lyon
1 LBBE, UMR 5558, CNRS, VAS,
Villeurbanne, France.
florian.lecorvaisier@gmail.com
florian.lecorvaisier@univ-lyon1.fr

L'utilisation répandue des vaccins a conduit à une diminution globale du nombre de cas dans de nombreux pays. Par exemple, en Angleterre et au Pays de Galles, le nombre de cas déclarés par an est progressivement passé de plus de 90 000 en 1956 à moins de 10 000 dans les années 1990, et à moins de 1 000 dans les années 2000 [2]. La vaccination a cependant eu des conséquences moins prévisibles, comme un changement dans le mode de transmission (c'est-à-dire qui infecte qui, et comment), ou l'âge moyen auquel les enfants contractent généralement la maladie [3] (→).

Ces conséquences de la vaccination peuvent être regroupées en « effets à court terme », qu'on pourrait aussi appeler « conséquences écologiques ». Les effets à plus long terme, ou « conséquences évolutives », de la vaccination anticoquelucheuse ont commencé à être observés plus récemment. Ces effets sont fortement liés à l'histoire particulière des vaccins développés contre la coqueluche.

(→) Voir la Synthèse de N. Guiso, m/s n° 4, avril 2007, page 399

Histoire de la vaccination anticoquelucheuse

L'histoire des vaccins anticoquelucheux commence dans les années 1930 lorsque deux scientifiques américaines, Pearl Kendrick (1890-1980) et Grace Eldering (1900-1988), développent le premier vaccin efficace contre *B. pertussis* [4]. À la suite de leurs travaux pionniers, la recherche sur les vaccins avancera dans de nombreux pays et, dans les années 1940-1950, la plupart des pays développés intégreront un vaccin anticoquelucheux dans leur calendrier vaccinal. Les pays en voie de développement les rejoindront quelques années ou décennies plus tard.

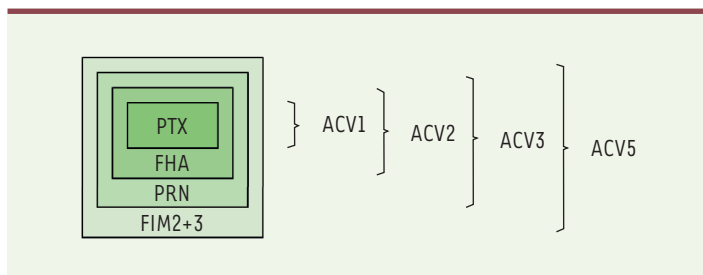


Figure 1. Composition en « poupées russes » des différents types de vaccins acellulaires utilisés. ACV : acellular vaccine ; PTX : toxine pertussique ; FHA : hémagglutinine filamenteuse ; PRN : pertactine ; FIM : fimbriae.

À cette époque, les vaccins qui sont élaborés sont des vaccins cellulaires, également appelés vaccins à cellules entières (ou WCV, pour *whole-cell vaccine*). Ce procédé implique l'inactivation de la bactérie qui est ensuite injectée dans l'organisme. Cette méthode est alors considérée comme hautement efficace, mais elle peut induire des effets secondaires parfois dramatiques. Parmi ceux-ci, on recense des réactions inflammatoires locales, une fièvre légère ou, plus rarement, une fièvre élevée pouvant être accompagnée de convulsions et de vomissements [5-7]. Ces vaccins cellulaires ont également été suspectés d'être à l'origine d'encéphalopathies et responsables de cas de mort subite du nourrisson. Cependant, ces effets ont par la suite été dissociés de l'emploi des vaccins cellulaires (pour une revue des symptômes induits par la vaccination cellulaire, voir [8, 9]).

Néanmoins, en 1974-1975, au Japon, deux décès suivant l'administration d'un vaccin anticoquelucheux conduiront le gouvernement à suspendre temporairement l'utilisation des WCV [10]. Des chercheurs japonais vont alors développer les premiers vaccins acellulaires (ACV, pour *acellular vaccine*). Ces vaccins entreront dans le calendrier vaccinal japonais en 1981 [10]. Les ACV se différencient des WCV par leur composition : alors que les WCV sont constitués de la bactérie entière ou de ses constituants, les ACV ne contiennent que quelques antigènes. Ces ACV sont ainsi considérés comme plus sûrs mais ils sont également plus onéreux que les WCV. Ils seront néanmoins adoptés par la majorité des pays développés dans les années 1990-2000, tandis que la plupart des pays en voie de développement continueront d'utiliser les traditionnels WCV. En France, les vaccins acellulaires seront d'abord introduits dans le calendrier vaccinal en 1998, pour être utilisés sous forme de rappels. Ils commenceront à être utilisés pour la primo-vaccination en 2001 et remplaceront totalement les WCV dès 2004 [11, 12].

Les antigènes inclus dans les ACV ont été sélectionnés pour leur rôle protecteur et leur immunogénicité. Ces antigènes sont au nombre de cinq : la toxine pertussique (PTX), l'hémagglutinine filamenteuse (FHA), la pertactine (PRN), et les fimbriae de type 2 et de type 3 (Fim2 et Fim3). Les vaccins actuellement sur le marché comprennent certains ou l'ensemble de ces antigènes : il existe ainsi des ACV contenant uniquement la PTX (ACV1), ou la PTX et la FHA (ACV2), ou la PTX, la FHA et la PRN (ACV3), ou les cinq antigènes (ACV5) (Figure 1). Les vaccins anticoquelucheux actuellement commercialisés portent également des antigènes d'autres agents pathogènes, notamment *Corynebacterium diphtheriae* (agent de la diphtérie) et *Clostridium tetani* (agent du tétanos) (vaccins acellulaires DTaP ou cellulaires DTwP) [11]. En France, depuis 2018, les nouveau-nés sont vaccinés par

un vaccin hexavalent protégeant contre la diphtérie, la coqueluche, le tétanos, la poliomyélite, l'infection à *Haemophilus influenzae* de type b et l'hépatite B [13]. Il existe donc deux grandes familles de vaccins anticoquelucheux : les vaccins cellulaires (historiques) et les vaccins acellulaires (plus récents) dont quatre types ont été développés. Les pays n'utilisent pas tous les mêmes vaccins, et cela a comme conséquence des pressions de sélection différentes menant à la sélection de nouvelles souches de *B. pertussis*.

Sélection de nouveaux allèles de *B. pertussis* après vaccination

Depuis environ deux décennies, plusieurs pays signalent une recrudescence des cas de coqueluche, malgré une couverture vaccinale généralement stable et élevée. Plusieurs études relient ces résurgences de l'infection à l'émergence de nouvelles souches de *B. pertussis* résultant de l'utilisation de vaccins [14-19].

Chez *B. pertussis*, la pertactine, une protéine de sa membrane externe, est codée par le gène *prn*. La toxine pertussique est quant à elle codée par différents gènes, dont *ptxA*. La majorité des vaccins utilisés dans le monde sont produits à partir de souches portant les l'allèle *prn1* de *prn*, et l'allèle *ptxA2* de *ptxA*. Il existe cependant des vaccins utilisant des souches pourvues d'autres allèles et l'allèle *ptxA2* ne constitue que 54 % des souches vaccinales [20]. Lors de leur isolement, ces souches vaccinales présentaient le profil allélique typique des souches sauvages alors en circulation dans les populations, ce qui n'est plus le cas aujourd'hui.

L'allèle majoritaire de *prn* pendant la période pré-vaccinale, qui est utilisé aujourd'hui pour la formulation des vaccins, est l'allèle *prn1*. Depuis quelques années, l'émergence et la propagation d'un nouvel allèle, *prn2*, absent des vaccins, ont été largement documentées dans plusieurs pays comme l'Australie [21], le Japon [19], la Russie [22], le Royaume-Uni [23], la Suède [24] ou encore la Finlande [25]. Le lien entre la vaccination et l'émergence de ce nouvel allèle fait peu de doute, ce dernier apparaissant dans chacun de ces



pays après l'introduction de la vaccination. Au Danemark, l'allèle *prn2* a été détecté pour la première fois au début des années 1970, après l'introduction d'un WCV dans le pays. Cet allèle deviendra largement majoritaire dans les années 1990, jusqu'à ce que l'allèle *prn1* disparaisse complètement en 2000. En 2003, le Danemark remplacera son vaccin cellulaire par un vaccin acellulaire, un ACV1, dépourvu de PRN. Quelques années plus tard, l'allèle *prn1* sera de nouveau retrouvé dans les données épidémiologiques [26]. Il apparaît donc que, lorsqu'une population est immunisée par un vaccin contenant l'allèle *prn1*, l'allèle *prn2* est favorisé. En revanche, en l'absence de pression vaccinale contre cet allèle, l'allèle *prn1* réapparaît. Ce dernier confère certainement un avantage adaptatif à la bactérie au sein d'une population non immunisée contre la PRN.

La toxine pertussique est constituée de deux domaines : le domaine B, responsable de l'adhérence de la toxine à la surface de la cellule cible, et le domaine A, qui supporte l'activité enzymatique (toxique) de la protéine et pénètre dans la cellule [27]. Le domaine A est codé par le gène *ptxA*. L'allèle pré-vaccinal qui a été utilisé dans de nombreuses préparations vaccinales est l'allèle *ptxA2*. Cependant l'émergence d'un nouvel allèle, *ptxA1*, a été signalée dans de nombreux pays, dont le Danemark [26], l'Italie [28], le Japon [19] et le Royaume-Uni [23]. Comme pour l'allèle *prn2* de *prn*, l'allèle *ptxA1* a émergé dans ces pays suite à l'introduction des vaccins. L'exemple de la Serbie est, à ce titre, éclairant. La particularité de la Serbie est en effet d'avoir modifié la composition de son WCV au cours de son histoire. Jusqu'en 1985, le WCV serbe contenait deux souches exprimant *ptxA2* et, au cours des années 1960 et 1970, on a pu observer l'émergence et la propagation de souches exprimant l'allèle *ptxA1*. En 1985, deux nouvelles souches contenant l'allèle *ptxA1* ont été ajoutées au vaccin et la proportion de souches portant l'allèle ancestral *ptxA2* a de nouveau augmenté [29]. L'utilisation de l'allèle *ptxA2* dans les vaccins semble donc avoir exercé une pression de sélection favorisant un nouvel allèle, *ptxA1*. Cependant, l'ajout de ce nouvel allèle dans le WCV a créé un environnement propice à la réémergence de l'allèle *ptxA2*, bien qu'il soit toujours présent dans le vaccin. La réémergence de l'allèle *ptxA2*, pourtant présent dans les vaccins, est un phénomène qui n'a pas encore été compris. Si l'on compare à la PRN, le vaccin serbe utilisé avant 1985 contenait l'allèle *prn1*, et on a observé l'émergence de souches *prn2*. En 1985, deux souches *prn2* ont été introduites dans le vaccin, causant alors l'émergence d'un nouvel allèle, *prn11* [29]. Les causes de la sélection d'un nouvel allèle non vaccinal pour la PRN mais pas pour la PTX restent à déterminer.

Perte d'expression de la pertactine par *B. pertussis*

Un phénomène autre que l'émergence d'allèles non vaccinaux a également été décrit : la perte d'expression de la pertactine (PRN). En effet, l'émergence de souches déficientes en PRN (ou PRN-) a été signalée dans de nombreux pays. Les vaccins ACV et WCV ne semblent cependant pas exercer la même pression de sélection sur les souches PRN-.

Le Brésil et l'Argentine sont deux pays utilisant toujours des vaccins de type WCV. Des ACV sont également disponibles dans ces deux pays mais

ils sont peu utilisés. En Argentine, une étude fondée sur 350 souches collectées entre 2000 et 2017 n'a mis en évidence que deux souches PRN- isolées en 2016, dont l'une provenait d'un bébé qui avait été vacciné par un ACV [30]. De même, au Brésil, l'étude de 511 souches collectées entre 2000 et 2016 n'a révélé que trois souches PRN- [31]. En Europe, la Russie et, dans une moindre mesure, la Pologne, utilisent toujours un WCV. En Russie, sur 75 souches collectées entre 2001 et 2009, aucune n'était PRN- [32]. En Pologne, cependant, parmi les 188 souches isolées entre 1959 et 2016, quatre d'entre elles étaient PRN-, ce qui représente 15,4 % des souches collectées sur la période 2010-2016. Notons que pendant cette période, la proportion d'ACV a dépassé celle des WCV pour la primo-vaccination dans ce pays. [33]. Les vaccins de type WCV ne semblent donc pas créer de pression de sélection en faveur des souches PRN-.

L'Australie et le Japon représentent deux exemples-type du rôle des ACV dans la sélection de souches PRN-. En Australie, les ACV (ACV3-5) ont remplacé totalement les WCV dès 1999. Les premières souches PRN- ont été identifiées en 2008, représentant alors 5 % des souches collectées [34]. Ce chiffre va rapidement augmenter pour atteindre 89,7 % sur la période 2013-2017 [35]. Le délai entre l'introduction de la vaccination par ACV et l'émergence des souches PRN- peut être dû au temps de renouvellement de la population et/ou au temps requis pour la perte de l'immunité acquise avec les vaccins WCV qui circulaient avant 1999.

Le Japon est le premier pays à avoir introduit la vaccination par ACV dans son calendrier vaccinal, dès 1981. C'est pendant la période 1995-1999 que les premières souches PRN- seront isolées, représentant alors 5 % des souches collectées. Cette proportion augmentera les années suivantes, pour culminer à 41 % des souches isolées entre 2005 et 2007 [36, 37]. Par la suite, la proportion de souches PRN- va continuellement diminuer, pour atteindre 8 % sur la période 2014-2016 [37]. Cette dynamique particulière peut s'expliquer par la composition des ACV utilisés au Japon. Avant 2012, cinq vaccins ACV circulaient au Japon, parmi lesquels seulement trois contenaient la PRN. En conséquence, une partie de la population n'a pas développé d'immunité vaccinale contre la PRN, expliquant la proportion faible de souches PRN- par rapport à l'Australie. En 2012 deux nouveaux vaccins dépourvus de PRN ont été mis sur le marché et largement distribués. Cela a conduit à ce que la plupart des enfants nés après cette date ne développent pas d'immunité contre la PRN, ce qui explique, en partie, la diminution de la proportion de souches PRN-,

la pression de sélection ayant été relâchée. En 2015, de nouveaux vaccins comprenant la PRN ont été mis sur le marché. Il est donc possible qu'on observe à nouveau une augmentation de la prévalence des souches PRN- mais le manque de données ne permet pas de conclure pour l'instant.

L'Australie et le Japon ne sont que deux exemples de pays dans lesquels on retrouve ce phénomène d'émergence de souches PRN- à la suite de l'introduction des ACV. Celle-ci a été documentée dans d'autres pays, comme en Nouvelle-Zélande [38], aux États-Unis [39] ou au Canada [40]. En Europe, une corrélation temporelle a même été établie entre la durée depuis l'introduction de la vaccination et la proportion de souches PRN- en circulation [11].

La perte de la PRN semble résulter de modifications génétiques plutôt que d'une plasticité phénotypique : les souches perdent leur capacité à produire la PRN, et pas simplement à contrôler sa production. Parmi les mécanismes identifiés, on peut citer l'introduction d'éléments transposables dans la séquence codante du gène *prn*, des mutations non-sens, des délétions de séquences de plusieurs milliers de bases, ou encore une perte complète du gène *prn* [11, 36, 39]. Cela signifie que l'émergence de souches PRN- n'est pas issue de la propagation clonale d'une même souche, mais que plusieurs événements indépendants d'acquisition de la déficience en PRN ont lieu dans différents pays, voire au sein d'un même pays. Il s'agit donc d'un exemple typique de convergence évolutive, produisant un même résultat sans que son origine soit commune, indiquant une forte pression de sélection contre la production de la PRN (Figure 2).

Surexpression de la toxine pertussique par *B. pertussis*

La toxine pertussique (PTX) est l'un des principaux facteurs de virulence de *B. pertussis*, et tous les vaccins, WCV comme ACV, contiennent la PTX. Dans la bactérie, la production de la toxine est contrôlée par l'opéron *ptx*, constitué de plusieurs gènes sous le contrôle du promoteur *ptxP*. En population pré-vaccinale, l'allèle majoritaire de *ptxP* était *ptxP1*. Cependant, depuis l'introduction de la vaccination, de nombreuses études ont signalé l'émergence d'un nouvel allèle, *ptxP3*, notamment en Australie [34], au Danemark [26], en Italie [28], au Japon [19], aux Pays-Bas [41] et en Suède [42].

L'allèle *ptxP3* serait associé à une plus forte production de PTX et donc à une plus grande virulence [41]. L'émergence de souches produisant d'importantes quantités de PTX s'expliquerait par le mode d'action des vaccins acellulaires. Les ACV permettent d'éviter les formes graves de la coqueluche, en particulier par la production d'anticorps spécifiques de la toxine. En revanche, ils échoueraient à bloquer la colonisation et la transmission de la bactérie. Ils induisent en effet une réponse de type Th1/Th2 alors que l'immunité acquise naturellement ou induite par les WCV induit une réponse Th1/Th17, cette dernière permettant de réduire la colonisation [43].

Les ACV réduisent donc la mortalité associée à la production de la PTX (puisque'ils diminuent les symptômes qu'elle induit) mais ne freinent pas la propagation de la bactérie. Celle-ci pourrait donc, en théo-

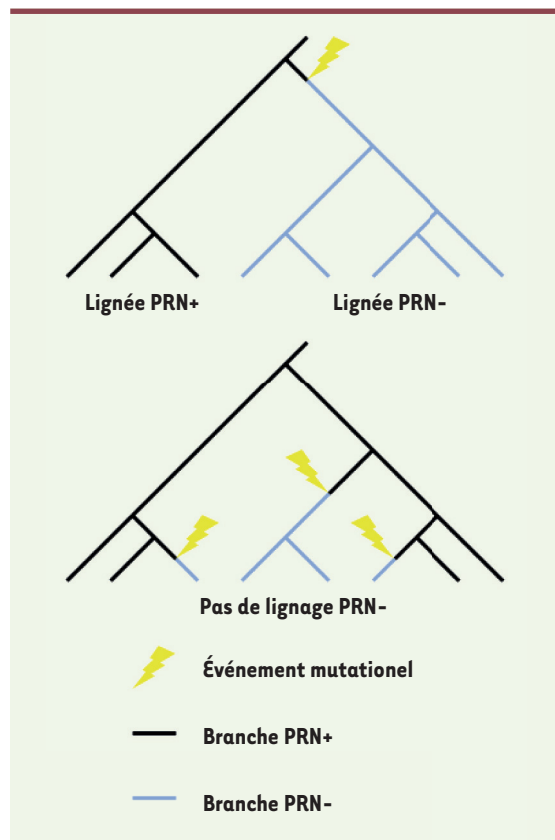


Figure 2. Deux scénarios d'apparition du phénotype « déficience en PRN ». En haut, le scénario de la dispersion monoclonale : la déficience en PRN (PRN-) apparaît une fois dans la phylogénie de *Bordetella pertussis*, et toutes les souches PRN- partagent le même ancêtre commun. En bas, le scénario de la convergence évolutive : la déficience en PRN apparaît plusieurs fois dans la phylogénie de *B. pertussis* par des mécanismes variés et les souches PRN- ne partagent pas d'ancêtre commun. PRN : pertactine.

rie, produire de très grandes quantités de toxine sans induire de formes graves de la maladie chez l'hôte vacciné. Des souches produisant plus de toxine - et donc, plus virulentes - pourraient donc être sélectionnées et émerger [44].

Course à l'armement contre les pathogènes

Il existe une théorie en biologie de l'évolution qui énonce que pour coexister dans un environnement, deux espèces antagonistes doivent continuellement s'adapter mutuellement. C'est la théorie dite de la « Reine Rouge » [45]. Dans le cas de la coqueluche, la sélection naturelle a fait son œuvre en sélectionnant des souches de *B. pertussis* capables de résister à l'immunité induite par les WCV. En revanche, pour l'espèce humaine, cible de la bactérie, aucune modification pouvant compenser cette évolution n'a pour l'instant été décrite.



Les vaccins WCV historiques ont progressivement été remplacés par les ACV. Ces nouveaux vaccins ne furent pas développés selon une approche évolutive mais plutôt selon une approche médicale. Leur objectif n'était pas de créer de nouvelles pressions contre la bactérie, étant donné que les souches employées pour produire les ACV sont généralement identiques à celles utilisées pour produire les WCV. L'objectif principal était donc plutôt d'obtenir des vaccins provoquant moins d'effets secondaires que les WCV tout en restant aussi efficaces. Alors que le premier point est globalement considéré comme une réussite par la communauté scientifique et médicale, le deuxième point a souvent été remis en question dans la littérature scientifique et reste encore débattu. Des expériences sur les babouins ont montré que les ACV échouent à protéger contre l'infection par la bactérie et sa transmission [43]. Des observations au sein de la population humaine remettent néanmoins en question cette conclusion, l'utilisation massive des ACV étant associée à une diminution des cas de coqueluche dans les pays utilisant cette forme de vaccin [46].

Conclusion

Au cours des décennies suivant l'introduction des vaccins anticoquelucheux, au milieu du ^{xx} siècle, le génotype et le phénotype des souches circulantes de *B. pertussis* ont été modifiés. L'allèle *prn1*, majoritairement exprimé par les souches avant l'avènement des vaccins, a été remplacé par l'allèle *prn2*, et l'allèle *ptxA2* par l'allèle *ptxA1*. Les vaccins acellulaires ACV entraînent la sélection de souches ne produisant plus la pertactine, l'un des antigènes communément inclus dans ces vaccins, et de nouvelles souches, produisant plus de toxine pertussique, ont émergé.

Les vaccins représentent certainement la plus puissante des armes dont dispose l'humanité pour combattre les agents pathogènes, suivie de près par les antibiotiques. L'efficacité de certains vaccins reçoit depuis quelques décennies le contre-coup de l'évolution. Face aux promesses d'un monde où il serait possible d'éliminer les maladies à prévention vaccinale, la sélection complique cette tâche prométhéenne et nous rappelle que le vivant est plongé dans une lutte constante pour survivre et se reproduire, favorisant les microbes capables d'évoluer afin d'échapper au filtre des antibiotiques ou des vaccins. ♦

SUMMARY

Impact of vaccination on the evolution of *Bordetella pertussis*

Vaccines against pertussis, or whooping cough, have been commercialized and used in most countries worldwide for decades. The history of these vaccines is distinctive, marked by the transition from whole-cell vaccines to acellular vaccines in many developed countries over the last two decades. This particular history has had a significant impact on the evolution of *Bordetella pertussis*, the etiological agent of whooping cough. Both genetic and phenotypic changes appeared, with the emergence of novel alleles for antigens targeted by the vaccines and changes in the expression of these antigens. The main consequence of these changes is the resurgence of whooping cough in many countries and the appearance of strains capable of evading vaccine-induced

immunity. The emergence of novel strains under vaccine pressure underscores the importance of considering biological evolution in the conception of new vaccines and vaccine strategies. ♦

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

- Lin Z, Ivanov YV, Preston A, et al. Acquisition and loss of virulence-associated factors during genome evolution and speciation in three clades of *Bordetella* species. *BMC Genomics* 2016 ; 17 : 767.
- Choi YH, Campbell H, Amirthalingam G, et al. Investigating the pertussis resurgence in England and Wales, and options for future control. *BMC Med* 2016 ; 14 : 121.
- Guiso N. Impact de la vaccination sur l'épidémiologie des maladies infectieuses : exemple de la coqueluche. *Med Sci (Paris)* 2007 ; 23 : 399-403.
- Shapiro-Shapin CG, Pearl Kendrick, Grace Eldering, and the Pertussis Vaccine. *Emerg Infect Dis* 2010 ; 16 : 1273-8.
- Miller CL, Pollock TM, Clewer ADE. Whooping-cough vaccination : An assessment. *Lancet* 1974 ; 304 : 510-3.
- Cody CL, Baraff LJ, Cherry JD, et al. Nature and Rates of Adverse Reactions Associated with DTP and DT Immunizations in Infants and Children. *Pediatrics* 1981 ; 68 : 650-60.
- Strom J. Further experience of reactions, especially of a cerebral nature, in conjunction with triple vaccination : a study based on vaccinations in Sweden 1959-65. *BMJ* 1967 ; 4 : 320-3.
- Boulesteix J. Tolérance et efficacité du vaccin anti-coquelucheux entier. *Médecine et Maladies Infectieuses* 1995 ; 25 : 1299-304.
- Trollfors B. *Bordetella Pertussis* Whole Cell Vaccines-Efficacy and Toxicity. *Acta Paediatrica* 1984 ; 73 : 417-25.
- Nakayama T. Vaccin chronique in Japan. *J Infect Chemother* 2013 ; 19 : 787-98.
- Barkoff A-M, Mertsola J, Pierard D, et al. Pertactin-deficient *Bordetella pertussis* isolates : evidence of increased circulation in Europe, 1998 to 2015. *Eurosurveillance* 2019 ; 24 : 1700 832.
- Belchior E, Guillot S, Poujol I, et al. Comparison of whole-cell versus acellular pertussis vaccine effectiveness in school clusters of pertussis, France, 2013. *Médecine et Maladies Infectieuses* 2020 ; 50 : 617-9.
- Ministère de la Santé et de la Prévention. Calendrier des vaccinations et recommandations vaccinales 2023. Paris, 2023. <https://sante.gouv.fr/prevention-en-sante/preserver-sa-sante/vaccination/calendrier-vaccinal>
- Lefrancq N, Bouchez V, Fernandes N, et al. Global spatial dynamics and vaccine-induced fitness changes of *Bordetella pertussis*. *Sci Transl Med* 2022 ; 14 : eabn3253.
- Brandal LT, Vestreheim DF, Bruvik T, et al. Evolution of *Bordetella pertussis* in the acellular vaccine era in Norway, 1996 to 2019. *Eur J Clin Microbiol Infect Dis* 2022 ; 41 : 913-24.
- Lesne E, Cavell BE, Freire-Martin I, et al. Acellular Pertussis Vaccines Induce Anti-pertactin Bactericidal Antibodies Which Drives the Emergence of Pertactin-Negative Strains. *Front Microbiol* 2020 ; 11 : 2108.
- Jayasundara D, Lee E, Octavia S, et al. Emergence of pertactin-deficient pertussis strains in Australia can be explained by models of vaccine escape. *Epidemics* 2020 ; 31 : 100 388.
- Etskovitz H, Anastasio N, Green E, et al. Role of Evolutionary Selection Acting on Vaccine Antigens in the Re-Emergence of *Bordetella Pertussis*. *Diseases* 2019 ; 7 : 35.
- Zomer A, Otsuka N, Hiramatsu Y, et al. *Bordetella pertussis* population dynamics and phylogeny in Japan after adoption of acellular pertussis vaccines. *Microb Genom* 2018 ; 4 : e000180.
- Bart MJ, Harris SR, Advani A, et al. Global Population Structure and Evolution of *Bordetella pertussis* and Their Relationship with Vaccination. *mBio* 2014 ; 5 : e01074-14.
- Octavia S, Sintchenko V, Gilbert GL, et al. Newly Emerging Clones of *Bordetella pertussis* Carrying *prn2* and *ptxP3* Alleles Implicated in Australian Pertussis Epidemic in 2008-2010. *J Infect Dis* 2012 ; 205 : 1220-4.

RÉFÉRENCES

22. Borisova O, Kombarova SY, Zakharova NS, et al. Antigenic Divergence between *Bordetella pertussis* Clinical Isolates from Moscow, Russia, and Vaccine Strains. *Clin Vaccine Immunol* 2007 ; 14 : 234-8.
23. Fry NK, Neal S, Harrison TG, et al. Genotypic Variation in the *Bordetella pertussis* Virulence Factors Pertactin and Pertussis Toxin in Historical and Recent Clinical Isolates in the United Kingdom. *Infect Immun* 2001 ; 69 : 5520-8.
24. Hallander HO, Advani A, Donnelly D, et al. Shifts of *Bordetella pertussis* Variants in Sweden from 1970 to 2003, during Three Periods Marked by Different Vaccination Programs. *J Clin Microbiol* 2005 ; 43 : 2856-65.
25. Elomaa A, Advani A, Donnelly D, et al. Strain Variation among *Bordetella pertussis* Isolates in Finland, Where the Whole-Cell Pertussis Vaccine Has Been Used for 50 Years. *J Clin Microbiol* 2005 ; 43 : 3681-7.
26. Petersen RF, Dalby T, Dragsted DM, et al. Temporal Trends in *Bordetella pertussis* Populations, Denmark, 1949-2010. *Emerg Infect. Dis* 2012 ; 18 : 767-74.
27. Scanlon K, Skery C, Carbonetti N. Role of Major Toxin Virulence Factors in Pertussis Infection and Disease Pathogenesis. In : Fedele G, Ausiello CM, editors. *Pertussis Infection and Vaccines*. Advances in Experimental Medicine and Biology. Cham : Springer International Publishing, 2019 : pp. 35-51.
28. Loconsole D, De Robertis AL, Morea A, et al. Resurgence of pertussis and emergence of the *ptxP3* toxin promoter allele in South Italy. *Pediat Infect Dis J* 2018 ; 37 : e126-31.
29. Dakic G, Kallonen T, Elomaa A, et al. *Bordetella pertussis* vaccine strains and circulating isolates in Serbia. *Vaccine* 2010 ; 28 : 1188-92.
30. Carriquiriborde F, Regidor V, Aispuro PM, et al. Rare Detection of *Bordetella pertussis* Pertactin-Deficient Strains in Argentina. *Emerg Infect Dis* 2019 ; 25 : 2048-54.
31. Leite D, Camargo CH, Kashino SS, et al. Prevalence and characterization of pertactin deficient *Bordetella pertussis* strains in Brazil, a whole-cell vaccine country. *Vaccine : X* 2021 ; 8 : 100103.
32. Kurova N, Njamkepo E, Brun D, et al. Monitoring of *Bordetella* isolates circulating in Saint Petersburg, Russia between 2001 and 2009. *Res Microbiol* 2010 ; 161 : 810-5.
33. Polak M, Zasada AA, Mosiej E, et al. Pertactin-deficient *Bordetella pertussis* isolates in Poland — a country with whole-cell pertussis primary vaccination. *Microbes Infect* 2019 ; 21 : 170-5.
34. Lam C, Octavia S, Ricafort L, et al. Rapid Increase in Pertactin-deficient *Bordetella pertussis* Isolates, Australia. *Emerg Infect Dis* 2014 ; 20 : 626-33.
35. Xu Z, Octavia S, Luu LDW, et al. Pertactin-Negative and Filamentous Hemagglutinin-Negative *Bordetella pertussis*, Australia, 2013-2017. *Emerg Infect Dis* 2019 ; 25 : 1196-9.
36. Otsuka N, Han H-J, Toyozumi-Ajisaka H, et al. Prevalence and genetic characterization of pertactin-deficient *Bordetella pertussis* in Japan. *PLoS ONE* 2012 ; 7 : e31985.
37. Hiramatsu Y, Miyaji Y, Otsuka N, et al. Significant Decrease in Pertactin-Deficient *Bordetella pertussis* Isolates, Japan. *Emerg Infect Dis* 2017 ; 23 : 699-701.
38. Ring N, Davies H, Morgan J, et al. Comparative genomics of *Bordetella pertussis* isolates from New Zealand, a country with an uncommonly high incidence of whooping cough. *Microb Genom* 2022 ; 8 : 000756.
39. Pawloski LC, Queenan AM, Cassidy PK, et al. Prevalence and Molecular Characterization of Pertactin-Deficient *Bordetella pertussis* in the United States. *Clin Vaccine Immunol* 2014 ; 21 : 119-25.
40. Tsang RSW, Shuel M, Cronin K, et al. The evolving nature of *Bordetella pertussis* in Ontario, Canada, 2009-2017 : strains with shifting genotypes and pertactin deficiency. *Can J Microbiol* 2019 ; 65 : 823-30.
41. Mooi FR, Loo IHM van, Gent M van, et al. *Bordetella pertussis* Strains with Increased Toxin Production Associated with Pertussis Resurgence. *Emerg Infect Dis* 2009 ; 15 : 1206-13.
42. Advani A, Gustafsson L, Åhrén C, et al. Appearance of *Fim3* and *ptxP3*-*Bordetella pertussis* strains, in two regions of Sweden with different vaccination programs. *Vaccine* 2011 ; 29 : 3438-42.
43. Warfel JM, Zimmerman LI, Merkel TJ. Acellular pertussis vaccines protect against disease but fail to prevent infection and transmission in a nonhuman primate model. *Proc Natl Acad Sci USA* 2014 ; 111 : 787-92.
44. Gandon S, Mackinnon MJ, Nee S, et al. Imperfect vaccines and the evolution of pathogen virulence. *Nature* 2001 ; 414 : 751-6.
45. Van Valen L. A new evolutionary law. *Evolutionary Theory* 1973 ; 1 : 1-30.
46. Domenech de Cellès M, Magpantay FMG, King AA, et al. The pertussis enigma : reconciling epidemiology, immunology and evolution. *Proc R Soc B* 2016 ; 283 : 20152309.

TIRÉS À PART

F. Lecorvaisier



Guide pratique de la théorie polyvagale

S'ancrer dans la sécurité de Deb Dana

La théorie polyvagale offre une approche inédite de la santé physique, mentale et du bien-être. Accessible à tous, ce livre de Deb Dana invite le lecteur à expérimenter la théorie plutôt que simplement la comprendre. Grâce à des outils concrets, il pourra passer de la théorie à la pratique en prenant conscience des divers états de son système nerveux autonome, trouvant ainsi son propre chemin vers la régulation des voies vagales ventrales...Un guide indispensable pour les thérapeutes, et pour tous ceux qui veulent mieux s'armer face aux défis quotidiens, découvrant leur système nerveux autonome et leur point d'ancrage dans la sécurité.

Ce livre vient en complément de 'Théorie polyvagale et sentiment de sécurité', de Stephen Porges, qui traite de façon accessible des aspects théoriques.

ISBN : 978-2-7598-3211-8 148 pages - 12 € TTC

En vente sur laboutique.edpsciences.org



B – Suppléments au Chapitre 2

B.1 – Critère d'invasion mutuelle

On s'intéresse en premier lieu aux conditions dans lesquelles la souche 1 peut envahir une population où la souche 2 est à l'équilibre endémique. Par définition, on a ainsi

$$\begin{aligned}\frac{dI_2^*(t)}{dt} &= 0 \\ \iff \beta_2 s_2^* I_2^*(t) + \beta_2' v_2^* I_2^*(t) - \gamma I_2^*(t) - d I_2^*(t) &= 0 \\ \iff \beta_2 s_2^* I_2^*(t) + \beta_2' v_2^* I_2^*(t) &= \gamma I_2^*(t) + d I_2^*(t) \\ \iff \beta_2 s_2^* + \beta_2' v_2^* &= \gamma + d,\end{aligned}\tag{S1}$$

où s_2^* , v_2^* et I_2^* sont la proportion d'individus susceptibles, la proportion d'individus vaccinés et le nombre d'individus infectés par la souche 2 à l'équilibre endémique avec la souche 2, respectivement.

La souche 1 peut envahir la population où la souche 2 est à l'équilibre si et seulement si le flux d'individus entrant dans le compartiment I_1 est positif, donc

$$\begin{aligned}\frac{dI_1(t)}{dt} &> 0 \\ \iff \beta_1 s_2^* I_1(t) + \beta_1 v_2^* I_1(t) - \gamma I_1(t) - d I_1(t) &> 0 \\ \iff \beta_1 s_2^* I_1(t) + \beta_1 v_2^* I_1(t) &> \gamma I_1(t) + d I_1(t) \\ \iff \beta_1 s_2^* + \beta_1 v_2^* &> \gamma + d.\end{aligned}\tag{S2}$$

En injectant la fin de la formule (S1) dans la formule (S2), on obtient finalement

$$\begin{aligned}\beta_1 s_2^* + \beta_1 v_2^* &> \beta_2 s_2^* + \beta_2' v_2^* \\ \iff (\beta_1 - \beta_2') v_2^* &> (\beta_2 - \beta_1) s_2^* \\ \iff \frac{(\beta_1 - \beta_2')}{(\beta_2 - \beta_1)} &> \frac{s_2^*}{v_2^*}.\end{aligned}\tag{S3}$$

On s'intéresse ensuite aux conditions dans lesquelles la souche 2 peut envahir une population

où la souche 1 est à l'équilibre. Par définition, on a ainsi

$$\begin{aligned}
& \frac{dI_1^*(t)}{dt} = 0 \\
\iff & \beta_1 s_1^* I_1^*(t) + \beta_1 v_1^* I_1^*(t) - \gamma I_1^*(t) - dI_1^*(t) = 0 \\
& \iff \beta_1 s_1^* I_1^*(t) + \beta_1 v_1^* I_1^*(t) = \gamma I_1^*(t) + dI_1^*(t) \\
& \iff \beta_1 s_1^* + \beta_1 v_1^* = \gamma + d,
\end{aligned} \tag{S4}$$

où s_1^* , v_1^* et I_1^* sont la proportion d'individus susceptibles, la proportion d'individus vaccinés et le nombre d'individus infectés par la souche 1 à l'équilibre endémique avec la souche 1, respectivement.

La souche 2 peut envahir la population où la souche 1 est à l'équilibre si et seulement si le flux d'individus entrant dans le compartiment I_2 est positif, donc

$$\begin{aligned}
& \frac{dI_2(t)}{dt} > 0 \\
\iff & \beta_2 s_1^* I_2(t) + \beta_2' v_1^* I_2(t) - \gamma I_2(t) - dI_2(t) > 0 \\
& \iff \beta_2 s_1^* I_2(t) + \beta_2' v_1^* I_2(t) > \gamma I_2(t) + dI_2(t) \\
& \iff \beta_2 s_1^* + \beta_2' v_1^* > \gamma + d.
\end{aligned} \tag{S5}$$

En injectant la fin de la formule (S4) dans la formule (S5), on obtient finalement

$$\begin{aligned}
& \beta_2 s_1^* + \beta_2' v_1^* > \beta_1 s_1^* + \beta_1 v_1^* \\
\iff & (\beta_2 - \beta_1) s_1^* > (\beta_1 - \beta_2') \\
& \iff \frac{s_1^*}{v_1^*} > \frac{\beta_1 - \beta_2'}{\beta_2 - \beta_1}
\end{aligned} \tag{S6}$$

Finalement, en combinant les inéquations (S3) et (S6), on retrouve la formule (2.11).

B.2 – Durée des simulations

Le choix de la durée des simulations dans le chapitre 2 n'a pas été fait au hasard. Comme expliqué dans le texte principal, des phénomènes s'inscrivant dans des temporalités tout à fait différentes sont inclus dans le modèle décrit dans la Figure 2.1 et dans les équations (2.2). En

conséquence, la résolution numérique du modèle doit être faite sur un temps considérable avant d'atteindre un état suffisamment proche de l'équilibre pour qu'on puisse considérer l'avoir atteint.

Dans les premiers temps de ce travail (avant l'optimisation du code), les simulations étaient arrêtées à 10^5 jours (environ 274 ans). L'équilibre dans chaque compartiment était ensuite calculé comme étant la moyenne des 10% des dernières valeurs enregistrées dans chaque compartiment. Prendre en considération les 10% de dernières valeurs permet de minimiser l'impact possible d'oscillations des effectifs en queue de simulations sur l'estimation des équilibres.

Cependant, considérer l'équilibre « atteint » après 10^5 jours s'avéra être une erreur. En réalité, pour de nombreuses combinaisons de valeurs de paramètres, l'équilibre semble bien être atteint après 10^5 jours, mais ensuite la dynamique de la population change avant d'atteindre un nouvel équilibre. La mise en évidence de ce problème a mené à rallonger le temps sur lequel est faite chaque simulation, et donc nécessairement le temps de calcul.

Ce problème de « faux équilibre » atteint après 10^5 jours peut partiellement s'expliquer par les conditions initiales utilisées pour les simulations. Comme expliqué tout au long de ce chapitre, la souche 2 est avantagée lorsque la population est majoritairement composée d'individus susceptibles (S), et la souche 1 est avantagée lorsque la population est majoritairement composée d'individus vaccinés (V). Avec les conditions initiales présentées ci-dessus, la population initiale a une structure très favorable à la souche 2, qui envahit alors très rapidement la population. Le compartiment I_1 va alors voir son effectif converger infiniment près de zéro. La population atteint un semblant d'équilibre. Cependant, la proportion d'individus vaccinés est désormais élevée, ce qui va permettre à la souche 1 de réémerger. Ce processus sera d'abord très lent, avant d'accélérer pour que finalement la population atteigne un nouvel équilibre.

B.3 – Couverture vaccinale et proportion de vaccinés

Soit δ la proportion relative d'adultes vaccinés. Nous avons $p_b = \delta p_p$. L'équation (2.9) nous donne la proportion d'individus vaccinés dans la population lorsque aucune souche ne circule (la population est saine). Cette proportion varie en fonction de la perte de l'immunité (η), des rappels (ω), et bien évidemment de la couverture vaccinale primaire (p_p). Nous nous intéressons

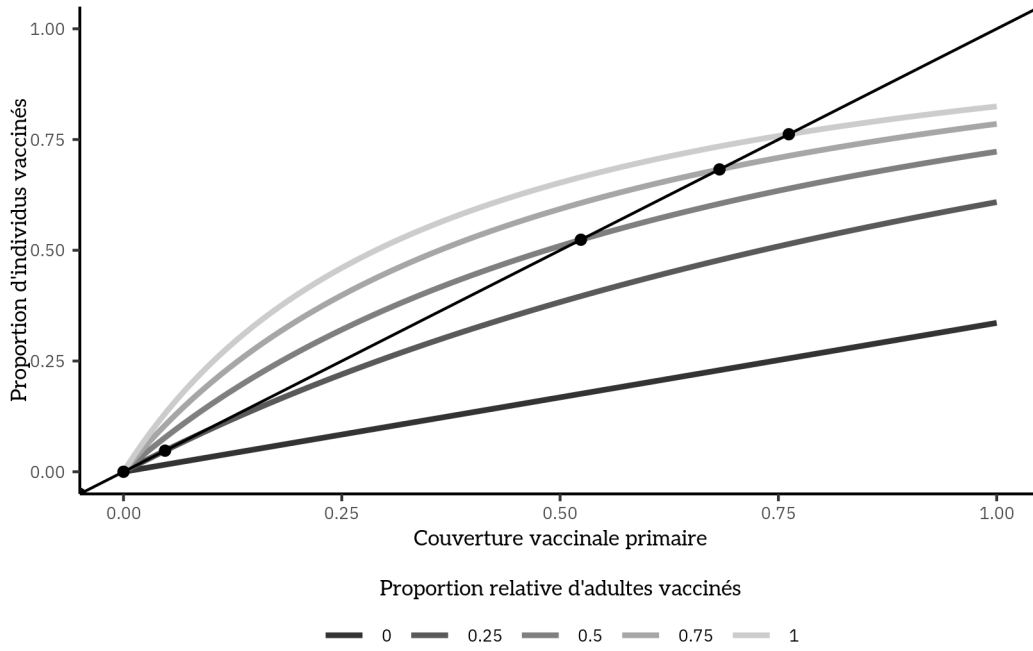


FIGURE S1 – **Proportion d'individus vaccinés (v_{DFE}) en fonction de la couverture vaccinale (p_p)** pour différentes valeurs de proportion relative d'adultes vaccinés (δ). La droite noire est la droite d'équation $v_{DFE} = p_p$ et les points noirs indiquent les valeurs de p_p pour lesquelles les courbes de v_{DFE} croisent cette droite.

ici aux conditions pour lesquelles $v_{DFE} > p_p$, donc

$$\frac{p_p + \delta p_p f}{\delta p_p f + \eta + d} > p_p. \quad (S7)$$

La résolution de l'inéquation (S7) nous permet de découvrir que v_{DFE} est supérieur à p_p lorsque $\frac{\delta f - \eta}{\delta f} > p_p$. La figure S1 montre ainsi que plus la couverture vaccinale primaire est élevée, plus il est probable que la proportion d'individus vaccinés soit supérieure à cette couverture vaccinale. Également, lorsque la proportion relative d'adultes vaccinés est élevée, la proportion d'individus vaccinés est supérieure à la couverture vaccinale primaire pour de plus hautes valeurs de celle-ci.

C – Suppléments au Chapitre 3

C.1 – Discrétisation du modèle SIR

Le flux d'individus infectés par la souche s est

$$\frac{dI_s(t)}{dt} = \beta_s S(t) I_s(t) - \gamma I_s(t). \quad (\text{S8})$$

On pose $f(t) = \beta_s S(t) - \gamma$, ce qui permet de réécrire (S8) sous la forme

$$\frac{dI_s(t)}{dt} = f(t) I_s(t). \quad (\text{S9})$$

En réorganisant les termes de (S9), on obtient

$$\frac{1}{I_s(t)} dI_s(t) = f(t) dt. \quad (\text{S10})$$

Ensuite, on intègre de chaque côté de (S10) pour avoir

$$\int \frac{1}{I_s(t)} dI_s(t) = \int f(t) dt. \quad (\text{S11})$$

L'intégrale du côté gauche de l'équation est $\ln |I_s(t)|$ tandis que l'intégrale du côté droit est une primitive de $f(t)$ qu'on notera $F(t)$. L'équation (S11) devient donc

$$\ln |I_s(t)| = F(t) + C \quad (\text{S12})$$

où C est la constante d'intégration.

On passe chaque côté de (S12) en exponentiel pour obtenir

$$|I_s(t)| = e^{F(t)+C} = e^C \cdot e^{F(t)} \quad (\text{S13})$$

avec $e^C > 0$.

Finalement, en posant $K = \pm e^C$ et en sachant que $F(t) = \int_0^t f(u) du$, on obtient

$$I_s(t) = K e^{\int_0^t f(u) du} \quad (\text{S14})$$

où $\int_0^t f(u) du$ est une primitive définie de $f(t)$ de 0 à t .

Le but ensuite est exprimer $I_s(t+1)$ en fonction de $I_s(t)$. Sachant que $F(t) = \int_0^t f(u) du$ On a

$$I_s(t+1) = K e^{\int_0^{t+1} f(u) du}. \quad (\text{S15})$$

On décompose l'intégrale de l'équation (S15) pour obtenir

$$\int_0^{t+1} f(u) du = \int_0^t f(u) du + \int_t^{t+1} f(u) du. \quad (\text{S16})$$

En injectant (S16) dans (S15) on obtient

$$I_s(t+1) = K e^{\int_0^t f(u) du + \int_t^{t+1} f(u) du}. \quad (\text{S17})$$

Sachant la propriété des exponentielles qui est que $e^{a+b} = e^a \cdot e^b$, (S17) devient

$$I_s(t+1) = K e^{\int_0^t f(u) du} \cdot e^{\int_t^{t+1} f(u) du}. \quad (\text{S18})$$

Enfin, on injecte (S14) dans (S18) et, en sachant que $f(u) = \beta_s S(u) - \gamma$, on obtient

$$I_s(t+1) = I_s(t) e^{\int_t^{t+1} (\beta_s S(u) - \gamma) du}, \quad (\text{S19})$$

ce qui correspond à l'équation (3.2).

C.2 – Formulation du modèle logit

L'équation (3.4) peut être développée en utilisant l'équation (3.2). On obtient alors

$$p_1(t+1) = \frac{I_1(t) e^{\int_t^{t+1} (\beta_1 S(u) - \gamma) du}}{I_1(t) e^{\int_t^{t+1} (\beta_1 S(u) - \gamma) du} + I_2(t) e^{\int_t^{t+1} (\beta_2 S(u) - \gamma) du}}. \quad (\text{S20})$$

On peut ensuite développer l'équation (S20) pour obtenir

$$\begin{aligned}
p_1(t+1) &= \frac{I_1(t) e^{\int_t^{t+1} (\beta_1 S(u) - \gamma) du}}{I_1(t) e^{\int_t^{t+1} (\beta_1 S(u) - \gamma) du} + I_2(t) e^{\int_t^{t+1} (\beta_2 S(u) - \gamma) du}} \\
&= \frac{I_1(t) e^{\int_t^{t+1} \beta_1 S(u) du} e^{-\gamma}}{I_1(t) e^{\int_t^{t+1} \beta_1 S(u) du} e^{-\gamma} + I_2(t) e^{\int_t^{t+1} \beta_2 S(u) du} e^{-\gamma}} \\
&= \frac{I_1(t) e^{\int_t^{t+1} \beta_1 S(u) du}}{I_1(t) e^{\int_t^{t+1} \beta_1 S(u) du} + I_2(t) e^{\int_t^{t+1} \beta_2 S(u) du}} \\
&= \frac{I_1(t) \frac{e^{\int_t^{t+1} \beta_1 S(u) du}}{e^{\int_t^{t+1} \beta_2 S(u) du}}}{I_1(t) \frac{e^{\int_t^{t+1} \beta_1 S(u) du}}{e^{\int_t^{t+1} \beta_2 S(u) du}} + I_2(t) \frac{e^{\int_t^{t+1} \beta_2 S(u) du}}{e^{\int_t^{t+1} \beta_2 S(u) du}}} \\
&= \frac{I_1(t) e^{\int_t^{t+1} \beta_1 S(u) du} - \int_t^{t+1} \beta_2 S(u) du}{I_1(t) e^{\int_t^{t+1} \beta_1 S(u) du} - \int_t^{t+1} \beta_2 S(u) du + I_2(t)} \\
&= \frac{\frac{I_1(t)}{I_1(t)+I_2(t)} e^{\int_t^{t+1} \beta_1 S(u) du} - \int_t^{t+1} \beta_2 S(u) du}{\frac{I_1(t)}{I_1(t)+I_2(t)} e^{\int_t^{t+1} \beta_1 S(u) du} - \int_t^{t+1} \beta_2 S(u) du + \frac{I_2(t)}{I_1(t)+I_2(t)}} \\
&= \frac{p_1(t) e^{\int_t^{t+1} \beta_1 S(u) du} - \int_t^{t+1} \beta_2 S(u) du}{p_1(t) e^{\int_t^{t+1} \beta_1 S(u) du} - \int_t^{t+1} \beta_2 S(u) du + (1 - p_1(t))} \\
&= \frac{\frac{p_1(t)}{1-p_1(t)} e^{\int_t^{t+1} \beta_1 S(u) du} - \int_t^{t+1} \beta_2 S(u) du}{\frac{p_1(t)}{1-p_1(t)} e^{\int_t^{t+1} \beta_1 S(u) du} - \int_t^{t+1} \beta_2 S(u) du + 1}.
\end{aligned} \tag{S21}$$

On pose $\text{logit}(p_1(t)) = \ln\left(\frac{p_1(t)}{1-p_1(t)}\right)$. On continue le développement de S21 :

$$\begin{aligned}
p_1(t+1) &= \frac{e^{\text{logit}(p_1(t)) + \int_t^{t+1} \beta_1 S(u) du - \int_t^{t+1} \beta_2 S(u) du}}{e^{\text{logit}(p_1(t)) + \int_t^{t+1} \beta_1 S(u) du - \int_t^{t+1} \beta_2 S(u) du} + 1} \\
\Leftrightarrow \frac{p_1(t+1)}{1-p_1(t+1)} &= e^{\text{logit}(p_1(t)) + \int_t^{t+1} \beta_1 S(u) du - \int_t^{t+1} \beta_2 S(u) du} \\
\Leftrightarrow \ln\left(\frac{p_1(t+1)}{1-p_1(t+1)}\right) &= \ln\left(e^{\text{logit}(p_1(t)) + \int_t^{t+1} \beta_1 S(u) du - \int_t^{t+1} \beta_2 S(u) du}\right) \\
\Leftrightarrow \text{logit}(p_1(t+1)) &= \text{logit}(p_1(t)) + \int_t^{t+1} \beta_1 S(u) du - \int_t^{t+1} \beta_2 S(u) du \\
&= \text{logit}(p_1(t)) + \int_t^{t+1} (\beta_1 - \beta_2) S(u) du
\end{aligned} \tag{S22}$$

On retrouve ainsi l'équation 3.5.

C.3 – Code du modèle bayésien

Le code présenté ci-dessous est utilisé dans les Chapitres 3 et 4 pour toutes les analyses bayésiennes. Le fonctionnement du code est expliqué dans la Section 3.7.

```
model {
  p[1] ~ dunif(0, 1)
  logitp[1] <- log(p[1] / (1 - p[1]))
  S[1] ~ dbin(p[1], n[1])
  for (t in 2:(tm + 1)) {
    epsilon[t - 1] ~ dnorm(0, 1 / (se ** 2))
    deltalogitp[t] <- a[1] + sum(c(0, a[(1:nv) + 1]) * X[t, ]) + epsilon[t - 1]
    logitp[t] <- logitp[t - 1] + deltalogitp[t]
    p[t] <- exp(logitp[t]) / (1 + exp(logitp[t]))
    S[t] ~ dbin(p[t], n[t])
  }
  for (f in 1:(nf + 1)) {
    mu[f] ~ dunif(-100, 100)
    sigma[f] ~ dunif(0, 50)
    a[f] ~ dnorm(mu[f], 1 / sigma[f] ** 2)
  }
  se ~ dunif(0, 50)
}
```