Modélisation et optimisation de la réponse à des vaccins et à des interventions immunothérapeutiques : application au virus Ebola et au VIH

Chloé Pasin

université de BORDEAUX

THÈSE PRÉSENTÉE POUR OBTENIR LE GRADE DE

DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE SOCIÉTÉS, POLITIQUE, SANTÉ PUBLIQUE
SPÉCIALITÉ SANTÉ PUBLIQUE, OPTION BIOSTATISTIQUE

Par Chloé PASIN

MODÉLISATION ET OPTIMISATION DE LA RÉPONSE À DES VACCINS ET À DES INTERVENTIONS IMMUNOTHÉRAPEUTIQUES. APPLICATION AU VIRUS EBOLA ET AU VIH.

Modeling and optimizing the response to vaccines and immunotherapeutic interventions. Application to Ebola virus and HIV.

Sous la direction de Rodolphe THIÉBAUT et François DUFOUR

Soutenue le 30/10/18 à Bordeaux

Membres du jury

MOLINA-PARÍS Carmen, Professor, University of Leeds (Leeds, UK)	Présidente
DE SAPORTA Benoîte, Professeure, Université de Montpellier (Montpellier, France)	Rapporteure
HENS Niel, Hoogleraar, Hasselt University (Hasselt, Belgique)	Rapporteur
CRAUSTE Fabien, Chargé de Recherche, CNRS (Bordeaux, France)	Examinateur
DUFOUR François, Professeur, Institut Polytechnique de Bordeaux (Bordeaux, France)	Co-directeur de thèse
THIÉBAUT Rodolphe, Professeur, Université de Bordeaux (Bordeaux, France)	Directeur de thèse

Contents

Remerciements

Je souhaite d'abord remercier Rodolphe Thiébaut. J'ai beaucoup appris pendant ces années de travail avec toi. Malgré ton emploi du temps de ministre, tu as toujours su te rendre disponible lorsqu'il le fallait. Merci pour ta bienveillance, ton énergie, ta confiance en moi et ta façon de trouver les bons mots lorsque j'avais besoin d'être remotivée. Je souhaite également te remercier pour les opportunités que tu m'as données, notamment par l'intermédiaire des nombreuses collaborations que j'ai eu depuis mon stage de master jusqu'à la fin de cette thèse. J'ai conscience de la chance que ça a été pour moi. J'espère sincèrement que nous continuerons à travailler ensemble sur d'autres projets.

Je remercie également François Dufour. Merci pour tes encouragements et ton enthousiasme constant durant ces 3 années de travail. Je pense que c'était un vrai défi de combiner un travail théorique avec une application clinique. C'était difficile pour moi de développer une application numérique et même si ça a pris du temps, tu n'as jamais perdu patience ni confiance en mon travail. J'espère que ce travail pourra engendrer d'autres collaborations et développements du contrôle optimal aux questions cliniques.

Fabien Crauste, merci pour ton intérêt pour mon travail et tes encouragements ces derniers mois. J'espère sincèrement avoir l'opportunité de travailler avec toi sur de futurs projets. Benoîte de Saporta, je te remercie d'avoir accepté de rapporter ma thèse. Merci également de m'avoir donné l'opportunité de venir présenter mon travail à Montpellier. I would also like to warmly thank Niel Hens for accepting to report my thesis. I am sure it will be a good opportunity to discuss the Ebola project. Carmen Molina-París, thank you very much for accepting to be part of my jury. I hope we will have lots to discuss on both IL-7 and Ebola.

Ce travail de thèse n'aurait pas pu être conduit sans les nombreux collaborateurs avec lesquels j'ai pu échanger et travailler. Tout d'abord, le projet EBOVAC a été l'opportunité pour moi d'échanger avec de nombreuses personnes et je suis très reconnaissante d'avoir pu participer aux discussions scientifiques qui ont enrichi mon travail et ma compréhension des enjeux liés au développement vaccinal. Je remercie donc IMI pour le financement de ce projet et tous les acteurs des projets EBOVAC1&2. Je souhaite d'abord remercier l'équipe de Bordeaux, Eugénie, Laetitia, Christine B, Christine S, Anton, en particulier ceux qui m'ont aidé à distinguer plus clairement tous les développements vaccinaux en cours. Je remercie les autres collaborateurs de ces projets, notamment Thierry Van Effelterre qui

a eu une attention particulière pour le travail de modélisation, ainsi que les principaux investigateurs des essais cliniques et les participants volontaires, sans qui je n'aurais pas pu réaliser mes travaux de modélisation. Je remercie également le Vaccine Research Institute, pour le soutien financier ainsi que l'accès aux données des études INSPIRE sans lesquelles le travail d'optimisation n'aurait pas été mené. Je remercie aussi Anne-Sophie Beignon, grâce à qui j'ai beaucoup appris en immunologie. Merci d'avoir répondu à mes nombreuses questions. Merci aussi à Jean-Louis, que j'ai apprécié recroiser à l'occasion de plusieurs meetings et conférences. I would also like to warmly thank Zoe Moodie. I enjoyed our collaboration before the start of my PhD and it has been quite long to value this work, so I would like to thank you for your constant support through the last years. I hope we will get other chances to work together in the future. I thank also Andy Yates, with who I will work on my postdoc projects. Thank you for welcoming me in New York and giving me the opportunity to attend the meeting in Iceland. I am looking forward to starting new projects with you. J'ai aussi une pensée pour les nombreuses personnes que j'ai rencontrées lors des conférences : il y a eu des rencontres très sympathiques et cela a aussi été une vraie source d'enrichissement professionnel.

Je souhaite également remercier les personnes avec qui j'ai été amenée à travailler et échanger à Bordeaux. Merci à Huilong Zhang de l'équipe CQFD, qui a su m'aider à mener à bien la programmation en Matlab. Je remercie aussi évidemment tous les membres des équipes biostatistiques et SISTM du centre INSERM U1219. En particulier, je remercie Daniel Commenges qui m'a co-encadrée lors de mon arrivée en 2014 pour mon stage de Master 2. J'ai beaucoup appris des nombreuses discussions entre Rodolphe et vous. Je voudrais également remercier Mélanie Prague : merci pour ta disponibilité lorsqu'il s'agissait de répondre aux questions sur NIMROD. J'apprécie beaucoup collaborer avec toi et j'espère que nous continuerons à travailler sur d'autres projets ensemble. Merci pour ces opportunités, en particulier le travail avec Alison, who I also thank for welcoming me a couple of days in Boston. Je remercie aussi Laura V, Irene, ainsi que Laura R, Lise et Edouard avec qui j'ai travaillé sur différents projets. J'ai beaucoup appris en travaillant avec vous tous. Je souhaite également dire un grand merci à Sandrine, sans qui je n'aurais pu effectuer autant de déplacements professionnels pour des meetings ou des réunions ou bien pour valoriser mon travail dans des conférences. Merci pour ta constante bonne humeur ainsi que ton professionnalisme et ton efficacité.

Lors de ces 2 dernières années, j'ai également obtenu un poste de moniteur à l'ISPED. Je souhaite donc remercier l'équipe pédagogique de m'avoir initiée à l'enseignement. J'espère avoir pu être à la hauteur de vos exigences. Je remercie notamment Pierre, pour ton aide organisationnelle et ta disponibilité lorsque j'avais des questions. Merci aussi à ceux avec qui j'ai été amenée à échanger, notamment Fleur, Valérie, Aurélie, Eric, Ka-

ren et Marta. Merci également à Marie Christine, Bénédicte et Carole pour la gestion administrative.

Je souhaite aussi remercier les doctorants investis dans l'association des doctorants, ainsi que les directeurs de l'EDSP2 et Ghyslaine, qui m'ont permis de participer à la vie événementielle et administrative de l'ED.

Je voudrais dire un grand merci aux personnes que j'ai rencontrées à l'ISPED et qui m'ont supportée pendant ces dernières années. Certains sont devenus bien plus que des collègues de travail. Merci au groupe des M... : ces quelques années ont été remplies d'innombrables moments de partages, bons repas, soirées, fou rires, karaokés, cours de kizomba (ou pas), sessions de running, playa, cafés aux Capu, potins dans tous les salons de thés de Bordeaux, et j'en passe. Merci donc à Robin, Anaïs et Boris pour votre bonne humeur. Merci à Mathilde pour ton oreille attentive. Merci à Henri et Loïc, je me souviens encore de mon 1er jour à l'ISPED avec vous. J'ai l'habitude de dire que vous me fatiguez mais vos zzzziiit incessants m'ont quand même manqués ces derniers mois au bureau. Et enfin un grand merci aux girls, les fleurs de ma vie, pour votre belle amitié, votre soutien et les nombreux bons moments passés avec vous : Astou, la mamacita, tu es toujours pleine de bonne énergie quand on se voit ; Alexandra, avec toi on peut refaire le monde pendant des heures ; Perrine, tu es une co-bureau irremplaçable ; et Emilie, on se suit depuis le début de la thèse, dans les hauts et les bas, et ces derniers mois auraient été beaucoup plus difficiles sans toi.

Je voudrais également remercier Maude, pour ton énergie débordante et Corentin, avec qui je ne me lasse pas de discuter. Merci pour nos petites pauses culturelles à l'opéra de Bordeaux. Je remercie également Bachirou et Majid avec qui c'est toujours agréable de discuter au détour d'un bureau ou d'un couloir. Je tiens à remercier également les nombreux collègues du bureau John Snow et ceux du 45, qui ont contribué à une bonne ambiance de travail durant ces années : Hadrien, Bruno, Chariff, Solenne, Myriam, Jean Noel, Anthony, Melany, Casimir, Viviane, et ceux que j'ai oublié.

Bordeaux a aussi été le lieu de belles retrouvailles et rencontres. Je pense particulièrement à Cécile. Quel plaisir de te retrouver ici et de partager de nombreux moments ces 3 années avec toi, Patrice et évidemment Louis que j'apprécie tellement voir grandir. Merci pour votre soutien, surtout ces dernières semaines. Je voudrais également dire un grand merci à Ashod, pour ta présence et tes encouragements ces dernières années. Tu sais à quel point cela compte pour moi. Et puis je pense aussi à Arnaud, merci pour nos bons moments et ton enthousiasme pour mon travail.

Je voudrais également remercier mes amis « d'ailleurs », que j'ai retrouvés lors de mes voyages ces dernières années. Certaines amitiés se fichent de la distance. Kaleena, l'américaine frenchie, que de bons moments passés ensemble, please move to NYC soon.

Bryan, it was great to see you during the past years. Chtimi, merci de m'avoir accueillie à Paris à maintes reprises ; tu as cette capacité à savoir me faire rire en quelques secondes. Anso, ma taupine préférée, je pense que ma meilleure soirée à Bordeaux c'était avec toi ! J'ai hâte de venir voir ton nouveau chez toi. Julien, où que l'on soit, c'est toujours un bon moment avec toi. Un immense merci pour ce beau voyage en Afrique du Sud. Et enfin un grand merci à Karthik, de supporter mon sens de l'humour inexistant et d'être un si bon compagnon de voyage.

Même si je suis partie il y a déjà quelques années, Toulouse reste ma maison et la ville où j'ai construit de belles amitiés qui durent maintenant depuis longtemps. J'ai une pensée affectueuse pour Julie et Morgane, c'est toujours un bon moment de se retrouver. Je dis un grand merci à ma médecin préférée Emilie ; grâce à toi Bon Entendeur m'a beaucoup accompagnée lors de l'écriture de cette thèse... En tout cas maintenant le nouvel an à NY nous tend les bras ! Je remercie aussi Iris, même si l'on ne se voit pas souvent, je sais que l'on peut se retrouver et parler pendant des heures. Je pense aussi à la bande de loulous, que je suis heureuse de voir s'agrandir avec Hugo et Lou : merci à Lucile, Kevin l'iron man, Pauline, Popow, et enfin Romain. Un grand merci parce que tu n'as jamais cessé de t'intéresser à mon travail et tu sais l'importance que cela a pour moi.

Je finis par remercier toute ma famille, je pense en particulier à ma cousine Clairou et ma tatie Denise. J'ai une pensée pour ma mamie Blanche, j'espère qu'elle serait fière de moi si elle était là aujourd'hui. Et puis je finis par remercier ceux qui font de Saint Jean mon chez moi, l'endroit où il fait toujours bon revenir : mon frère Nico, Sarah, mes deux petits cœurs Clémence et Emma qui j'espère auront le courage d'ouvrir la thèse de leur tata un jour, mon frérot David et évidemment mes parents. Vous êtes mes piliers. Vous m'avez toujours soutenue dans mes choix et dans mon travail, et je sais que vous continuerez à le faire dans les années futures. Pour ça je vous dis un grand merci.

Scientific production

Articles

Thesis publications

• Pasin C., Dufour F., Villain L., Zhang H., Thiébaut R. Controlling IL-7 injections in HIV-infected patients, *Bulletin of Mathematical Biology*, 80(9): 2349-2377, 2018. https://doi.org/10.1007/s11538-018-0465-8

• Pasin C., Balelli I., Van Effelterre T., Bockstal V., Solforosi L., Prague M., Douoguih M., Thiébaut R., with the EBOVAC1 consortium. Dynamics of the humoral immune response to a prime-boost Ebola vaccine: quantification and sources of variation, *In preparation for submission.*

Related publications

• Villain L., Commenges D., Pasin C., Prague M., Thiébaut R. Adaptive protocols based on predictions from a mechanistic model of the effect of IL7 on CD4 counts, *Statistics in Medicine*, 2018. https://doi.org/10.1002/sim.7957

• Gross L., Lhomme E., Pasin C., Richert L., Thiébaut R. Ebola vaccine development: Systematic review of preclinical and clinical studies, and meta-analysis of determinants of antibody response variability after vaccine, *International Journal of Infectious Diseases*, 74: 83-96, 2018. https://doi.org/10.1016/j.ijid.2018.06.022

Other publications

• Pasin C., Halloran M.E., Gilbert P.B., Langevin E., Ochiai R.L., Pitisuttithum P., Capeding M.R., Carrasquilla G., Frago C., Cortés M., Chambonneau L., Moodie Z. Periods of high dengue transmission defined by rainfall do not impact efficacy of dengue vaccine in regions of endemic disease, *Under revision in PLOS ONE.*

• Lhomme E., Richert L., Moodie Z., Pasin C., Kalams S.A., Morgan C., Self S., De Rosa S.C., Thiébaut R. Early CD4+ T cell responses are associated with subsequent CD8+ T cell responses to an rAd5-based prophylactic prime-boost HIV vaccine strategy, *PLOS*

ONE. 2016 Apr 28;11(4):e0152952. https://doi.org/10.1371/journal.pone.0152952

• Steinwart I., Pasin C., Williamson R., Zhang S. Elicitation and identification of properties, *Conference on Learning Theory*. 2014 May 29 (pp. 482-526).

Book chapter

• Thiébaut R., Villain L., Pasin C., Commenges D. Modelling the response to Interleukin-7 therapy in HIV-infected patients. In *Mathematical, Computational and Experimental T Cell Immunology, Springer. Under revision.*

Communications

Oral communications at international conferences

• Pasin C., Villain L., Dufour F., Commenges D., Thiébaut R. Use of mathematical modeling for optimizing and adapting immunotherapy protocols in HIV-infected patients. *Population Approach Group in Europe (PAGE) meeting*, Montreux, Switzerland, 2018.

• Pasin C., Prague M., Eggo R., Van Effelterre T., Balelli I., Snape M., Anazala O., Praygod G., Anywaine Z., Solforosi L., Verbruggen N., Bockstal V., Watson-Jones D., Edmunds J., Douoguih M., Thiébaut R. Modelling the humoral immune response to Ebola vaccine. Results from EBOVAC1 project. *Systems Immunology & Vaccine Design symposium*, Heidelberg, Germany, 2017.

• Pasin C., Richert L., Commenges D., Thiébaut R. Modelling the immune response to HIV vaccine. *British Society in Immunology (BSI) meeting: Mathematical modelling in Immunology*, Cambridge, United Kingdom, 2015.

Oral communications at French conferences

• Pasin C., Dufour F., Thiébaut R. Optimisation des stratégies d'injection d'interleukine 7 pour des patients infectés par le virus de l'immunodéficience humaine. *SMAI 2017, 8e Biennale Française des Mathématiques Appliquées et Industrielles*, Ronce-les-Bains, France, 2017.

• Pasin C., Richert L., Commenges D., Thiébaut R. Modélisation de la réponse à un vaccin VIH. *Groupe de Recherche Statistiques et Santé*, Paris, France, 2015.

Invited talks at French conferences

• Pasin C., Thiébaut R. Modelling the humoral immune response to Ebola vaccine. *Bordeaux Modelling Workshop*, Bordeaux, France, 2016.

• Pasin C., Dufour F., Thiébaut R. Optimal administration of IL-7 in HIV-infected patients. *Bordeaux Modelling Workshop*, Bordeaux, France, 2016.

• Pasin C., Commenges D., Jarne A., Richert L., Lhomme E., Thiébaut R. Mechanistic modelling of CD4+ T cell response to rAd5/DNA HIV vaccine. *Bordeaux Modelling Workshop*, Bordeaux, France, 2014.

Invited talks at French seminars

• Pasin C., Dufour F., Thiébaut R. Optimal administration of IL-7 in HIV-infected patients. *Séminaire de Probabilités et Statistiques*, Montpellier, France, 2017.

Written communications (posters) at international conferences

• Pasin C., Villain L., Dufour F., Thiébaut R. Optimal administration of IL-7 in HIV-infected patients, *Systems Approaches in Immunology and Infectious Diseases*, Santa Fe, NM, USA, 2016.

• Prague M., Pasin C., Wittkop L., Duffau P., Lazaro E., Cazanave C., Vareil M., Bonnet F., Thiébaut R. In silico clinical trials for evaluation of HIV short-cycle strategies. *Conference on Retroviruses and Opportunistic Infections (CROI)*, Boston, MA, USA, 2018.

• Gerold J.M., Pasin C., Balelli I., Lim S., Osuna C., Whitney J.B., Barouch D.H., Prague M., Hill A. SIV rebound kinetics following TLR7-agonist & therapeutic vaccine administration. *Conference on Retroviruses and Opportunistic Infections (CROI)*, Boston, MA, USA, 2018.

Abbreviations and notations

Abbreviations

Ad: Adenovirus

AIC: Akaike criterion

AIDS: Acquired Immune Deficiency Syndrome

ASC: Antibody-secreting cell

cART: Combination Antiretroviral Therapy

ChAd: Chimpanzee adenovirus

DC: Dendritic cell

DNA: Deoxyribonucleic acid

DRC: Democratic Republic of Congo

EBOV: Ebola virus

ELISA: Enzyme-linked immunosorbent assay

EVD: Ebola virus disease

FDA: Food and Drug Administration

GP: Glycoprotein

GRN: Gene regulatory network

HAV: Hepatitis A virus

HBV: Hepatitis B virus

HIV: Human Immunodeficiency Virus

HPV: Human Papillomavirus

IFN: Interferon

Ig: Immunoglobulin

IL: Interleukin

LCVa: Approximation of the leave-one-out cross validation criterion

MAP: Maximum a Posteriori

MARV: Marburg virus

MCMC: Markov chain Monte Carlo

MHC: Major Histocompatibility complex

mRNA: Messenger Ribonucleic acid

MVA: Modified Vaccinia Ankara

NHP: Non-Human Primate

NK: Natural Killer

NP: Nucleoprotein

ODE: Ordinary Differential Equation

PEB: Parametric Empirical Bayes

PDMP: Piecewise Deterministic Markov Process

RNA: Ribonucleic acid

RVS: Robust-variance scoring

rVSV: Recombinant Vesicular Stomatitis Virus

SDE: Stochastic Differential Equation

SUDV: Sudan ebolavirus

TAFV: Tai Forest ebolavirus

TNF: Tumor Necrosis Factor

VV: Vaccinia Virus

ZEBOV: Zaire ebolavirus

Notations

— In chapter 3 "**Modeling the immune response to Ebola vaccine**":

- x: scalar x
- $\mathbf{x}$: vector $\mathbf{x}$
- $\mathbf{X}$: vector $\mathbf{X}$
- $\mathbf{x}^T$: $\mathbf{x}$ transpose
- $p(A|B)$: probability of A conditional on B
- $\mathbb{E}_\theta[X]$: expectation of X under the distribution of θ
- $Tr(M)$: trace of matrix M

— In chapter 4 "**Optimizing immune therapies in HIV-infected patients**":

- x: scalar x
- $\mathbf{x}$: vector $\mathbf{x}$
- $\mathbf{X}$: space $\mathbf{X}$
- $\mathbb{E}^u_{x_0}$: expectation under strategy u with starting point x_0

Résumé substantiel

Introduction

Les vaccins ont constitué une avancée majeure de la médecine des dernières décennies et ont permis l'éradication de certaines maladies telles que la variole ou la rougeole. Le principe de la vaccination est basé sur la mémoire immunitaire : après exposition à un pathogène, l'organisme est capable de générer une meilleure réponse en cas de ré-exposition. Cette réponse est spécifique au pathogène et se produit de manière plus rapide et plus efficacement, en termes qualitatif et quantitatif. Cependant, les mécanismes permettant de générer et maintenir cette mémoire immunitaire ne sont pas encore totalement connus, et les connaissances immunologiques à propos de la vaccination sont principalement empiriques. Cela pose donc problème quant au développement de vaccins efficaces contre certaines maladies infectieuses plus complexes, telles que le VIH, Ebola ou bien le paludisme. Certaines stratégies vaccinales récentes ont engendré des résultats encourageants : ces stratégies, dites "prime-boost", consistent à combiner plusieurs produits en injections répétées. Cependant, l'utilisation de ces nouveaux vaccins soulève de nouvelles questions : en particulier, combien d'injections sont nécessaires ? Dans quel ordre ? A quel délai ? En effet, on considère que si l'injection de boost est effectuée trop tôt, les cellules sont trop sollicitées et la différentiation en cellules mémoires n'est pas encore terminée. La réponse secondaire n'est donc pas optimale. Si l'on attend trop longtemps, la quantité de cellules mémoires aura déjà commencé à diminuer et la réponse ne sera pas optimale non plus. Il est difficile de mettre en place un essai clinique pour répondre à chacune de ces questions, car ceux-ci sont très longs et coûteux. De plus, un autre défi dans le développement vaccinal réside dans la variabilité populationnelle de la réponse immunitaire à une stimulation antigénique. En effet, de nombreux facteurs peuvent influencer la réponse immunitaire, que ce soit des facteurs génétiques, démographiques, environnementaux, ou bien liés au microbiome. Ces facteurs ne sont pas indépendants les uns des autres et il est donc encore difficile de les prendre en compte et de quantifier leur impact sur la réponse immunitaire. Pour répondre à ces questions, de nombreuses données sont générées dans le cadre d'essais cliniques vaccinaux chez des humains. Ces données sont de différents types : génomiques, protéomiques, métaboliques, ... Cependant, il est aussi difficile d'intégrer toutes ces don-

nées et d'en retirer l'information nécessaire pour comprendre les mécanismes d'action des vaccins.

L'approche dite de biologie systémique a pour objectif de mieux comprendre le fonctionnement du système immunitaire en analysant sa dynamique dans son ensemble, grâce à l'intégration de données de multiples marqueurs de la réponse immunitaire. Cela passe en particulier par la modélisation mathématique du système immunitaire. En effet, les interactions biologiques entre les acteurs du système immunitaire sont généralement complexes et non linéaires. Le comportement global est donc difficile à prédire. La modélisation mathématique permet de prendre en compte cette complexité. L'intérêt des modèles réside aussi et surtout dans leur capacité à quantifier les dynamiques du processus biologique étudié et à capturer l'impact de certains facteurs sur la variabilité du processus. Les modèles, bien calibrés et estimés, représentent également un vrai outil de prédiction. De nombreux modèles du système immunitaire ont déjà été proposés dans la littérature et leur intérêt a largement été discuté. Il y a cependant peu de modèles concernant spécifiquement la réponse à un vaccin.

Le travail de cette thèse s'inscrit dans l'approche de biologie systémique, avec deux objectifs particuliers : le premier est de modéliser la dynamique de la réponse immunitaire à un vaccin, et le suivant est de proposer un outil numérique pour optimiser les protocoles d'injections répétées. En pratique, le travail est divisé en deux projets. Dans le premier, nous proposons l'application d'un modèle de la réponse immunitaire humorale, basé sur un système d'équations différentielles et nous estimons les paramètres du modèle en utilisant des données provenant d'essais cliniques de phase 1 sur un vaccin contre Ebola. L'estimation du modèle permet de quantifier la dynamique du système immunitaire, de prédire la durabilité de la réponse, ainsi que de déterminer l'impact de facteurs environnementaux et liés au vaccin sur la variabilité de cette réponse. Dans le deuxième projet, nous nous intéressons à des problèmes d'optimisation. En effet, l'idée principale est d'utiliser les modèles mathématiques de la réponse vaccinale pour déterminer le schéma optimal de prime-boost, et en particulier le délai optimal entre les injections. Nous proposons donc un outil numérique, basé sur la théorie du contrôle optimal et permettant d'optimiser des schémas d'injections. En particulier, cet outil est appliqué à des protocoles d'immunothérapie injectée à des patients atteints par le VIH.

Modélisation de la réponse immunitaire à un vaccin contre Ebola

L'épidémie d'Ebola de grande envergure qui a eu lieu en Afrique de l'Ouest entre 2014 et 2016 a mis en évidence le manque de produits thérapeutiques et/ou vaccinaux efficaces contre le virus Ebola. Cela a engendré la mise en place de nombreux projets visant à accélérer le développement de vaccins ou médicaments contre le virus. En particulier, un vaccin prime-boost consistant en une injection avec le vecteur adénovirus 26 (Ad26) et le vecteur Modified Vaccinia Ankara (MVA) est évalué dans des essais cliniques de phase 1 à 3. Certains de ces essais sont réalisés dans le cadre du consortium EBOVAC, qui est inclus dans le programme Ebola+ de l'Innovative Medicines Initiative (IMI). Ce consortium réunit des partenaires académiques avec le laboratoire pharmaceutique fabriquant le vaccin. L'INSERM faisant partie de ce consortium, nous avons eu accès aux données de 3 essais cliniques de phase 1 réalisés sur des adultes volontaires sains dans 4 pays : Royaume Uni, Kenya et Ouganda/Tanzanie. Dans ces essais cliniques, les participants ont été randomisés pour recevoir soit Ad26 puis MVA ou MVA puis Ad26 à 28 ou 56 jours d'intervalle. Des mesures des marqueurs de la réponse immunitaire ont été effectuées à des temps consécutifs jusqu'à 1 an après la première injection vaccinale. En particulier, le niveau d'anticorps a été mesuré ; on ne sait pas encore si un niveau donné d'anticorps engendre une protection contre l'infection par le virus Ebola, mais des études chez des primates non humains ont montré que la survie après une injection intramusculaire du virus était associée à un niveau élevé d'anticorps. C'est donc actuellement le marqueur préférentiel pour évaluer l'immunogénicité des vaccins candidats dans les essais cliniques. Une question majeure concerne la durabilité de la réponse immunitaire ainsi que les facteurs pouvant influencer cette durabilité. En particulier, certaines études ont montré que les anticorps étaient maintenus dans l'organisme grâce à une population de cellules B productrices d'anticorps ayant une longue demi-vie. Cependant, il semblerait qu'une autre population de cellules B soit capable de réagir rapidement après rencontre avec l'antigène pour produire un certain nombre d'anticorps avant de rapidement mourir.

Nous avons donc utilisé un modèle pour la dynamique de la réponse humorale après l'injection de boost, constitué de deux populations distinctes de cellules productrices d'anticorps, ayant des demi-vies différentes et des taux de production d'anticorps différents. La dynamique de chaque compartiment du modèle est décrite à l'aide d'une équation différentielle ordinaire. L'intérêt étant de quantifier la dynamique de la réponse immunitaire humorale après l'injection, les paramètres du système d'équations différentielles ont été estimés en utilisant les données des 3 essais cliniques de phase 1 réalisés en Europe et

Afrique de l'Est. Cette estimation est basée sur une approche populationnelle, utilisant des modèles linéaires mixtes sur chacun des paramètres. Cela permet d'estimer une valeur moyenne du paramètre dans la population, ainsi que l'impact de covariables (facteurs liés au vaccin ou à l'environnement géographique) sur ces paramètres et également la variabilité intra-individuelle, induite par un effet aléatoire normalement distribué autour de la valeur moyenne. De plus, l'estimation est basée sur un modèle d'observation, qui suppose que les observations cliniques correspondent à une fonction des compartiments du modèle mathématique à laquelle s'ajoute une erreur d'observation normalement distribuée. Concrètement, l'estimation est effectuée en maximisant la vraisemblance globale, obtenue en calculant les vraisemblances individuelles et en intégrant sur les effets aléatoires. La maximisation est ensuite effectuée en utilisant un algorithme de type Newton, qui est basé sur une approximation de la matrice hessienne utilisant seulement les dérivées premières de la log vraisemblance, ce qui facilite les calculs numériques. De plus, plusieurs critères sont utilisés afin de s'assurer de la convergence de l'algorithme. Un autre aspect du programme de maximisation est qu'il permet d'utiliser des connaissances biologiques obtenues à partir d'expérimentations ou d'autres estimations en définissant des distributions a priori sur les paramètres. Dans ce cas, on réalise une approximation normale de la distribution a posteriori et on estime le maximum a posteriori du paramètre en question. Cela se traduit numériquement par la maximisation d'une vraisemblance pénalisée par les connaissances a priori.

Après sélection et estimation du modèle, les résultats suivants ont été obtenus : la demi-vie moyenne des anticorps a été estimée à 24 jours (intervalle de confiance [22,26] jours). Cette estimation semble cohérente avec des études précédentes ayant estimé la demi-vie des anticorps entre 3 semaines et 2 mois. De plus, deux populations de cellules productrices d'anticorps ont pu être bien distinguées : la première a une demi-vie variant de 1 à 5 jours, selon le régime de vaccination. Cette estimation est également cohérente avec d'autres études qui montrent que les cellules productrices d'anticorps sont sujettes à un pic autour de 7 jours après injection vaccinale et disparaissent après 10 à 14 jours. La deuxième population de cellules a une demi-vie de plusieurs années. Comme les données ne sont disponibles que jusqu'à 1 an après la première injection vaccinale, il est difficile d'identifier avec précision un intervalle de confiance autour de la demi-vie de cette population à longue durée de vie. Cependant, un profil de vraisemblance a été effectué et a permis de déterminer une valeur minimale de 5 années. Cela signifie que la moitié des cellules présentes 7 jours après l'injection de boost persiste au moins 5 ans dans l'organisme, tout en continuant à produire des anticorps. D'autres études concernant des vaccins différents ont également permis d'identifier une persistance des anticorps pendant plusieurs années, suggérant le maintien par une population de cellules B capables de vivre

longtemps dans l'organisme. Nous avons également identifié comment certains facteurs influencent la dynamique de la réponse immunitaire humorale. Le régime vaccinal semble notamment modifier la demi-vie des cellules productrices d'anticorps qui répondent rapidement. Cette modification n'a pas d'impact sur la persistance des anticorps à long terme. En revanche, la localisation géographique a un impact sur les niveaux de persistance des anticorps. En effet, les participants d'Europe ont des niveaux d'anticorps à long terme plus élevés que les participants d'Afrique de l'Est. Dans le modèle, cela est dû à une différence significative de valeur d'un paramètre signifiant que les cellules ayant une longue durée de vie produisent plus d'anticorps et/ou sont présentes à un niveau plus élevé 7 jours après l'injection de boost chez les européens que chez les africains. Cette différence pourrait être liée à l'environnement immunitaire, les participants africains étant plus sujets à un environnement immunitaire activé par d'autres co-infections ou parasites. Nous avons également examiné l'impact potentiel de la réponse cellulaire de lymphocytes T CD4 produisant des cytokines sur la dynamique humorale. Cependant, nous n'avons pas pu identifier d'effet significatif de la réponse cellulaire. Cela peut être dû au fait que la réponse mesurée correspond aux cellules circulant dans le sang, alors que les interactions entre les cellules T et B se produisent généralement dans les centre germinatifs dans les organes lymphoïdes.

Ce premier modèle de la réponse immunitaire humorale à un vaccin contre Ebola a donc engendré des résultats intéressants, tant sur l'aspect quantitatif de la dynamique que sur l'identification des facteurs de variabilité de la réponse immunitaire. Il peut cependant être amélioré, notamment en prenant en compte la mémoire immunitaire. En effet, c'est la génération de la mémoire immunitaire qui est d'intérêt principal lors d'une vaccination. Pour cela, un travail a été commencé dans l'équipe afin de proposer des modèles pour la réponse immunitaire dès la première injection vaccinale dans lesquels les cellules B mémoires sont générées. Ces cellules sont rapidement capables de se différencier en cellules productrices d'anticorps après la deuxième injection vaccinale. Un premier travail a déterminé l'identifiabilité du modèle, la sensibilité de la dynamique des compartiments par rapport aux paramètres, ainsi que la calibration du modèle. Des données supplémentaires, en particulier celles concernant les cellules B, doivent être utilisées pour pouvoir estimer ce modèle.

Dans le cadre d'une modélisation qui s'inscrit dans une approche de biologie systémique, il serait également intéressant d'intégrer d'autres marqueurs de la réponse immunitaire, et en particulier des acteurs de la réponse innée. Cela pourrait être effectué en utilisant des valeurs de certains marqueurs majeurs à certains moments de la réponse comme covariable dans un modèle de la réponse adaptative. Il serait également envisageable d'utiliser un système d'équations différentielles modélisant les dynamiques de tous

les acteurs d'intérêt mais cela engendre plus de difficultés en terme d'estimation des paramètres. Une autre possibilité serait d'intégrer les données d'expression génique dans un modèle mécaniste de la réponse immunitaire. A l'heure actuelle, cela représente un vrai défi méthodologique. Le choix du type de modèle est également crucial, car les mécanismes de transcription et expression génique contiennent une stochasticité intrinsèque qu'il est difficile d'ignorer. Des modèles de réseaux de gènes ainsi que des méthodes d'inférence ont déjà été développés et il serait intéressant d'évaluer la possibilité d'utiliser ces méthodes et de les intégrer dans un modèle mécaniste de la réponse immunitaire. Cela permettrait à long terme de définir des modèles intégratifs de la réponse vaccinale, permettant d'aider à la mise en place de futurs essais cliniques.

Optimisation d'immunothérapies pour des patients infectés par le VIH

Un autre aspect de la thèse consiste à développer des méthodes d'optimisation de régimes d'injections répétées. Cela a été effectué en particulier dans le cadre clinique de patients atteints par le VIH. Ces patients reçoivent un traitement antirétroviral, ce qui leur permet de contrôler le virus et d'avoir une charge virale indétectable. Cependant, leur système immunitaire n'est pas totalement reconstruit suite à la prise du traitement, et les niveaux de lymphocytes T CD4 sont effectivement trop bas, inférieurs à 500 cellules par μL de sang. Des études ont montré que des patients infectés par le VIH avec des niveaux de CD4 plus élevés que cette limite ont un état de santé aussi satisfaisant qu'une personne saine. Il est donc crucial de développer des thérapies permettant d'augmenter les niveaux immunitaires de ces patients. L'immunothérapie par injections d'une cytokine, l'interleukine 7 (IL-7) est donc envisagée, cette cytokine stimulant la prolifération des lymphocytes T CD4 et augmentant potentiellement leur production thymique, leur survie et leur maturation. Des essais cliniques ont évalué l'effet d'injections répétées d'IL-7 sur la reconstitution de l'ensemble des lymphocytes T CD4 et ont montré que des injections réalisées en cycles de 3 injections espacées d'une semaine pouvaient aider à maintenir les niveaux de CD4 au-dessus de 500 cellules par μL de sang.

Afin de mieux comprendre et quantifier les mécanismes d'action de l'IL-7, des modèles ont déjà été développés et estimés sur les données des essais cliniques précédemment évoqués. Un modèle simple contient deux populations de CD4, une population étant au repos et l'autre étant en train de proliférer. Ces deux populations ont des taux de mort différents et les cellules au repos peuvent entrer en prolifération au taux π alors que celles en prolifération arrêtent de proliférer après une dernière division au taux ρ. Ce

modèle a permis d'évaluer l'effet des différentes injections dans un même cycle sur la prolifération des cellules, ainsi que leur effet sur la survie des cellules. Ce modèle ayant également montré un pouvoir prédictif certain, il est possible de l'utiliser pour simuler l'effet de différents protocoles d'injections sur des patients infectés par le VIH. L'étape suivante est donc d'optimiser les protocoles d'injections, c'est-à-dire utiliser un minimum d'injections d'IL-7 tout en maximisant le temps passé avec le nombre de lymphocytes T CD4 au-dessus de 500. Pour cela nous avons développé une méthode basée sur la théorie du contrôle optimal, et cette méthode a été évaluée sur un ensemble de pseudo-patients. Ce sont des patients fictifs générés en tirant aléatoirement un ensemble de paramètres suivant la loi a posteriori estimée sur les données des essais cliniques. Cela permet d'avoir un ensemble de patients représentatif de la population d'étude.

Afin d'appliquer des résultats récents de la théorie du contrôle optimal, nous avons d'abord décrit le processus à l'aide d'un modèle spécifique : un processus de Markov déterministe par morceaux (PDMP). Cette classe de modèles correspond à un processus qui suit une trajectoire déterministe ponctuée de sauts aléatoires. Un PDMP peut être défini de manière itérative : à partir d'un point de l'espace d'état, le processus suit une trajectoire définie par le flot (par exemple la solution d'un système d'équations différentielles) jusqu'à ce qu'un saut se produise. Cela peut arriver de manière aléatoire, selon une certaine intensité, ou bien de manière déterministe lorsque le processus atteint une frontière de l'espace d'état. Dans les deux cas, la mesure de transition permet de déterminer l'état à partir duquel le processus reprend. Dans le cas particulier du contrôle impulsionnel à la frontière, il est possible d'effectuer des actions ponctuelles lorsque le processus atteint la frontière de l'espace, ce qui peut modifier l'état à partir duquel le processus recommence. Dans notre cas particulier, nous suivons la trajectoire des CD4 et les injections d'IL-7 peuvent modifier la valeur du paramètre de prolifération des cellules pendant un temps aléatoire de plusieurs jours. Une stratégie (ici un protocole d'injections) correspond à un ensemble d'actions réalisées jusqu'à un certain horizon. A chaque stratégie, il est possible d'associer un critère de performance qui compile en fait l'ensemble des coûts engendrés par chacune des actions. Ici, le critère combine le nombre d'injections d'IL-7 effectuées ainsi que le temps passé avec un nombre de cellules CD4 inférieur à 500 cellules par μL. L'objectif est de minimiser ce critère de performance et de déterminer la stratégie correspondante. Pour cela, un opérateur intégro-différentiel, aussi appelé opérateur de Bellman dans la littérature, est défini à partir des caractéristiques du PDMP. En itérant l'opérateur de Bellman, on obtient une suite de fonctions qui converge théoriquement vers la valeur minimale du critère de performance, encore appelée la fonction valeur. Celle-ci permet alors de déterminer la stratégie optimale.

A l'heure actuelle, il n'y a pas de méthode générale pour résoudre les problèmes de

contrôle optimal. La résolution du problème optimal par l'itération d'une suite nous a permis de développer une méthode numérique basée sur la programmation dynamique. Pour cela, la suite est approchée sur une grille de l'espace d'état. La difficulté numérique réside non seulement dans l'organisation de la grille pour permettre de calculer la suite itérative sur la grille, mais également dans la taille de la grille qui peut engendrer des temps computationnels assez élevés. L'algorithme itératif a été développé sur le logiciel de calcul Matlab, et a été appliqué à un ensemble de 50 pseudo-patients afin de vérifier l'efficacité de la méthode. Le critère de performance a été calculé sur la stratégie optimale ainsi déterminée, et comparé à d'autres protocoles cliniques "naïfs". Les résultats obtenus sur les pseudo-patients ont montré que la stratégie optimale déterminée avait un coût moins élevé que les 5 autres protocoles cliniques envisagés. En effet, même si le nombre moyen de CD4 sur un horizon d'1 an était plus faible que celui obtenu avec des protocoles contenant plus d'injections, le temps passé en dessous de 500 était similaire, tout en utilisant moins d'injections. Cela montre que la stratégie ainsi déterminée est bien capable de réaliser un équilibre entre ces 2 quantités. De plus, la stratégie optimale ainsi déterminée est assez intuitive, puisqu'il s'agit d'utiliser 2 injections par cycle tant que les niveaux de CD4 du patient sont faibles (inférieur à 500), puis des injections seules permettant de maintenir le patient au-dessus de la limite de 500. Les résultats suggèrent donc que la méthode de détermination d'un protocole optimal d'injection fonctionne sur ces pseudo-patients, et pourrait être utilisée dans le cadre d'optimisation de protocoles de futurs essais cliniques.

Une limitation majeure de cette méthode est qu'elle suppose que les paramètres du patient sont parfaitement connus. Même si les méthodes d'estimation se sont montrées efficaces dans ces cas de modélisation, il y a cependant de l'incertitude lorsque les paramètres d'un nouveau patient inclus dans l'essai clinique sont estimés. Il est néanmoins difficile de gérer les problèmes d'estimation et d'optimisation de manière simultanée. L'estimation engendre une stochasticité due à l'incertitude autour de la valeur des paramètres biologiques du patient étudié, tandis que lors de l'optimisation du PDMP, la stochasticité est intrinsèque au modèle biologique en lui-même. Dans ce dernier cas, il pourrait être intéressant d'appliquer la méthode provenant de la théorie du contrôle optimal à d'autres processus biologiques. En particulier, les réseaux de gènes, déjà évoqués dans le cadre d'une approche de biologique systémique, peuvent être modélisés par des PDMP. L'expression génique pourrait être intégrée dans un modèle de la dynamique de marqueurs majeurs de la réponse immunitaire. Il pourrait alors être envisagé d'utiliser les méthodes de contrôle pour optimiser l'expression de ces gènes. Cela permettrait de contrôler de manière précoce la réponse vaccinale.

Conclusion

Dans cette thèse, nous avons donc réalisé des travaux à la fois mathématiques, immunologiques et biostatistiques afin de comprendre, quantifier et optimiser la réponse immunitaire à des interventions préventives et thérapeutiques contre des maladies infectieuses. Cela montre que des outils méthodologiques complexes peuvent être utilisés pour répondre à des questions cliniques concrètes et analyser des données longitudinales obtenues lors d'essais cliniques. Ces outils permettent la mise en place d'essais cliniques dits *in silico*, qui consistent à utiliser des simulations computationnelles spécifiques à chacun des patients étudiés, afin d'améliorer les développements cliniques. Ces essais pourraient à terme réduire le nombre de sujets recrutés dans les essais cliniques ou bien même remplacer des études animales ou humaines. Concernant les essais vaccinaux, une approche *in silico* pourraient également être proposée grâce aux outils de modélisation et d'optimisation : après avoir développé et estimé un modèle mécaniste intégrant toutes les informations disponibles (génomiques, protéomiques, microbiome, facteurs environnementaux), des pseudo-patients peuvent être simulés en utilisant les distributions des paramètres dans la population d'étude. Ensuite, des méthodes d'optimisation peuvent être utilisées pour déterminer, à titre individuel ou populationnel, quelle(s) serai(en)t la (les) meilleure(s) stratégie(s) optimale(s) à tester dans un futur essai clinique. Des choix devront être effectués dans la modélisation, notamment concernant la complexité du modèle au regard des données disponibles, mais également la stochasticité qui ne peut être négligée lorsque la dynamique de certaines acteurs, notamment à l'échelle génomique, est modélisée. Ces méthodes pourront permettre d'adopter une approche de biologie systémique pour de futurs développements vaccinaux.

1 Introduction

Even though vaccines have contributed to a major improvement of global public health, the development of effective immune interventions against infectious diseases such as Human Immunodeficiency Virus (HIV), Ebola or malaria is more difficult. These interventions can be preventive – to avoid infection by the virus, or therapeutic – to help the immune system to get rid of the virus. The development of vaccines remains a challenge not only because we need a better understanding of the mechanisms of protection generated by immunological memory after vaccination, but also because clinical trials assessing the safety, tolerability, immunogenicity and efficacy of vaccines are very long and expensive. Moreover, recent developments in vaccines have been based on "prime-boost" regimens, which combine several products in consecutive immunizations. This has induced new questions regarding the design of clinical trials, in particular: how many immunizations should be made ? In which order ? How long should we wait between the immunizations ? The "systems biology" approach aims at addressing these questions and understanding the whole operating mode of the immune system by integrating data from several markers of the immune response. This is difficult, as the immune system is composed of a large number of actors, connected by complex, non linear interactions. Mathematic models and computational methods are major useful tools in this approach.

The work presented in this thesis aims to fit in this approach, with two particular objectives: the first is to model the dynamic of the immune response following vaccine immunizations. It helps quantifying the biological process, especially by estimating the parameters associated to the model, based on data generated in human clinical trials. Determining a suitable model could also help predicting the immune response of a newly studied participant and to numerically compare several vaccine regimens. Defining a model complex enough to capture the dynamics of the immune response, but also not too complicated regarding the availability of the data for parameter estimation represents a challenge in itself. The second objective of the work is to optimize the clinical protocol for generating an efficient immune response. This is obtained by the development of a numerical tool based on optimal control theory, to determine the best product combinations that should be tested in a protocol of clinical interventions.

In practice, my work was separated into two main projects. The first is part of a European project, funded by IMI (Innovative Medicines Initiative) and based on interna-

tional collaborations with both private and academic partners. This project consists in developing a mathematical model for the humoral immune response to an Ebola vaccine, tested in humans in phase 1 to 3 clinical trials. The models allows not only to evaluate the durability of the immune response, as measured by the persistence of antibody concentrations, but also to determine some factors explaining the variability of the response across the studied population. The model is mechanistic: it aims at translating the knowledge about the biological process into mathematical equations. In particular, we worked on a basic mechanistic model, based on a system of ordinary differential equations (ODEs). The estimation of parameters relies on a population approach: each of the parameters is assumed to admit a mean value in the population, and individual variation can be observed around this value. This variability can also be explained thanks to environmental or vaccine-related factors. The parameters are estimated through a statistical method based on likelihood maximization. The second project consists in developing optimization tools for protocols based on consecutive injections. A first application of interest consists in the optimization of immunotherapeutic interventions in HIV-infected patients. The optimization of these injections (time and doses) is obtained by solving an optimal control problem. The method requires to model the process with another mathematical tool, piecewise deterministic Markov models (PDMP), where the system follows a deterministic trajectory and changes in the system can occur discretely after some random time periods. These models have been widely studied and some recent theoretical results allow to solve the optimal control by computing an iterative sequence based on an integro-differential operator. All together, these projects aim at developing mathematical and computational methods for the analysis of the dynamics of the immune system, in order to improve the understanding of the mechanisms of action of the immune system and to propose optimized and/or personalized immune interventions that should be tested in future clinical trials.

The thesis is organized as follows: as the modeling work is based on immunological knowledge, the first chapter aims at introducing the major notions of immunology, necessary to understand the principle of vaccinations. We focus in particular on the interest of prime-boost regimens. We also emphasize the different factors of variability influencing the immune response to vaccination. In this chapter, we also underline the role of mathematical modeling for understanding the key mechanisms of the immune response, quantifying the dynamics of the different actors and predicting the outcome of some interventions on the immune system.

The second chapter is devoted to the first project of modeling the humoral response to Ebola vaccine. We review in this chapter the latest clinical developments of Ebola vaccines, which were accelerated by the recent large West Africa epidemic: this helps

understanding the context of the generation of the data used for the estimation of the model. We also remind the existing methods of modeling the humoral immune response, to justify the choice of the model. Then, we present the model and its possible extensions.

In the third chapter, we develop the second project of optimizing the design of clinical trials. The theory of optimal control is applied to a particular biological framework (IL-7 immune therapy), which results are reminded, to understand the context of the study and the necessity for developing sophisticated methods of optimization. The theoretical results previously obtained for controlling PDMP are presented and the numerical method is applied on a number of pseudo-patients.

Finally, we conclude on how both modeling and optimizing approaches coul be combined in a systems vaccinology approach relying on *in silico* trials.

2 Immunological challenges

Abstract: In this chapter, we introduce some key notions in immunology. In particular, we focus on the establishment and maintenance of the immunological memory, which is crucial in vaccine development. We explain the principle of vaccination and the challenges that still remain in developing efficient vaccines against infectious diseases. Finally, we underline the role of mathematical modeling in understanding the mechanisms of the immune response and quantifying the dynamics of the biological processes. These models, coupled with optimization tools, could help improving the design of future clinical trials and accelerate the clinical development of vaccines.

Key Words: Immunological memory; vaccine; clinical trial; prime-boost regimens; factors of variability; mechanistic models; ordinary differential equations.

2.1 Generalities on the human immune system

2.1.1 Actors of the immune response

This introduction on the immune response is mostly adapted from the book of Abbas et al. [2010]. The immune system is composed of cells and molecules able to detect and react to different varieties of pathogens, inducing in this way immunity. Their organized response to protect the organism against the pathogens constitutes the immune response. The initial response is generated by the innate immunity, followed by the response of adaptive immunity. The innate response is non-specific and provides a first quick response against infection by microbes. In particular, it includes the physical and chemical barriers through the action of the skin and the mucous membranes, the activity of phagocytes (neutrophils, macrophages) which ingest pathogens and of natural killer (NK) cells, which trigger death of viral-infected cells through the release of some specific cell-secreted proteins, called cytokines.

The adaptive immunity, for its part, adapts to the infection by developing better response abilities, acquired with repeated exposures to the pathogen. It consists in the recognition of the pathogen, the development of a specific response and the generation of immune memory. It is mediated by the cell population of lymphocytes, produced from dif-

ferentiation of stem cells into the bone marrow. The adaptive response is itself composed of two types of responses: the humoral immunity and the cellular immunity. Humoral immunity is generated by antibodies, circulating in the blood and mucosal secretions and produced by the B lymphocytes. Antibodies are proteins with the shape of a Y, also called immunoglobulins (Ig), with one constant region at their base and one variable region able to recognize the antigen. They are composed of two heavy chains and two light chains. The heavy chains can be classified in 5 categories, determining 5 types of antibodies: IgA, IgD, IgE, IgG and IgM. Antibodies can either be found on the membrane of B lymphocytes and they act in this case as surface receptors for recognizing the antigen, or they are secreted by plasma cells (coming from activated B cells) and in this case they reside in the circulation, tissue and mucosal sites to eliminate the pathogen. Antibodies act against the pathogens by different mechanisms. They can bind to the antigen and promote its elimination by phagocytes. This way, they represent the first line of defense against extracellular antigens. They can also block the binding of viruses to target cells and neutralize in this way their infectivity.

Cellular immunity is achieved by the action of T lymphocytes. They are specifically committed when the antigen is intracellular and proliferates inside host cells, such as viruses: in that case, antibodies cannot access the antigen and T lymphocytes can enhance its elimination by killing infected cells. T lymphocytes can be divided into several categories of cells with different functions. Helper T cells, expressing the CD4 glycoprotein at their surface (also written $CD4^+$ T cells), secrete cytokines inducing proliferation and differentiation of the main actors of the immune response (T cells, B cells, macrophages), and stimulating antibodies production by B cells. Cytotoxic T cells, expressing the CD8 glycoprotein at their surface (also written $CD8^+$ T cells or CTLs), release cytokines able to eliminate virally infected cells and tumor cells. Finally, regulatory T cells (written Treg) are active in suppressing inflammatory responses and reducing the risk of autoimmune diseases. NK cells, which are part of the innate response also constitute a class of lymphocytes. The different categories of lymphocytes and their role in the immune response are summarized in figure 2.1.

The functions of the lymphocytes are mainly mediated through the presence of cytokines, which are proteins secreted by the cells playing a role in the immune response. Cytokines impact on the differentiation of lymphocytes and their effector functions, as well as on the hematopoiesis (formation and development of blood cells). Some are also called interleukins (IL), as they are produced by leukocytes (macrophages or T lymphocytes) and act on other leukocytes, with a standard nomenclature written with a number (IL-1, IL-2, ..). We do not intend here to review the role of all cytokines, but we will focus on some cytokines of interest, which will be useful for the work done in this thesis:

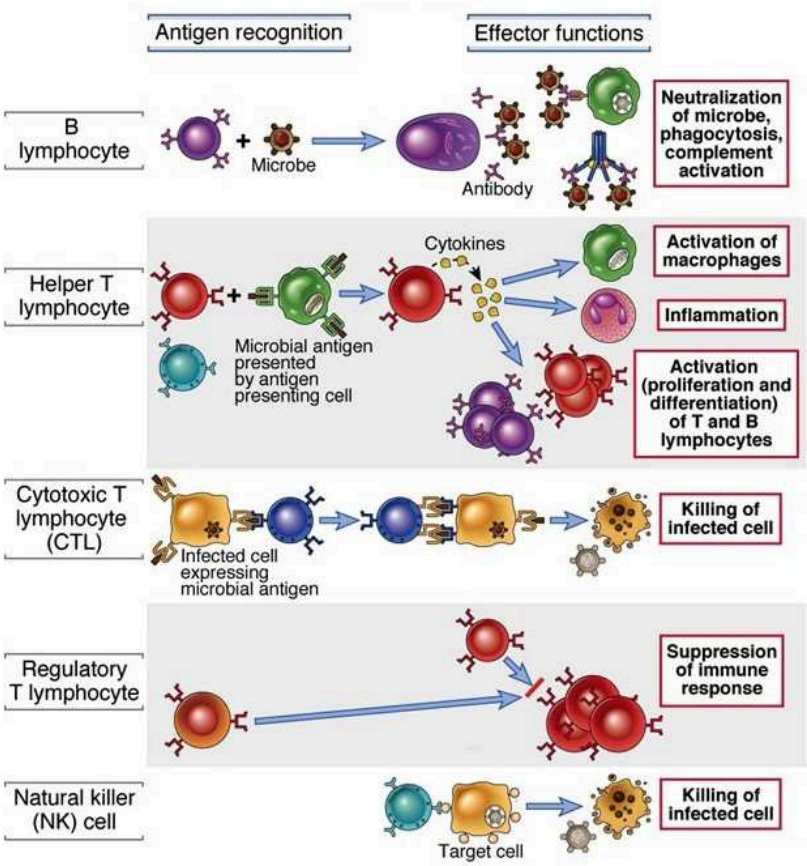


Figure 2.1 – Classes of lymphocytes. Taken from Abbas AK, Lichtman AJ, Pillai S, Cellular and Molecular Immunology, edition 11, Copyright Elsevier 2011.

— Interleukin-2 (IL-2): IL-2 plays a major role in the adaptive immune response. In particular, it is mainly produced by $CD4^+$ T lymphocytes and enhances the survival, proliferation and differentiation of activated T lymphocytes. It also has an impact on the proliferation of B cells and stimulates the production of antibodies. Moreover, IL-2 also has a role in the innate response, by stimulating the proliferation and differentiation of NK cells and Tregs.

— Interferon γ (IFN-γ): IFN-γ is produced by both NK cells and T lymphocytes and has functions in both innate and adaptive immunity. First, IFN-γ enhances phagocytose by macrophages. It also stimulates the differentiation of naive $CD4^+$ T cells and the recognition of antigens by T lymphocytes. Finally, it has an action on B cells and promotes the antibody response.

— Tumor Necrosis Factor (TNFα): TNFα is a cytokine of the innate immunity, mainly produced by phagocytes, but can also be produced by T lymphocytes and NK cells. Its main role is to improve the recruitment and the activation of neutrophils and monocytes (precursors of macrophages) to eliminate the pathogen at the site of infection. It can also induce the programmed death of some cell types.

— Interleukin-7 (IL-7): IL-7 is a cytokine produced by stromal cells in the bone marrow and the thymus. It stimulates the early development of B and T lymphocytes (before any immune response occurs) and is necessary for the survival of mature cells, including $CD4^+$ naive and memory T cells [Seddon et al., 2003] and $CD8^+$ naive T cells.

2.1.2 Phases of the adaptive immune response

Adaptive immune response is characterized by several phases: recognition of the antigen, activation of lymphocytes, elimination of the pathogen, before a return to baseline state and maintenance of the generated memory. This phases are represented in figure 2.2.

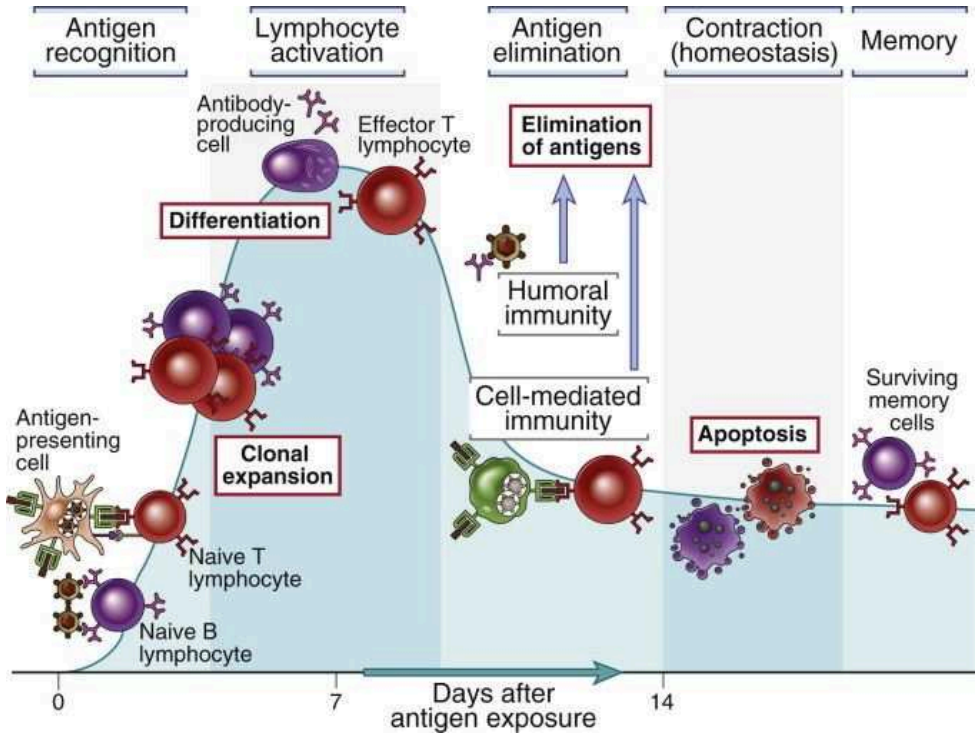


Figure 2.2 – Phases of the adaptive immune response. Taken from Abbas AK, Lichtman AJ, Pillai S, Cellular and Molecular Immunology, 3rd edition, Copyright Elsevier 2008.

The adaptive immune response is initiated by antigen recognition. T lymphocytes can only recognize peptide antigens expressed on the cell surface and encoded by genes in a particular locus defined as the major histocompatibility complex (MHC), but not soluble antigens. The MHC molecules are recognized by a receptor on the T cell surface called the T cell receptor (TCR). T cells expressing the same TCR correspond to a same clone of T cells. The possibility to recognize antigen is ensured by the large diversity of clones within

a given individual, each clone corresponding to a distinct antigen. To enable recognition by T cells, antigens should be captured and presented to their corresponding specific lymphocytes. This role is played by antigen-presenting cells, in particular by dendritic cells (DCs), able to transport the antigen to lymphoid organs and present it to naive T lymphocytes to trigger the immune response. In particular, DCs express at their surface class II molecules of the MHC, which are recognized by $CD4^+$ T cells. On the other side, $CD8^+$ T cells are able to recognize class I molecules of the MHC, expressed by any cells. Therefore, $CD8^+$ T cells can target any virus-infected cell for elimination. Naive B cells are situated in peripheral lymphoid tissues and can be activated when the antigen for which they are specific bind to their membrane receptor. Naive B cells express first the IgM and IgD antigen receptors, but can produce other antibodies after being activated by a pathogen, such as IgA, IgE and most particularly IgG. This process is called isotype switching. A selection of the B cells producing antibodies with higher affinity to the antigen is made before expansion of these cells. Most of the time, the activation of B cells and antibody production is dependent on the $CD4^+$ T cell response [Crotty, 2015].

Antigen recognition by the lymphocytes induces a phase of activation, during which lymphocytes undergo a large clonal expansion due to proliferation of the activated cells. For the T lymphocytes, differentiation into effector and memory cells occurs simultaneously. These T cells can either stay in the lymphoid organs or migrate to non-lymphoid tissues. For the humoral response, B cells can differentiate into antibody-secreting plasma cells or into memory cells. Most of the affinity maturation and memory generation happen through B-T cells interaction in germinal centers, a region in the lymphoid follicles. Antibodies can enter the circulation, while plasma cells migrate from the peripheral lymphoid organs to the bone marrow. Both B and T effector cells can act quickly to eliminate the pathogen during the effector phase, while memory cells will remain in the organism, ready to respond at the next encounter with the pathogen.

Most of the effector lymphocytes are short-lived and die by apoptosis at the end of the immune response (after elimination of the pathogen). Apoptosis is a programmed cell death, induced by different pathways. In the case of an immune response, it is due to the fact that the survival of lymphocytes can depend on the presence of antigen and also because the organism is auto-regulated and limits its own number of specific cells by homeostasis process.

A small proportion of the activated lymphocytes with memory phenotype survive after the end of the immune response and sustain in the organism. These cells are able to respond quicker and more intensely in case of re exposure to the pathogen, as explained in the following section.

Lymphocytes can be classified in different subsets depending on their phenotypes

(such as effector and memory phenotypes). These phenotypes are defined by specific combinations of markers. Markers are molecules expressed at the surface of the cells. For example, as previously defined, T helper cells express the CD4 marker on their surface and are then written $CD4^+$ T cells. New technologies such as flow cytometry and more recently mass cytometry allow the identification of the co-expression of a large number of markers at the surface of the cells, and the level of expression can also be measured. Combination of markers help defining a number of cell subtypes (e.g., central and effector within the memory cells), but will not be detailed in this thesis. We can although mention the Ki67 marker, which is used for detecting the proliferation of cells. We will consider in this thesis that cells expressing Ki67 are in a proliferating state.

2.1.3 Immunological memory and secondary responses

After primary exposure to a pathogen, the organism acquires the ability to generate a better response in case of a secondary exposure: it is called immunological memory [Ahmed and Gray, 1996]. Several definitions can be given to immunological memory [Farber et al., 2016], but most of them agree to say that this antigen-specific response occurs faster and is quantitatively and qualitatively better at eliminating the pathogen compared to the primary response. This is due to two main reasons. First, the number of antigen-specific lymphocytes increases at each encounter with the pathogen, and then more cells can recognize the pathogen and expand. Moreover, as mentioned in section 2.1.2, the immune response generates a number of memory cells which have different characteristics than naive ones, including more efficiency in eliminating the pathogen. As an example, after a first encounter, B cells can produce antibodies with better affinity and binding properties than antibodies produced by naive B cells activated during a primary response. Thus, the antibody level increases after repeated immunization, as well as the affinity of these antibodies, as shown in figure 2.3. For T lymphocytes, although there are differences in the differentiation and memory generation of $CD4^+$ and $CD8^+$ lymphocytes [Seder and Ahmed, 2003], both memory populations show an increased sensitivity to the antigen and require less stimulation pathways for activation than naive cells. It should be noted here that there is still poor understanding of human memory T cells, whether it be the diversity of the memory T cell subsets, the mechanisms involved in the maintenance of the memory, their migration in the organism or their reactivation, and there is a large literature on these questions [van Leeuwen et al., 2009; MacLeod et al., 2010; Martin and Badovinac, 2014; Fraser et al., 2013]. Similarly, the relative roles of long-lived plasma cells and memory B cells after re-encounter of a pathogen are not completely clear and more studies should be realized to address these questions [Pape et al., 2011; Nutt et al.,

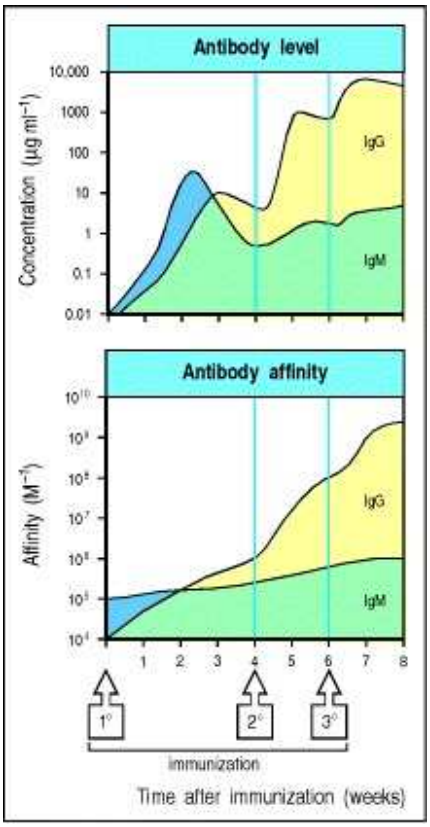


Figure 2.3 – Evolution of the number of antibodies and their affinity through repeated immunizations. Antigen-specific IgGs have more affinity than IgM to the pathogen. Taken from Janeway et al..

2015; Taylor et al., 2012]. However, it is beyond the scope of this work to develop these points in details. We will focus on vaccine development, as the principle of vaccination relies on immunological memory.

2.2 Vaccine development

2.2.1 Principle of vaccination

The development of vaccines has been empirical, following Edward Jenner's discovery at the end of the 18th century: he observed that humans infected by an animal poxvirus underwent an attenuated disease and could be protected against smallpox [Plotkin, 2014]. This led to the eradication of smallpox by 1980 and the development of a large number of successful vaccines: polio is almost completely eliminated, and the incidence of diphteria, tetanus, pertussis has been reduced by more than 95% in the last decades [Rappuoli et al., 2011]. Globally, 2-3 millions of deaths are currently prevented every year thanks to immunization against diphteria, tetanus, pertussis and measles [WHO, 2018d]. For

example, measles vaccination is estimated to having helped reducing the number of deaths by 20.4 millions between 2000 and 2016 [WHO, 2018e]. In the specific case of Human Papillomavirus (HPV), an elimination sounds possible [Brisson et al., 2016] with the vaccine as reported especially in Australia [Patel et al., 2018].

The basis of vaccination relies on the immunological memory: the immune system is able to generate a better response to a pathogen after a primary exposure, inducing protection and preventing the recipient from infection by the pathogen. Exposure to an attenuated virus can therefore develop an immunological memory to the pathogen without putting the recipient at risk of developing the disease. This approach has been very successful for some viruses but revealed insufficient and unsafe for other pathogens for which there is no natural recovery after infection, with high levels of mutations, or for which the antibody response is not sufficient to protect against the disease [Germain, 2010]. New technologies have allowed developing new vaccine platforms, such as DNA vaccines, virus-like particles, viral subunits, fusion proteins and peptids, or viral vector-based vaccines; they rely on the same initial idea of vaccination and aim at presenting one particular part of the antigen to the immune system to generate a strong and long-lived immunological memory able to protect against infection in case of encounter with the real antigen. In particular, viral-based vaccines are gene-depleted viruses in which are inserted genes of the targeted virus. These vaccines are effective as they can generate high immunogenicity through both humoral and cellular responses [Ura et al., 2014]. For safety purpose, non replicative and low pathogenic viruses are usually selected. However, their efficacy can be limited by the pre-existing immunity to the vector: previous exposures can have induced the development of neutralizing antibodies specific to the viral vector [Mast et al., 2010; Priddy et al., 2008]. In particular, the class of adenoviruses (Ad) is widely evaluated in clinical trials [Hammer et al., 2013; Gurwith et al., 2013]; there exists a large number of human serotypes causing, among others, cold and sore throats. Vaccinia virus is also generally used: it is a member of the poxvirus family, used in the smallpox vaccine. More specifically, modified vaccinia Ankara (MVA), an attenuated strain of this virus, has been evaluated in several clinical trials for vaccines against several diseases, e.g. HIV [Gómez et al., 2011] and malaria [Bejon et al., 2007].

2.2.2 Challenges in vaccine development for infectious diseases

Even though vaccines have allowed major progresses in reducing the incidence of some diseases worldwide, we are still lacking effective vaccine against some infectious diseases such as HIV, malaria and Ebola virus disease. This represents a major public health challenge. The current difficulties to develop effective vaccines against these infectious

diseases is due to our lack of deep knowledge in immunology and the mechanisms of action of immune memory [Hagan et al., 2015; Pulendran and Ahmed, 2011; Germain, 2010]. We will develop in this section some specific difficulties faced by researchers in the development of effective vaccines.

2.2.2.1 Clinical development

The development of vaccine is a long and expensive process. As every other drug, it needs to be tested in several phases before approval process by the health authorities, such as the Food and Drug Administration (FDA) in the USA or the European Medicines Agency (EMA) in Europe. Before evaluating the vaccine in human studies, it is first tested on animals in preclinical studies, such as mice or non-human primates (NHP). Clinical development undergoes a process in three phases. Phase 1 clinical trials recruit a small number of subjects (less than one hundred), usually healthy, to assess the safety and tolerability of the vaccine. These trials can also be dose-escalating, where increasing doses are tested until finding a dose inducing an acceptable level of the immune response while limiting the amount of adverse events. Phase 2 trials are realized on more subjects (100-300), representative of the vaccine target population, and aim at evaluating the immunogenicity of the vaccine, while still monitoring its safety. They can last several months to years. Phase 2 trials are sometimes divided in phase 2a and 2b trials, where phase 2b trials are specifically designed to evaluate the efficacy of the vaccine. Phase 3 trials aim at evaluating the effectiveness of the vaccine intervention. For that, a large number of subjects are recruited (300 to thousands) and are randomized in a placebo arm (or other reference if another vaccine already exists) and a vaccine recipient arm. For a prophylactic vaccine, effectiveness corresponds to the ability to prevent from infection when at risk. Finally, after the vaccine has been licensed and is made available, it undergoes a so-called phase 4, corresponding to the long-term surveillance of the product. So far, only a few HIV vaccine candidates have been evaluated in efficacy trials [Stephenson et al., 2016], but encouraging results were recently obtained from a phase 1/2 clinical trial [Bekker et al., 2018] and initiated a phase 2b/3 efficacy trial to assess the capacity of the vaccine to confer protection against HIV infection in South Africa [Barouch, 2018]. The beneficial effect of some vaccine, such as HIV or Ebola vaccine are very difficult to assess in phase 3 trials: indeed, due to its mode of transmission and the preventive actions taken in parallel to vaccine studies, it is difficult to assess the real amount of protection induced by an HIV vaccine candidate. For Ebola, it would be necessary to observe the potential protection induced during an epidemic, during which vaccine transmission is active. These diseases require then the use of surrogates of protection, which can be used to replace the disease incidence criterion. Although some surrogates of protection exist

for other diseases, as reviewed in Plotkin [2010], it is not yet the case for HIV and Ebola. This will be developed in more details for Ebola in section 3.1.2.2. This issue constitutes a real challenge in developing and evaluating new vaccines.

2.2.2.2 Prime-boost regimens

In the last few years, there has been some growing interest in the so-called "prime-boost" regimens for developing new vaccines [Ramshaw and Ramsay, 2000]. These regimens combine several products in distinct, consecutive immunizations. They can either be homogeneous, by using the same product, or heterogeneous when different products are injected. They are expected to produce a better immune response, both quantitatively and qualitatively. The only clinical trial that showed moderate classical efficacy (31.2 % with confidence interval 1.1-52.1) of a vaccine against HIV infection was based on a prime-boost regimen, consisting of four priming immunizations of a first vaccine followed by two booster immunizations of another type of vaccine [Rerks-Ngarm et al., 2009]. Following this study, there has been a wide range of clinical trials evaluating prime-boost regimens for HIV or Ebola vaccines [Baden et al., 2016; De Rosa et al., 2011; Tapia et al., 2016; Milligan et al., 2016]. However, these regimens raise a number of questions that are still to be answered by improving our immunological knowledge [Sallusto et al., 2010]. In particular, we have not yet determined how many immunizations are necessary, and how the order of administration has an impact on the immune response, in the case of heterogeneous regimens. Moreover, the interval of time that should be considered between immunizations is still an open question of research. There is some consensus on the fact that boosting the immune response too early is sub-optimal [Sallusto et al., 2010], as some time is needed for the cells to acquire a memory phenotype [Wherry and Ahmed, 2004], and the cells of the immune system have a limited capacity of expansion and could be exhausted with intense stimulation [Pollard et al., 2009]; waiting too long before the next immunization can also be sub-optimal as the immunity wanes with time as the number of immune cells decreases. Determining the optimal window of time for booster immunizations can also depend on other factors, such as the type of vaccine, the strength of the primary immunization or the baseline state of the immune system [Wherry and Ahmed, 2004]. It represents then a real challenge in the development of new efficient vaccines.

2.2.2.3 Variability of the immune response

The variability of the immune response at both individual and population levels has been observed in several vaccination studies. It can be induced by a large number of factors, including genetic factors, demographic factors, environmental factors and micro-

biome. These factors were widely covered in Van Loveren et al. [2001] and Pulendran [2014], and their effect in the case of HIV vaccine was reviewed in de Bruyn [2010]. This section does not intend to establish a review of all factors impacting the immune response, but rather to give some insights on previously conducted studies and the difficulty to assess the relative and causal effect of all these factors. The influence of genetic factors was described in particular in studies on identical twins: they have shown that the heritability for antibody response to several vaccines was at least of 40% [Lee et al., 2006; Newport et al., 2004]. A summary of the studies correlating genomic and transcriptomics variations to the immune response can be found in O'Connor and Pollard [2013]. In term of demographic factors, sex has been regularly shown to induce different immune response to both infectious diseases and vaccines. In a review on human vaccines, Cook [2008] shows that a large number of studies have found differences in antibody response between men and women, associated to a difference in the clinical efficacy in some vaccines - against influenza, hepatitis A virus (HAV), hepatitis B virus (HBV), pneumococcal polysaccharide, diphteria and to a difference of adverse effects - for rubella, measles and yellow fever vaccines. The difference was not consistent among all vaccines, meaning that in some cases female had greater antibody concentrations than male and it was the other way around in other cases. These observed differences could be due to genetic factors (genes from the X chromosome), hormones levels or anatomic differences [Fish, 2008]. In addition to sex, age is also known to be associated with vaccine efficacy, especially for influenza vaccine [Seidman et al., 2012; Nakaya et al., 2015] and dengue vaccine, where efficacy was observed to increase with age [Capeding et al., 2014], especially against severe disease. This variation could however be partly explained by increased prevalence of baseline immunity with age [Dans et al., 2018; Sridhar et al., 2018]. Moreover, higher variations in the immune response seem to be observed in aging populations [Shen-Orr and Furman, 2013]. The effect of nutrition was also studied as a potential factor impacting on the immune response [Savy et al., 2009], and oral polio vaccine was shown to be less effective in infants with malnutrition [Haque et al., 2014]. Of interest is also the observed difference of immune response between different populations (due to geographic settings and/or ethnicity). A study of demographic factors influencing the immune response to an HIV vaccine candidate showed that African Americans secreted more neutralizing antibodies after vaccination compared to White Americans, suggesting that ethnicity itself could affect the immune response to some immunogens [Montefiori et al., 2004]. In this study, the effects of age and sex on the immune response were not found significant. In other studies comparing African subjects to Western countries subjects, the immune response was mostly found lower in Africa. A study assessing the immunogenicity of Ad-based strategies for prophylactic HIV vaccine showed that T cell responses were lower in East

Africa compared to South Africa and the United States, although no difference in antibody concentrations was observed [Baden et al., 2016]. An interesting study on yellow fever vaccine showed the induced $CD8^+$ T cell and B cell responses were lower in subjects recruited in Uganda compared to subjects in Switzerland [Muyanja et al., 2014]. Prior to vaccination, the immune environment was more activated in African subjects, as shown by higher frequencies of differentiated T and B cell subsets and proinflammatory monocytes, as well as exhausted and activated NK cells. This immune activation at baseline was negatively correlated with yellow fever specific neutralizing antibody concentrations after vaccination. Altogether these results showed that the state of the immune environment before vaccination was different between African and European subjects and could impair the efficacy of the vaccine. Some studies showed the efficacy of Bacille Calmette-Guérin vaccination was lower in African infants compared to European ones [Black et al., 2002; Lalor et al., 2009]. This difference was actually explained by pre-existing immunity, as the percentage of subjects with IFNγ response at baseline (before vaccination) was higher in Malawi than in the UK and the fold-increase of IFNγ response after vaccination was higher in the UK compared to Malawi. Vector-based approaches for developing vaccines are also affected by this point, as pre-existing immunity to the vector can also have influence on the outcome of the vaccination. It was especially shown in the Step trial, which was interrupted after an increased rate of HIV infection was observed in vaccinated subjects who were seropositive for Ad5 at baseline compared to placebo recipients [Buchbinder et al., 2008]. Other factors such as co-existing infections are also considered as factors influencing the immune response to vaccination, in particular with infectious diseases (HIV, HBV, cytomegalovirus [Nielsen et al., 2015]) or parasites [Da'dara and Harn, 2010]. In addition to all these factors, the microbiome has been recently suggested to affect the response to some vaccines [Ferreira et al., 2010], possibly due to cross-reaction between some microbiota peptides and agents of the immune system. In particular, the impact of microbiome on vaccine response was studied on HIV vaccine [Williams et al., 2018] and oral rotavirus vaccines [Magwira and Taylor, 2018]. The effect of gut/intestinal microbiome could also reflect the effect of nutrition on the vaccine response. It also suggests the interest of assessing the effect of prebiotics, probiotics and antibiotics on the immunogenicity of vaccines: some studies have shown that probiotics could increase the immunogenicity of polio vaccination [de Vrese et al., 2005] and oral rotavirus vaccination [Isolauri et al., 1995]. However, in a study on Indian infants, an antibiotic therapy modified the intestinal environment without improving the immunogenicity of oral poliovirus vaccine [Grassly et al., 2016]. Finally, all these factors (genetic, environmental, microbiome) are not independent from each other and the way they affect the immune response still remains a key question in vaccinology.

2.3 The role of mathematical modeling

Mathematical modeling of the dynamics of the immune system is part of the effort of systems biology, mentioned in the introduction, to understand more deeply the whole process of immunity and the mechanisms establishing the immune memory. The necessity for using mathematical tools comes from both the limitation in immunological knowledge and the availability of more and more immune data that should be precisely analyzed to improve our understanding of the immune system. Modeling offers the opportunity to integrate the information of several datasets. Moreover, biological interactions are usually very complex and lead to non-linear behaviors of the systems: it can be due to feedbacks processes or amplification of effects due to high proliferation [Germain, 2001]. It can make it difficult to experimentally predict the reaction of the immune system to stimulations. Models can help understanding the underlying mechanisms of the biological process and generating new hypotheses. More than that, the value of the models reside in their ability to quantify the dynamics of biological process [Germain, 2017] and the factors impacting on these dynamics, but also the variability of dynamics within a given population. Finally, well calibrated and estimated models can be useful for predicting the outcome of interest using observed factors.

There exist different types of models, but we can mainly distinguish descriptive from mechanistic models. Descriptive models are data-driven and aim at best fitting the data and trying to explain an outcome (such as a biological marker) from other factors. They are mainly composed of statistical models, such as, for example, linear or non linear mixed models when the data is constituted of repeated measurements. On the other hand, mechanistic models are based on biological knowledge and aim at describing the mechanisms of action of a process: the initial knowledge of how the process evolves is translated into mathematical equations which can then be applied and compared to experimental results [Vodovotz et al., 2017]. These models can better account for the complex nonlinear relationships between biological components than descriptive models. We will focus here on mechanistic models composed of systems of ODEs, as it will be the main tool in this thesis. However, other tools can be used such as agent-based models, where each agent is an individual model entity and evolves in a complex system by interacting with other agents and following established rules [Bonabeau, 2002]. This type of models is relevant when studying spatial and local interactions between cells and they were for example developed for antigen recognition and activation of lymphocytes [Seiden and Celada, 1992], interactions between innate and adaptive responses [Folcik et al., 2007] or in multi-scale models of the cellular immune response [Prokopiou et al., 2014]. However, it presents some limitations, as the outcome is usually only a computational simulation of the evolution

of the system and it is difficult to obtain a mechanistic interpretation of the dynamics of the system.

Rather than considering individual agents, systems of ODEs model population sizes and concentrations. They constitute a widely-used, flexible tool for modeling the dynamics of the immune system and their utility in immunology have been reviewed in several papers [Yates et al., 2001; Callard and Yates, 2005; Castro et al., 2016]. One of the main challenges when modeling the immune system is to find a good balance between the complexity of the model, the scientific question and the availability of the information/data. For example, there are quite a number of studies aiming at modeling the whole immune system using systems of ODEs. These models account for the dynamics of a large number of actors of the immune system and include the spatial dimensions [Lee et al., 2009; Kim et al., 2007]. However this leads to systems with many equations and parameters, and we do believe that the complexity of these models is too high compared to the amount of data available to correctly estimate, or at least calibrate the model. Smaller systems of ODEs have been successfully used for modeling the dynamics of infectious diseases. In particular, HIV infection has been substantially studied [Perelson et al., 1993; Perelson, 2002; Perelson and Ribeiro, 2013] which allowed, among others, better understanding of the CD4-HIV interactions, quantification of rates of HIV production, prediction of the effect on some antiretroviral treatments and offered more recently the possibility to individualize the treatment strategies [Prague et al., 2013b]. This subject is however beyond the scope of the thesis.

Regarding the establishment of the immune memory, ODE models have been useful for quantifying the generation of immune cells and determining the pathway of differentiations leading to different subset of immune cells. In particular, a model of the $CD8^+$ T cells dynamics following infection by lymphocytic choriomeningitis virus in mice was first developed in de Boer et al. [2001] and allowed quantifying the magnitude of the response to different epitopes of the virus. It consisted in antigen-specific naive cells becoming activated cells after encounter with pathogen, a proportion of them able to differentiate into memory cells after some given time. This model was then used in several studies [de Boer et al., 2003; Kohler, 2007; Graw et al., 2012] and helped predicting the generation of memory cells. This model was also extended in Antia et al. [2003] and Antia et al. [2005] to determine the pathway of differentiation from naive to memory cells, by comparing two models (one where proliferating effector cells differentiate into memory cells and one where proliferating memory cells differentiate into effectors) and trying to fit mice data with both models. This allowed to identify a preferential pathway of differentiation, as the other model was not able to fit the data within biologically reasonable ranges of parameters values. Pathways of differentiation were also studied in Crauste et al. [2017], by

considering more subsets of effector and memory cells, defined by combinations of markers. This model allowed predictions of the quantity of memory cells generated during the immune response using early measurements during the effector phase.

Mechanistic models have proved valuables in both understanding the mechanisms of the immune response and predicting the quantity of memory response generated after infection. However, there is less literature on mechanistic models for the response to vaccine immunizations and the study of the interval between prime-boost immunizations. Some papers studied the immune response to yellow fever and vaccinia virus vaccination in humans [Le et al., 2015], and hepatitis A vaccination [Andraud et al., 2012]. They will be described in more details in section 3.1.3. For the study of the interval between prime-boost immunizations, we can mention the work of Castiglione et al. [2012], where authors developed an agent-based model for the dynamics of innate and adaptive immunity during the different phases of antigen recognition and response to pathogen, accounting for affinity between the different actors of the immune response. The model was calibrated on real data and allowed identification of an optimal time window for a boost immunization. A combined agent-based and ODE-based model accounting for the spatial aspects of the immune response, with actors in the lymph nodes and the blood, was also developed by Gong et al. [2014] and allowed the simulation of the response to a boost immunization. Both models gave insights into the determination of optimal schedule of secondary immunization, but were lacking a clinical application, due to the fact that the model was not estimated on real data. More generally, there is some literature on determining optimal schedules of repeated injections in other applications than vaccine, as will be developed in section 4.2.1.1, and these type of methods could be applied in other frameworks, such as optimization of vaccine regimens. Control theory applied to immunology should help designing future immune interventions for preventing and curing infectious diseases.

This justifies our approach in this thesis, of first developing a model of the immune response to an Ebola vaccine and estimating its parameters on data from clinical trials, and then applying the theory of optimal control to establish a tool for the optimization of repeated injections, applied in particular to the HIV framework. Both aspects of this work are developed in the following chapters.

3 Modeling the immune response to Ebola vaccine

Abstract: In this chapter, we show how Ebola vaccine development has accelerated in the past few years. In particular, prime-boost strategies combining Ad26 and MVA platforms have been tested in phase 1 to 3 clinical trials in Europe, Africa and the USA. We present a model for the dynamics of the humoral response following the boost immunization, in order to quantify the contribution of two populations of ASC (differing by their half-life) and estimate the factors impacting the immune response. The parameters are estimated using binding antibody concentrations data from 177 subjects in Europe and East Africa, with a population approach accounting for the effect of covariates and unexplained inter-individual variability. In particular, the different vaccine regimens seem to affect only the peak of the antibody response, but the geographical location has an impact on the dynamics of the long-lived ASCs: it induces a persistence of antibodies at higher levels in European subjects compared to East African ones. This could have an impact in the implementation of future clinical vaccination strategies. Models of the immune response could be improved by integrating more data and compartments.

Key Words: Ebola virus disease; vaccine; EBOVAC consortium; humoral response; antibodies; antibody-secreting cells; mechanistic model; ordinary differential equations; population approach; linear mixed models; parameters estimation; likelihood maximization; factors of variability.

3.1 Biological and clinical context

3.1.1 General introduction on Ebola

3.1.1.1 Ebola virus disease

Ebola virus disease (EVD) is a pathology that has repeatedly caused deadly epidemics of hemorrhagic fever in African countries [Peters and Peters, 1999] since its discovery in 1976 in the Democractic Republic of Congo (DRC) and Sudan [Johnson et al., 1977]. EVD is due to Ebola virus (EBOV), a filovirus of the genus Ebolavirus. Filoviruses also include

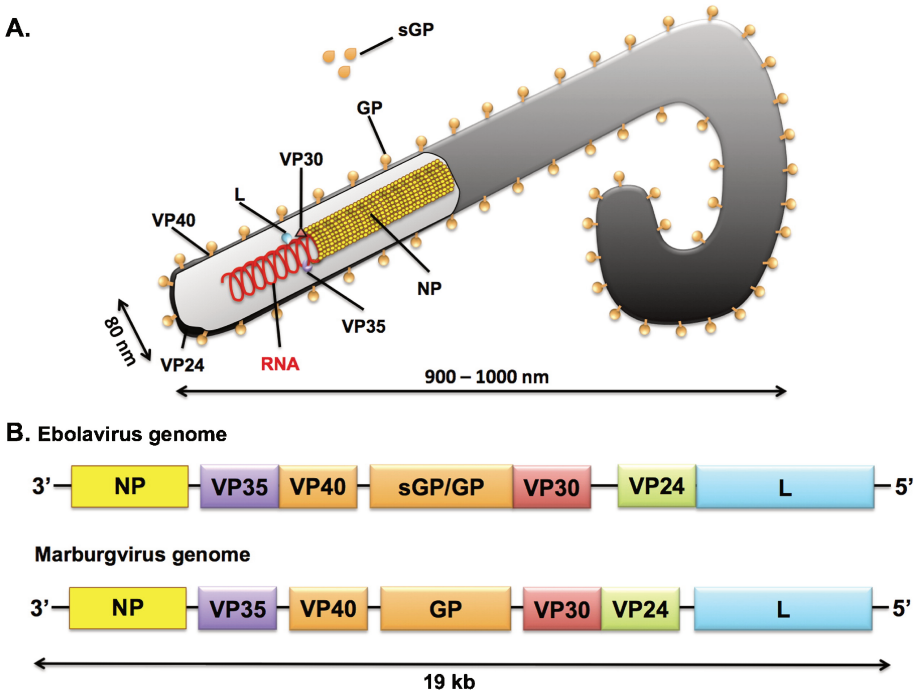


Figure 3.1 – Figure from Rougeron et al. [2015]. (A) Schematic illustration of a filovirus. sGP is specific to EBOV. (B) Schematic representation of Ebola virus and Marburg virus genomes.

the genus Marburgvirus, with Marburg virus (MARV) discovered in Marburg, Germany in 1967 and causing also an hemorragic fever in humans, and the genus Cuevavirus, with Lloviu virus, mainly affecting bats. Ebolavirus genus contains 5 species, named after the region where they were first identified: Bundibugyo ebolavirus (BDBV), Reston ebolavirus (RESTV), Sudan ebolavirus (SUDV), Tai Forest ebolavirus (TAFV) and Zaire ebolavirus (ZEBOV) - the most virulent. ZEBOV contains itself several strains, differing by phenotypic characteristics, such as Mayinga, Kikwit and Makona strains [Kuhn et al., 2013].

Filoviruses are characterized by a lipid envelope shaped as a long filament and a RNA genome encoding for 7 structural proteins - one glycoprotein (GP) appearing repeatedly on the outer viral envelope, a nucleoprotein (NP), virion protein 24, 30, 35, 40 (VP24, VP30, VP35, VP40) and an RNA-dependent RNA polymerase (L), with also two non-structural proteins - one soluble (sGP) and one small soluble (ssGP), as shown in Figure 3.1 [Rougeron et al., 2015].

Although the host reservoir of EBOV is not fully determined, fruit bats are thought to be asymptomatic hosts. Primates are susceptible hosts, but they are unlikely to be the true reservoir, due to the high pathogenicity of EVD among them. Transmission from animals to human population is realized through contact with body secretions, in particular blood

of infected animals found in the rainforest or during hunt-related activities. Human-to-human transmission of EBOV occurs via direct contact with body fluids of infected people and surfaces and/or materials contaminated with these fluids. High risk of transmission are also induced by burial rituals involving direct contact with the body of the deceased. This close-contact mode of transmission explains why relatives and health-care workers are frequently infected [WHO, 2018a]. Sexual transmission from survival patients could induce EBOV infections even after West African countries are declared Ebola-free [Butler, 2015]. Studies have shown viable Ebola virus can persist in semen for months and there is a risk of sexual transmission [Thorson et al., 2016]. Additional surveillance data and research are necessary to better assess this risk.

EBOV penetrates the body through lymphatic and blood vessels by direct contact with broken skin of mucous membrane. The main target cells of the virus are the dendritic cells and macrophages. These cells circulate in the body and allow the virus to spread in the body. It induces a systemic inflammatory response syndrome and multiple tissue damages, especially in the liver.

The incubation period can last up to 21 days. The first symptoms are not specific and include headache, fatigue and muscle pain. It is usually followed by vomiting, diarrhea, symptoms of impaired kidney and liver function and in some cases, internal and external bleeding (especially from the digestive system). Laboratory tests are necessary to confirm diagnosis [Liu et al., 2015].

EVD is often fatal in humans, with a case fatality rate around 50%, varying from 25% to 90 % in previous outbreaks. Filoviruses have been classified as Category A potential bioterrorism agents by the Centers for Disease Control and Prevention [CDC, 2018], which corresponds to the highest risk pathogens for public health.

3.1.1.2 West Africa epidemic

The recent outbreak that occured in 2014-2016 in West Africa caused 28 616 cases and 11 310 fatalities [WHO, 2016b]. The epidemic was due to an outlier strain (Makona) of ZEBOV, sharing a common ancestor with the known DRC and Gabon strains, but in a different clade [Baize et al., 2014]. It started in Gueckedou rainforest region, on the east of Guinea, in December 2013, closely followed by infections in Macenta region. A first press release was issued by the World Health Organization on March 23, 2014. Cases then spread to Sierra Leone and Liberia, which are bordering countries of Guinea. These 3 countries were the most affected by the epidemic, but some travel-related cases were reported in other African countries (Mali, Nigeria, Senegal), Europe (UK, Italie, Spain) and the USA. The Public Health Emergency of International Concern was started by WHO on August 8, 2014 and terminated on March 29, 2016. The unprecedented scale

of the epidemic was not found to come from higher rates of infection and transmissibility than already observed [WHO Ebola Response Team, 2014]. The infection was easily spread due to the large number of connections between Guinea, Liberia and Sierra Leone populations in the border area, with high traffic between rural and urban areas. Moreover, control measures were not set quickly enough to contain the epidemic.

An additional outbreak started in the western part of DRC on April 4, 2018. This outbreak was declared by the Ministry of Health on May 8, 2018. A total of 54 cases were reported, including 33 fatalities. The last laboratory confirmed EVD case was found on June 2, 2018 and the end of the outbreak was declared on July 24, 2018 [WHO, 2018b]. However, another (unrelated) outbreak in the eastern part of the country was reported to WHO by the Ministry of Health on August 1, 2018. As of August 20, 2018, a total of 102 cases were reported, including 59 deaths [WHO, 2018c]. Both outbreaks were caused by the ZEBOV species.

3.1.1.3 Prevention measures and therapeutic developments

There is no licensed therapeutic treatment for EVD at the moment, and most of the patient care constitutes of rehydration with oral or intravenous fluids, and treatment of his/her specific symptoms. To limit outbreaks, prevention and control measures should be implemented. They mainly concern the reduction of the risk of human-to-human transmission in households and health-care facilities, but also the reduction of wildlife-to-human transmission and of the risk of sexual transmission [WHO, 2018a]. Isolation of patients with suspected or confirmed EVD is necessary to allow them access to care and prevent them from transmitting the disease. Persons who have been in contact with an infected patient should be monitored for 21 days to detect a possible infection. As burial traditions induce high transmission risks, precautions should be taken with dead bodies and communicated to the community. The population should be informed of simple hygiene measures, such as hand hygiene. It is also necessary to train all health-care workers to use personal protective equipment [WHO, 2016a]. However, in the case of large-scale epidemics, it can be problematic to establish rapidly and effectively these control and prevention measures.

Having developed preventive vaccines or efficient treatments before the next epidemic could help preventing a similar large-scale epidemic by reducing both transmission rate and fatality rate of the virus. In this section, we will only speak of the therapeutic treatments. Several products are currently being evaluated, but no antiviral drug could demonstrate a significant effect on the survival rate in humans. Most of the products have been tested in animals during the past years (mice or NHP), but not in humans. The West Africa outbreak accelerated the evaluation of some candidate products, which have

been undergoing phase 1 to 3 trials since then [WHO, 2015]. However, these products have experienced low enrollments of patients (due to late initiation during outbreaks), making it difficult to assess their efficacy with enough statistical power and to validate their effect on different populations. The treatment candidates can be divided in two categories [Liu et al., 2017]: small molecule inhibitors (including the antiviral molecule Favipiravir and the small inhibitory RNA TKM-100802) and immune-based therapeutics (including interferons, convalescent plasma, combination of monoclonal antibodies ZMapp). In particular, the antiviral Favipiravir has shown activity against EBOV in mice and tolerance in phase 1 clinical trials on healthy humans and phase 2/3 clinical trials on humans infected with influenza. A clinical efficacy trial was started in Guinea in December 2014 and enrolled 126 patients but showed no efficacy in reducing mortality of patients infected with EVD [Sissoko et al., 2016]. Recent encouraging results from a macaque study suggest the drug efficacy could be improved at higher doses [Guedj et al., 2018]. The combination of monoclonal antibodies ZMapp showed 100% survival in NHP when administered 5 days after virus challenge and safety was evaluated in phase 1 clinical trials. A randomized controlled efficacy study was run in 2015 in Liberia, Sierra Leone and Guinea but showed no efficacy on survival in infected humans [PREVAIL II Writing Group, 2016]. The TKM-100802 inhibitory RNA treatment showed some limited efficacy in NHP and some related side effects were found in phase 1 trials. It was tested in a single arm phase 2 clinical trial in Sierra Leone, showing no efficacy [Dunning et al., 2016b]. The Brincidofovir, a small antiviral molecule, was also tested with no efficacy on 4 patients in 2015 [Dunning et al., 2016a] but trial was incomplete and the product was withdrawn by the company. Overall, treatments against EVD have not yet showed convincing results of efficacy in reducing mortality in infected patients, but some are still under investigation and development.

3.1.2 Ebola vaccine development

3.1.2.1 Clinical state of the art

Similar to therapeutic development, Ebola vaccine research has been accelerated following the West Africa outbreak. A large number of platforms have been considered as vaccine candidates against EVD. All contain a viral component of EBOV, usually the GP and/or NP of the virus, coming from potential different species and strains. Reviews of Ebola vaccine development can be found in Venkatraman et al. [2018], Keshwara et al. [2017], Lambe et al. [2017] and Wang et al. [2017]. We also realized a review and a meta-analysis in the team [Gross et al., 2018]; details will be given in the following section. In addition, we searched for Ebola vaccine in clinicaltrials.gov to determine a list of the

different platforms and their state of development.

Similar to hepatitis A or flu vaccine, the first attempt was the use of an inactivated whole virus, which raised safety concerns, and other types of vaccines were developed. DNA vaccines consist in a circular DNA molecule (plasmid) encoding one or several genes of viral proteins. The first human clinical trial tested in 2003 a DNA vaccine and showed that 3 immunizations were safe and immunogenic [Martin et al., 2006]. It was followed a few years after by a phase 1 [Sarwar et al., 2014] and a phase 1b trial [Kibuuka et al., 2015], showing that multiple doses were needed to sustain the immune response. Several other platforms have also been tested in phase 1 clinical trials, including a vaccine based on human parainfluenza virus 3 (HPIV3) in the USA, human monoclonal antibodies in the USA or an EBOV GP nanoparticle vaccine (with or without adjuvant) in Australia. In parallel, a replication-defective recombinant Ad5 platform (rAd5) was first tested in the USA [Ledgerwood et al., 2010], then in China [Zhu et al., 2015; Li et al., 2017] and Sierra Leone [Zhu et al., 2017]. However, some concern was raised on the pre-existing immunity in the population to Ad5 virus. Research on replication-defective recombinant vectors focused then on serotypes with low human seroprevalence. It included chimpanzee adenovirus (ChAd), and in particular ChAd3 alone [Ledgerwood et al., 2015, 2017; Rampling et al., 2015; De Santis et al., 2016] or combined with MVA [Tapia et al., 2016; Ewer et al., 2016] or with Ad26. The ChAd3 vaccine has been evaluated in phase 2 trials in West Africa. It has also been investigated simultaneously with the VSV platform in a phase 2 trial under the Partnership for Research on Ebola Virus in Liberia (PREVAIL) [Kennedy et al., 2017]. Adenoviruses Ad26 and Ad35 were also considered as potential vectors for Ebola vaccine. After showing its beneficial boosting effects, MVA platform has also been tested in combination with adenoviruses [Tapia et al., 2016; Ewer et al., 2016], and in particular, Ad26/MVA regimens have been undergoing phase 1 to 3 trials. Some of these trials have been realized in the EBOVAC consortium, as part of the Innovative Medicines Initiative Ebola+ program [Eurosurveillance editorial team, 2015], which aims to assess a novel prime-boost preventive vaccine regimen against EVD. This consortium associates academic European partners (University of Oxford, London School of Hygiene and Tropical Medicine, INSERM and INSERM transfert) with the manufacturer of the vaccine (Janssen). In particular, as part of this consortium, we had access to the data of three phase 1 trials realized on healthy adult volunteers in four countries - United Kindgom [Milligan et al., 2016; Winslow et al., 2017], Kenya and Uganda/Tanzania. In these trials, participants were randomized to received either Ad26 then MVA or MVA then Ad26 with a delay of 28 or 56 days. Another arm tested a 15 days delay in the UK only. In addition, phase 2 trials have been conducted under this consortium in Europe (France, UK) and Africa (Burkina Faso, Uganda, Kenya, Ivory Coast), and a phase 2b

trial is ongoing in Sierra Leone. All assess the Ad26 then MVA regimen at 28, 56 or 84 days of delay. In parallel, a recombinant replication-competent platform, using the recombinant vesicular stomatitis virus (rVSV) has been developed [Regules et al., 2017; Agnandji et al., 2016; Huttner et al., 2015; Heppner Jr et al., 2017] and tested during the 2014-2015 epidemic using a ring-vaccination approach: after diagnosis of a new infected patient, all people who were in contact with that case in the previous 21 days were defined as a ring. A total of 117 rings were identified, containing a mean number of 80 people. Randomization was applied to the rings in a 1:1 ratio to receive the vaccine either one day after identification or 21 days after. The outcome was considered to be only cases of Ebola virus disease with an onset 10 or more days from randomization. In that sense, vaccine efficacy was estimated 100% (95% confidence interval 68.9–100.0), but this perfect efficacy was questioned due to differences of interventions in the placebo and treated clusters [Metzger and Vivas-Martínez, 2018]. The ring vaccination was also implemented in DRC during the epidemics of 2018. In the first epidemic, vaccination was realized between May 21 and June 26, 2018 with a total of 3481 people vaccinated, including health professionals, contacts of confirmed EVD cases and contacts of these contacts. In the second epidemic, vaccination started on August 8, 2018. As of August 19, 2018, health care workers were first vaccinated and immunizations were realized in 10 vaccination rings around 28 recently confirmed cases. Additional investigations on potential Ebola vaccines are necessary and still undergoing, especially under the Partnership for Research on Ebola VACcinations (PREVAC), a phase 2b trial on healthy individuals in Guinea, Liberia, Mali and Sierra Leone [Lévy et al., 2018]. In this trial, participants are randomized in a 2:1:2:1:1 ratio to receive either Ad26/MVA or placebo/placebo or rVSV/placebo or rVSV/rVSV or placebo/placebo, all with a 56 days delay. If funding permits, participants will be followed up to 5 years under the PREVAC-UP project. Finally, longer follow-ups will also be realized for some participants of EBOVAC1 and EBOVAC 2, and an additional phase 2 trial will be conducted in Sierra Leone, Guinea and DRC to evaluate the safety and immunogenicity of the Ad26/MVA vaccine candidate in children, under the EBOVAC3 project. A summary of the trials recorded in clinicaltrials.gov and assessing platforms that reached the phase 2 clinical development can be found in table 3.1.

Table 3.1 – Clinical trials assessing Ebola virus vaccine platforms. N indicates the number of patients enrolled in trial (including placebos) if recruitment is completed. If the trial is still recruiting or has not yet started recruitment, N* indicates the estimated targeted number of enrolled participants. Regarding the status, "ongoing: long follow-up" means that all participants were recruited, vaccination was administered and only long-term (more than 1 year) measurements still need to be made. Grey colored trials correspond to EBOVAC1, blue to EBOVAC2, green to EBOVAC3, yellow to PREVAC.

Platform	Immunogen	Phase	Country	Population (N)	Published results/status
rAd5	EBOV GP, SUDV GP	1	USA	Adults (31)	Ledgerwood et al. [2010]
		1	China	Adults (120)	Li et al. [2017]
		2	Sierra Leone	Adults (500)	Zhu et al. [2017]
ChAd3	EBOV GP, SUDV GP	1	USA	Adults (20)	Ledgerwood et al. [2015]
	EBOV GP	1	UK	Adults (60)	Rampling et al. [2015]
		1	Switzerland	Adults (120)	De Santis et al. [2016]
		2	Cameroon, Mali, Nigeria, Senegal	Adults (3013)	Completed
		2	Mali, Senegal	Children (600)	Completed
ChAd3/MVA	ChAd3: EBOV GP (or EBOV, SUDV GP). MVA: EBOV GP	1a	UK	Adults (38)	Completed
		1b	Senegal	Adults (40)	Completed
		1b	Uganda	Adults (90)	Completed
		1b	USA	Adults (143)	Completed
	ChAd3: EBOV GP. MVA: TAFV NP, EBOV GP, SUDV GP	1b	Mali	Adults (91)	Tapia et al. [2016]
		1	UK	Adults (60)	Ewer et al. [2016]
		1	USA	Adults (60*)	Not yet recruiting
ChAd3/Ad26	EBOV GP	1	UK	Adults (32)	Completed
ChAd3 // rVSV	ChAd3: EBOV GP. rVSV: EBOV GP	2	Liberia	Adults (1500)	Kennedy et al. [2017]

		1	USA	Adults (65)	Ongoing: long follow-up
		1	USA	Adults (164)	Completed
		3	USA	Adults (525)	Completed
		3	USA	Adults (329)	Completed
		2	USA, Kenya, Mozambique, Nigeria, Tanzania, Uganda	Adults (578)	Ongoing: long follow-up
Ad26/MVA	Ad26: EBOV GP. MVA: TAFV NP, EBOV GP, SUDV GP	1	UK	Adults (87)	Winslow et al. [2017]
		1	Kenya	Adults (72)	Ongoing: long folllow-up
		1	Uganda, Tanzania	Adults (72)	Ongoing: long follow-up
		2b	Sierra Leone	Adults, children (1019)	Ongoing: long follow-up
		2	UK, France,	Adults (290)	Ongoing: long follow-up
		2	Burkina Faso, Uganda,	Adults: healthy (616), HIV+ (141)	Ongoing: long follow-up
			Kenya, Ivory Coast	Adolescents (129), children (131)	+additional immunization
		2	Guinea, DRC, Sierra Leone	Adolescents & children HIV -/+ (600*)	Not yet recruiting
rVSV	EBOV GP	1	USA	Adults (78)	Regules et al. [2017]
		1	Germany, Switzerland,	Adults (158)	Agnandji et al. [2016]
		1	Gabon, Kenya		
		1	USA	Adults (513)	Heppner Jr et al. [2017]
		1/2	Switzerland	Adults (56)	Huttner et al. [2015]
		2	Canada, Burkina Faso, Senegal	Adults & adolescents HIV+ (200*)	Ongoing: recruitment
		2	USA, Canada	Adults at occupational risk (18*)	Ongoing: recruitment
		3	Guinea	Adults (5643), children (194)	Henao-Restrepo et al. [2017]
Ad26/MVA // rVSV	Ad26: EBOV GP. MVA: TAFV NP, EBOV GP, SUDV GP. rVSV: EBOV GP	2b	Guinea, Liberia, Mali, Sierra Leone	Adults (1400*) Adolescents (466*) Children (934*)	Ongoing: recruitment

3.1.2.2 Immune marker of interest

A prophylactic (preventive) vaccine is assessed on its ability to prevent the infection by a pathogen. Vaccine efficacy is estimated by comparing the incidence of the disease among vaccinated subjects to equally exposed unvaccinated subjects and computing the percentage of reduction of risk between these two populations. However, EVD epidemics are usually localized and induce moderate numbers of fatalities (< 100), making it difficult to evaluate the efficacy of a vaccine candidate during active transmission. A substitute endpoint should then be used; it is a quantity (e.g., a biological marker) that can be measured instead of the clinical endpoint. More precisely, in vaccine development, a surrogate of protection is needed. It has to be statistically associated with the occurrence of the disease and situated on the causal pathway between vaccination and protection [WHO, 2013].

No immune surrogate of protection has been identified so far for the evaluation of Ebola vaccine. Some studies on animals have shown that humoral response and survival were associated. In particular, protection against lethal challenge of EBOV in rVSV vaccinated macaques was found to be correlated with high titers of EBOV-specific IgG [Wong et al., 2012; Marzi et al., 2015], although a large variation of level value was observed, possibly due to differences in assays. Also, heterologous prime-boost regimens combining Ad26 vector with either Ad35 or MVA induced protection in NHP after intramuscular challenge, and survival was particularly associated with high humoral response and less so with cellular response [Callendret et al., 2018]. With rVSV vaccine, neutralizing antibodies were not always detected after immunization, meaning that they may not correlate with protection. In the case of non-replicating, Ad-based vaccines injected to NHP, antibody response was associated to protection against EBOV challenge [Sullivan et al., 2011] but the cellular response may also play a role, especially for long-term protection [Stanley et al., 2014]. Although no surrogate of protection has been identified and bridging studies between NHP models and humans are undergoing but have not been completed yet [Golding et al., 2018; Sullivan et al., 2009], the community has focused on using the antibody response as main criterion for assessing Ebola vaccine candidates in phase 1/2 trials [Krause et al., 2015].

The binding antibody response is measured with the enzyme-linked immunosorbent assay (ELISA). This assay is used to detect the presence of antigens or antibodies in a sample. When used to assess the binding antibody levels, the protocol is the following: Ebola protein samples are fixed in probes on a microplate. The analysis sample is then added to the protein mixture. If the sample contains Ebola-specific antibodies, they will bind to the proteins situated on the microplate. A step of washing is realized, to make

sure that only specific binding relationships are maintained. Secondary antibodies are then added to the plate. These antibodies are labeled with enzymes which are activated with an additional substrate and can be detected using fluorescence technologies. The signal detected is proportional to the number of binding antibodies in the sample. The optical density of the sample is compared to a standard curve, which allows to determine an equivalent concentration of binding antibodies (in ELISA units).

As mentioned in section 2.2.2.3, many factors can have an impact on the quality and quantity of the immune response to vaccine immunizations. In order to assess the factors influencing the antibody response variability after Ebola vaccination, a meta-analysis was conducted in the team. This work, **Ebola vaccine development: Systematic review of pre-clinical and clinical studies, and meta-analysis of determinants of antibody response variability after vaccination**, L. Gross, E. Lhomme, C. Pasin, L. Richert, R. Thiébaut was published in the International Journal of Infectious Diseases (september 2018), volume 74, pp 83-96 [Gross et al., 2018]. The article can be found in appendix A. In short, a review of Ebola vaccine studies was conducted and a meta-regression was estimated on human groups including at least 8 subjects. Among different factors, related to the vaccine (platform, route of administration, insert), to the measurement (delay between immunization and measure, method, antigen, strain, similarity between strain and vaccine insert) or to the population (geographic location, age, sex), only vaccine platforms and viral strains used for antibody detection were associated with antibody response. However most of the heterogeneity of the response remained unexplained (95%), suggesting that other factors could impact the antibody response, such as genetics, as previously mentioned, or the measurement technique for ELISA. It justifies the interest of randomized clinical trials for formal comparisons of vaccine immunogenicity. This study also underlined the opportunity we had to work on data generated in the context of the EBOVAC1 consortium, as the clinical trials were realized with very similar study protocols, reducing the risk of variation induced by factors related to the vaccine and the measurement technique.

3.1.3 Modeling the humoral immune response: state of the art

3.1.3.1 Dynamics of the humoral immune response

As previously shown, antibodies represent the marker of interest when evaluating vaccines against Ebola virus. They actually play a crucial role in preventing many infections and represent a good correlate of protection for a lot of vaccines [Plotkin, 2010]. Antibodies are actors of the humoral response, which also include B cells. After activation, B cells differentiate into either ASCs producers of antibodies, or into memory B cells

with an ability of quick reaction during secondary encounters. It has been observed that ASCs peak a few days (7-10) after encounter with the pathogen, and levels drop after a few weeks [Lanzavecchia and Sallusto, 2009; Fink, 2012; Huang et al., 2014; Frölich et al., 2010; Odendahl et al., 2005]. In the meantime, antibody concentrations also increase until reaching a peak around 20-30 days after encounter, before decreasing with time. Maintenance of antibodies has been studied and can last several years [Amanna et al., 2007; Amanna and Slifka, 2010; Pool et al., 2018]. Yet, antibodies are known to be short-lived, with a half-life depending on the antibody subtype: around 2-3 days for IgDs and IgEs [Abbas et al., 2010], 5-8 days for IgMs and IgAs [Abbas et al., 2010; Brekke and Sandlie, 2003] and estimations vary between 20 and 50 days in different studies of intravenous IgG preparations [Berkman et al., 1990; Brekke and Sandlie, 2003] and passive immunity through maternal antibodies [O'Dempsey et al., 1996; Leuridan et al., 2011; Brinkhof et al., 2013; Vilajeliu et al., 2016; Voysey et al., 2017]. The maintenance of the antibodies is actually explained by the existence of a population of long-lived ASCs which produce and sustain antibodies over time [Slifka et al., 1998; Radbruch et al., 2006; Hammarlund et al., 2017]. In parallel, memory B cells remain in the organism, ready to react faster and to differentiate into ASCs producing antibodies with higher affinity to the antigen [Tarlinton and Good-Jacobson, 2013; Inoue et al., 2018]. Modeling the whole process of the humoral response is very challenging, and most of the models in the literature have focused on modeling the antibody decay, in order to predict the duration of the response and the time at which a threshold of interest is reached (e.g., the value corresponding to a correlate of protection).

3.1.3.2 Models for the antibody decay

One of the first modeling work for the decrease of antibodies was realized in the case of hepatitis B vaccine. A threshold of antibody concentrations under which subjects are at risk of infection was defined in Jilg et al. [1984]. Below this value, an additional immunization was required to boost the immune system and the time of this additional immunization needed to be determined. A simple model was used to quantify the variability of antibody decline among the population [Nommensen et al., 1989] and led the authors to recommend that the duration of protection against HBV should be individually estimated with a second measurement of antibody concentrations. In this paper, authors assumed an exponential decline on the antibodies. It actually comes from the assumption that after the observed peak, antibodies decline with a time-independent decay rate δ. In that case, the differential equation verified by the antibodies (Ab) is the following:

$$\frac{dAb(t)}{dt} = -\delta Ab(t), \tag{1}$$

which can be solved in:

$$Ab(t) = Ab_p e^{-\delta(t-t_p)}, \tag{2}$$

with t_p the time of peak and Ab_p the antibody concentration at t_p. As antibody concentrations are usually transformed using the $\log_{10}$, equation (2) can be written as:

$$\log_{10}(Ab(t)) - \log_{10}(Ab_p) = -\frac{\delta}{\log(10)}(t - t_p), \tag{3}$$

which corresponds to a linear decrease of the $\log_{10}$-transformed antibodies. By using a measurement at t_p and an additional measurement at a time after t_p, parameter δ can be estimated for each subject and the time at which antibodies will reach the given threshold can also be determined.

Basically, assuming an exponential decrease of the antibodies corresponds to using a linear model on the log-transformed antibodies. This linear decay has also been applied to geometric mean titers following Hepatitis A Vaccine injections [Van Damme et al., 1994]. Similar to Nommensen et al. [1989], individual estimates of the decline of antibodies following hepatitis A vaccine immunization were estimated to account for the between-subjects variability [Wiens et al., 1996]. Accounting for this variability by treating individual data instead of using the mean titers over the studied population can actually be addressed by using linear mixed models: these statistical tools allow to evaluate the inter-individual variability through random effects and the possible influence of other factors, such as age, gender or environmental factors on the antibodies dynamics, as detailed in section 2.2.2.3. In that case, if we consider that $t = 0$ corresponds to the time at which antibodies reach their peak, the antibody concentration for individual i at time j can be written:

$$\log_{10}(Ab(t_{ij})) = \beta_0 + \gamma_{0i} + \beta^T Z_i + (\beta_1 + \gamma_{1i})t_{ij} + \beta_s^T Z_{si} t_{ij} + \epsilon_{ij}, \tag{4}$$

with β_0 the population mean value at time of peak, β_1 the population mean value of the decreasing slope. Z are covariates modifying the value at time of peak and β their associated effects, and Z_s are covariates affecting the decreasing slope, with β_S their associated effects. Z and Z_s can be similar and share common variables. Random effects γ on the intercept and the slope are such that:

$$\begin{pmatrix} \gamma_{0i} \\ \gamma_{1i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix} \right), \tag{5}$$

Finally, ϵ is a normally distributed error $\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$. Linear mixed models were used in several studies to predict the persistence of antibodies following vaccinations or natural infections. In particular, they were applied in studies with a follow-up data between 3 and 6 years after the last vaccine immunization to model the decline of antibodies and predict

their persistence up to 10 to 20 years. Applications include hepatitis A vaccine [Bovier et al., 2002], diphteria, tetanus and pertussis vaccine [Cheuvart et al., 2004; Bailleux et al., 2008] and Japanese Encephalitis vaccine [Abe et al., 2007]. Linear mixed models were also used in the case of hepatitis B vaccine [Renard et al., 2001], by including data up to 11 years after the vaccine immunization and predicting the antibodies concentrations at year 12. Another interesting application was proposed by Amanna et al. [2007] to assess the long-term duration of humoral immunity to common antigens.

However, using a linear mixed model to describe the decrease of antibodies after their peak relies on the strong assumption that the decay rate is independent of time and induce a single-slope decrease. Some studies have shown that this assumption is simplistic. In particular, a biphasic decay of the antibodies has been widely observed in experiments, with a strong decrease right after the peak of the antibody response followed by a slower decline [Vidor, 2010; White et al., 2015]. Some early mention of the necessity to model the biphasic decay of the antibody response can be found in Wiedermann et al. [1997], as authors found that classical models could not fit their data and proposed a "two-component model" accounting for the two phases of the dynamic. Several ways to improve the linear modeling have been proposed. One natural method to account for time-changing slope values and still consider the between-subject variability is to use piecewise linear mixed models: it corresponds to using linear mixed models with parameters differing on some time intervals, in particular slope values. Other methods consider non-linear models. Indeed, instead of considering the log-transformed antibody concentrations as a linear function of time, "exponential-type" models assume a power of time. In that case, the modeling equation is of the form:

$$\log(Ab(t)) = c + K(t - t_0)^a. \tag{6}$$

Note that when $a = 1$, we obtain the particular case of a linear model. These exponential-type models were applied in particular in the case of Haemophilus influenzae type b infections, with data up to 4 years and predictions up to 10 years [Leino et al., 2000]. Linear-mixed models, piecewise linear-mixed models and non linear model based on an exponential-type function were compared in some studies: in particular, estimations on data from Japanese Encephalitis vaccine [Desai et al., 2012] and hepatitis A [López et al., 2015; Theeten et al., 2015] showed that piecewise linear-mixed models induce better fit and can help predict the persistence of antibodies after immunization. Another non-linear approach is the "power-law" model [Fraser et al., 2007]: it accounts for the heterogeneity in the rate of decay of several populations of ASCs by using a Gamma distribution, which leads to the following equation for the antibody dynamics:

$$\log(Ab(t)) = k - a\log(c + t). \tag{7}$$

This model was proposed to be extended in the same article by considering two distinct populations of ASCs: one population is active, with an heterogeneous decay rate modeled by a Gamma distribution and a given proportion of the cell population is memory, with a null decay. This model induces a long-term antibody plateau. It is written as:

$$\log(Ab(t)) = k + \log[(1-\pi)^{-a}(c+t) + \pi], \tag{8}$$

with $\pi > 0$ inducing a long-term antibody persistence. Modified power-law model was shown to fit better the data from HPV vaccination compared to the simple power-law model [Fraser et al., 2007]. Still in the HPV context, the data was better fitted by a modified power-law model compared to a power-law model and piecewise linear models [David et al., 2009]. In both cases, models were fitted on follow-up data up to 6 years and used for predictions between 20 and 30 years after vaccine immunization. Power-law and modified power-law models were also fitted on hepatitis E vaccine data up to 6 years after immunization [Chen et al., 2015]. Both models allowed predictions of persistence up to 30 years, but their performances could not be clearly distinguished with the available data. Finally, models of the antibody dynamics using fractional polynomial models were proposed and estimated on hepatitis A vaccine data [Hens et al., 2014]: a piecewise linear-mixed model with 3 slopes was found to fit the best the data up to 17 years and allowed prediction of persistence up to 25 years.

Overall these models represent a good tool to predict the long-term persistence of antibodies after vaccination, when data is already available up to a few years after the immunization. However, they only model the decreasing phase of the antibody dynamics and do not account for the previous temporal phase or the dynamics of ASCs. Moreover, the biological interpretation of complex descriptive statistical models is difficult, as mentioned in section 2.3. A good alternative to model the dynamics of the humoral response, accounting for both ASCs and antibodies dynamics, is to use mechanistic, ODE-based models.

3.1.3.3 Mechanistic models for the humoral response to vaccine

Mechanistic models are based on the knowledge from the biological process, and are able to account for the non-linear interactions between the actors of the process. They can be based on ODEs from which parameters can be estimated to quantify the characteristics of the process. In the case of the humoral immune response following vaccination, only a few ODE-based models were used. In Wilson et al. [2007], the dynamics of circulating antigen, immunological memory and antibody concentrations after hepatitis B vaccine immunization were modeled with ODEs to quantify the mechanisms of the immune response. The interest is focused on immunological memory, as defined by the ability to

produce circulating antibody. Protection from infection is driven by this immunological memory, as some patients with low anti-HBVs antibody concentrations can also be protected, thanks to reactivation of their immunological memory. Immunological memory does not correspond to a given population of cells but more to the expansion capacity of memory B and T lymphocytes, changes in cytokine production, and affinity phenomena. Vaccine antigen (V) is supposed to decrease at constant rate σ. Memory (M) is assumed to be generated at both antigen-dependent and independent rates, with a limited capacity N. Memory has the capacity to produce antibodies (A) in presence of vaccine antigen at rate δ, with a maximum level N. Antibodies decrease at a time-dependent rate μ/T, with T the time since last vaccination. The equations of the dynamics following a booster immunization are the following:

$$\begin{cases} \dfrac{dV}{dt} &= -\sigma V, \\ \dfrac{dM}{dt} &= (\gamma V + \beta M)\Big(1 - \dfrac{M}{N}\Big), \\ \dfrac{dA}{dt} &= \delta M V\Big(1 - \dfrac{A}{N}\Big) - \dfrac{\mu A}{T}. \end{cases} \tag{9}$$

The model was used to fit antibodies concentration data from several clinical trials assessing the effect of different vaccines (three generations of vaccines have been used up to now). The estimation showed that quantity of memory and the time before its generation significantly varied between the different vaccines. As the amount of memory (and not the antibody concentrations) is supposed to generate protection, predictions of the model supported the hypothesis that a single dose of vaccination could be sufficient to induce protective immunological memory.

Another study by Le et al. [2015] modeled the dynamics of ASCs after vaccinia virus (VV) vaccination in human volunteers with ODEs. The model was actually an extension of a widely known model for the $CD8^+$ T cell response, developed in de Boer et al. [2001] and Antia et al. [2003], as mentioned in 2.3. After VV vaccination, it takes some time T_{on} until ASCs (written B in the article) are activated and start proliferating. Then, after a few days, a contraction phase arises and most of the ASCs die at rate δ_B; it happens during some given period of time. Some memory cells remain in the organism and die with a much lower rate δ_M. The equation of the dynamics can be written as:

$$\frac{dB(t)}{dt} = \begin{cases} 0 & \text{if } t < T_{on}, \\ \rho B(t) & \text{if } T_{on} \leq t < T_{off}, \\ -\delta_B B(t) & \text{if } T_{off} \leq t < T_{mem}, \\ -\delta_M B(t) & \text{if } T_{mem} < t. \end{cases} \tag{10}$$

Antibodies (Ab) are produced by ASCs cells at rate $\theta(t)$ and die at rate δ_{Ab}:

$$\frac{dA(t)}{dt} = \theta(t)B(t) - \delta_{Ab}Ab(t). \tag{11}$$

A first model considered a rate θ constant with time and another rate linearly increasing with time after T_{on}. Finally, an improvement of this model was also proposed in the same paper, in order to account for the hypothesis that antibodies are produced by several populations of ASCs. Authors included in the model circulating ASCs, that were able to migrate to the bone marrow and differentiate at rate m into long-lived plasma cells. Antibodies are then produced at constant rate θ by both circulating cells and cells from the bone marrow. A patient-by-patient estimation method allowed good fit of the data by the models. Unfortunately, due to the few number of data measurements and the increased number of parameters in the alternative models, statistical differences between the three models could not be assessed. The models helped quantifying the kinetics of the immune response following vaccination, but could not help distinguishing the main mechanisms involved in the production of antibodies by ASCs.

Finally, in Andraud et al. [2012], authors aimed at applying the "plasma-cell imprinted lifespan" model developed by Amanna and Slifka [2010]: it assumes that antibodies are maintained by a population of long-lived plasma cells, located in survival niches and with an "imprinted" lifespan, independent from the replenishment from memory B cells (due to boost immunizations of possible re infections). In the model, the plasma cells are then divided in two populations with two different lifespans (one short and one long). These populations are assumed not to be renewed and to decline with time at two different decay rates. Model equations are then:

$$\begin{cases} \dfrac{dP_s}{dt} &= -\mu_s P_s, \\ \dfrac{dP_l}{dt} &= -\mu_l P_l, \\ \dfrac{dAb}{dt} &= \phi_s P_s + \phi_l P_l - \mu_{Ab} Ab. \end{cases} \tag{12}$$

The system has an analytic solution, as plasma cells admit an exponential decline. By writing $\Phi_s = \phi_s P_{s0}$ and $\Phi_l = \phi_l P_{l0}$, we obtain:

$$Ab(t) = \frac{\Phi_s}{\mu_{Ab} - \mu_s} e^{-\mu_s t} + \frac{\Phi_l}{\mu_{Ab} - \mu_l} e^{-\mu_l t} + \left(A_0 - \frac{\Phi_s}{\mu_{Ab} - \mu_s} - \frac{\Phi_l}{\mu_{Ab} - \mu_l} \right) e^{-\mu_{Ab} t}. \tag{13}$$

Additional hypothesis can simplify the model; especially, a model where $\mu_l = 0$ was tested, as long-lived cells are expected to have a very long lifespan. Moreover, a model assuming that antibodies die much faster than ASCs was also tested (with $\mu_{Ab} \gg \mu_s$ and μ_l). These models were used to fit data of antibody concentrations from 1 to 10 years after

hepatitis A vaccine immunizations and compared with power-law decay models. They were estimated with a population approach and allowed both individual and population prediction of immunity waning. The asymptotic model, assuming that $\mu_l = 0$ was found to be the best for fitting the data. However, additional measurement of the early antibody response (before one year) were found to help improving parameters estimation, and in particular those related with the long-lived ASCs. It could mean that early measurements would help better distinguishing between the role of short-lived and long-lived ASCs. This modeling work succeeded in estimating three scales of the humoral response dynamics, corresponding to the lifespan of antibodies (around 20-30 days), short-lived plasma cells (several months) and long-lived ones (decades).

There has been a large interest in modeling the humoral immune response to vaccine immunizations, but only a few studies accounted for the dynamics of the ASCs in addition to the antibodies. This is also related to the availability of the data, as most of the time only the antibody concentrations are measured. Moreover, we have seen that both individual and population fitting are used to estimate the key parameters of the dynamics of the response. In our work, we have focused on using a population approach to estimate the parameters of an ODE-based mechanistic model for the dynamics of the humoral response to Ebola vaccine. We will detail in the next section the method for the estimation.

3.2 Method: parameters estimation

3.2.1 The population approach

A first way to use ODE-based mechanistic models is to simulate trajectories with a certain range of parameters values to assess the ability of the model to reproduce qualitatively the dynamics of the biological process. These simulations also allow to get an idea of how parameters may impact the trajectories of the compartments of the system. Parameters of the model can also be chosen by calibration: it consists in determining some well-chosen combinations of the parameters that allow to reproduce quantitatively the clinical data. This is done by comparing the predicted compartment trajectories to the data. Both approaches can be valuable for running simulations and comparing several models related to different biological hypotheses. They actually represent a first step of study when developing a model. However, ODE systems can almost never be solved with an analytic solution and the output is usually non-linear in the parameters. Simulations with a given number of parameter combinations does not allow to explore the whole space of possible solutions and do not ensure to determine the optimal values corresponding to the clinical data. One way to deal with this issue is to estimate the parameters of the

ODEs. However, due to the complexity of the models and the clinical constraints on the data, estimation of parameters is not that easy and represents a real statistical challenge. Patients-by-patients fits have been widely used [Ribeiro et al., 2002; Dixit and Perelson, 2005; Le et al., 2015]: parameters are determined for each patient, by minimizing a given criterion – usually the mean squared error on the patient's data and population parameters are generally obtained by computing empirical means of the patients' estimated values. Our approach is based on population estimation, for several reasons. First, data can be unbalanced in biological experiments. It means that measurement timings and availability can differ between subjects. This additional constraint can make it difficult to estimate parameters of some individual with only a few measurement. Using the whole population data can handle that issue [Thiébaut and Walker, 2008]. Another argument for the statistical power of the approach relies on a simple analogy with statistical regression models. Indeed, if several groups of individuals can be defined based on a covariate (e.g., immune intervention), it is possible to stratify the estimation within each intervention group, but more powerful to use the data on all subjects and include the covariate in the model. Finally, the variability between patients adds information to the model; for example, it can help constraining the parameters value of the model. Indeed, if a model is considered to be acceptable for both a placebo and an immune intervention group, with the intervention affecting only the value of some parameters, the estimation on the placebo dynamics will help constraining the values of other parameters that are supposed to be shared between the two groups. The population approach can also help determining which parameters induce more or less variability between subjects behaviors. Here, linear-mixed models are applied to the parameters of the ODE system; covariate effects can be estimated on these parameters, and the unexplained inter-individual variability is modeled with random effects.

3.2.2 Estimation with NIMROD

Parameter estimation of ODE systems was mostly realized using NIMROD (normal approximation inference in models with random effects based on ordinary differential equations) tool, which was previously developed in the team [Prague et al., 2013a]. This method of estimation was implemented in Fortran. In this section, we present the general model of ODEs handled by NIMROD and the statistical method of estimation. It relies on likelihood maximization using a Newton-like algorithm which approximates the Hessian using first derivatives. Acceptable computation times are achieved thanks to parallel computation.

3.2.2.1 General model

The general model is based on three layers. First, we consider the mathematical model, which consists of a system of ODEs modeling the biological compartments (e.g. population of cells), represented by the vector $\mathbf{X}(t) = (X_1(t), \ldots, X_K(t))$. In a given population of n independent subjects, the ODE system for subject i is written:

$$\begin{cases} \dfrac{d\mathbf{X}^i(t)}{dt} &= f(\mathbf{X}^i(t), \boldsymbol{\xi}^i(t)), \\ \mathbf{X}^i(0) &= h(\boldsymbol{\xi}^i(0)). \end{cases} \tag{14}$$

$\boldsymbol{\xi}$ is the vector of n_b biological parameters. The parameters can be transformed to account for biological constraints (e.g., logarithmic transformation if parameters correspond to rates and are positive). We use one-to-one transformation functions ψ_l, $l = 1 \ldots n_b$:

$$\tilde{\xi}_l^i(t) = \psi_l(\xi_l^i(t)). \tag{15}$$

From now, we will consider that $\psi = \log$. The population approach includes statistical linear-mixed models on the biological parameters. This allows to introduce covariates on parameters and to account for the between-subjects variability. For every parameter ξ_l and patient i, we write:

$$\tilde{\xi}_l^i(t) = \tilde{\xi}_{l_0} + \boldsymbol{\beta}_l^T \mathbf{z}_l^i(t) + u_l^i, \tag{16}$$

where $\tilde{\xi}_{l_0}$ is the intercept and represents the mean value of parameter over the population, $\boldsymbol{\beta}_l$ is a vector of regression coefficients, $\mathbf{z}_l^i$ is a vector of n_e explanatory variables and u_l^i is an individual random effect, following a centered normal distribution with variance ω_l^2. Random effects are independent and applied on a subset of q biological parameters. The last aspect of the model is the observation model, as in practice not all compartments of the model are directly observed and we only have access to discrete time observations $\mathbf{Y}^i(t_{ij})$ of some function of $\mathbf{X}^i(t)$. We suppose there are known link functions g_m, $m = 1, \ldots, M$ allowing an additive measurement error model:

$$Y_m^i(t_{ij}) = g_m(\mathbf{X}^i(t_{ij})) + \epsilon_{ijm}, \tag{17}$$

with $\epsilon_{ijm} \sim \mathcal{N}(0, \sigma_m^2)$. The estimation problem corresponds to the determination of all parameters: intercepts, regression coefficients, variance of random effects and variance of measurement errors. It corresponds to the vector of parameters θ such that:

$$\theta = \left[(\xi_{l_0})_{l=1..n_b}, (\beta_l)_{l=1..n_e}, (\omega_l)_{l=1..q}, (\sigma_l)_{l=1..M}\right]. \tag{18}$$

3.2.2.2 Likelihood and scores computation

For each individual i at time j and each link function m, we assume that the error follows a normal distribution: $\epsilon_{ijm} \sim \mathcal{N}(0, \sigma_m^2)$. The individual likelihood of subject i

given the random effects can be computed as in Guedj et al. [2007a]:

$$\mathcal{L}_{F_i|u^i} = \prod_{\substack{m=1..M \\ j=1..n_i}} \frac{1}{\sigma_m\sqrt{2\pi}} \exp\left[\frac{1}{2}\left(\frac{Y_m^i(t_{ij}) - g_m(\mathbf{X}^i(t_{ijm}, \tilde{\boldsymbol{\xi}}^i))}{\sigma_m}\right)^2\right]. \tag{19}$$

The observed individual likelihood is obtained by integration over the random effects:

$$\mathcal{L}_{\mathcal{O}^i} = \int_{\mathbb{R}^q} \mathcal{L}_{F_i|\mathbf{u}^i}(\mathbf{u})\Phi(\mathbf{u})\,\mathrm{d}\mathbf{u}, \tag{20}$$

where Φ is the normal density of $\mathcal{N}(0, I_q)$. Then the full (given random effects) and observed individual log likelihood are : $L_{F_i|u^i} = \log \mathcal{L}_{F_i|u^i}$ and $L_{\mathcal{O}^i} = \log \mathcal{L}_{\mathcal{O}^i}$. Finally, the global observed log likelihood is obtained by summing all individual contributions:

$$L_{\mathcal{O}} = \sum_{i=1}^{n} L_{\mathcal{O}^i}. \tag{21}$$

The likelihood is computed thanks to adaptative gaussian quadrature [Guedj et al., 2007a]. Indeed, it has been shown that this method may be more precise than others, and especially Laplacian methods [Lesaffre and Spiessens, 2001]. The likelihood is then maximized through a Newton-like algorithm using only the first derivatives of the log likelihood (scores). The observed individual scores are deduced by Louis' formula [Louis, 1982] :

$$U_{\mathcal{O}_i} = \frac{\partial L_{\mathcal{O}_i}}{\partial \theta} = (\mathcal{L}_{\mathcal{O}_i})^{-1} \int_{\mathbb{R}^q} \mathcal{L}_{F_i|u^i}(u) U_{F_i|u^i}(u)\Phi(u)\,\mathrm{d}u. \tag{22}$$

These scores have analytic expressions. As an example, we will present here the formula for the fixed effects ϕ and refer to Guedj et al. [2007a] for the other parameters. The score is computed given the random effect of subject i:

$$U^{\phi_l}_{F_i|u^i}(\theta) = \frac{\partial L_{F_i|u^i}}{\partial \tilde{\xi}_l^i} = \sum_{m,j} \frac{1}{\sigma_m^2} \frac{\partial g_m(\mathbf{X}^i(t_{ijm}, \tilde{\boldsymbol{\xi}}^i))}{\partial \tilde{\xi}_l^i} \left[Y_{ijm} - g_m(\mathbf{X}(t_{ijm}, \tilde{\boldsymbol{\xi}}^i)\right]. \tag{23}$$

This computation requires the determination of the sensitivity equations of the ODE system, $\dfrac{\partial X_k(t, \tilde{\boldsymbol{\xi}}^i)}{\partial \tilde{\xi}_l^i}$, as:

$$\frac{\partial g_m(\mathbf{X}^i(t, \tilde{\boldsymbol{\xi}}^i))}{\partial \tilde{\xi}_l^i} = \sum_{k \le K} \frac{\partial g_m(\mathbf{X}(t_i, \tilde{\boldsymbol{\xi}}^i))}{\partial X_k} \frac{\partial X_k(t, \tilde{\boldsymbol{\xi}}^i)}{\partial \tilde{\xi}_l^i}. \tag{24}$$

The observed individual scores are obtained by numerical integration, using an adaptive Gaussian quadrature, similar to the computation of the individual log likelihood.

3.2.2.3 Newton-like algorithm for likelihood maximization

The maximization of the likelihood is usually realized using a Newton method. The method implemented in NIMROD is an extension of the Newton-Raphson algorithm using

an approximation of the Hessian matrix based on the scores only. Consider a model $(P_{\boldsymbol{\theta}})_{\boldsymbol{\theta}\in\Theta}$ with Θ an open subset of $\mathbb{R}^m$. In a population of n independent subjects, $L_{\mathcal{O}^i} = \log \mathcal{L}_{\mathcal{O}^i}$ is the observed individual log likelihood for subject i and $L^{\boldsymbol{\theta}} = \sum_{i=1}^n L_i^{\boldsymbol{\theta}}$ is the global observed log likelihood. The maximum likelihood estimator is defined by :

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}\in\Theta} L^{\boldsymbol{\theta}}. \tag{25}$$

The method requires the computation of the score and the Hessian matrix (assuming that the log likelihood is twice differentiable):

$$\mathbf{U}_i^{\boldsymbol{\theta}} = \frac{\partial L_i^{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}}, \tag{26}$$

$$\mathbf{U}^{\boldsymbol{\theta}} = \sum_{i=1}^n U_i^{\boldsymbol{\theta}}, \tag{27}$$

$$H(\boldsymbol{\theta}) = \frac{\partial^2 L_i^{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}^2}. \tag{28}$$

The information matrix is

$$I(\boldsymbol{\theta}) = \mathbb{E}_\theta[-H(\boldsymbol{\theta})] = \mathbb{E}_{\boldsymbol{\theta}}[-\mathbf{U}^{\boldsymbol{\theta}}(\mathbf{U}^{\boldsymbol{\theta}})^T]. \tag{29}$$

Assume that the $L_i^{\boldsymbol{\theta}}$ are iid, then we can write $I(\boldsymbol{\theta}) = n\mathcal{I}(\boldsymbol{\theta})$, with $\mathcal{I}(\boldsymbol{\theta})$ independent from n. Moreover, assuming that $\exists \boldsymbol{\theta}_* \in \Theta$ such that $P_{\boldsymbol{\theta}_*} = P_*$, the maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ verifies the following:

$$n^{-1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_*) \to \mathcal{N}(0, \mathcal{I}(\boldsymbol{\theta}_*)^{-1}). \tag{30}$$

The likelihood maximization can be realized with the Newton-Raphson iterative algorithm, defined by the following iteration:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - H^{-1}(\boldsymbol{\theta}_k)\mathbf{U}(\boldsymbol{\theta}_k). \tag{31}$$

The idea of the algorithm comes from Taylor's formula. We will explain the principle for a function $f : \mathbb{R} \to \mathbb{R}$ twice differentiable. For x and $x_0 \in \mathbb{R}$, we can write:

$$f'(x) = f'(x_0) + (x - x_0)f''(x_0) + O((x - x_0)^2). \tag{32}$$

As the aim is to maximize f, we are looking for x such that $f'(x) = 0$. It comes from equation (32) that if x and x_0 are close enough:

$$x \approx x_0 - (f''(x_0))^{-1} f'(x_0). \tag{33}$$

The iterative sequence such that:

$$x_{k+1} = x_k - (f''(x_k))^{-1} f'(x_k) \tag{34}$$

Algorithm	Value of G
Marquart	$G = H + \lambda diag(H)$
Levenberg	$G = H + \lambda Id$
RVS algorithm	$G(\boldsymbol{\theta}_k) = \sum_{i \leq n} \mathbf{U}_i(\boldsymbol{\theta}_k)\mathbf{U}_i^T(\boldsymbol{\theta}_k) - n^{-1}\mathbf{U}(\boldsymbol{\theta}_k)\mathbf{U}^T(\boldsymbol{\theta}_k)$

Table 3.2 – Different types of Newton-like algorithms

converges to x such that $f(x)$ is a maximum of f. A classical Newton-Raphson algorithm, as expressed in (31), requires an inversion of matrix H. Some extensions of the algorithm have been proposed to handle more general cases where H is not positive definite. The iterative step is of the form:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - G^{-1}(\boldsymbol{\theta}_k)\mathbf{U}(\boldsymbol{\theta}_k), \tag{35}$$

where $G(\boldsymbol{\theta})$ has to be positive definite for all $\boldsymbol{\theta}$ and close to $H(\boldsymbol{\theta}_k)$ when $\boldsymbol{\theta}_k \to \hat{\boldsymbol{\theta}}$. As shown in table 3.2, Marquart and Levenberg algorithms [Marquardt, 1963] propose to use matrix H with modifications of its diagonal. The Robust-variance scoring (RVS) algorithm, developed in Commenges et al. [2006], is based on an expression of G from Berndt et al. [1974]. This algorithm was shown to have a good convergence rate and to run faster that the Marquardt algorithm, as it only requires the computation of the scores, which is less computationally demanding than computing the Hessian matrix. This advantage is even more interesting when the number of parameters is large. Convergence of the RVS algorithm is ensured by three criteria. A stopping criterion is a condition of the form $C_k < c$, computed at each iteration of the algorithm, with c a given stopping value. The first two criteria are based on the evolution of the algorithm on the parameter space and on the log likelihood value (with d a distance defined on both spaces):

$$d(\boldsymbol{\theta}_{k+1}, \boldsymbol{\theta}_k) \leq \eta_1 \tag{36}$$

and

$$d(L(\boldsymbol{\theta_{k+1}}), L(\boldsymbol{\theta_k})) \leq \eta_2. \tag{37}$$

However, both criteria could be small even when the algorithm is fixed on a wrong direction (close to local maxima for example). A criterion was proposed in Commenges et al. [2006], with a stopping value independent from the problem of estimation. It corresponds to the ratio of the numerical approximation error $d(\boldsymbol{\theta}_k, \hat{\boldsymbol{\theta}})$ and the statistical error $\mathbb{E}_{\boldsymbol{\theta}_*}[d(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}_*)]$:

$$C_k = \frac{d(\boldsymbol{\theta}_k, \hat{\boldsymbol{\theta}})}{\mathbb{E}_{\boldsymbol{\theta}_*}[d(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}_*)]}. \tag{38}$$

By using a distance of the form $d(x, y) = (x - y)^T M(x - y)$ and taking $M = G$, it can be shown that:

$$C_k \simeq m^{-1} U(\boldsymbol{\theta}_k)^T G(\boldsymbol{\theta}_k)^{-1} U(\boldsymbol{\theta}_k). \tag{39}$$

In practice, we use $C_k < c$, with $c < 1$ and as close as possible to 0 (taken if possible equal to 0.1).

3.2.2.4 Prior distributions

Depending on the number and schedule of observations available, the number of subjects in the study, the measurement precision and the observed components of the model, estimation identifiability issues can arise [Guedj et al., 2007b]. One method to deal with this problem is to reduce the number of parameters to be estimated by fixing some of them, using values from previous studies. Another more flexible way to improve the accuracy obtained on the estimation of parameters relies on bayesian approaches, which help constraining the space of exploration of the algorithm. Bayesian inference is usually implemented with Markov chain Monte-Carlo (MCMC) method, as in the Stan software [Gelman et al., 2015]. However, these methods can be time-consuming in complex models. An alternative was implemented in NIMROD, by using a normal approximation of the posterior distribution and estimating the Maximum A Posteriori (MAP) [Drylewicz et al., 2012]. This method is easy to include in the model, with a faster computation time than MCMC. This approach is justified by the fact that the posterior distribution can be approximated. In particular, Bernstein-Von Mises theorem [Van der Vaart, 2000] states than under weak conditions, the posterior distribution converges to a normal distribution. Parameter $\boldsymbol{\theta}$ is initially defined by a probability density $\pi(\boldsymbol{\theta})$. The posterior distribution is obtained by accounting for the observations: $\pi(\boldsymbol{\theta}|Y)$. As shown in Drylewicz et al. [2012], applying Bayes formula gives:

$$\hat{\boldsymbol{\theta}}_{MAP} = \arg\max_{\boldsymbol{\theta} \in \Theta} L^{\boldsymbol{\theta}}_{MAP}, \tag{40}$$

with

$$L^{\boldsymbol{\theta}}_{MAP} = L(\boldsymbol{\theta}) + \log(\pi(\boldsymbol{\theta})). \tag{41}$$

Numerically, it corresponds to maximizing the penalized log-likelihood:

$$L^{(P)}(\boldsymbol{\theta}) = L(\boldsymbol{\theta}) - J(\boldsymbol{\theta}), \tag{42}$$

with $J(\boldsymbol{\theta}) = -\log(\pi(\boldsymbol{\theta}))$. The same algorithm (RVS) can be used for likelihood maximization, but with a modified function G such that:

$$G(\boldsymbol{\theta}_k) = \sum_{i=1}^{n} U_i(\boldsymbol{\theta}_k) U_i^T(\boldsymbol{\theta}_k) - n^{-1} U(\boldsymbol{\theta}_k) U^T(\boldsymbol{\theta}_k) - \frac{\partial^2 J}{\partial \boldsymbol{\theta}^2}(\boldsymbol{\theta}_k), \tag{43}$$

and $G(\boldsymbol{\theta}_k) \simeq \dfrac{\partial^2 L^{(P)}}{\partial \boldsymbol{\theta}^2}(\boldsymbol{\theta}_k)$ near the maximum [Prague et al., 2012]. In practice, J is easy to compute as only normal priors are used on the biological parameters $\boldsymbol{\phi}$. In that case:

$$J(\boldsymbol{\theta}) = \sum_j \frac{(\phi_j - \mathbb{E}(\phi_j))^2}{2var(\phi_j)}. \tag{44}$$

It allows to determine an approximately normal posterior distribution on the biological parameters.

3.2.2.5 Individual parameters

In addition to providing population estimation of the parameters, the method also allows the computation of individual parameters. Using some individual data, parameters of a given patient can be computed using parametric empirical bayes estimators (PEB) [Morris, 1983; Kass and Steffey, 1989]. The individual estimator can be written as:

$$\hat{\boldsymbol{\xi}}^i_{|F^i_j} = \hat{\boldsymbol{\xi}}_0 + \hat{\boldsymbol{\beta}}^T \boldsymbol{z}(t)^i + \hat{\boldsymbol{u}}^i_{|F^i_j}, \tag{45}$$

where $\hat{\boldsymbol{\xi}}_0$ and $\hat{\boldsymbol{\beta}}$ are the MAP values obtained from the estimation on a dataset excluding observations 1 to j of patient i (written as F^i_j). Estimation of the individual random effects $\hat{\boldsymbol{u}}^i_{|F^i_j}$ can be computed by maximizing the individual likelihood of patient i, based on observations $\boldsymbol{Y}_{i1}, .., \boldsymbol{Y}_{ij}$:

$$\hat{\boldsymbol{u}}^i_{|F^i_j} = \underset{\boldsymbol{u}\in\mathbb{R}^q}{\arg\max} \left\{ \log[p(\boldsymbol{Y}_{i1}, ..\boldsymbol{Y}_{ik}|\hat{\boldsymbol{\theta}}, \boldsymbol{z}^i, \boldsymbol{u})] - J(\hat{\boldsymbol{\theta}}, \boldsymbol{u}) \right\}, \tag{46}$$

with $\hat{\boldsymbol{\theta}}$ the MAP value. Similar to the likelihood maximization, computation of individual random effects can be realized with different methods, but the one implemented in NIMROD is a Newton-like algorithm. This allows to compute individual trajectories from the population estimation.

3.2.2.6 Model selection

In practice, a number of statistical models are evaluated to fit the data, by testing different combinations of random and fixed effects on several parameters. Models are compared and selected using several criteria: first, quality of fit, which is a visual checking that the trajectories correspond to the expected dynamics, regarding the data. For example, the Visual Predictive Check (VPC) is a tool comparing the percentiles of the real data to the percentiles of data simulated with the estimated parameters [Post et al., 2008]. Moreover, the likelihood is used to compare models (a higher likelihood is expected for better models). We also use an approximation of the leave-one-out cross-validation criterion (LCVa), developed in Commenges et al. [2007], which is an extension of Akaike

criterion (AIC), accounting for the number of parameters but also for the penalization and the number of observations [Commenges et al., 2008, 2015]. The LCVa estimates a risk, so better models are expected to obtain lower values of LCVa. It is defined as :

$$LCVa = -n^{-1}\big[L(\hat{\theta}) - Tr(H_{L^P}^{-1}(\hat{\theta})H_L(\hat{\theta}))\big], \tag{47}$$

where H_{L^P} and H_L corresponds to the Hessians of minus the penalized log-likelihood and the log-likelihood respectively.

Several methods of model selection can be used. Our method of selection can be considered as a backward stepwise approach. The model is first estimated by considering random effects on all parameters. Then, for each random effect, we can test if the variance is significantly different from 0, using a Wald test. Non significant random effects can be removed from the model, one by one, by starting by the less significant one and estimating again the model at each removal. The same method can be applied to the selection of fixed effects (covariates) on the parameters. Significance of the fixed effects is also tested using a Wald test and by checking that the variance of random effect is usually reduced. We can process with the same removing method with the covariates. Each time a parameter is removed, the LCVa and likelihood of the models with and without the parameter are checked to make sure the parameter can actually be removed without reducing the performance of the model. In the end, we obtain a model with selected random and fixed effects on the parameters. However, it should be underlined that it can be computationally difficult to estimate a model including random effects on all parameters, as the likelihood is integrated on each random effect and the integration dramatically increases the computational time. It can then be more convenient to evaluate combinations of random effects on several subsets of parameters to determine the best fitting one. Also, it should be kept in mind that the method of model selection should be adapted to the modeling question. Indeed, the method described here corresponds more to an explanatory approach and a statistical selection. As we are using mechanistic models, some mechanisms can already be known from previous work and some of the effects (either random and/or fixed) can be applied to the model because it makes sense in a biological point of view.

3.2.3 Estimation with other tools

In this work, parameters estimation was mainly realized using NIMROD, as described in the previous sections. However, other algorithms proposed different approaches for likelihood maximization. In particular, the Stochastic Approximation Expectation-Maximization (SAEM) algorithm is widely used, especially in the pharmacokinetics/pharmacodynamics (PK/PD) field, as implemented in NONMEM [Beal et al., 1992] and

MONOLIX [Kuhn and Lavielle, 2005]. This algorithm is a stochastic version of the EM algorithm developed by Dempster et al. [1977]: it is an iterative procedure, and each iteration is composed of two steps. The first one simulate the random effects with the conditional distribution using a MCMC procedure and the second one updates the values of the parameters of the model. This method was shown to converge to the maximum likelihood estimate and MONOLIX is used in a wide range of non linear mixed models estimated with a population approach, in particular in the PK/PD field [Lavielle and Mentré, 2007; Chan et al., 2011]. This software can also be used to perform statistical tests used in model selection such as Wald test and likelihood ratio test [Samson et al., 2007]. However, there is a risk of convergence to local maxima, because the number of iterations is limited and the parameters distance between two iteration steps decreases. As the convergence is ensured with more criteria in NIMROD, we have observed in our general use of these softwares that convergence was more robust in NIMROD in the context of our applications. Yet, it should be underlined here that MONOLIX has been used punctually during my PhD, to explore the space of estimated values, choose well initial values to be tested in NIMROD, or to compare estimations obtained with NIMROD.

3.3 Application of the mechanistic modeling to Ebola vaccine trial data: "Dynamics of the humoral immune response to a prime-boost Ebola vaccine: quantification and sources of variation"

The following paper is under preparation for submission.

Dynamics of the humoral immune response to a prime-boost Ebola vaccine: quantification and sources of variation

Chloé Pasin,[a,b] Irene Balelli,[a,b] Thierry Van Effelterre,[c] Viki Bockstal,[c] Laura Solforosi,[c] Mélanie Prague,[a,b] Macaya Douoguih,[c] Rodolphe Thiébaut,[a,b#] with the EBOVAC1 consortium

[a] INSERM U1219, Univ. Bordeaux, Bordeaux, France ;
[b] INRIA SISTM team, Talence, France ;
[c] Janssen

ABSTRACT The Ebola vaccine based on Ad26.ZEBOV/MVA-BN-Filo prime-boost immunizations is being evaluated in multiple clinical trials. The long-term immune response to the vaccine is unknown, including factors associated with the response and variability around the response. We analyzed data from three phase I trials performed by the EBOVAC1 consortium in four countries - UK, Kenya, Tanzania and Uganda. Participants were randomized in four groups according to the interval between prime and boost immunization (28 or 56 days) and the sequence in which Ad26.ZEBOV and MVA-BN-Filo were administered. Consecutive ELISA measurements of the IgG binding antibody concentrations against the Kikwit glycoprotein (GP) were available in 177 participants to assess the humoral immune response up to 1 year post prime. Using a mathematical model for the dynamics of the humoral response, from 7 days after the boost immunization up to 1 year after the prime immunization, we estimated the durability of the antibody response and the influence of different factors on the dynamics of the humoral response. Ordinary differential equations (ODEs) described the dynamics of antibody response and two populations of antibody-secreting cells (ASC), short-lived (SL) and long-lived (LL). Parameters of the ODEs were estimated using a population approach. It has been estimated that half of the LL ASCs could persist at least five years. The vaccine regimen significantly affected the SL ASCs and the antibody peak but not the long-term response. The LL ASCs compartment dynamics differed significantly by geographic regions analyzed, with a higher long-term antibody persistence in European subjects. These differences could not be explained by the observed differences in cellular immune response.

IMPORTANCE The Ebola virus disease epidemic of 2014-2016 has caused 11 310 deaths in Guinea, Sierra Leone and Liberia. As prevention measures did not completely control the disease and no treatment is available yet, research has focused on accelerating the development of preventive vaccines. Combining different vector-based vaccines in prime-boost immunizations could induce a long-term response, assessed through binding antibody concentrations to Ebola virus GP. Using data from phase I trials in East African and European subjects, the dynamics of the humoral immune response following a boost immunization were modeled. We found that half of the LL plasma cells, which are responsible for the persistence of antibodies, could sustain at least five years after the boost immunization. Moreover, the vaccine regimens affect only the early humoral immune response after the boost immunization, while long-term antibody concentrations are significantly higher in European subjects than in East African ones.

KEYWORDS: Ebola, vaccine, mechanistic modeling, antibody response

Following the recent outbreak of Ebola virus disease (EVD) in West Africa that caused 28 616 cases and 11 310 fatalities (1), the clinical development of several Ebola vaccine candidates has been accelerated. Among the vaccine candidates, a heterologous prime-boost strategy combining immunizations with adenovirus type 26 (Ad26.ZEBOV, Janssen Vaccines and Prevention) and modified vaccinia Ankara (MVA-BN-Filo, Bavarian Nordic®) is being developed by Janssen. Prime-boost regimens are expected to be more immunogenic than prime-only vacci-

nation strategies (2) (3) (4) (5). In the case of the heterologous Ad26.ZEBOV/MVA-BN-Filo prime-boost vaccination regimen, non-human primate studies have shown full protection in vaccinated animals following lethal Ebola virus challenge (6). Different immunization regimens using Janssen's vaccine candidate have been evaluated in phase 1 to 3 clinical trials. In particular, we will focus here on three phase 1 trials performed by the EBOVAC1 consortium on healthy adult volunteers in four countries - United Kingdom (7) (8), Kenya (9), Uganda and Tanzania (10). The consortium is part of the Innovative Medicines Initiative Ebola+ programme (11), which aims to assess a novel prime-boost preventive vaccine regimen against EVD. Results of the three phase 1 trials showed no vaccine-related serious adverse events and persistent levels of IgG binding antibodies in all vaccine recipients.

One of the potential assets of the Ad26.ZEBOV/MVA-BN-Filo vaccine regimens is the establishment of a long-term immune response, which is in part characterized by the Ebola virus GP-specific binding antibody response after vaccination. Although no immune correlate of protection has been identified yet, preclinical studies have shown that the peak antibody concentrations post-vaccination are correlated with survival after intramuscular challenge in a non-human primate model, which is the closest model to humans (12) (13) (14). Whether circulating antibody concentrations also correlate with long term protection is not established, however it is of particular interest to quantify the dynamics of the humoral immune response and to estimate the durability of the antibody response. We proposed to use a mathematical model to address these questions. We had a unique opportunity to analyze the data from the three trials in the context of EBOVAC1, because they were conducted almost simultaneously with very similar study protocols. The uniqueness of the data also relied on the large number of consecutive immunogenicity measurements following the boost immunization.

Most of the models that were already developed for the dynamics of the antibody response focused on the decline of the antibody concentrations after the peak response. Linear or piecewise-linear decreases of the antibody response were fitted to data from a large number of vaccines, including Hepatitis B vaccine (15), combined diphteria, tetanus and pertussis vaccine (16) (17), Japanese encephalitis chimeric virus vaccine (18), Hepatitis A vaccine (19) (20) and human papillomavirus-16/18 vaccine (21) (22) (23). However, linear mixed models are limited in term of biological interpretation. The structure of mechanistic models is based on biology and is able to capture non-linear interactions. The estimation of the model parameters gives a quantification of the biological phenomenon. Only a few within-host models were developed to describe the humoral immune response following vaccination. The dynamics of ASCs after vaccinia virus vaccination of human volunteers were described (24) by extending a widely known model for the CD8 T cell response (25) (26). However, this model did not account for the immunologic hypothesis that antibodies are produced by several populations of ASCs. Indeed, it has been suggested that the vast majority of plasma cells generated through immunization are SL cells (27) (28), peaking 7 days after the immunization and lasting very shortly in the organism (28) (29) (30) (31). However, the half-life of antibodies was estimated between 20 and 50 days in several studies (32) (33) (34) (35) (36) (37) (38). Therefore, the persistence of antibody response, observed to last for several years (39), is expected to be generated by LL plasma cells (28) (40) (41) (42) (43). Using long-term data following hepatitis A vaccination (up to 10 years after the boost immunization), an ODE-based mechanistic model helped quantifying 3 scales of the humoral response dynamics (44), corresponding to the life spans of antibodies (around 20-30 days), and two populations of ASCs (one living several months and the other one decades).

Here, we used the same mechanistic model for the humoral immune response, with two populations of ASCs (SL and LL) and the antibody population. Parameters were estimated on data available from three trials of the EBOVAC1 consortium, with a 1 year follow-up for participants of the study including up to 9 consecutive measurements of antibody concentrations. This model allowed to quantify the dynamics of the humoral immune response following different prime-boost vaccine regimens.

RESULTS

Mechanistic model of the immune response. A preliminary analysis was performed to estimate linear trends of the antibody concentrations decrease from 21 days after the boost immunization onwards. The method and results of this analysis are detailed in Appendix A3. This analysis showed in particular the need to model two phases of antibody decline. A mechanistic model was used to fit these dynamics. Based on previous work in immunology (45) and modeling (24) (44), we made the hypothesis that antibodies are produced by two distinct

populations of ASCs, which can be distinguished by two different half-lives: some are assumed SL and others LL. We made the assumption that from 7 days after the boost immunization, both populations of cells decay with time. This decay is applied to the whole compartment of cells, and could either mean that these cells are still generated but their death rate is higher than their proliferation rate or that cells are not generated anymore; in this case the decay corresponds to their net loss. In any case, the assumption can be justified by some experimental evidence that ASCs peak a few days after reaction to pathogen and decrease then after (13) (30) (31) (46). A recent review also suggested that the kinetics of ASCs were similar among different pathogens, with a peak response around 7 or 8 days following infection (47). As ASCs peak around 7 days after immunization, it could then reasonably be assumed that they decay after this time. LL cells are expected to play a role on a longer time scale as these cells are the ones supposed to sustain antibodies (42) (43). The model of the dynamics of the humoral immune response from 7 days after boost immunization and its parameters are represented in Fig 1. SL and LL ASCs decrease respectively at rates δ_S and δ_L and produce antibodies at rates θ_S and θ_L. Antibodies decrease at rate δ_{Ab}. SL and LL values 7 days after the boost immunization are unknown and written respectively as S_0 and L_0. For identifiability issues, we defined two new parameters : $\phi_S = \theta_S S_0$ and $\phi_L = \theta_L L_0$. They correspond to the rate of antibody production times the ASC baseline level, which we call the influx. We used a population approach to estimate parameters δ_{Ab}, δ_S, δ_L, ϕ_S and ϕ_L, assess the effect of the different factors on these parameters and also their inter-individual variability (see equation 5 in materials and methods section).

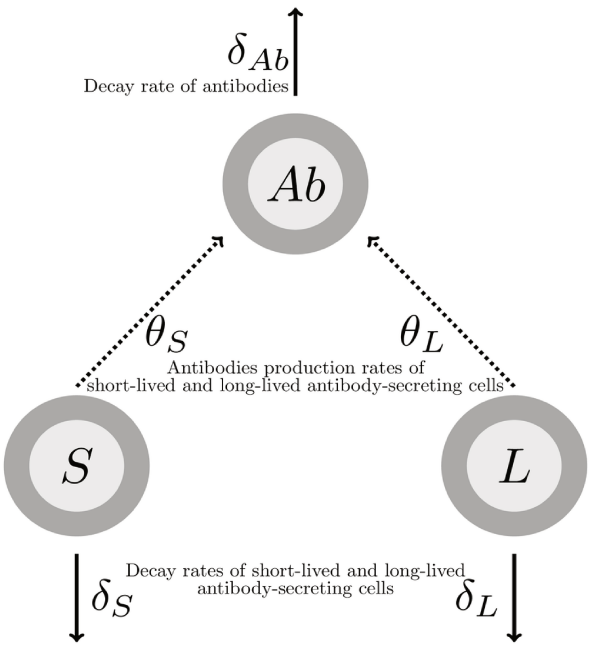


FIG 1 Model of the humoral immune response from 7 days after the boost immunization. S corresponds to the SL ASCs, L corresponds to the LL ASCs and Ab corresponds to the antibodies.

Descriptive analysis of the data. A summary of the characteristics of the dataset is given in Table 1. As shown in Fig 2, dynamics of antibodies from 7 days after the boost immunization were similar across all groups and geographic regions: a peak of the antibody response was observed 21 days after the boost immunization, then antibody concentrations followed a biphasic decay with a first sharp decrease followed by a slower decline. Antibody concentrations at specific time points are detailed in Table 1 and Fig 3. There was no statistical difference of antibody concentrations at the peak of the response between European and East African subjects (p-value of t-test on $\log_{10}$-transformed antibody concentrations = 0.76). However, one year after the prime immunization, the antibody concentrations of European subjects were statistically significantly higher than those of East African subjects (p-value of t-test on $\log_{10}$-transformed antibody concentrations $< 3.10^{-15}$).

TABLE 1 Summary of data characteristics

	Europe		East Africa				Total	
	UK		Kenya		Uganda/Tanzania			
Number of participants (N)	59		59		59		177	
Group MVA/Ad26 D29	15		14 (1 non completed)		15		44	
Group MVA/Ad26 D57	15		15		14 (1 non completed)		44	
Group Ad26/MVA D29	15		15		15		45	
Group Ad26/MVA D57	14 (1 lost of follow-up)		15		15		44	
Sex								
Men	21 (36%)		42 (71%)		47 (80%)		110 (62%)	
Women	38 (64%)		17 (29%)		12 (20%)		67 (38%)	
Age (years)								
Mean (sd)	35.5 (9.9)		25.9 (6.2)		26.4 (6.5)		29.7 (9.1)	
BMI (kg/m^2)								
Mean (sd)	25.2 (4.0)		23.1 (3.6)		22.3 (3.7)		23.6 (4.0)	
Antibody concentrations (in log$_{10}$ ELISA units/mL)	Mean (sd)	N	Mean (sd)	N	Mean(sd)	N	Mean (sd)	N
7 days post boost	2.83 (0.69)	59	3.01 (0.65)	58	2.75 (0.71)	59	2.86 (0.69)	176
21 days post boost (peak)	3.95 (0.43)	59	4.01 (0.39)	59	3.85 (0.43)	59	3.94 (0.42)	177
1 year post prime	3.38 (0.40)	51	2.70 (0.43)	58	2.79 (0.39)	59	2.94 (0.50)	168

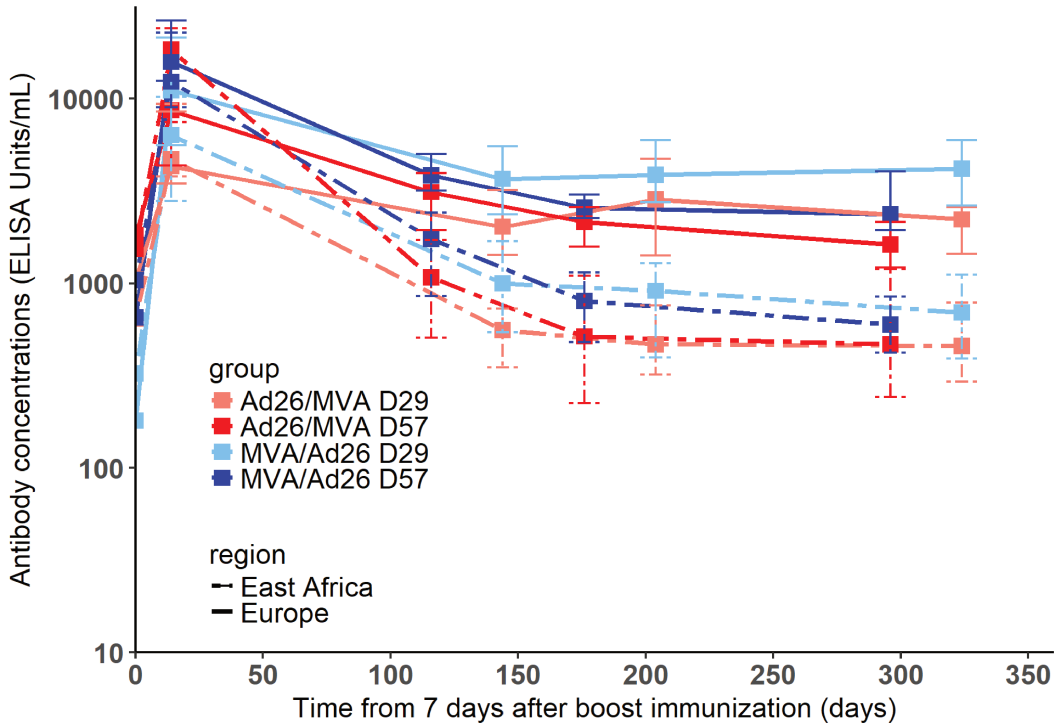


FIG 2 Dynamics of antibody concentrations (in log$_{10}$ scale) from 7 days post boost in European and East African subjects of each group of vaccination. Each color corresponds to a group of vaccination: light blue is MVA/Ad26 D29, dark blue is MVA/Ad26 D57, light red is Ad26/MVA D29, dark red is Ad26/MVA D57. Solid lines correspond to medians in European subjects and dashed lines to medians in East African subjects. 25-75 quantiles are also represented.

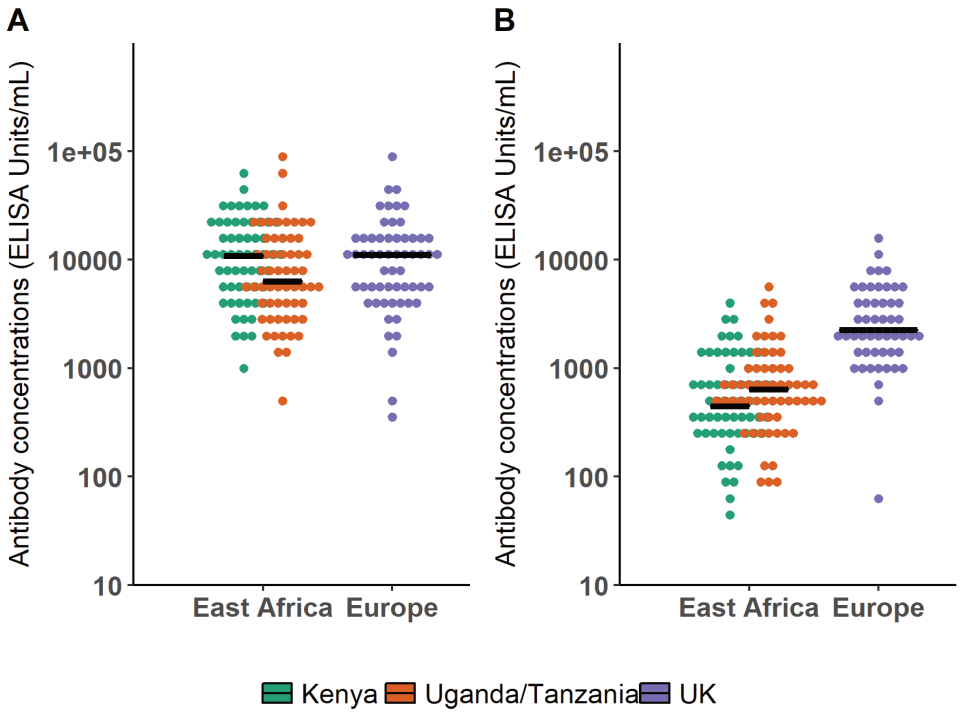


FIG 3 **Comparison of antibody concentrations (in $\log_{10}$ scale) in European and East African subjects. Horizontal lines correspond to median values within each study. A: Boxplot of antibody concentrations at time of the observed peak (21 days after the boost immunization) in European and East African subjects. B: Boxplot of antibody concentrations 1 year after the prime immunization in European and East African subjects.**

The cellular response was also studied from the time of the boost immunization onwards. In particular, we focused on the total percentage of stimulated $CD4^+$ T cells producing at least one of the three cytokines interleukin-2 (IL2), interferon-γ (IFNγ) or tumor necrosis factor α (TNFα). The dynamics of these cytokine-secreting $CD4^+$ T cells after the boost immunization are shown in Fig 4. Dynamics were very similar across all groups, with a peak response measured between 7 and 21 days after the boost immunization. We plotted in Fig 5 the distributions of the percentage of $CD4^+$ T cells in European and East African subjects at specific time points and tested if there was a difference with a two-sided Wilcoxon rank sum test. There was a significant difference between European and East African subjects after prime prior to boost immunization and 7 days after the boost immunization (p-values = 0.015 and < 0.001 respectively), with higher percentages of $CD4^+$ T cells in European than in East African subjects. At 21 days after the boost immunization, this difference was no longer significant (p-value=0.23).

Parameters estimation and goodness of fit of the model. Using a population approach, we estimated the value of the parameters of the model and assessed the effect of different factors on these parameters. Prior distribution were used to constrain the parameters values, according to previous biological knowledge. This is detailed in the method section. After selection of the model, we found that the vaccine regimens induced different mean decay rates of SL ASCs (δ_S). We also estimated different mean values of the LL influx parameter ($\phi_L = \theta_L L_0$) between East African and European subjects. The variation of three parameters (ϕ_S, ϕ_L and δ_{Ab}) could not be fully explained by the measured factors (geographic region, vaccine regimen) and was handled with normal random effects. This model allowed to fit well the data of participants, as shown in Fig 6. Parameters estimation is displayed in Table 2 and allowed quantification of the humoral immune response to the prime-boost regimen. Mean value of circulating antibodies half-life across participants was estimated at 24 days with 95% confidence interval [22,26] days, and 5-95 quantiles of the distribution of individual values ranging from 18 to 36 days. The histogram of the estimated antibody half-life in all participants is shown in Fig 7. This estimation was lower than values estimated through passive immunity (36) (37) (38) but still seemed biologically plausible and consistent with some previous

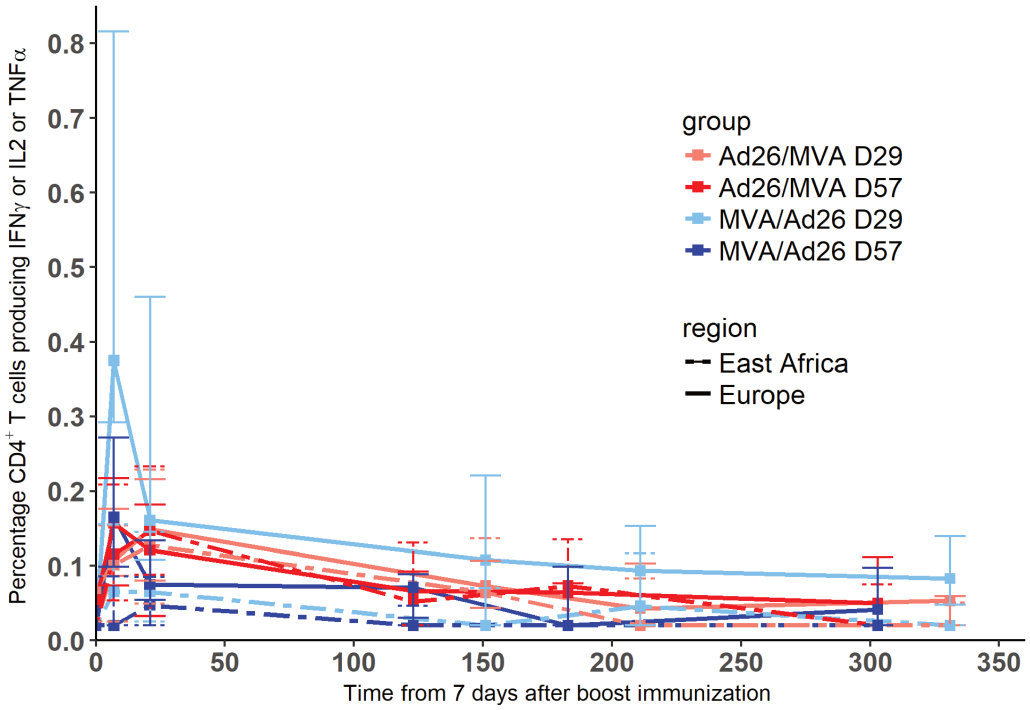


FIG 4 Dynamics of percentages of CD4⁺ T cells producing at least one of the three cytokines IL2, IFNγ or TNFα from the time of boost immunization. The lower limit of quantification was 0.04, and values under this limit were imputed to half of it (=0.02). Each color corresponds to a group of vaccination: light blue is MVA/Ad26 D29, dark blue is MVA/Ad26 D57, light red is Ad26/MVA D29, dark red is Ad26/MVA D57. Solid lines correspond to medians in European subjects and dashed lines to medians in East African subjects. 25-75 quantiles are also represented.

studies (34) (45). Sensitivity analyses were realized on the estimation of that parameter: in particular, we estimated the parameters by using a more constraining prior distribution on the decay rate of antibodies δ_{Ab}, corresponding to a mean half-life value of 42 days with 5-95 quantiles of the parameter distribution at 34-51 days. Compared to the previous estimation of a 24 days half-life, new estimation of the antibody half-life with the stronger prior barely changed (27 days with confidence interval [25;29] days), showing robustness in the estimation of this parameter. Moreover, estimation of ASCs half-life allowed to distinguish two populations of ASCs, one with a very short half-life (only a few days) and one with a much longer half-life (a few years). The difference in value between ϕ_S and ϕ_L, with $\phi_S > \phi_L$, also suggested that either SL cells were present at a much higher level 7 days after the boost immunization and/or they produced many more antibodies than LL cells. These estimations supported the hypothesis of the generation of a very quickly reactive population of cells with a short lifespan, and a long-term antibody production sustained by another population of cells able to last several years in the organism (28). We estimated the half-life of these LL cells to be 6.0 years with a 95% confidence interval of [2.7;13] years. The upper bound of this confidence interval is actually artificially introduced by the normal approximation of the parameter distribution, but we did not have enough information to determine with precision this upper bound. Indeed, data were only available up to 1 year after the prime immunization and the decline of this population is slow. We performed a profile likelihood to explore the identifiability of parameter δ_L, as shown in the figure in Appendix A1. For several values of δ_L, the model was estimated and model criteria were computed (non-penalized log-likelihood and likelihood cross-validation criteria (LCVa) (48)): the resulting profile showed a flat behavior after a half-life value of 5 years, meaning that data up to 1 year did not allow to distinguish between an estimated half-life of 5 years or more. The estimation based on the currently available data suggested that half of the long-lived ASCs generated at 7 days after the boost immunization would persist for at least 5 years.

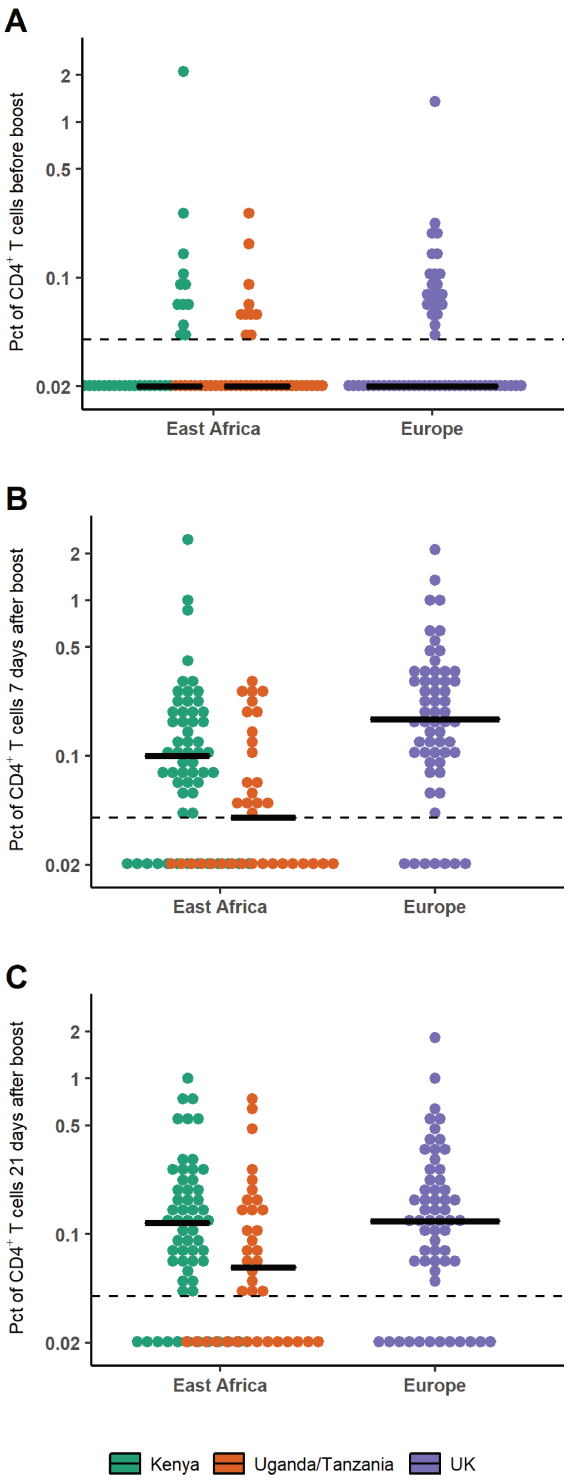


FIG 5 Comparison of percentages (pct, in log scale) of $CD4^+$ T cells producing at least one of the three cytokines IL2, IFNγ or TNFα in European and East African subjects after prime prior to boost immunization (A), 7 days after boost immunization (B) and 21 days after boost immunization (C). Horizontal lines correspond to the median value within each trial. Dashed lines correspond to the lower limit of quantification (=0.04); values under this limit were imputed to half of it (=0.02).

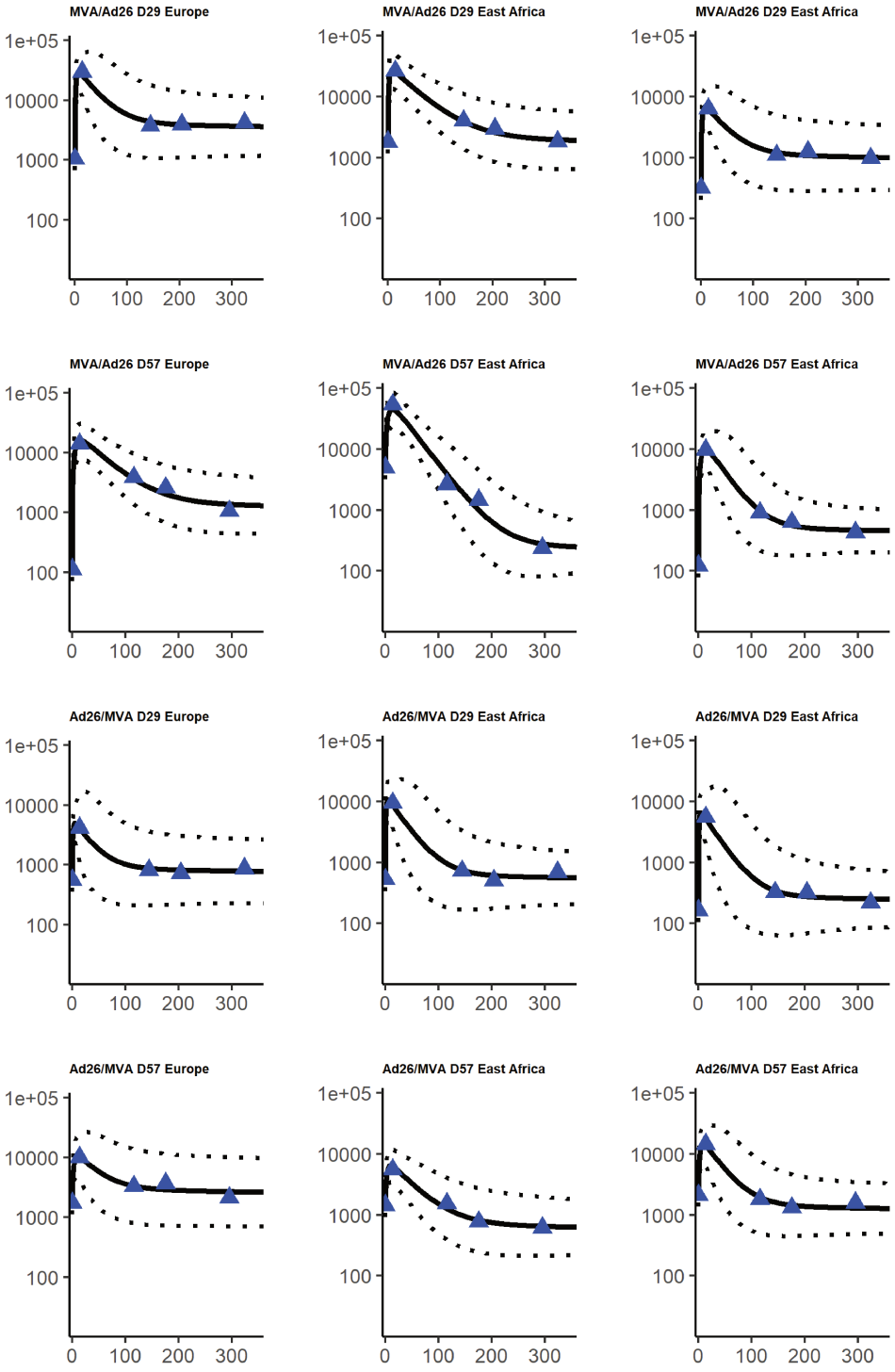


FIG 6 Fits of a random sample of subjects. The x-axis corresponds to the time from 7 days after boost immunization (in days) and the y-axis corresponds to the antibody concentrations (in ELISA units/mL, log10 scale). For each subject, blue triangles correspond to the observed data, solid line corresponds to the prediction from the model and dashed line corresponds to the 95% prediction interval, accounting for the uncertainty on parameters estimation and the measurement error.

TABLE 2 Parameters estimation

Parameter	Mean	95% Confidence interval
Antibodies half-life (days): $log(2)/\delta_{Ab}$	24	[22;26]
Long-lived cells half-life (years): $log(2)/\delta_L$	6.0	[2.7;13]
Short-lived cells half-life (days): $log(2)/\delta_S$		
MVA/Ad26 D29 group	2.0	[1.3;3.0]
MVA/Ad26 D57 group	4.9	[3.1;7.7]
Ad26/MVA D29 group	1.2	[0.8;1.9]
Ad26/MVA D57 group	3.0	[1.9;4.7]
ϕ_S (ELISA Units/mL.days^{-1})	2755	[1852;4100]
ϕ_L (ELISA Units/mL.days^{-1})		
East African participants	16.6	[13.7;20.1]
European participants	70.7	[54.0;92.7]
σ_{ϕ_S} (inter individual standard deviation on ϕ_S)	0.92	[0.83;1.01]
σ_{ϕ_L} (inter individual standard deviation on ϕ_L)	0.85	[0.78;0.92]
$\sigma_{\delta_{Ab}}$ (inter individual standard deviation on δ_{Ab})	0.30	[0.24;0.36]
σ_{Ab} (standard deviation on observations)	0.10	[0.10;0.10]

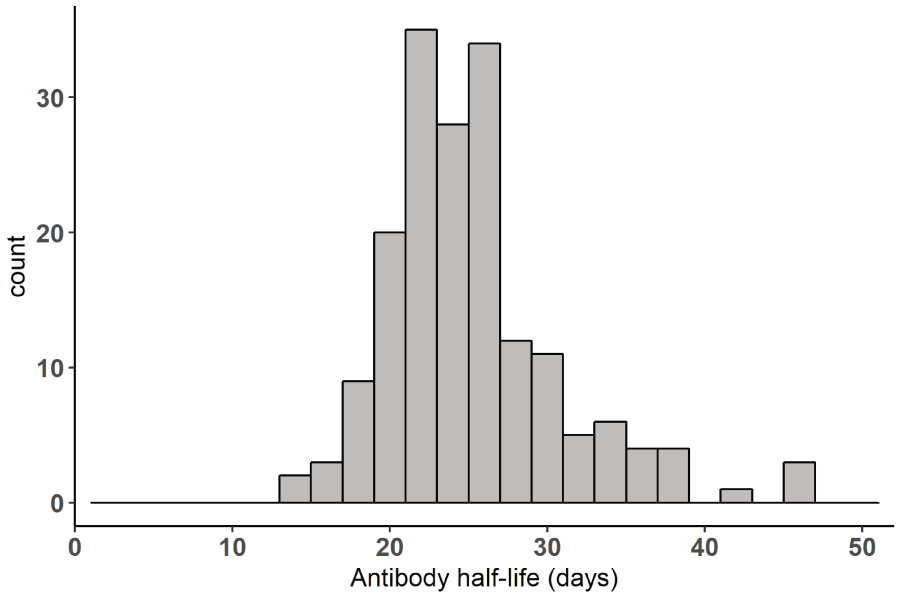


FIG 7 **Histogram of estimated antibody half-life in all participants.**

Factors influencing the dynamics of the humoral response. As detailed in Table 2, we estimated that the mean SL ASCs half-life varies from 1.2 days to 4.9 days, depending on the vaccine regimen. This estimation was consistent with findings on the kinetics of ASCs following infection/vaccination, suggesting that circulating plasmablasts peak at 7 days after boost immunization but are absent after 10 to 14 days (47). The model estimates suggested that Ad26/MVA induced a shorter half-life compared to MVA/Ad26 and boost immunization at day 56 induced a longer half-life compared to boost immunization at day 28. These differences had an impact on the early antibody dynamics: the longer the half-life of SL ASCs the higher the peak of antibody concentrations. The impact of the value of the SL ASCs half-life was negligible in the longer term, as shown in Fig 8. This suggests that the different prime-boost regimens induced different early responses after the boost immunization but no difference in the duration of the antibody responses.

Descriptive analysis of the data suggested a significant difference of mean antibody concentrations 1 year after the prime immunization between the East African and European trials. Estimation of the model helped understanding and quantifying this difference. We found a significant different value of parameter ϕ_L between East African and European subjects. This meant that either European subjects had more LL ASCs 7 days after the boost immunization and/or these cells produced more antibodies than in East African subjects. We see in Fig 8 that the difference in ϕ_L value explains why European subjects reached higher antibody concentrations 1 year after the immunization than East African subjects. Moreover, we did not find any difference of the value of the LL ASCs decay δ_L across all subjects. This meant that in the longer term, antibody concentrations should decline in the same way in both East African and European subjects, as we expect the decay of antibodies to be driven by the decay of LL ASCs. From the model estimations, we expected the LL ASCs to persist as long in European subjects as in East African ones, but to be produced in a higher number after the boost immunization and/or to secrete more antibodies in European subjects compared to East African ones. This difference could result from a more activated immune environment at baseline in East African subjects (49). From this hypothesis, as environmental characteristics differ between Kenya and Uganda/Tanzania, we could expect more inter-individual variability of the value of ϕ_L within the East African group of participants compared to European group. A Fisher test for equality of variances showed that there was no significant difference of variance for ϕ_L between East African and European values (p-value=0.30), meaning that there was no additional unexplained variability in the East African group compared to the European one.

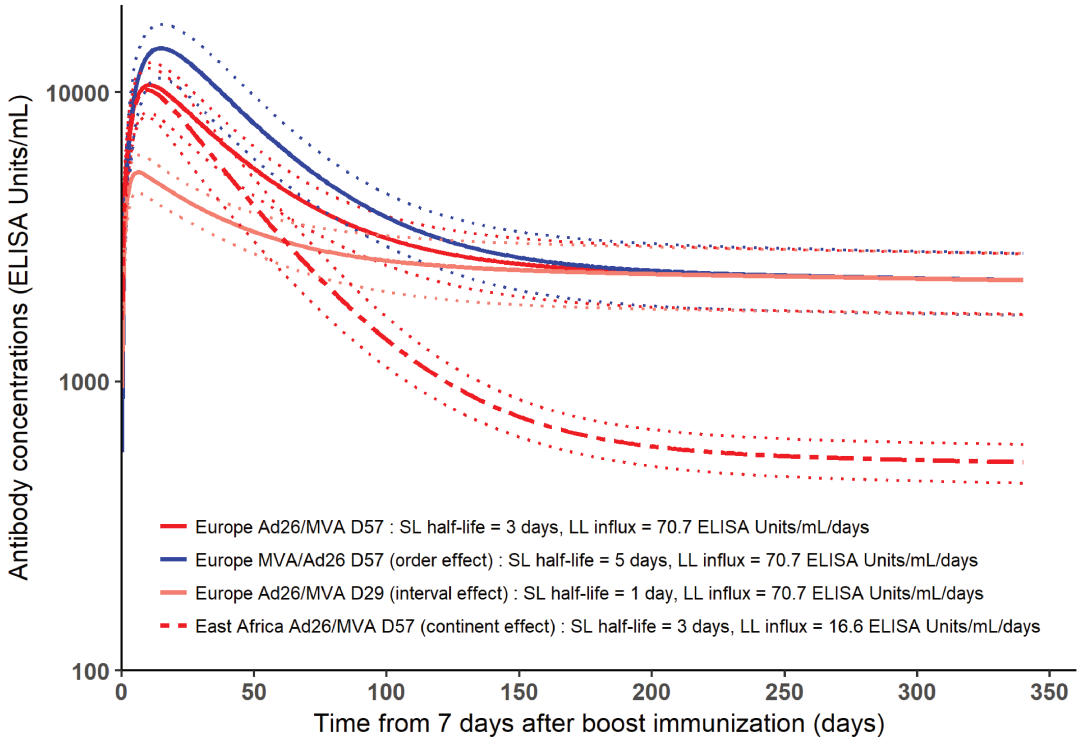


FIG 8 Marginal predictions of antibody concentrations dynamics from the model (in $\log_{10}$ scale). Marginal predictions show the effect of covariates on antibody concentrations dynamics. Dark red plain line (Europe Ad26/MVA D57 group) differs from dark blue plain line (Europe MVA/Ad26 D57) only by order of immunizations, showing that MVA/Ad26 order implies a higher peak of antibody concentrations than Ad26/MVA order. Dark red plain line (Europe Ad26/MVA D57 group) differs from light red plain line (Europe Ad26/MVA D29) only by interval between prime and boost immunizations, showing that a boost at day 56 induces a higher peak of antibody concentrations than a boost at day 28. Finally, dark red plain line (Europe Ad26/MVA D57 group) differs from dark red dashed line (East Africa Ad26/MVA D57 group) only by geographic region, showing that European subjects have a similar antibody peak as East African ones, but higher sustained antibody concentrations. For all curves, light dashed lines correspond to 95% confidence intervals accounting for the uncertainty on parameters estimation.

Following this result, we explored if the estimated difference between East African and European subjects could be explained by the magnitude of the cellular $CD4^+$ T cell response. It came from the hypothesis that differences in the pathogens to which individuals are exposed during everyday life could have an effect on the cellular response (50). As $CD4^+$ T cells are required for the humoral immune response, we made the hypothesis that the difference between East African and European subjects could be mediated by a difference in the T helper response early after the boost immunization. In the mechanistic model, the difference was estimated on parameter ϕ_L: the mean value over European subjects was higher than in East African subjects. As there was also a random effect on ϕ_L, we were able to compute the individual estimated value of this parameter. We computed the correlation between the value of ϕ_L and the percentage of $CD4^+$ T cells producing at least one of the three cytokines IL2, IFNγ or TNFα at different time points: after prime prior to boost immunization, 7 and 21 days after the boost immunization. Results are displayed in Fig 9. We did not observe any clear relationship between the $CD4^+$ T cell percentages and ϕ_L values. Pearson correlation coefficients were only significantly different from zero at 7 days after boost immunization, with a moderate value of 0.2. To further explore the hypothesis that the difference of value of ϕ_L could be mediated by the T helper response, we introduced the percentage of the $CD4^+$ T cell producing cytokines 7 days after the boost immunization in the mechanistic model as a covariate on ϕ_L. Effect of the covariate was added and tested separately on ϕ_L, with or without the geographic region variable, as shown in equation 6 from the materials and methods section. Without the geographic region variable, the estimated effect of the $CD4^+$ T cell was significant (p-value = 0.03), but the likelihood of the model was much lower than for the model including the geographic region variable without the CD4 variable (136.34 versus 171.97). In a model including both geographic region and CD4 variables, the estimated effect of the $CD4^+$ T cell response was not significant (p-value 0.64). Overall, these results suggested that the difference of ϕ_L value between the geographic regions could not be explained by the measure of the percentage of $CD4^+$ T cells producing at least one of the cytokines IL2, IFNγ or TNFα 7 days after the boost immunization.

DISCUSSION

The mechanistic model accounting for two populations of ASCs allowed to quantify the dynamics of the antibody response following different prime-boost vaccine regimens. In particular, it allowed to estimate a lower bound of the durability of the antibody response through LL plasma cells. Moreover, we were able to identify several factors influencing the response to vaccine. We found that vaccine regimen impacts the magnitude of the early antibody response through the dynamics of the SL ASCs, but has no effect on the LL ASCs and thus on the long-term persistence of antibodies. It suggests a minor impact of the interval between the prime and the boost immunizations on the long-term level of the binding antibodies.

The dynamics of LL ASCs were estimated to differ by geographic region, inducing a higher long-term level of antibodies in European subjects compared to East African ones. Several factors could contribute to the geographic effect, such as HLA subtypes, nutritional status, co-infections or pre-existing immunity. Demographic factors could also play a role in this difference, although no significative effect of sex and age was found on the decrease of the antibody concentrations in the linear mixed model or on the parameter ϕ_L (see appendix A3 for details). The absence of association between this difference and circulating $CD4^+$ T cells producing cytokines does not exclude alternative effects of the $CD4^+$ T cells on the humoral response, for example a link with plasma cells and antibody production at the level of the lymphoid organs. The difference of immune response between different geographic regions has already been identified in some other vaccination studies, even if the vast majority of vaccination programs in Africa have had a tremendous positive public health impact. The efficacy of Bacille Calmette-Guérin vaccination was observed to be lower in African infants compared to European ones (51). West Africans showed lower T-cell response following vaccination which an HIV vaccine candidate compared to South Africans and North Americans (5). The efficacy of the licensed yellow fever vaccine 17D was also found to be lower in African population compared to European one; an activated immune environment prior to vaccination was hypothesized (49). In the case of Ebola vaccine, as the protective level has not been determined yet, we do not know if the difference in antibody concentrations has implications on the efficacy of the vaccine. Yet, the observed difference in long-term antibody responses between East African sites and the UK site is an interesting outcome that would justify

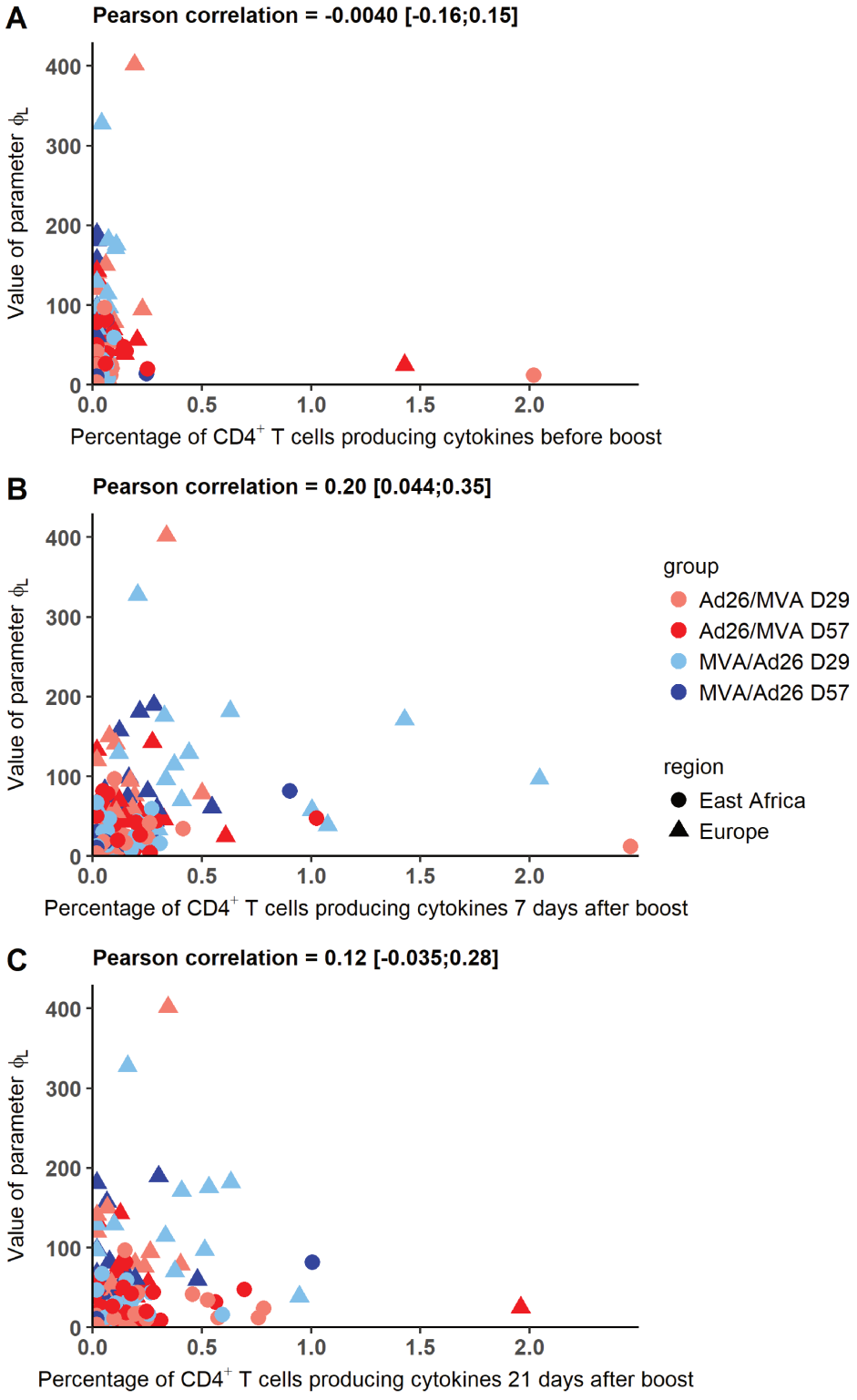


FIG 9 Value of parameter ϕ_L versus the percentage of CD4$^+$ T cells producing IL2, IFNγ or TNFα after prime prior to boost, 7 days after the boost immunization and 21 days after the boost immunization in European and East African subjects of each group of vaccination. Each color corresponds to a group of vaccination: light blue is MVA/Ad26 D29, dark blue is MVA/Ad26 D57, light red is Ad26/MVA D29, dark red is Ad26/MVA 57.

additional mechanistic studies to identify which factors contribute to these differences.

The fact that immune memory is not considered in the model represents a limitation, especially in term of prediction of the response to exposure to wild type virus. However, the role of the memory response and the immune response levels required for protection are not known at the moment. Moreover, the main limitations for the estimation of the model are the low number of subjects as the data was generated in phase 1 trials, the lack of data on the number of plasmablasts, and the lack of measurements beyond 1 year. However, the statistical analysis using a population approach allowed to determine a lower bound of the long-term response.

These results will benefit from additional data coming from phase 2 studies to confirm the robustness of the long-term response. Several studies showed that antibody responses in humans do not reach steady state levels until approximately 2-3 years after infection or vaccination (45). More data should also allow a better identification of the half-life of the two ASC populations and will increase the statistical power of the analysis. Moreover, the differences between geographic regions will be refined using data from West African subjects. Additional studies looking at the effect of other factors on the immune response, such as malaria co-infection, will help explaining these potential differences.

In conclusion, this first modeling study estimates promising binding antibody responses to prime-boost regimens combining Ad26 and MVA in an Ebola vaccine. While the long-term antibody persistence is not found to be influenced by the vaccine regimen in the model, the geographic region could potentially impact the long-term dynamics by its effect on dynamic parameters associated to the LL ASCs.

MATERIALS AND METHODS

Ethics statement. The UK trial protocol and study documents were approved by the UK National Research Ethics Service. The Kenya trial protocol and study documents were reviewed and approved by the local Ethics Committee, and the Kenyan regulatory authority. The Uganda/Tanzania trial protocol and study documents were reviewed and approved by the Tanzanian Medical Research Coordinating Committee of the National Institute for Medical Research; the Tanzania Food and Drugs Authority; the Uganda Virus Research Institute Research and Ethics Committee; the Uganda National Council for Science and Technology; the Uganda National Drug Regulatory Authority and the Ethics Committee of the London School of Hygiene and Tropical Medicine. These trials were conducted in accordance with the principles of Good Clinical Practice and the Declaration of Helsinki, and all participants gave formal, written consent before undergoing any trial-related procedure.

Immunogenicity measurements. We analyzed data from three randomized, observer-blind, placebo-controlled, phase I trials in four countries on healthy volunteers aged 18 to 50. They aimed at assessing the safety and tolerability of two novel candidate Ebola vectors combined in different prime-boost regimens. The first vector is a monovalent, recombinant, E1/E3-deleted, replication-defective, adenovirus type 26 vector vaccine encoding Ebola virus Mayinga variant GP (Ad26.ZEBOV). It was produced in PER.C6 human cells and injected in single dose at concentration of 1×10^{11} viral particles/mL. The second vector is a recombinant, replication-defective, modified vaccinia Ankara vector vaccine (MVA-BN-Filo) expressing Mayinga variant GP, Sudan virus Gulu variant GP, Marburg virus Musoke variant GP, and Tai Forest virus nucleoprotein. It was produced in chicken embryo fibroblasts and injected at a concentration of 2×10^8 median tissue culture infective dose ($TCID_{50}$)/mL.

Trials were realized in UK, Kenya and Uganda/Tanzania. Results of the UK trial were described in (7) and (8). Within each trial, eligible participants were equally randomized into four vaccination regimens (within each they received active vaccine or placebo in a 5:1 ratio): two with MVA-BN-Filo as a prime vaccine on day 1 followed by Ad26.ZEBOV on day 29 or day 57 (MVA/Ad26 D29 and MVA/Ad26 D57) and two with a prime immunization of Ad26.ZEBOV at day 1 boosted by MVA-BN-Filo on day 29 or day 57 (Ad26/MVA D29 and Ad26/MVA D57). In UK, there was an additional open-label group receiving Ad26.ZEBOV on day 1 followed by MVA-BN-Filo on day 15. This arm was not included in the analysis as this regimen was not realized in East African countries. We included in the analysis only subjects who received both prime and boost immunizations, which corresponded to a total of 177 subjects over all groups and countries. Subjects were followed up to 1 year after receiving the prime immunization, with consecutive immunogenicity assessments performed on blood samples. These samples were taken before prime and boost immunizations, 7 days after prime and boost immunizations and 21 days after the boost immunization. Subjects allocated to groups receiving a boost immunization at day 57

had an additional sample taken at day 29. Further samples were taken at day 180, 240 and 360 after the prime immunization. Design of the trials is summarized in Fig 10. We analyzed antibody concentrations as the total IgG response against Ebola virus Kikwit variant GP: this was assessed by enzyme-linked immunosorbent assay (ELISA) from BBRC (Battelle) in the UK and Uganda/Tanzania trials and Q2 Solutions in the Kenya trial. Moreover, cellular data obtained from intracellular cytokine staining (ICS) at the HIV Vaccine Trials Network laboratory on frozen peripheral blood mononuclear cells (PBMC) were explored. PBMC were stimulated with one of 2 peptide pools covering the GP from Mayinga variant of Ebola virus (Think Peptides, UK). A 16-color staining panel was used and our analysis was based on the total percentage of $CD4^+$ T cells producing IL2, TNFα or IFNγ. Further details of immunogenicity measurements are given in the Appendix of (7). We focused on the data sampled after the boost immunization in subjects who received both prime and boost immunizations with active components, since we are mainly interested in the duration of the antibody response and its decrease after the observed peak following the boost immunization.

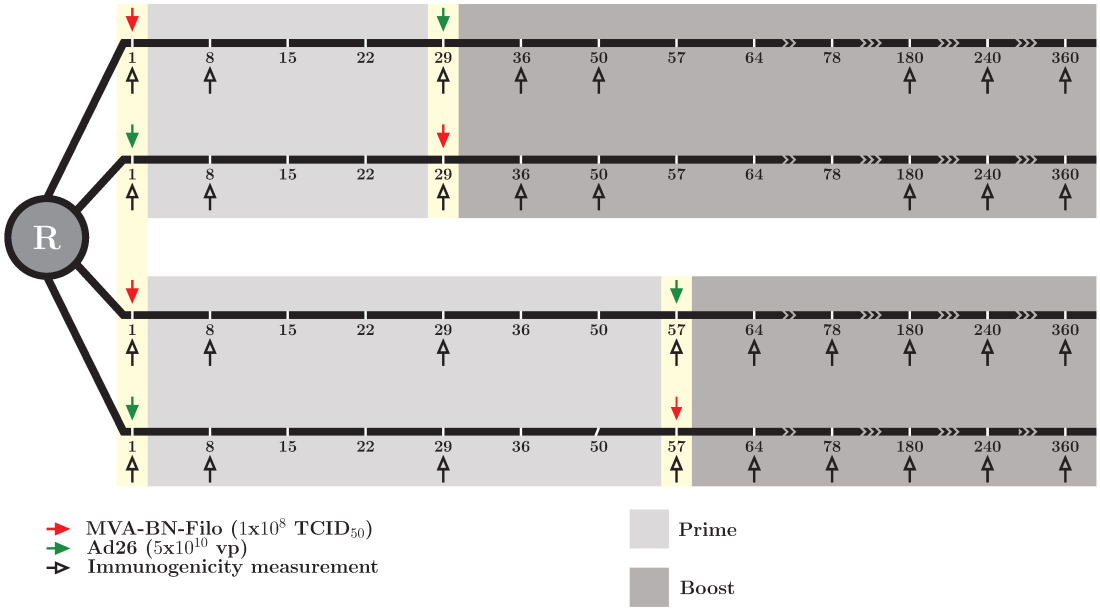


FIG 10 Design of EBOVAC1 trials.

Mechanistic model A first preliminary analysis of the decrease of antibody concentrations was realized using linear mixed models. It is described in Appendix A3. However, our main approach relies on mechanistic models divided in three layers, described in (52): first, we used a mathematical model, based on ordinary differential equations and describing the dynamics of the biological process, as was done for hepatitis A vaccine (44). Then we used a statistical model accounting for the inter-individual variability and the effect of covariates on the parameters. Finally, we considered an observation model, as immunological measurements do not cover all compartments of the mathematical model.

The mathematical model, represented in Fig 1, relies on the hypothesis that antibodies are produced by two distinct populations of ASCs, differing by their decay rate (44). It contains three compartments, the SL cells (S), the LL cells (L) and the antibodies (Ab). Time was rescaled in order to consider only the dynamics of antibody concentrations from 7 days after boost immunization, after which both populations of ASCs decrease with time.

The corresponding ordinary differential equations are the following:

$$\frac{dS}{dt} = -\delta_S S \quad (1)$$

$$\frac{dL}{dt} = -\delta_L L \quad (2)$$

$$\frac{dAb}{dt} = \theta_S S + \theta_L L - \delta_{Ab} Ab \quad (3)$$

with δ corresponding to decay rates and θ production rates. The equation for the antibodies dynamics can be written as:

$$\frac{dAb}{dt} = \phi_S e^{-\delta_S t} + \phi_L e^{-\delta_L t} - \delta_{Ab} Ab \quad (4)$$

with $\phi_S = \theta_S S_0$ and $\phi_L = \theta_L L_0$, where $S_0 = S(t = 0)$ and $L_0 = L(t = 0)$ are the initial conditions at 7 days after the boost immunization. As SL and LL ASCs populations were not observed, θ_S and S_0 could not be identified separately (and so were θ_L and L_0). The initial condition $Ab(t = 0)$ is given by the data (measure at 7 days after the boost immunization). Among the 177 subjects, only 1 did not have a measure of the antibody concentration 7 days after the boost immunization. The value was imputed by using the mean value of his/her group of vaccination in his/her trial, i.e., the mean value of Kenyan subjects in group MVA/Ad26 D29. Finally, we estimated the five following biological parameters: $\xi = (\phi_S, \delta_S, \phi_L, \delta_L, \delta_{Ab})$.

For the statistical model, as described in (52), the parameters ξ_l, $l = 1..5$ are transformed using a logarithm transformation to ensure positivity of production and decay rates. Moreover, a mixed-effect model was introduced on each parameter to account for between-subject variations and possible covariates. Value of parameter $\tilde{\xi}_l = ln(\xi_l)$ for each subject i can be written:

$$\tilde{\xi}_l^i(t) = \tilde{\xi}_{l_0} + \beta_l z_l^i + u_l^i \quad (5)$$

where $\tilde{\xi}_{l_0}$ is the intercept and represents the mean ln-transformed value of parameter ξ_l across the population, β_l is a vector of regression coefficients, z_l^i is a vector of n_e explanatory variables and u_l^i is an individual random effect, following a centered normal distribution with variance ω_l^2. Random effects were independent from each other and applied on a subset of q biological parameters. In practice, after selection (see section Parameters estimation for the model selection method), we applied random effects on the following parameters: ϕ_S, ϕ_L and δ_{Ab}. We assessed the effect of $n_e = 3$ explanatory variables on all parameters except δ_{Ab}: the order of immunization (binary variable equal to 0 when the subject receives a prime with MVA-BN-Filo boosted by Ad26.ZEBOV, 1 if the subject receives Ad26.ZEBOV then MVA-BN-Filo), the interval between the two immunizations (binary variable equal to 0 when the subject receives a prime-boost regimen with an interval of 28 days, 1 when the interval is of 56 days), and the geographic region (binary variable equal to 0 in Europe and 1 in East Africa). Additionally, we also assessed the effect of the cellular response as an explanatory variable. This was done by considering the percentage of CD4$^+$ T cells producing cytokines 7 days after boost immunization. The variable $CD4^i(boost + 7days)$ was added to the vector z_l of explanatory variables, and its effect was estimated on parameter ϕ_L. Values of β_{gr} and β_{CD4} were estimated in :

$$\tilde{\phi}_L^i(t) = \tilde{\phi}_{L_0} + \beta_{gr} geographic\_region^i + \beta_{CD4} CD4^i(boost + 7days) + u_l^i \quad (6)$$

with $CD4^i(boost + 7days)$ the percentage of CD4$^+$ T cells producing cytokines 7 days after the boost immunization in participant i.

For the observation model, we had access to immunological measurements of IgG binding antibodies concentrations against the Kikwit GP in all studies. We assumed there was a measurement error normally distributed on the log_{10} value of the antibody concentrations. In practice, we assumed we observe for patient i at discrete time j, :

$$Y(t_{ij}) = log_{10}(Ab(t_{ij})) + \epsilon_{ij} \quad (7)$$

with

$$\epsilon_{ij} \sim \mathcal{N}(0, \sigma_{Ab}^2). \quad (8)$$

ϵ being an additive normally distributed measurement error.

Parameters estimation With the three layers of the mechanistic model, the estimation problem corresponds to the determination of parameters intercepts, regression coefficients, standard deviations of random effects and standard deviations of measurement errors. The vector of parameters θ can be written as :

$$\theta = [(\tilde{\xi}_{l_0})_{l=1..n_b}, (\beta_l)_{l=1..n_e}, (\omega_l)_{l=1..q}, (\sigma_l)_{l=1..M}] \quad (9)$$

Estimation was made using NIMROD software, available at http://etudes.isped.u-bordeaux2.fr/BIOSTATISTIQUE/NIMROD/documentation/html/index.html. It uses a maximum-likelihood approach (53) with a Newton-like algorithm (54) which approximates the Hessian by using first derivatives of the likelihood. Several criteria ensured the convergence of the algorithm. Moreover, we could account for information on parameters, obtained from biological knowledge and previous estimations in the literature, by adding a prior distribution on these parameters. This led to the determination of the Maximum a Posterior (MAP) estimator through the maximization of a penalized likelihood (55). In practice, we used a normal prior distribution on the ln-transformed population mean value of biological parameters $\tilde{\xi}_{l_0}$. Some previous work showed that antibodies half-life could vary between a few weeks to a couple of months. Studies of intravenous IgG preparations reported half-life around 20 to 30 days (32) (33), while studies of passive immunity with maternal transmission of antibodies to infants have reported half-life varying from 20 days (34) to 35-50 days (35) (36) (37) (38). These studies also highlighted the inter-individual variability over the half-life of antibodies, as well as the possible effect of geographic regions. We used an informative prior distribution on $\tilde{\delta}_{Ab_0}$ such that mean antibody half-life would be 45.2 days, and the variance was chosen such that the 5-95 quantiles of the distribution were 6 days - 9 months. Additional sensitivity analyses were performed with a much lower variance on the prior distribution implying 5-95 quantiles of the a priori distribution to be 34-51 days. We used non informative prior distributions on parameters $\tilde{\phi}_{S_0}$ and $\tilde{\phi}_{L_0}$ as we did not have any information on their possible value: mean value of the ln-transformed parameters is taken equal to 0, with standard deviation equal to 10. We used prior distributions on $\tilde{\delta}_{S_0}$ and $\tilde{\delta}_{L_0}$. It helped constraining the estimation such that $\delta_{S_0} > \delta_{L_0}$ as expected by the definition of the SL and LL populations. We used a large prior distribution on $\tilde{\delta}_{S_0}$ as we did not know exactly the time scale of their half-lives. Mean value corresponded to a half-life of 1.88 days, with 5 - 95 quantiles equal to 0.0005 day and 7000 days. Parameter δ_L was expected to be close to 0, but as data were collected up to 1 year after the prime immunization, we did not expect the model to be able to distinguish a half-life of more than a few years. To account for this constraint, we used a prior distribution with a mean value corresponding to a half-life of 1.2 year, and 5 - 95 quantiles corresponding to half-lives of 40 days and 14 years. Table in Appendix A2 sums up the information on the prior normal distributions.

Selection of the model random effects and covariates was realized by performing estimation on several models that were compared according to two criteria: log-likelihood (to be maximized) and approximation of the likelihood based cross-validation criterion (LCVa) (48) (to be minimized). We proceeded in the following way: we first estimated the model parameters using several combinations of 2 random effects (one on the SL compartment, i.e., either on ϕ_S or δ_S and one on the LL compartment). We selected the best combination and then added a random effect on δ_{Ab}, which considerably improved the model. The variability on parameter δ_L was complicated to capture: δ_L has an effect mainly on the late dynamics of the antibodies and data is not available beyond 1 year after the prime immunization. It brought us to compare only two combinations of three random effects: on ϕ_S, ϕ_L and δ_{Ab} and on δ_S, ϕ_L and δ_{Ab}. Using model criteria, we kept the combination corresponding to the best model, namely the one with random effects on ϕ_S, ϕ_L and δ_{Ab}. For the covariate selection, we proceeded with a backward stepwise approach. First, the model was estimated with all covariates (order, interval and geographic region) on all parameters except δ_{Ab}. Covariates were removed one by one: in particular, at each iteration i, the less significant covariate Z_k was determined using the p-value of the Wald test and removed. Model criteria ensured that the model was not worse without the covariate Z_k compared to the model including Z_k. At the next iteration, the model did not contain covariate Z_k. The least significant covariate Z_{k+1} was removed in a similar way. These steps were repeated until only significant covariates that could not be removed without altering the performance of the model were kept. Sensitivity analyses were realized: in particular, we estimated first the model with only the geographic region covariate on all parameters and applied the backward stepwise approach. Then we added the

order and interval covariates and performed the same approach. Interactions between order and interval were added and tested but were not significant and did not improve the model.

APPENDICES

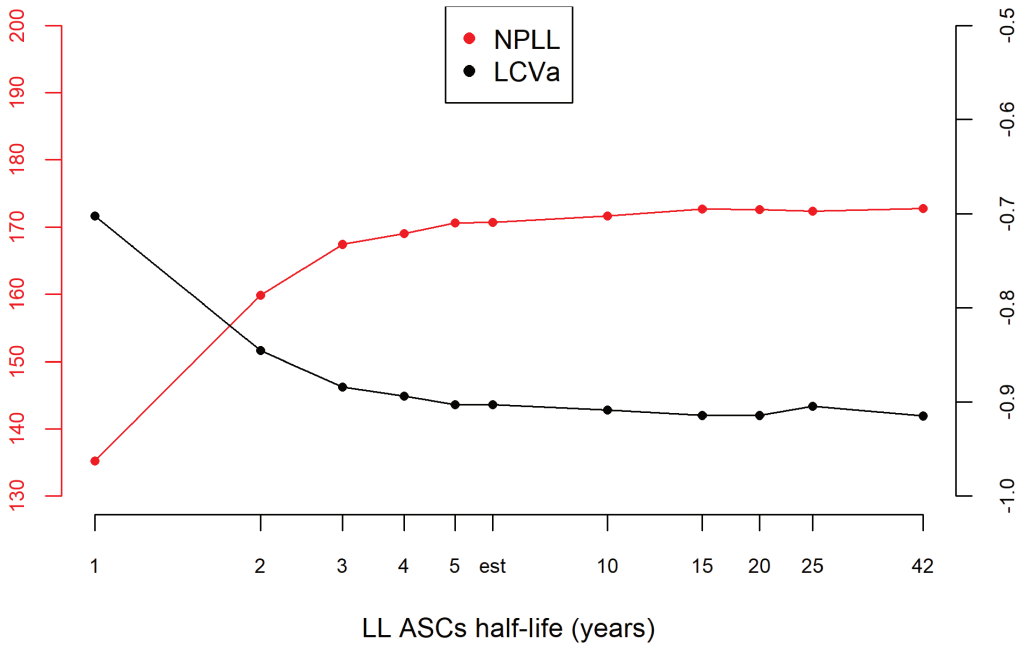


FIG A1 Profile likelihood on parameter δ_L. Left axis corresponds to the non penalized log likelihood in red (NPLL), which needs to be maximized. Right axis corresponds to the LCVa in black, which needs to be minimized. Both criteria are computed for several values of long-lived cells half-life, including the estimated (est) one, and represented on a log-scaled axis.

TABLE A2. A priori distributions of the parameters of the mechanistic model.

	Log scale		Natural scale			Half-life		
Parameter	Mean	sd	Mean	Q5	Q95	Mean	Q5	Q95
ϕ_S	0	10	1	7.10^{-8}	1.10^{8}	NA	NA	NA
ϕ_L	0	10	1	7.10^{-8}	1.10^{8}	NA	NA	NA
δ_{Ab}	-4.1	1.0	0.017	0.0032	0.086	41 days	8 days	216 days
δ_S	-1.0	5.0	0.37	1.10^{-4}	1372	1.88 day	5.10^{-4} day	7029 days
δ_L	-6.5	1.5	0.0015	1.10^{-4}	0.018	1.3 year	40 days	15 years

APPENDIX A3. Linear mixed models Before using the mechanistic model, a preliminary analysis was realized with linear mixed models to model the decrease of antibodies concentrations from 21 days after the boost immunization (which corresponds to the observed peak). The aim was to estimate linear trends and their variation according to vaccine regimen, geographic regions and clinical characteristics such as age, sex and BMI. Two models were estimated : single-slope (SS) model and change-of-slope (CS) model, with a change of slope at time τ in order to distinguish the early strong decrease following the peak of antibody to the lighter one at the end of follow-up. Time was rescaled in order to consider only the dynamics of antibody concentrations from 21 days after boost

immunization: this time point was redefined as the origin of time. More precisely, for groups receiving boost at day 29, data were rescaled from day 50 and available measurements were then at day 0, 130, 190 and 310. For groups receiving boost at day 57, data were rescaled from day 78 with available measurements at day 0, 102, 162, 282. As two observation points (at least) were needed before and after the value of τ to estimate the two slopes in all groups, we chose $\tau = 150$ days on the rescaled time. Covariates such as age and BMI were centered around the mean value of the study population. We also used the variable relative to vaccine regimens (order and interval) and geographic settings. This last categorical variable was either the geographic region (=0 for Europe and 1 for East Africa) or the trial (=0 for UK, 1 for Kenya, 2 for Uganda/Tanzania). Finally, the vector of covariates was :

$$Z = (age, sex, BMI, order, interval, geographic\_setting, order * interval) \tag{10}$$

We estimated the effect of covariates Z on the peak value of antibodies (intercept) and on the decreasing slopes of antibody concentrations. For individual i at rescaled time j, we write the corresponding antibody concentration Ab_{ij}. Linear mixed models can be written as :

$$(SS): \quad log_{10}(Ab_{ij}) = \beta_0 + \gamma_{0i} + \beta_1 t_{ij} + \beta_{cov}^T Z_i + \beta_{cov_t}^T Z_i t_{ij} + \epsilon_{ij} \tag{11}$$

$$\begin{aligned}(CS): \quad log_{10}(Ab_{ij}) &= \beta_0 + \gamma_{0i} + \beta_{cov}^T Z_i + \beta_b t_{ij} 1_{\{t_{ij}<\tau\}} + \beta_{cov_b}^T Z_i t_{ij} 1_{\{t_{ij}<\tau\}} \\ &\quad + \beta_a t_{ij} 1_{\{t_{ij}\geq\tau\}} + \beta_{cov_a}^T Z_i t_{ij} 1_{\{t_{ij}\geq\tau\}} + \epsilon_{ij}\end{aligned} \tag{12}$$

where $1_{\{t<\tau\}}$ and $1_{\{t\geq\tau\}}$ are equal to 1 when $t < \tau$ and $t \geq \tau$ respectively, 0 otherwise. In both cases, $\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$. We first realized backward selection on SS model using the geographic region variable. At each step, covariate with the highest p-value of Student test for β (> 0.05) was removed from the model. Performance of the models was assessed with Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). At the end of this first selection, CS model was estimated, using only selected variables. At this point, no additional selection was needed. In the final selected model, we also evaluated the trial variable instead of the geographic region.

After selection process, the best SS model was the following:

$$\begin{aligned}log_{10}(Ab_{ij}) &= \beta_0 + \gamma_{0i} + \beta_{age} age_i + \beta_{order} order_i \\ &+ \beta_{interval} interval_i + \beta_{gr} geographic\_region_i \\ &+ \beta_1 t_{ij} + \beta_{interval_t} interval_i t_{ij} \\ &+ \beta_{gr_t} geographic\_region_i t_{ij} + \epsilon_{ij}\end{aligned} \tag{13}$$

Variables age, order, interval and geographic region have a statistically significant effect on the value of antibody concentration 21 days post boost, and only variables interval and geographic region have a significant effect on the decreasing slope of antibody concentration. Using a CS model significantly improved the BIC criterium (BIC of SS model = 660.5, BIC of CS model = 532.7). However, using the trial variable instead of the geographic region variable improved the AIC criterium but not the BIC one (AIC/BIC of CS model with geographic region variable = 473.8/532.7, AIC/BIC of CS model with trial variable = 463.6/536.0). Table 3 shows the results of the CS linear mixed model using the trial variable. The biphasic decay is well captured by this model, as it can be seen that the decrease is estimated to be stronger before 150 days post boost than after, for all groups in all trials. Overall, we see that antibody concentrations have similar values 21 days after boost immunization across countries, with higher values when subjects are boosted at day 57. The decrease is lower in European subjects compared to East African ones, both before and after 150 days after boost immunization. It can be noted that subject characteristics BMI and sex were not statistically associated to antibody concentrations, and age was only associated to the concentration 21 days after boost immunization but not the decrease. Adjusted on other covariates, an increase of 10 years in age induces a reduction of 0.10 $\log_{10}$ of antibody concentration at 21 days after boost immunization (confidence interval -0.17;-0.038), and in the trials population, 50% of the subjects were aged 22-35 years. It is clinically less important than the order of vaccine immunizations, as the MVA/Ad26 regimen compared to the Ad26/MVA induces higher concentrations at 21 days after boost immunization of 0.18 $\log_{10}$ (confidence interval 0.086;0.28) and a boost at day 57 compared to boost at day 29 induces higher concentrations at 21 days after boost immunization of 0.27 $\log_{10}$ (confidence interval 0.15;0.38).

TABLE 3 Results of the CS linear mixed model

	Europe	East Africa	
	UK	Kenya	Uganda/Tanzania
Antibody concentrations 21 days post boost (in $\log_{10}$ ELISA units/mL)			
Group MVA/Ad26 D29	3.94 [3.81;4.07]	3.80 [3.67;3.93]	3.69 [3.67;3.93]
Group MVA/Ad26 D57	4.21 [4.08;4.34]	4.07 [3.94;4.20]	3.96 [3.94;4.20]
Group Ad26/MVA D29	3.76 [3.63;3.89]	3.62 [3.49;3.75]	3.51 [3.49;3.75]
Group Ad26/MVA D57	4.03 [3.90;4.16]	3.89 [3.76;4.01]	3.78 [3.76;4.01]
Slope value before 150 days post boost (in $\log_{10}$ ELISA Units/mL per 30 days)			
Group MVA/Ad26 D29	-0.075 [-0.10;-0.048]	-0.20 [-0.23;-0.17]	-0.17 [-0.19;-0.14]
Group MVA/Ad26 D57	-0.15 [-0.18;-0.12]	-0.28 [-0.31;-0.25]	-0.24 [-0.27;-0.21]
Group Ad26/MVA D29	-0.075 [-0.10;-0.048]	-0.20 [-0.23;-0.17]	-0.17 [-0.19;-0.14]
Group Ad26/MVA D57	-0.15 [-0.18;-0.12]	-0.28 [-0.31;-0.25]	-0.24 [-0.27;-0.21]
Slope value after 150 days post boost (in $\log_{10}$ ELISA Units/mL per 30 days)			
Group MVA/Ad26 D29	-0.038 [-0.049;-0.027]	-0.12 [-0.13;-0.11]	-0.089 [-0.10;-0.078]
Group MVA/Ad26 D57	-0.086 [-0.098;-0.074]	-0.16 [-0.18;-0.15]	-0.14 [-0.15;-0.12]
Group Ad26/MVA D29	-0.038 [-0.049;-0.027]	-0.12 [-0.13;-0.11]	-0.089 [-0.10;-0.078]
Group Ad26/MVA D57	-0.086 [-0.098;-0.074]	-0.16 [-0.18;-0.15]	-0.14 [-0.15;-0.12]

This preliminary analysis showed the importance of modeling the biphasic decay of antibody concentrations, as a CS model was better than a SS one. Moreover, it highlighted the differences in immune response between East African and European subjects, especially on the decreasing slope of antibody concentrations. Finally, no subject-specific factors had an effect on the dynamics of antibody concentrations except for age, but with a lower impact than geographic region and vaccine-related factors. Only these last factors were considered to potentially affect the dynamics of the humoral immune response in the mechanistic model.

A final check was realized after parameter ϕ_L of the mechanistic model was estimated to be significantly different between East African and European subjects. As the proportion of women included in the UK trial is higher than the one in East Africa (64% versus 29% and 20%) and the average age is 10 years higher in the UK trial, as seen in Table 1, the variables age and sex were tested separately as additional covariates on the parameter ϕ_L, with or without the geographic region variable. Without the geographic region variable, the estimated effect of age and sex was not significant (p-value = 0.54 and 0.23 respectively). With the geographic variable, the estimated effect of sex was not significant either (p-value = 0.46) and the effect of age was significant (p-value = 0.045) but with a low magnitude compared to the effect of geographic region ($\beta = 0.024$ for a 10-years difference versus a difference of $\beta = 1.36$ between European and East African subjects). These results suggested that the difference of ϕ_L value between the geographic regions could not be explained by potential confounding demographic factors.

ACKNOWLEDGMENTS

We thank the members of the EBOVAC1 consortium, in particular the principal investigators of the EBOVAC1 trials, Matthew Snape, Omu Anzala, George Praygod, Zacchaeus Anywaine. We also thank Houreratou Barry, John Edmunds, Rosalind Eggo, Mario Jendrossek, David C. Kaslow, Christine Lacabaratz, Bailah Leigh, Jerry Sadoff, An Vermeulen, Deborah Watson-Jones for scientific discussions.

Computer time for this study was provided by the computing facilities MCIA (Mésocentre de Calcul Intensif Aquitain) of the Université de Bordeaux and of the Université de Pau et des Pays de l'Adour.

This work has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under Grant Agreement EBOVAC1 (No. 115854). This Joint Undertaking receives support from the European Union's Horizon

2020 research and innovation programme and EFPIA. The funder of the study had no role in the study design, data collection, data analysis, data interpretation, or writing the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

This work was supported by the Investissements d'Avenir program managed by the ANR under reference ANR-10-LABX-77. The funder had no role in the study design, data collection, data analysis, data interpretation, or writing the report.

Competing interests: TVE, VB, LS and MD are employees of Janssen Pharmaceuticals and may be Johnson & Johnson stockholders.

REFERENCES

1. World Health Organization, Situation report. Ebola virus disease. 10 June 2016. Geneva, Switzerland: WHO, 2016; 2017.
2. Lu S. Heterologous prime–boost vaccination. Curr. opinion immunology 2009; 21(3):346–351.
3. Barefoot B, Thornburg NJ, Barouch DH, Yu Js, Sample C, Johnston RE, Liao HX, Kepler TB, Haynes BF, Ramsburg E. Comparison of multiple vaccine vectors in a single heterologous prime-boost trial. Vaccine 2008; 26(48):6108–6118.
4. Reyes-Sandoval A, Berthoud T, Alder N, Siani L, Gilbert SC, Nicosia A, Colloca S, Cortese R, Hill AV. Prime-boost immunization with adenoviral and modified vaccinia virus Ankara vectors enhances the durability and polyfunctionality of protective malaria CD8+ T-cell responses. Infect. immunity 2010; 78(1):145–153.
5. Baden LR, Karita E, Mutua G, Bekker LG, Gray G, Page-Shipp L, Walsh SR, Nyombayire J, Anzala O, Roux S, et al. Assessment of the Safety and Immunogenicity of 2 Novel Vaccine Platforms for HIV-1 PreventionA Randomized TrialAssessment of 2 Novel Vaccine Platforms for HIV-1 Prevention. Annals internal medicine 2016; 164(5):313–322.
6. Callendret B, Vellinga J, Wunderlich K, Rodriguez A, Steigerwald R, Dirmeier U, Cheminay C, Volkmann A, Brasel T, Carrion R, et al. A prophylactic multivalent vaccine against different filovirus species is immunogenic and provides protection from lethal infections with Ebolavirus and Marburgvirus species in nonhuman primates. PloS one 2018; 13(2):e0192312.
7. Milligan ID, Gibani MM, Sewell R, Clutterbuck EA, Campbell D, Plested E, Nuthall E, Voysey M, Silva-Reyes L, McElrath MJ, et al. Safety and immunogenicity of novel adenovirus type 26–and modified vaccinia Ankara–vectored Ebola vaccines: a randomized clinical trial. Jama 2016; 315(15):1610–1623.
8. Winslow RL, Milligan ID, Voysey M, Luhn K, Shukarev G, Douoguih M, Snape MD. Immune Responses to Novel Adenovirus Type 26 and Modified Vaccinia Virus Ankara–Vectored Ebola Vaccines at 1 Year. Jama 2017; 317(10):1075–1077.
9. Mutua G, Anzala O, Luhn K, et al. Randomized clinical trial examining safety and immunogenicity of heterologous prime-boost Ebola vaccines, Ad26.ZEBOV and MVA-BN-Filo: 12-month data from Nairobi, Kenya. Manuscript submitted for publication; 2018.
10. Anywaine Z, Whitworth H, Kaleebu P, et al. Randomized clinical trial examining safety and immunogenicity of heterologous prime-boost Ebola vaccines, Ad26 and MVA: 12-month data from Uganda and Tanzania. Manuscript submitted for publication; 2018.
11. Eurosurveillance editorial team. First Innovative Medicines Initiative Ebola projects launched. Eurosurveillance 2015; 20(3):21014.
12. Sullivan NJ, Martin JE, Graham BS, Nabel GJ. Correlates of protective immunity for Ebola vaccines: implications for regulatory approval by the animal rule. Nat. Rev. Microbiol. 2009; 7(5):393–400.
13. Wong G, Richardson JS, Pillet S, Patel A, Qiu X, Alimonti J, Hogan J, Zhang Y, Takada A, Feldmann H, et al. Immune parameters correlate with protection against ebola virus infection in rodents and nonhuman primates. Sci. translational medicine 2012; 4(158):158ra146–158ra146.
14. Dye JM, Herbert AS, Kuehne AI, Barth JF, Muhammad MA, Zak SE, Ortiz RA, Prugar LI, Pratt WD. Postexposure antibody prophylaxis protects nonhuman primates from filovirus disease. Proc. Natl. Acad. Sci. 2012; 109(13):5034–5039.
15. Nommensen F, Go S, MacLaren D. Half-life of HBs antibody after hepatitis B vaccination: an aid to timing of booster vaccination. The Lancet 1989; 334(8667):847–849.
16. Bailleux F, Coudeville L, Kolenc-Saban A, Bevilacqua J, Barreto L, André P. Predicted long-term persistence of pertussis antibodies in adolescents after an adolescent and adult formulation combined tetanus, diphtheria, and 5-component acellular pertussis vaccine, based on mathematical modeling and 5-year observed data. Vaccine 2008; 26(31):3903–3908.
17. Cheuvart B, Burgess M, Zepp F, Mertsola J, Wolter J, Schuerman L. Anti-diphtheria antibody seroprotection rates are similar 10 years after vaccination with dTpa or DTPa using a mathematical model. Vaccine 2004; 23(3):336–342.
18. Desai K, Coudeville L, Bailleux F. Modelling the long-term persistence of neutralizing antibody in adults after one dose of live attenuated Japanese encephalitis chimeric virus vaccine. Vaccine 2012; 30(15):2510–2515.
19. López EL, Contrini MM, Mistchenko A, Kieffer A, Baggaley RF, Di Tanna GL, Desai K, Rasuli A, Armoni J. Modeling the long-term persistence of hepatitis A antibody after a two-dose vaccination schedule in Argentinean children. The Pediatr. infectious disease journal 2015; 34(4):417–425.
20. Theeten H, Van Herck K, Van Der Meeren O, Crasta P, Van Damme P, Hens N. Long-term antibody persistence after vaccination with a 2-dose Havrix™(inactivated hepatitis A vaccine): 20 years of observed data, and long-term model-based predictions. Vaccine 2015; 33(42):5723–5727.
21. Schwarz T, Spaczynski M, Kaufmann A, Wysocki J, Gałaj A, Schulze K, Suryakiran P, Thomas F, Descamps D. Persistence of immune responses to the HPV-16/18 AS04-adjuvanted vaccine in women aged 15–55 years and first-time modelling of antibody responses in mature women: results from an open-label 6–year follow-up study. BJOG: An Int. J. Obstet. & Gynaecol. 2015; 122(1):107–118.
22. Fraser C, Tomassini JE, Xi L, Golm G, Watson M, Giuliano AR, Barr E, Ault KA. Modeling the long-term antibody response of a human papillomavirus (HPV) virus-like particle (VLP) type 16

prophylactic vaccine. Vaccine 2007; 25(21):4324–4333.
23. David MP, Van Herck K, Hardt K, Tibaldi F, Dubin G, Descamps D, Van Damme P. Long-term persistence of anti-HPV-16 and-18 antibodies induced by vaccination with the AS04-adjuvanted cervical cancer vaccine: modeling of sustained antibody responses. Gynecol. oncology 2009; 115(3):S1–S6.
24. Le D, Miller JD, Ganusov VV. Mathematical modeling provides kinetic details of the human immune response to vaccination. Front. cellular infection microbiology 2015; 4:177.
25. De Boer RJ, Oprea M, Antia R, Ahmed R, Perelson AS, Murali-krishna K. Recruitment Times, Proliferation, and Apoptosis Rates during the $CD8^+$ T-Cell Response to Lymphocytic Choriomeningitis Virus. J. Virol. 2001; .
26. Antia R, Bergstrom CT, Pilyugin SS, Kaech SM, Ahmed R. Models of CD8+ Responses: 1. What is the Antigen-independent Proliferation Program. J. Theor. Biol. 2003 Apr; 221(4):585–598. http://linkinghub.elsevier.com/retrieve/pii/S0022519303932085.
27. Lanzavecchia A, Sallusto F. Human B cell memory. Curr. opinion immunology 2009; 21(3):298–304.
28. Hammarlund E, Thomas A, Amanna IJ, Holden LA, Slayden OD, Park B, Gao L, Slifka MK. Plasma cell survival in the absence of B cell memory. Nat. communications 2017; 8(1):1781.
29. Huang KYA, Li CKF, Clutterbuck E, Chui C, Wilkinson T, Gilbert A, Oxford J, Lambkin-Williams R, Lin TY, McMichael AJ, et al. Virus-specific antibody secreting cell, memory B-cell, and sero-antibody responses in the human influenza challenge model. The J. infectious diseases 2014; 209(9):1354–1361.
30. Odendahl M, Mei H, Hoyer BF, Jacobi AM, Hansen A, Muehlinghaus G, Berek C, Hiepe F, Manz R, Radbruch A, et al. Generation of migratory antigen-specific plasma blasts and mobilization of resident plasma cells in a secondary immune response. Blood 2005; 105(4):1614–1621.
31. Frölich D, Giesecke C, Mei HE, Reiter K, Daridon C, Lipsky PE, Dörner T. Secondary immunization generates clonally related antigen-specific plasma cells and memory B cells. The J. Immunol. 2010; 185(5):3103–3110.
32. Berkman SA, Lee ML, Gale RP. Clinical uses of intravenous immunoglobulins. Annals internal medicine 1990; 112(4):278–292.
33. Brekke OH, Sandlie I. Therapeutic antibodies for human diseases at the dawn of the twenty-first century. Nat. Rev. Drug Discov. 2003; 2(1):52–62.
34. O'dempsey TJ, McArdle T, Ceesay SJ, Secka O, Demba E, Banya WA, Francis N, Greenwood BM. Meningococcal antibody titres in infants of women immunised with meningococcal polysaccharide vaccine during pregnancy. Arch. Dis. Childhood-Fetal Neonatal Ed. 1996; 74(1):F43–F46.
35. Leuridan E, Hens N, Hutse V, Aerts M, Van Damme P. Kinetics of maternal antibodies against rubella and varicella in infants. Vaccine 2011; 29(11):2222–2226.
36. Brinkhof MW, Mayorga O, Bock J, Heininger U, Herzog C. Kinetics of maternally acquired anti-hepatitis A antibodies: prediction of waning based on maternal or cord blood antibody levels. Vaccine 2013; 31(11):1490–1495.
37. Vilajeliu A, Ferrer L, Munrós J, Goncé A, López M, Costa J, Bayas JM, Group PW, et al. Pertussis vaccination during pregnancy: Antibody persistence in infants. Vaccine 2016; 34(33):3719–3722.
38. Voysey M, Pollard AJ, Sadarangani M, Fanshawe TR. Prevalence and decay of maternal pneumococcal and meningococcal antibodies: A meta-analysis of type-specific decay rates. Vaccine 2017; 35(43):5850–5857.
39. Amanna IJ, Carlson NE, Slifka MK. Duration of humoral immunity to common viral and vaccine antigens. New Engl. J. Medicine 2007; 357(19):1903–1915.
40. Ho F, Lortan JE, MaClennan I, Khan M. Distinct short-lived and long-lived antibody-producing cell populations. Eur. journal immunology 1986; 16(10):1297–1301.
41. Slifka MK, Antia R, Whitmire JK, Ahmed R. Humoral immunity due to long-lived plasma cells. Immunity 1998; 8(3):363–372.
42. Manz RA, Hauser AE, Hiepe F, Radbruch A. Maintenance of serum antibody levels. Annu. Rev. Immunol. 2005; 23:367–386.
43. Radbruch A, Muehlinghaus G, Luger EO, Inamine A, Smith KG, Dörner T, Hiepe F. Competence and competition: the challenge of becoming a long-lived plasma cell. Nat. Rev. Immunol. 2006; 6(10):741–750.
44. Andraud M, Lejeune O, Musoro JZ, Ogunjimi B, Beutels P, Hens N. Living on three time scales: the dynamics of plasma cell and antibody populations illustrated for hepatitis A virus. PLoS Comput. Biol 2012; 8(3):e1002418.
45. Amanna IJ, Slifka MK. Mechanisms that determine plasma cell lifespan and the duration of humoral immunity. Immunol. reviews 2010; 236(1):125–138.
46. Mitchell R, Kelly DF, Pollard AJ, Trück J. Polysaccharide-specific B cell responses to vaccination in humans. Hum. vaccines & immunotherapeutics 2014; 10(6):1661–1668.
47. Carter MJ, Mitchell RM, Meyer Sauteur PM, Kelly DF, Trück J. The antibody-secreting cell response to infection: Kinetics and clinical applications. Front. immunology 2017; 8:630.
48. Commenges D, Joly P, Gégout-Petit A, Liquet B. Choice between Semi-parametric Estimators of Markov and Non-Markov Multi-state Models from Coarsened Observations. Scand. J. Stat. 2007; 34(1):33–52.
49. Muyanja E, Ssemaganda A, Ngauv P, Cubas R, Perrin H, Srinivasan D, Canderan G, Lawson B, Kopycinski J, Graham AS, et al. Immune activation alters cellular and humoral responses to yellow fever 17D vaccine. The J. clinical investigation 2014; 124(7):3147.
50. LaBeaud AD, Malhotra I, King MJ, King CL, King CH. Do Antenatal Parasite Infections Devalue Childhood Vaccination? PLOS Neglected Trop. Dis. 2009 05; 3(5):1–6. https://doi.org/10.1371/journal.pntd.0000442.
51. Lalor MK, Ben-Smith A, Gorak-Stolinska P, Weir RE, Floyd S, Blitz R, Mvula H, Newport MJ, Branson K, McGrath N, et al. Population differences in immune responses to Bacille Calmette-Guerin vaccination in infancy. The J. infectious diseases 2009; 199(6):795–800.
52. Prague M, Commenges D, Guedj J, Drylewicz J, Thiébaut R. NIMROD: A program for inference via a normal approximation of the posterior in models with random effects based on ordinary differential equations. Comput. methods programs biomedicine 2013; 111(2):447–458.
53. Guedj J, Thiébaut R, Commenges D. Maximum likelihood estimation in dynamical models of HIV. Biometrics 2007; 63(4):1198–1206.
54. Commenges D, Jacqmin-Gadda H, Proust C, Guedj J. A newton-like algorithm for likelihood maximization: The robust-variance scoring algorithm. arXiv preprint math/0610402 2006; .
55. Drylewicz J, Commenges D, Thiebaut R. Maximum a posteriori estimation in dynamical models of primary HIV infection. Stat. Commun. Infect. Dis. 2012; 4(1).

3.4 Discussion

3.4.1 Additional insights

We have shown that a simple model of the humoral immune response is able to capture the dynamics of the antibody concentrations and some of the factors associated with their variability. The obtained results could be strengthened in several ways. Additional measurements with a longer follow-up could help refining the estimation of the LL ASCs half-life. This could also help confirming the different levels at which antibodies are sustained in European and African participants. This estimation refinement should be made possible by a follow-up of 4 years on subjects enrolled in the EBOVAC1 Sierra Leona trial and some of the African subjects enrolled in EBOVAC2.

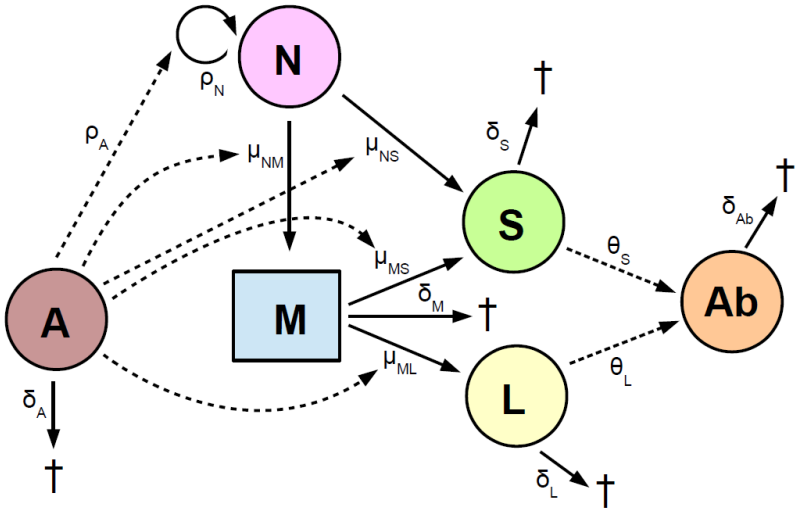
Moreover, it would be interesting to apply the same model to data collected in the consortium (to maintain consistency between the clinical protocols), but on other subjects. In particular, the Ad26/MVA was tested on a number of subjects enrolled in phase 2 trials in Europe and Africa. Using the antibody concentrations data from all these subjects will help refining the estimation of the parameters of the model and the effect of the factors on the variability of the response. Also, a clinical trial was conducted in Sierra Leone under the EBOVAC1 consortium. As we found differences of dynamics between European and East African subjects, it would be beneficial to check if West African subjects also have different antibodies dynamics induced by the parameters values by using the data generated in both the EBOVAC1 Sierra Leone and the EBOVAC2 trials. It would help determining if the observed dissimilarities in antibody concentrations after one year of follow-up are mainly between European and African subjects, or reveal differences at a more regional level. This could be explained by other region-specific environmental factors, and/or genetic and/or co-infection parameters. To assess the effect of co-infection, it would be useful to work on the sub-study of the Sierra Leone measuring the presence of malaria infection at baseline. This could be added as a covariate which effect can be tested on parameters of the mechanistic model. The effect of HIV infection could also be assessed by using data generated in the EBOVAC2 clinical trial, as an HIV+ cohort has been recruited. Moreover, children have been recruited in both the EBOVAC1 Sierra Leone clinical trial and EBOVAC2 in West Africa. Using these data would help estimating the effect of age on the immune response to immunizations. Finally, some participants enrolled in the phase 2 trials in Africa received their boost immunizations at different times than the one initially planned by the original clinical design; this heterogeneity could be useful in the modeling approach to understand more deeply the impact of the delay between immunizations on the immune response.

3.4.2 Through more complete models of the humoral immune response

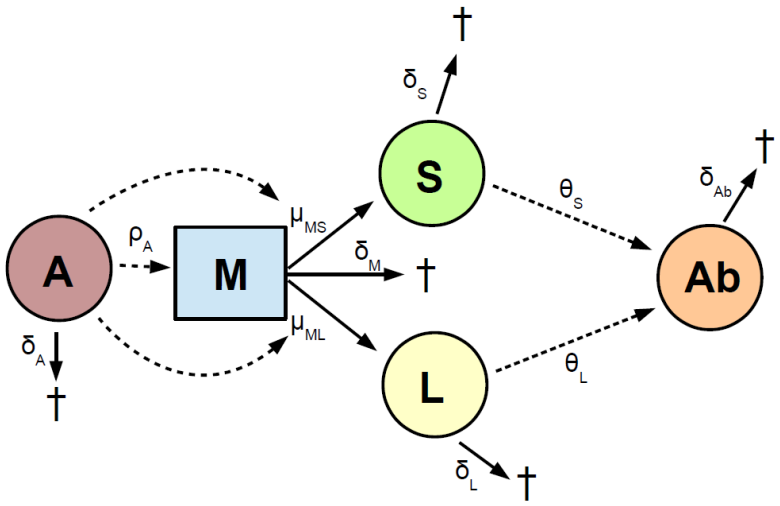
The modeling work of the immune response to Ebola vaccine was initiated by using a mechanistic model with only 3 compartments (two compartments of ASCs and a compartment of antibodies). Although simple, this model induced good fit of the antibody concentrations data, helped quantifying the dynamics of the humoral response and gave insights into the potential variability factors. However, we are aiming for a more complete model of the immune response. This would help refining the quantification of the dynamics of the response and understanding the role of the different populations of B lymphocytes in the establishment and maintenance of the response. In particular, it would be interesting to capture the proliferation phase of the B lymphocytes following a vaccine immunization and their differentiation into memory cells. Overall, a more complete model would give better understanding of the mechanisms of action of the immunological memory and the impact of the different vaccine regimens on the immune response: this could help determining optimized vaccine regimens and predicting the effect of a third injection/a natural infection by EBOV, especially in African populations.

Investigation of more complete models of the humoral immune response has been realized with Irene Balelli, during her post-doc project. First, due to its role in secondary responses, immunological memory should be included in the model. Moreover, heterologous prime-boost regimens rely on the use of different vaccine platforms, which could induce different reactions of the immune system; antigen should also be included in the model. Finally, as ASCs are produced by differentiation of the naive B cells, these should be added to the model as well. A graphic representation of a more complete model of the humoral immune response is shown in figure 3.2A. This model is however very complex due to its large number of compartments and parameters. A simplified version without the naive compartment was considered (see figure 3.2B): this model was shown to induce similar dynamics with less parameters, so it can improve the initial model without adding too much complexity in it. The corresponding equations are the following:

$$\begin{cases} \dfrac{dA}{dt} = -\delta_A A, \\ \dfrac{dM}{dt} = \rho_A A - (\mu_{MS} + \mu_{ML})AM - \delta_M M, \\ \dfrac{dS}{dt} = \mu_{MS} AM - \delta_S S, \\ \dfrac{dL}{dt} = \mu_{ML} AM - \delta_L L, \\ \dfrac{dAb}{dt} = \theta_S S + \theta_L L - \delta_{Ab} Ab. \end{cases}$$



A: Complete mechanistic model of the humoral immune response



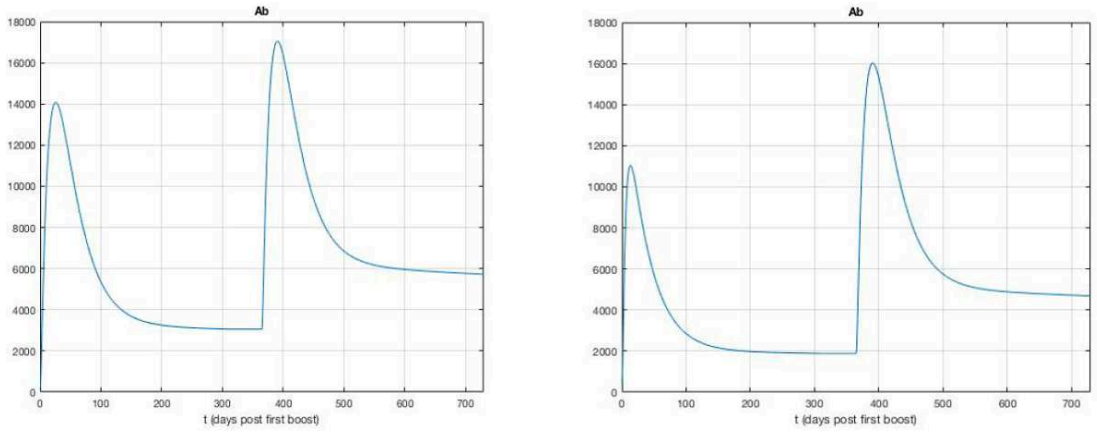
B: Simplified mechanistic model of the humoral immune response

Figure 3.2 – Models of the humoral immune response. A: antigen, N: naive B cells, M: memory B cells, S: short-lived ASCs, L: long-lived ASCs, Ab: antibodies. Courtesy of I. Balelli.

In this model, as the vaccine platforms are not recombinant, antigen decrease exponentially. The presence of antigen (A) artificially triggers the generation of memory cells (M). This compartment contains actually both activated B cells specializing against the antigen and memory B cells when the reaction is over. Under the presence of the antigen, memory B cells differentiate into either short-lived ASCs or long-lived ones (L). The compartments of short and long-lived ASCs are not renewed. These cells produce antibodies and die at different rates, in the same way as in the model presented in the article section 3.3.

Initial analyses of this model included identifiability investigation, sensitivity analysis and calibration of the model using the antibody concentrations data from EBOVAC1 clinical trials conducted in UK, Kenya and Uganda/Tanzania. The next step is the estimation of the parameters using additional data to quantify the dynamics of the model. In short, structural identifiability analysis showed that we can reach local identifiability by observing compartments M, S+L and Ab. It means that the antibody concentrations data only is not sufficient to correctly estimate the parameters of the model. However, this identifiability analysis makes the hypothesis that we wish to estimate the values of all parameters using "perfect" data (observable at every time with no measurement error). This is obviously not the case in real data, available at sparse time points with a measurement error. Moreover, a population approach aims at estimating distributions of parameters, which is different from a patient-by-patient approach where trajectories are estimated one-by-one. Sensitivity analysis has allowed to determine the relative variation of the antibody concentrations trajectory with respect to the parameters of the model during time. It helps understanding which parameters have more impact on the variation of the antibody concentrations, and at which moment of the time frame.

Finally, calibration was realized under consideration of previous knowledge on some parameters and additional hypothesis. Indeed, parameters δ_S, δ_L and δ_{Ab} were already estimated. Moreover, some biodistribution information were used for calibrating parameter δ_A [Sheets et al., 2008; Hanke et al., 2005]. Parameters ρ_A, μ_{MS} and μ_{ML} were expected to be platform-dependent. In particular, calibrated values are such that μ_{MS} and μ_{ML} are lower with MVA than with Ad26. It can explain why after a single injection of MVA, antibodies are not detectable, as there are not enough memory cells who differentiate into ASCs. This calibrated model was also used to predict the effect of a third immunization one year after the boost immunization and showed for example that in any case the antibody concentrations were increased compared to the second immunization, due to the presence of LL plasma cells and memory cells, and that the MVA/Ad26/Ad26 regimen could induce a higher antibody peak than the Ad26/MVA/Ad26 (see figure 3.3). This work should definitely be confirmed by parameters estimation using additional data generated in EBOVAC1 and EBOVAC2 consortium, in particular some B cells data generated from the EBOVAC2 UK subjects.



A: MVA/Ad26/Ad26 regimen

B: Ad26/MVA/Ad26 regimen

Figure 3.3 – Prediction from the calibrated model of the antibody concentrations under two regimens of vaccination consisting of 3 immunizations at 56 then 365 days of delay. In the plot, time starts from the second immunization. Courtesy of I. Balelli.

3.4.3 Through a systems vaccinology approach

The development of vaccine has mainly been empirical until now and there is a consensus on the need for a better understanding of the mechanisms involved in the establishment of the immune response to drive a more rationale vaccine development. Systems biology is a promising approach to improve this field of research and has broadly emerged in the last years [Pulendran et al., 2010; Germain, 2010; Hagan et al., 2015]. It aims at understanding the immune system as a whole entity by accounting for a large number of its actors and integrating all the newly available data generated thanks to biotechnology progress (genomic, expression profiling, proteomic, RNA sequencing, ..). Moreover, this approach is supposed to be quantitative and to be able to assess the respective contributions of the compartments of the immune response [Germain, 2017]. This kind of approach raises many challenges, including both the methods for data collection and for the computational analysis of the complex data sets to identify relationships between the numerous markers involved.

In vaccine development, aiming for a so-called systems vaccinology approach involves the inclusion of the information from genes, the microbiome and the environment: as shown in figure 3.4 and discussed in section 2.2.2.3, they are determinants of the human physiology, inducing different response to immunizations. In particular, we have shown in our modeling work how environment variables could be included in a mechanistic model of the immune response using covariates on the parameters. This approach should also aim

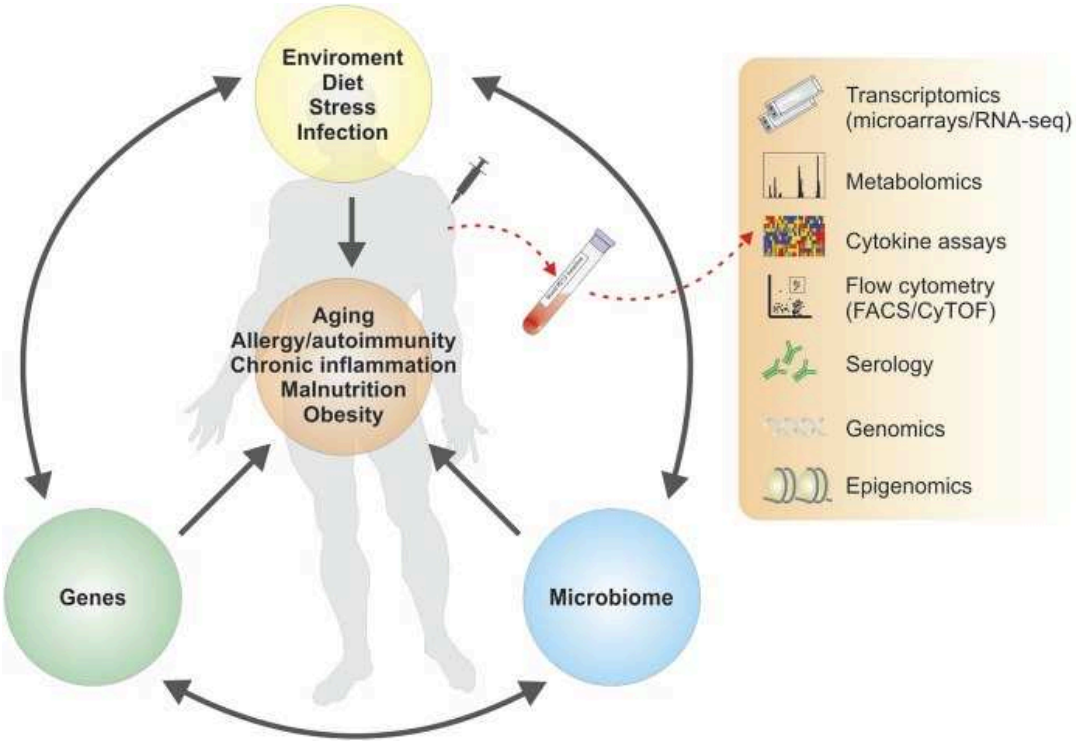


Figure 3.4 – Systems vaccinology approach. Figure taken from Pulendran [2014]

at including the early immune response following vaccination, for two main reasons. First, there has been increasing interest in understanding how the innate immune response can be modified by repeated immunizations and its role in the immune response. The review of Adams et al. [2016] shows that secondary NK cell responses, which are antigen-specific and produce more IFNγ, are very likely to happen and affect the immunity. Moreover, Palgen et al. [2018] have shown that prime and boost vaccine immunizations induced different phenotypic profiles of innate responses. Also, even though there is still work to do to understand the mechanisms by which innate immunity affects the generation of both memory T and B cells, there is evidence that the innate immune system can modify the protective immunity [Pulendran and Ahmed, 2006]. In addition, there is growing evidence on the fact that early markers of the immune response can be predictive of the efficacy of vaccines, or at least of the levels of biomarkers of interest. Only a few studies identified early innate immune signatures to predict the antibody response, in the case of yellow fever vaccine [Querec et al., 2009] and influenza vaccine [Nakaya et al., 2011]. More specifically on Ebola, a study on the rVSV-ZEBOV vaccine showed that 5 early innate markers were correlated with antibody concentrations at day 28 and after [Rechtien et al., 2017]. These markers included cytokines levels and specific cluster of innate cells (identified with phenotypic markers). These early signatures identifications could help rapid screening of vaccines with a quick determination of non responders and acceleration of the vaccine development pipelines. However, these identifications do not

necessary explain the mechanisms of immunity establishment.

In an effort of integrative analyses and quantification of the mechanisms of the immune response, it would be interesting to include the information provided by early markers in mechanistic models. Several methods could help the integration of this information: a simple way would be to add covariates on the parameters of the model, as done with the $CD4^+$ data on the antibody response in the article presented in section 3.3. For example, after identifying that a cytokine's concentration, a gene expression level, or a specific cell type's percentage at a given time point is associated to the antibody response, we could use this value as a covariate and test its effect on the parameters of the model. Let us imagine that the effect is significant on the half-life of the ASCs, but not on the antibody production rate, then it would mean that the presence cytokine or the innate cell population in question is associated to a better survival and/or proliferation of the ASCs, but not to their ability to generate antibodies. This method does not add complexity to the model itself, but makes a linear hypothesis regarding the association between the innate response marker levels and the parameters affecting the dynamics of the adaptive immune response, which could be a limitation of the modeling method. Another possibility is to model the dynamics of the innate marker and its underlying mechanisms of the effect on the compartments of the adaptive immune response: this can be realized by adding an equation to the ODE system, corresponding to the innate marker compartment, and accounting for its interactions with the other compartments of the system in the equations. It could eventually lead to a large system of ODEs for the dynamics of the immune response. However, this option has several drawbacks: the processes at different levels (genes, molecules, cell populations...) occur at different time scales. It could induce numerical issues. Moreover, there is some intrinsic stochasticity at the single-level that cannot be neglected and would not be accounted for in a deterministic ODE-based model. Finally, this option adds complexity to the model and can induce identifiability issues, if the data available is not rich enough compared to number of parameters added to the initial ODE model. Another way to use the gene expression data would be to obtain information on some unmeasured cell populations: if some genes are known to be expressed in majority by a population of cell X, their observed expression Y could be written as in equation (17), by:

$$Y^i(t_{ij}) = f(X^i(t_{ij})) + \epsilon_{ij}. \tag{48}$$

However, this method would request additional estimation techniques to determine function f and this adds a challenge to the parameters estimation method.

The transcriptomic data could be included by modeling the dynamics of gene expression after immunizations. Indeed, gene expression happens in two steps: first DNA is

transcribed into mRNA, then protein production is induced by the translation of the mRNA. The transcription of the gene DNA can be inhibited or activated by regulatory proteins (called transcription factors). Gene regulatory networks (GRN) have been widely used to capture these dynamics and interactions: their structure is constituted of nodes (genes, proteins) and edges (molecular interactions and relationships). This tool could be used to analyze the transcriptomic data. However, this kind of analysis raises several challenges. First, there is some large variability in gene expression, and it cannot be accounted simply by a perturbation around a population mean. There is a first difficulty in modeling this system with embedded stochasticity, and several methods have been developed, including bayesian networks, boolean networks, information theory, ODEs and piecewise deterministic Markov processes with deterministic trajectories defined with ODEs [de Jong, 2002; Hecker et al., 2009; Wu et al., 2014]. A major additional difficulty resides in estimating and inferring the model, both in statistical and computational terms. Indeed, there has been some work on how to infer GRN [Yu et al., 2004; Herbach et al., 2017] but it still represents a challenge [Marbach et al., 2010]. Moreover, even if a method is able to infer the gene network, the information provided should then be included in a mechanistic model of the biomarkers of interest. This question is far from trivial, and even if there has been some growing interest on methods for developing models integrating several types of data [Joyce and Palsson, 2006] there is still no general method and poor literature on how to integrate GRN modeling into a large-scale mechanistic model of the immune response.

In any case, the choice of the methodological tools for developing integrative models is directly related to the data available for the study. In the specific case of Ebola vaccine and more particularly the EBOVAC consortium, several datasets could provide information that would feed this kind of integrative model. In particular, NK cells subsets have been measured in the UK subjects of EBOVAC1 trial. Moreover, omic data will be obtained from European subjects included in EBOVAC2 trials.

Overall, our computation and modeling tools should be constantly improved to tackle this effort on integrating the different types of data generated during clinical trials. However, it should be kept in mind that in humans, the data is collected from blood samples and corresponds to circulating actors of the immune response. We know that major aspects of the immune response occur in lymphoid organs [Rappuoli et al., 2017]: for example, the differentiation of activated B cells into memory cells happens in the germinal centers and long-lived plasma cells are believed to migrate quickly to the bone marrow where they reside. Also, a study in nonhuman primates showed that most of the protective $CD8^+$ T cells were present in the liver compared to blood [Ishizuka et al., 2016]. This induces limitations in our understanding of some key mechanisms of the immune response

in humans. It could be included in the model, by modeling both circulating and tissue-residing cells in the ODE system, and accounting only for the circulating compartments in the observation part of the statistical model, but would induces parameters identifiability issues. Challenges remain in identifying blood signatures which are good predictors of the immune response occurring in other spatial compartments of the organism [Hagan and Pulendran, 2017].

4 Optimizing immune therapies in HIV-infected patients

Abstract: In this chapter, we explain why immunotherapy is considered for some HIV-infected patients. In particular, we focus on interventions based on interleukin 7 (IL-7) injections to help patients recovering acceptable levels of $CD4^+$ T lymphocytes when antiretroviral therapy is not sufficient. We recall the previous modeling work that was realized for quantifying the dynamics of $CD4^+$ T lymphocytes following IL-7 injections. As the mechanistic model has good ability to fit the data and predict the dynamics for new included patients, tools can be developed for optimizing and adapting the schedules of injections. In particular, we show that the process can be modeled using a piecewise deterministic Markov process (PDMP) and the control problem can be reduced to an impulse problem at the boundary of the state space. Based on theoretical results, this problem can be solved numerically using the dynamic programming approach by iterating an integro-differential operator, the so-called Bellman operator, leading to a sequence of functions converging to the value function. We show how our numerical method can help determining an optimal protocol of injections which allows administering fewer IL-7 injections than other naive clinical protocols but still maintains the patient at acceptable levels of $CD4^+$ T lymphocytes.

Key Words: Human Immunodeficiency Virus; immunotherapy; interleukin-7; immune reconstitution; repeated injections; piecewise deterministic Markov process; impulse control; dynamic programming; optimal protocols.

4.1 Biological and clinical context

4.1.1 General introduction on HIV

4.1.1.1 HIV epidemic

Since the start of the HIV epidemic more than 30 years ago, there has been 76.1 million of infections of people worldwide and 35 million of deaths related to AIDS [UNAIDS, 2017]. Prevention and treatment measures have allowed reducing by 16% the number

of transmission events between 2010 and 2016 and by 48% the number of deaths from AIDS-related causes between 2005 and 2016. However, HIV/AIDS still represent a major public health concern. At the end of 2016, approximately 36.7 million people were living with HIV; during that year, 1.8 million people were newly infected and 1.0 million people died of AIDS-related causes.

4.1.1.2 HIV infection

HIV is a retrovirus, characterized by a long period between the time of infection and the development of AIDS during which the immune system fails and eventually allows opportunistic infections to occur in the organism. Two types of HIV were identified, themselves containing several groups. The classification is mainly based on genetic differences. HIV-1 is present all over the world while HIV-2 is mainly present in West Africa. HIV-1 is related to SIV, which was transmitted either by chimpanzees (groups M, N and O) or gorillas (group P), probably first to bush meat hunters. Globally, most of the cases of HIV/AIDS are due to group M virus, which can itself be divided in nine subtypes or clades, from A to K.

HIV has a spherical structure. The viral envelope is composed by the glycoproteins gp120 and gp41. It contains a capsid with two single strands of RNA and enzymes allowing the replication of the virus. This replication requires human cells. In particular, HIV targets cells from the immune system, especially $CD4^+$ T lymphocytes but also macrophages and dendritic cells. HIV binds to receptors on the surface of the target cell through the CD4 receptor, and enters the cell after fusion of the envelope of the virus with the membrane of the cell. Reverse transcription allows then the conversion of the genetic material contained in HIV RNA into HIV DNA, which enters the cell nucleus. This viral DNA is integrated into the DNA of the cell. It allows the virus to use the cell system to generate HIV proteins. These proteins are assembled with HIV RNA at the surface of the cell to form new immature HIV. When out of the cell, protease release allows maturation to a new infectious virus. The whole cycle of replication of HIV is represented in figure 4.1.

Both humoral and cellular responses play their role in the attempt to eradicate HIV, thanks to antibody production and cytotoxic lymphocytes action. However, the immune system itself is weakened by the virus, as activated $CD4^+$ T lymphocytes are the most likely to be infected by HIV. After proliferation, some of these infected cells migrate to secondary lymphoid organs and can live there for several years. It creates the so-called HIV reservoirs. As HIV is integrated to their nucleus, it is impossible for the immune system to detect and eliminate the virus unless it starts to replicate.

The dynamics of the interaction between HIV and the immune system follow three

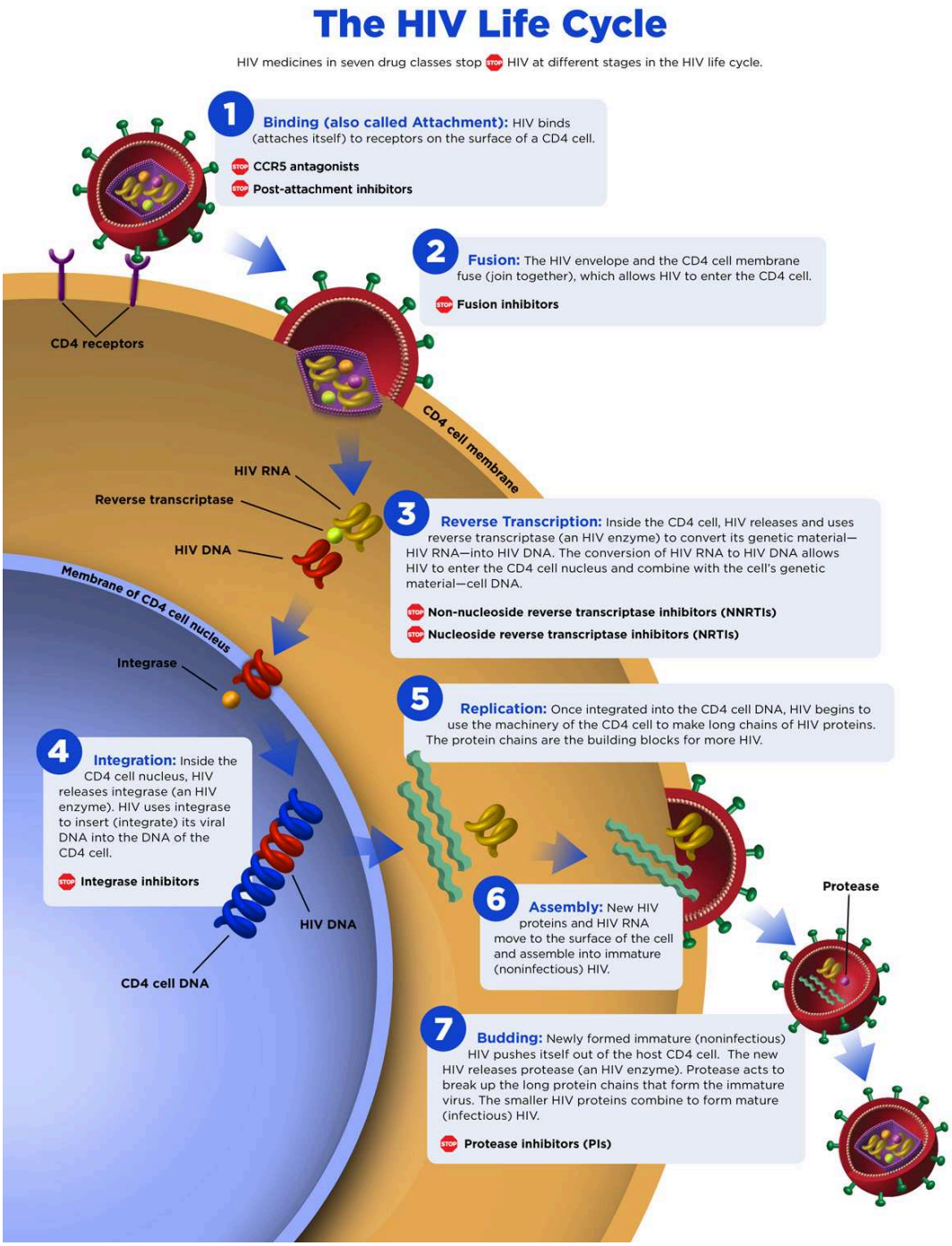


Figure 4.1 – HIV replication cycle and type of therapies targeting different steps of this cycle (AIDS info - NIH, 2017).

clinical phases: a primary infection phase lasting a few weeks after the infection, followed by an asymptomatic phase that can last several years and finally the AIDS phase. Primary infection is characterized by an initial burst of viremia and a fall of the $CD4^+$ T lymphocyte counts. During the asymptotic phase, viral load remains stable and the number of $CD4^+$ T lymphocytes is small but still acceptable (generally around 500-600

cells per μL of blood). Without any treatment, patients evolve to the AIDS phase after an average of 10 years. During this final phase, the immune system collapses, with a particular drop of the CD4 counts, and the viral load sharply increases. Patients usually die from opportunistic infections.

4.1.1.3 Combination Antiretroviral Therapy

Since the approval of the first antiretroviral molecule in 1987 by the FDA, a large number of treatments with different mechanisms of action have been introduced, now offering many possibilities to clinicians for treating HIV-infected patients. The first treatments were monotherapies based on Nucleotide Reverse Transcriptase Inhibitors (NRTI), a class of molecules interrupting HIV replication cycle via competitive inhibition of HIV reverse transcriptase and termination of the DNA chain (step 3 of figure 4.1) [Fischl et al., 1987]. They were shown to help decreasing the rate of progression to AIDS and increasing the CD4 counts in asymptomatic adults [Volberding et al., 1990] but drug resistances were also observed in early and late stages patients [Larder et al., 1989; Kozal et al., 1994]. Later, two new classes of antiretroviral molecules were introduced: the Protease Inhibitors (PI) which prevent viral formation after viral budding from infected cells (step 7 of figure 4.1), and the Non-Nucleotide Reverse Transcriptase Inhibitors (NNRTI), blocking step 3 of figure 4.1 as NRTIs. This significantly improved treatment strategies against HIV thanks to combination Antiretroviral Therapy (cART), which uses a minimum of two drugs from different classes and can help reducing the risk of mutations. More recently, two other types of molecules were made available: the Fusion Inhibitors (FI), in 2003, which prevent the entry of the virus in the target cell (step 2 of figure 4.1) and the Integrase Inhibitors (INI), in 2007, preventing viral DNA to integrate the host DNA (step 4 of figure 4.1). Several clinical trials have shown the increased efficacy of cART compared to monotherapies [Hammer et al., 1996; Delfraissy et al., 2008; Bierman et al., 2009]. The choice of the combination has evolved with time, and recommendations are currently published annually by health institutions, giving guidelines for clinical practice. In France, tritherapy is recommended to contain 2 NRTIs and a third agent, either a NNRTI, an INI, a PI alone or combined with ritonavir to improve pharmacokinetics properties [Morlat, 2017]. These regimens are preferred because of their efficacy, tolerability and they are easy to use.

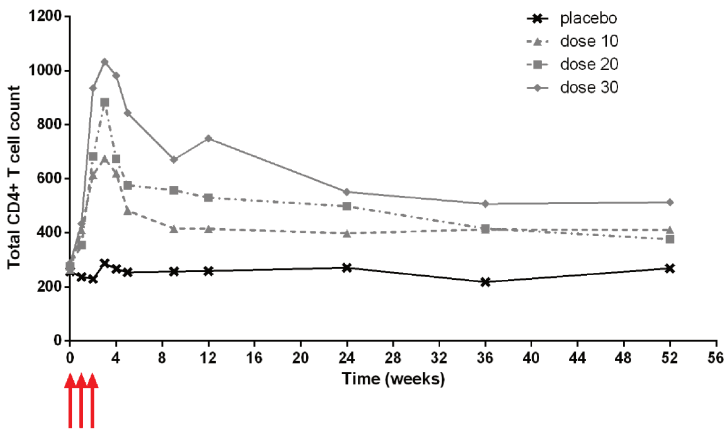
Overall, cART have improved life expectancy [Hammer et al., 1996; Autran et al., 1997; Palella Jr et al., 1998] and a large number of the patients who benefit from cART manage to control the virus and reconstitute their immune system [Battegay et al., 2006]. This has led to major improvement in reducing morbidity and mortality in HIV-infected patients [Young et al., 2012]. However, in some cases, we can observe either a virological

failure or an immunological failure. The virological failure is defined as the inability of the patient to control the virus and to reduce the viral load. In that case, the treatment should be modified, depending on some other parameters, such as the value of given biomarkers, the presence of other co-morbidities or the patient's habits. The reason for this failure should also be investigated: it could be due to non-adherence of the patient, meaning that medicine and doses prescriptions are not respected. Adherence constitutes a real challenge in the treatment of HIV-infected patients, but will not be developed in this thesis. We will focus here on the immunological failure: some patients are able to control the virus replication but cannot recover normal levels of $CD4^+$ T cells [Lange and Lederman, 2003]. Even if the percentage varies between studies, a significant subset of HIV-infected patients under treatment do not reach normal levels of $CD4^+$ T cells [Julg and Walker, 2009] – which lie between 500 and 1500 cells/μL in healthy individuals. In several European studies, this percentage was found to vary between 16 and 33% [Thiébaut et al., 2005]. In a study on 366 HIV-infected patients from cohorts in the USA, 44% of patients starting cART with CD4 counts below 100 cells/μL and 25% of the patients starting cART with CD4 counts between 100 and 200 cells/μL were found to have CD4 count below 500 cells/μL even after 7 years of cART [Kelley et al., 2009]. Among 400 patients from the UK Collaborative HIV Cohort Study, 7%, 33% and 65% of the participants were found to present CD4 counts below 200, 350 and 500 cells/μL respectively, after at least 2 years of treatment [O'Connor et al., 2014]. To evaluate the $CD4^+$ T cell response after treatment initiation, some studies have focused on providing curves of the evolution of the $CD4^+$ T cell response in large cohorts in Europe [Bouteloup et al., 2017] and South Africa [Yotebieng et al., 2015]. They showed that the median CD4 number after 12 months of cART was below 500 cells/μL and this number depends on the CD4 number at treatment initiation. A low CD4 gain was also found associated to an increased risk of death, consistent with other studies showing that low $CD4^+$ T cell numbers are associated with higher mortality, opportunistic infections or AIDS event [Thiébaut et al., 2005; Kitahata et al., 2009; Opportunistic Infections Project Team Of The Collaboration Of Observational HIV Epidemiological Research In Europe (COHERE) In EuroCoord et al., 2012]. It has also been shown that in HIV-infected patients, $CD4^+$ T cell counts above 500 cell/μL are associated with a nearly healthy clinical status [Lewden et al., 2007]. Immunological failure is then a major concern in HIV-infected patients. In that case, the treatment is not necessarily modified but co-infections are investigated and immunotherapies are considered to complement the cART. In the following section, we will introduce in particular the immune therapy with the cytokine interleukin-7 (IL-7).

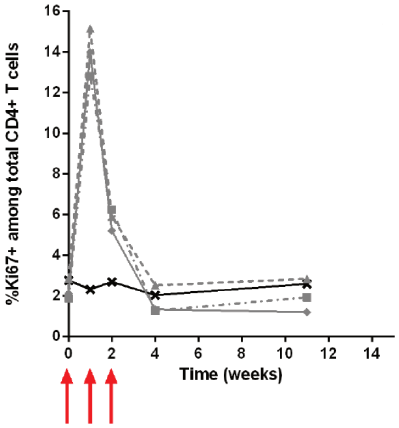
4.1.1.4 Immune therapy with interleukin-7

IL-7 is a cytokine produced by non-marrow-derived stromal and epithelial cells. It has been considered as a good candidate for intervention in HIV-infected patients with immunological failure because it may improve thymic production [Mackall et al., 2001; Okamoto et al., 2002] and cell survival, as shown *in vivo* [Tan et al., 2001; Vella et al., 1998; Leone et al., 2010]. A review showed that IL-7 has a role in thymic T cell development as well as survival, proliferation and maturation outside of the thymus [Beq et al., 2004]. Moreover, IL-7 was shown to regulate proliferation and survival of antiviral $CD4^+$ T cells [Lenz et al., 2004]. The safety and beneficial effect of injections of exogenous IL-7 was first established in an observational study [Camargo et al., 2009] and a phase I trial [Sereti et al., 2009]. Following encouraging results, the safety and immunogenicity of repeated administrations of recombinant human IL-7 were assessed in a prospective open-label phase 1/2a trial [Levy et al., 2009]. 14 HIV-infected patients under cART with $CD4^+$ counts between 100 and 400 cells/μL received 8 injections every two days in a 16 days period at dose 3μg/kg (6 patients) or 10μg/kg (8 patients) and were followed up to 48 weeks after the first injection. IL-7 was well tolerated and a significant dose-dependent increase of naive and central memory $CD4^+$ and $CD8^+$ T cells was observed, even at long term. Proliferation was also enhanced after injections, as observed thanks to increased expression of the Ki67 marker. These results suggested that IL-7 could have an effect on T cell cycling, thymic output and/or T cell survival. Then, a dose escalation study of repeated administration of a glycosylated rhIL7 (CYT107 - equivalent to IL-7) was realized in a multicenter phase 1/2a placebo controlled trial (INSPIRE study) [Levy et al., 2012]. Patients were divided in 3 groups receiving three weekly injections of IL-7, at day 0, 7 and 14 at doses 10μg/kg (7 patients + 2 placebos), 20μg/kg (8 patients + 2 placebos) or 30μg/kg (6 patients + 2 placebos). They were followed up to 52 weeks after the first injection. This trial confirmed the tolerability of IL-7 injections, as well as the dose-dependent increase of CD4 levels, up to 52 weeks after the first injection, as seen in figure 4.2A. In parallel with this increase, the measurements of Ki67 marker showed it underwent a peak of expression during IL-7 administration, with an observed maximum 7 days after the first injection (first available measurement) and a return to baseline expression between 7 and 14 days after the last injection of the cycle, as seen in figure 4.2B. It means that IL-7 stimulate cellular proliferation, which induces a peak of $CD4^+$ T cells shortly after the injections. Other potential effects on thymic output and cell survival are suggested by additional biomarkers measurements, but will not be developed here. Details can be found in Levy et al. [2012] and Thiébaut et al. [2014]. From tolerance and biologic activity, the dose of 20μg/kg was selected to be the most

clinically relevant and used in future trials. The following section presents the two other clinical trials realized and the associated modeling work.



A: Dose-dependent increase of total $CD4^+$ T cell counts



B: Percentage of $Ki67^+$ cells among $CD4^+$ T cells

Figure 4.2 – Dynamics of $CD4^+$ T cells following IL-7 injections in INSPIRE study. The weekly injections are represented with red arrows in the figures. Figures taken from Thiébaut et al. [2014].

4.1.2 Mathematical modeling of IL-7 immunotherapy

4.1.2.1 Effect of IL-7 injections on $CD4^+$ T cell dynamics

Following the analysis of the data from the INSPIRE study, a remaining question was to quantify the contribution of thymic production, peripheral proliferation and survival to the dynamics of $CD4^+$ T cells following IL-7 injections. Some modeling work was initiated in Thiébaut et al. [2014] to determine if peripheral cell proliferation was the only consequence of IL-7 injections, or if survival and thymopoiesis were also affected. A simple mechanistic model was developed : it consists in two compartments of $CD4^+$ T cells. One corresponds to proliferating (P) cells and the other one to resting (R), non-proliferating cells. Non-proliferating cells are produced by the thymus at constant rate λ. They enter proliferation at rate π and are lost at rate μ_R. Proliferating cells have a loss rate μ_P. When a cell cycle is over, the final division produces two non-proliferating daughter cells which enter compartment R at rate ρ. The loss rates μ_R and μ_P correspond to a combination of death rate and migration from peripheral blood to other compartments such as lymph nodes. The model is represented in figure 4.3 and the corresponding equations can be

written as:

$$\begin{cases} \dfrac{dR}{dt} &= \lambda + 2\rho P - \mu_R R - \pi R \\ \dfrac{dP}{dt} &= \pi R - \rho P - \mu_P P \end{cases} \tag{49}$$

The initial condition is taken at $t = 0$, before any injection and it was assumed that both

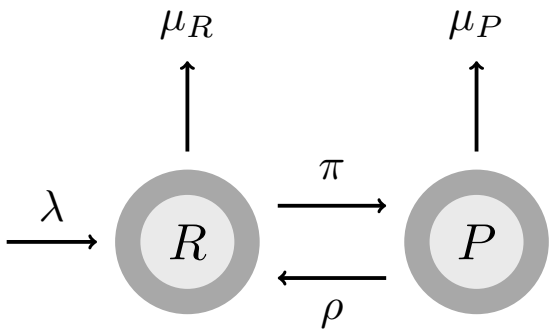


Figure 4.3 – Simple model of the CD4$^+$ dynamics. P=proliferating cells. R=resting cells.

populations were at equilibrium, meaning that $dR/dt = 0$ and $dP/dt = 0$. Estimation of parameters (λ, π, ρ, μ_R, μ_P) was realized using a population approach on INSPIRE data, using the estimation method described in section 3.2.2 and the discrete observations of both CD4$^+$ T cell counts and Ki67$^+$ percentages. A statistical model of the form:

$$\tilde{\xi}_l^i(t) = \tilde{\xi}_{l_0} + \boldsymbol{\beta}_l^T \mathbf{z}_l^i(t) + u_l^i \tag{50}$$

was used on every log-transformed parameter ξ_l for each individual i. In particular, the covariates vector $\mathbf{z}$ included the treatment effect (=0 if the patient received placebo or 1 if the patient received IL-7 injections) and the dose effect, assumed to be linear. For example, parameter π was written as:

$$\tilde{\pi} = \begin{cases} \tilde{\pi}_0 & \text{if } t = 0, \\ \tilde{\pi}_0 + \beta_0 trt + \beta_1 dose & \text{if } 0 < t \leq \tau, \\ \tilde{\pi}_0 & \text{if } t > \tau. \end{cases} \tag{51}$$

The time τ of IL-7 effect on proliferation rate was fixed to 16 days after sensitivity analyses. Several models were tested and compared :

- model 1 with only an effect of IL-7 on proliferation (π)
- model 2 with an effect of IL-7 on proliferation (π) and resting cells survival (μ_R), during or after IL-7 administration
- model 3 with an effect of IL-7 on proliferation (π) and thymic output (λ), during or after IL-7 administration

The potential additional effects of IL-7 on survival and thymic output were tested separately as a model including both effects and the effect on proliferation was not identifiable. After model selection, random effects were estimated on both λ and ρ rates and a statistically significant effect of dose was estimated on the proliferation rate. Model 1 was found to be worse than both models 2 and 3, indicating that proliferation is not the only effect induced by IL-7 on $CD4^+$ T cells. Model 2 was found slightly better than model 3 in terms of statistical criterion but the difference was not striking in term of patient fits. Using these estimation results, simulations suggested that strategies including repeated cycles of IL-7 injections could be used to maintain CD4 levels above 500 cells/μL and helped designing the clinical trial INSPIRE 3.

4.1.2.2 Effect of repeated cycles of IL-7 injections on $CD4^+$ T cell dynamics

Repeated cycles of IL-7 injections at dose 20μg/kg were assessed in two multicenter phase II trials. INSPIRE 2 was a single-arm trial including 23 patients and INSPIRE 3 was a 2 arms trial with a 3:1 randomization to IL-7 versus placebo, with a total of 84 patients treated and included in the analysis. Results of both studies were presented in Thiébaut et al. [2016]. HIV-infected patients were included if they were under cART, had low $CD4^+$ T cell counts (between 100 and 400 cells per μL of blood) and controlled the virus (undetectable viral load for at least 6 months). The design of INSPIRE 3 study (treatment group only) is represented in figure 4.4. A first cycle of injections was administered at the beginning of the study, then patients were followed up to 2 years. Repeated visits were made every 3 months, and if the CD4 T cell counts were measured < 550 cells/μL at the visit, a new cycle was administered to the patient. Patient in placebo group received no treatment until 1 year after enrollment and started the same design after. These trials showed that repeated cycles of IL-7 were well tolerated and allowed the maintenance of $CD4^+$ T cell counts above 500/μL in most of the study participants.

The modeling work was continued, in particular to quantify the effect of different injections in a cycle, the effect of repeated cycles on the $CD4^+$ T cell dynamics and the long-term efficacy of the therapy in maintaining $CD4^+$ T counts above 500 cells/μL. Several models, including more compartments or other processes such as feedback terms, were developed and tested on INSPIRE data in Jarne Munoz [2015], but will not be detailed here. However, we will summarize the results obtained in Jarne et al. [2017], using the same mathematical model as described in figure 4.3 with some statistical improvements. The parameters of the model were estimated on grouped data from all INSPIRE studies, including a total of 128 patients (21 from INSPIRE, 23 from INSPIRE 2 and 84 from INSPIRE 3), using the same population approach as presented in sections 3.2.2 and 4.1.2.1. All patients had repeated measurements of total $CD4^+$ T cell counts during the follow-up

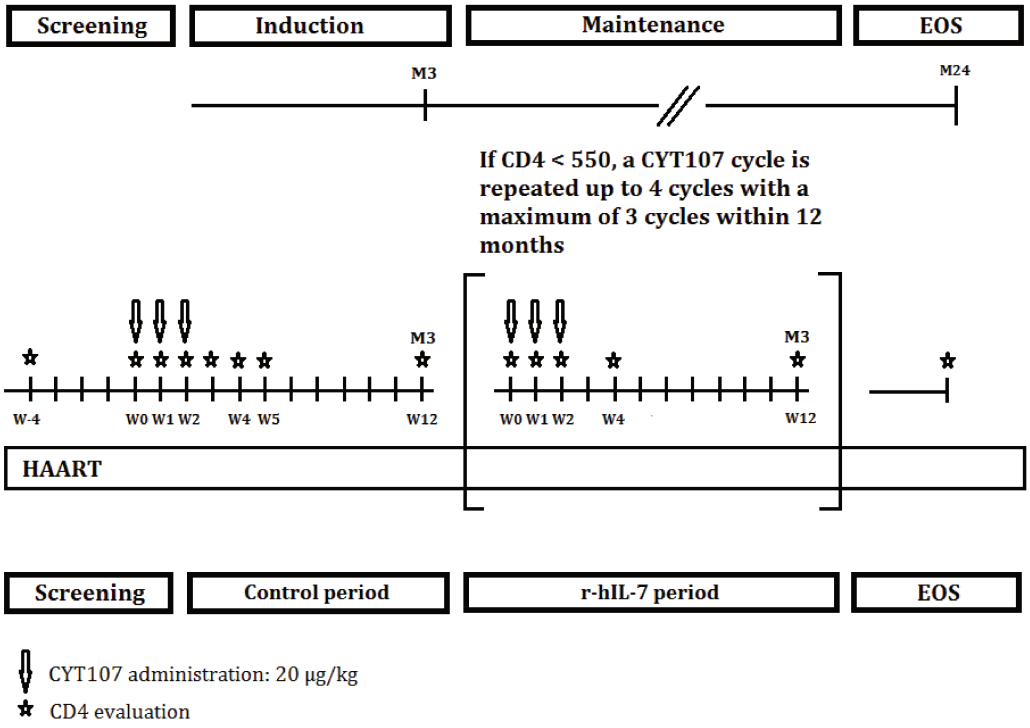


Figure 4.4 – Design of INSPIRE 3 study (treatment group). Adapted from Thiébaut et al. [2016].

and some Ki67 counts during the first weeks of the studies. It is important to underline the interest of the population approach here, due to the high heterogeneity of the data, as patients had different measurement time points, different number of measurements, received different number of cycles and injections (some cycles were incomplete), depending on their individual $CD4^+$ T cell dynamics. The final statistical model presented in Jarne et al. [2017] is the following:

$$\begin{cases} \tilde{\lambda}^i(t) &= \tilde{\lambda}_0 + u^i_\lambda, \\ \tilde{\pi}^i(t) &= \tilde{\pi}_0 + \left[\beta_C \mathbb{1}_{\{C(t)>1\}}, + \sum_{k=1}^{3} \mathbb{1}_{\{N^i_t=k\}} \beta_{\pi_k} d_i^{0.25}\right] \mathbb{1}_{\{N^i_t - N^i_{t-7}=1\}}, \\ \tilde{\rho}^i(t) &= \tilde{\rho}_0 + u^i_\rho, \\ \tilde{\mu}^i_R(t) &= \tilde{\mu}_{R0} + \beta_{\mu_R} d_i^{0.25} f(t), \\ \tilde{\mu}^i_P(t) &= \tilde{\mu}_{P0}, \end{cases} \tag{52}$$

with u^i_λ and u^i_ρ the individual random effects (normally distributed), d_i the dose received by patient i, N^i_t the number of injections patient i has received until time t, $\mathbb{1}_{\{N^i_t - N^i_{t-7}=1\}}$ an indicator function taking value 1 if an injection has been administered in the last 7 days, $\mathbb{1}_{\{C(t)>1\}}$ an indicator function equal to 1 if a cycle has been received before time t. The function f is such that the effect of IL-7 injection on the decay rate of resting cells

μ_R is assumed to be constant from 2 days after the first injection, to last 12 months and to linearly decreases to 0 during the following 12 months. Random effects were estimated on production rate λ and reversion rate ρ, meaning that these two parameters induce the observed inter-individual variations. The effect of IL-7 was found to be dose-dependent and to increase the value of the proliferation rate π during 7 days (the duration of the effect was determined by profile likelihood). Moreover, within a cycle, the quantitative effect of the injections on proliferation rate π value differs: in particular $\beta_{\pi 1} \geq \beta_{\pi 2} > \beta_{\pi 3}$ with the first and second injections having a similar effect, but the third injection has a much weaker effect than the previous ones. The effect of all cycles on the proliferation rate π were found to be lower than the effect of the first cycle. That could be either due to the fact that the immune system reacts to IL-7 by generating antibodies against the cytokine or simply because of homeostatic regulation, as the number of $CD4^+$ T cells is lower at the moment of the first cycle compared to the other ones. However, we were not able to distinguish these effects by modeling the available data. The model induced good fits of the data as shown in figure 4.5. The mechanistic modeling and parameters estimation in

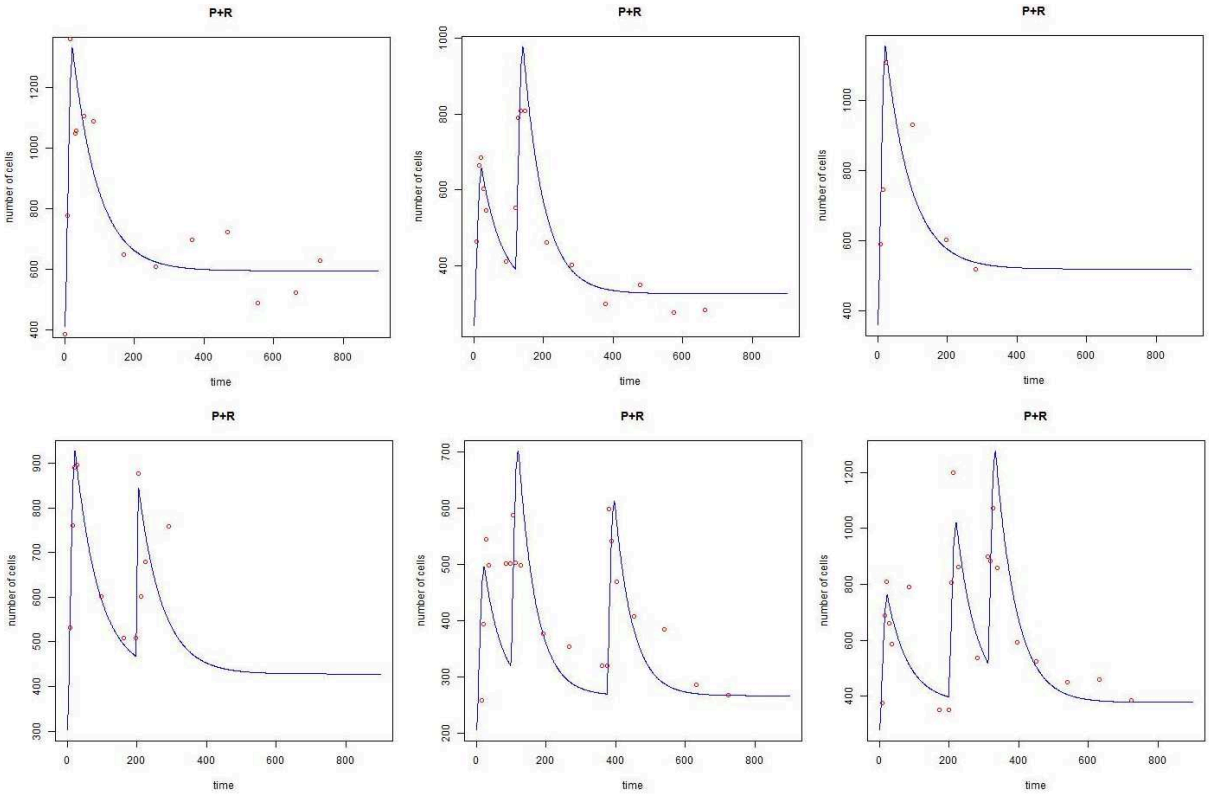


Figure 4.5 – Fits of some patients from INSPIRE studies. Time is in days. Red points correspond to the data and blue line corresponds to the fit obtained through the estimation of the mechanistic model.

the population approach allowed then the quantification of the effect of repeated cycles

of IL-7 on the dynamics of $CD4^+$ T cells. A major interest in the modeling approach is that the model can be used to simulate other potential designs of clinical trials and help determining better protocols of injections.

4.1.2.3 Simulation and comparison of clinical protocols

From the estimation of a good predictive model, simulations were realized in Jarne et al. [2017] to evaluate the possibility to reduce the number of injections in protocols. Four protocols were compared, all based on the same original design: visits are realized every 3 months and a new cycle is administered if the $CD4^+$ T cell counts is measured below 550 cells/μL. The variation between the protocols is due to the number of injections realized in each cycle. Protocol A corresponds to the original one, with cycles of 3 injections. Protocol B contains a first cycle of 3 injections, followed by cycles of 2 injections, while in protocol C it is followed by cycles of 1 injection. Finally, protocol D contains cycles of 2 injections only. The comparison of these four protocols was based on three criteria of clinical interest: the number of injections received, the mean $CD4^+$T cell counts and the time spent below 500 cells /μL. These protocols were applied to pseudo-patients, generated using the (normally approximated) posterior distribution of parameters estimated on INSPIRE data with the final model. The whole target population is considered as "low" responders to cART and have their $CD4^+$ T cell counts below 400 cells/μL when included in the trial. Two other sub-groups of interest were studied, in particular "very low" responders with $CD4^+$ T cell counts between 100 and 200 cells/μL, and "not too low" responders with $CD4^+$ T cell counts between 300 and 400 cells/μL. Protocols were studied and compared on these pseudo-patients. Results are detailed in Jarne et al. [2017]. In short, protocols B and D allow to reduce the number of injections compared to protocol A without significantly reducing the mean CD4 count or increasing the time spent under 500 cells/μL, for all patients. Protocol C could be used for "not too low" responders, as it dramatically reduces the number of injections compared to the other protocols, without reducing the mean CD4 count or increasing the time spent under 500 cells/μL. However, this protocol is not convenient for "very low" responders, for which it increases very much the time spent with low CD4 counts. These results are mainly due to the fact that the effect of a third injection in a cycle on cells proliferation was estimated to be much lower than the effect of the two previous injections.

This comparison of protocols suggested the possibility to reduce the number of injections in clinical trials, while still maintaining the patients' $CD4^+$ T cell count above 500 cells/μL. However, the simulation was still based on the criterion that all patient would undergo follow-up visits every 3 months. This may not be clinically relevant for all patients. Indeed, as shown by the data and estimated by the model, there is inter-individual

variability in the response to IL-7 injections: in particular, following the peak induced by the repeated injections of a cycle, the $CD4^+$ T cell counts decrease much faster in some patients compared to others. For patients with a strong decrease, the limit of 500 cells/μL could be reached quickly and before the following control visit at 3 months, whereas for patients with slower decline, the control visit could be realized much later than 3 months after the administration of the cycle of injections. Thus, it could be more efficient to adapt visits and measurements of $CD4^+$ T cell counts by patients or group of patients; this could reduce the number of visits and injections for patients while making sure that $CD4^+$ T cell levels are maintained over time. This question led us to develop methodological tools to optimize the response to immune interventions such as IL-7 injections. The pipeline is the following: when a new patient is included in a study, its individual parameters could be estimated using population distributions previously estimated in INSPIRE (or other) data. Then, optimization methods could be used on this patient. Figure 4.6 shows the possible pipeline for optimizing and adapting injections protocols. In particular, we will mention a Bayesian approach (3') and develop more precisely an approach based on the optimal control theory (3). However, as our main goal was to assess the feasibility of these approaches, step 2 was replaced by the generation of pseudo-patients using the posterior distribution estimated on the population of INSPIRE. These simulated data were used in the evaluation of the optimal control method.

4.2 Method: optimal control

4.2.1 Optimizing clinical protocols

4.2.1.1 Possible approaches

In control theory, two main methods have been developed separately to solve optimization problems: Pontryagin's maximum principle [Pontryagin, 1987] and Bellman's dynamic programming [Bellman, 1957]. Both methods can be applied to deterministic and stochastic processes. Pontryagin's maximum principle provides, under suitable assumptions, a set of necessary conditions that should be satisfied by an optimal control. The problem is reduced to solving a so-called Hamiltonian system including the maximum condition equation. In the dynamic programming method, the optimal control problem can be solved using an integro-differential equation, known as the Hamilton-Jacobi-Bellman equation.

A large number of biological applications have been modeled using an ODE system of

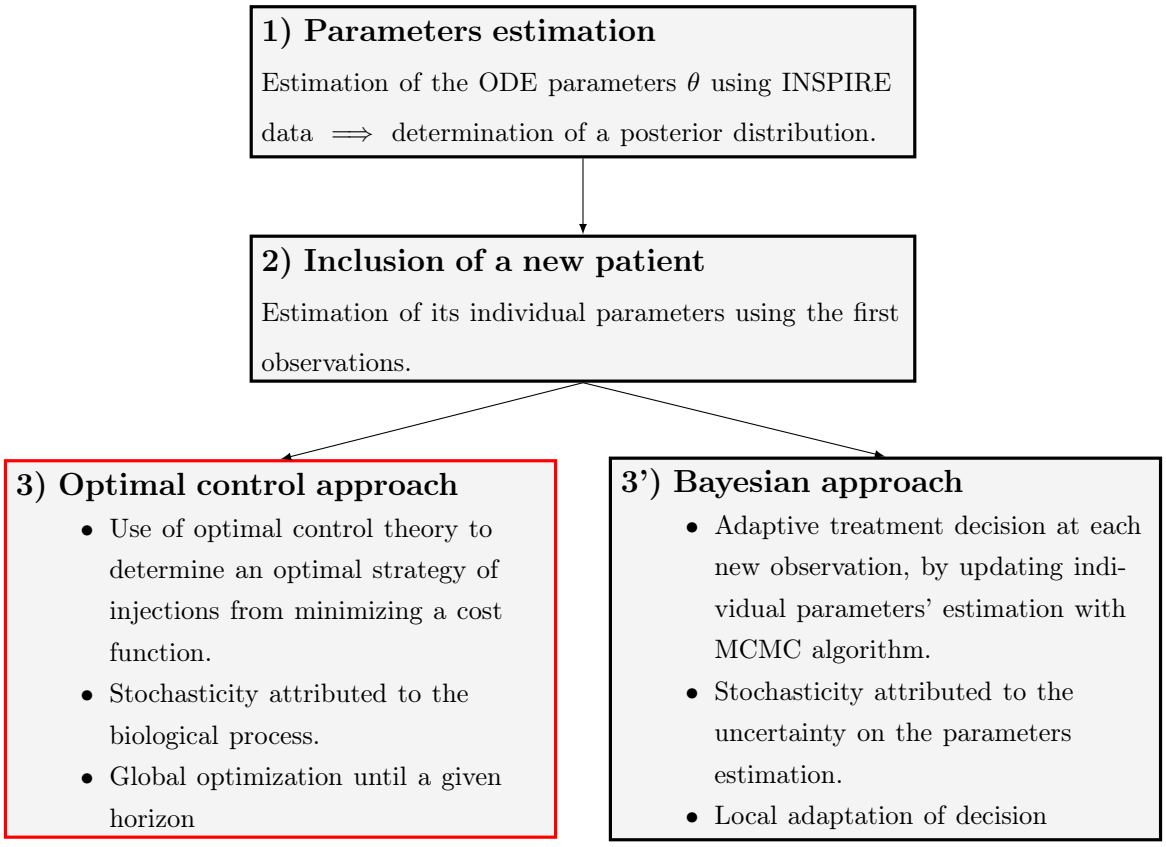


Figure 4.6 – Pipeline for optimizing schedule of IL-7 injections in a new patient

the form:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \tag{53}$$

where the solution of the ODE system depends on the dynamics of the control function $\mathbf{u}(t)$. The problem corresponds to minimizing a cost function depending on the control $\mathbf{u}$. In most of the cases Pontryagin's maximum principle was used to address this question. In the HIV field, Stengel [2008] and Yang et al. [2013] were interested in determining an optimal HIV treatment to minimize the viral load and maximize the number of uninfected $CD4^+$ T cells taking into account the occurrence of viral mutations. Solving this issue relied on using a model for the dynamics of viral infection and adding a control representing the effect of treatment. In Stengel [2008] the treatment was represented by terms corresponding to the mechanisms of the different possible treatments, e.g., the protease inhibitor reduces the rate of infection of $CD4^+$ T cells by the virus. In Yang et al. [2013], the control corresponded to the increase of number of $CD8^+$ T cells. Pontryagin's principle was used and adapted in both works. A similar method was used for determining optimal therapeutic protocols in cancer immunotherapy [Castiglione and Piccoli, 2006; Cappuccio et al., 2007; Castiglione and Piccoli, 2007; Pappalardo et al., 2010]. The first intervention corresponds to the injection of tumor specific DCs which allows activation of

helper and cytotoxic cells targeting the cancer cells. Other therapeutic scenarios, including $CD8^+$ T lymphocytes therapy and immunotherapy with cytokines were also explored. The cost function that should be minimized combines the size of tumor at the end of experiment and the potential toxicity of repeated treatment injections. Even if this tool has been widely used in biological applications, Pontryagin's maximum principle only provides necessary conditions and may only help to explicitly determine an optimal control strategy. Moreover, it can be applied to models including either ODEs or stochastic differential equations (SDEs), for which the theory is well developed [Yong and Zhou, 1999]. However, to the best of our knowledge, no maximum principle exists for PDMPs which are very well adapted for modeling most of the dynamical systems undergoing stochasticity, as will be described in section 4.2.2.1. In biology, stochasticity is difficult to account for, especially in inference. However, it is inherent to many processes and cannot be neglected at some levels of precision, for example in single-cell analysis where processes such as gene activation and protein production are modeled [Herbach et al., 2017; Cloez et al., 2017; Puszynski et al., 2016]. Markov Decision Processes, a subclass of constant piecewise PDMPs, have also been developed in the special case of a countable space and punctual observations [Winkelmann et al., 2014] to determine optimal treatment strategies in HIV by accounting for the potential mutation of the virus and the emergence of resistant strains to some treatments [Duwal et al., 2015]. The stochasticity can also be due to the variability between individuals, with a population distribution of some values of the key parameters affecting the biological process and inducing heterogeneity in the population. Theses sources of stochasticity were not accounted for in the previous cited examples using Pontryagin's maximum principle. Also, the dynamic of the control is not always known and cannot always be explicitly included in the ODE system as in equation (53). It is especially the case in immunology, as highlighted by Eftimie et al. [2016] in their review on mathematical modeling for immunology. In this paper, authors underline the crucial need for developing complex optimal control approaches in parallel to immunology experiments to improve clinical interventions. As part of this effort, we aimed at developing a tool based on optimal control theory and using dynamic programing for optimizing schedules of IL-7 injections in HIV-infected patients by accounting for some stochasticity in the biological process. This work was published [Pasin et al., 2018] and will be presented in section 4.2.2.

4.2.1.2 Bayesian approach for adapting protocols

In this section, we will present a Bayesian approach for adapting protocols of IL-7 injection in a succinct way. This approach was developed in parallel to the optimal control work by Laura Villain during her PhD, supervised by Daniel Commenges and

Rodolphe Thiébaut. For more details, the interested reader can refer to the manuscript of the article **Adaptive protocols based on predictions from a mechanistic model of the effect of IL7 on CD4 counts**, L. Villain, D. Commenges, C. Pasin, M. Prague, R. Thiébaut, accepted in Statistics in Medicine and presented in Appendix B. The general idea of this method relies on the fact that the model used for the dynamics of the $CD4^+$ T lymphocytes, as detailed in section 4.1.2 and Thiébaut et al. [2014], Thiébaut et al. [2016], Jarne et al. [2017] is deterministic; when a distribution of the parameters of a given patient is known, the corresponding trajectory of the $CD4^+$ T cells can be computed, as well as the distribution of any related quantity. Decisions to adapt the treatment can be taken using these predicted quantities. Moreover, each time a measurement is made on a patient, more data is available and the parameter estimation can be updated, with more precision. In practice, the individual parameters are estimated with an MCMC algorithm. Two strategies were evaluated: one, the adaptive criterion of injection (ACI) is based on the predicted risk to have CD4 counts below 500 cells/μL before the next visit (at 3 months). The other one, the adaptive time of injection (ATI) is based on the predicted time at which CD4 counts will reach 500 cells/μL. Within both criteria, the possibility to administer less than 3 injections per cycle was also assessed. It should be underlined that here, both criteria are computed from the estimation of individual parameters, obtained thanks to the observations, but not from the observations themselves as in the original protocol. In short, the results showed that all adaptive protocols had the ability to reduce the time spent with CD4 levels below 500 cells/μL without increasing much the number of injections for each protocol. The protocols based on the ATI criterion actually helped in reducing the number of visits of the patient. All protocols maintained the CD4 counts at higher numbers than the original protocol. This approach offers clinical prospects, as it could be applied in a larger trial to evaluate the impact of adaptive strategies on other clinical outcomes. Its feasibility relies on the good capacity of the deterministic model to capture the biological process and to be estimated using a limited number of markers and measurements. In parallel on this work, we developed the optimal control approach, as described in the following sections.

4.2.2 Optimal control on piecewise deterministic Markov processes

In this section, we will introduce the notations and definitions related to uncontrolled and controlled PDMPs. We will also remind the main results from Costa et al. [2016] and adapt them to our particular context. The terminology will be mostly taken from this article. From now, we will use the term of "impulse control" problem to refer to an

optimal control problem with a possibility of punctual actions but only on the boundaries of the system.

4.2.2.1 Uncontrolled PDMPs

In his paper, Davis [1984] observed that almost all continuous stochastic models could be defined as combination of diffusion, deterministic motion, and/or random jumps. Diffusions correspond to continuous time stochastic processes and include stochastic differential equations, which are ordinary differential equations with a white noise perturbation. The theory of control was well developed in this framework [Bensoussan, 1982; Peng, 1990] and numerical tools were then developed to solve the control problems. However, the number of applications was limited, because SDEs are not adapted for modeling a large number of physical and biological processes. Now, control of SDEs is particularly developed in finance [Framstad et al., 2004; Shen et al., 2014]. By developing PDMPs, Davis [1984] proposed a general framework for modeling and eventually optimizing continuous-time dynamical systems including uncertainty due to random occurrences, but without diffusion. Since this paper, PDMPs have been used to model processes in many applications including questions in biology such as neuronal membranes [Buckwar and Riedler, 2011] or gene expression [Bobrowski et al., 2007] or in physics such as maintenance of metallic structures [de Saporta et al., 2012].

A PDMP is characterized by local characteristics defined in a given state space $\mathbf{X}$: the flow ϕ, the jump rate η and the transition measure Q. The trajectory of a PDMP can be described by iteration: starting from a point $x_0 \in \mathbf{X}$, the process follows the flow $\phi(x_0, t)$ until a jump occurs at time T_1. This jump can either be spontaneous, following a Poisson-like law defined by rate η or deterministic when the flows hits the boundary of the state space. In both cases, the process starts again from a point determined by the transition measure $Q(.|\phi(x_0, T_1))$ and follows the flow until a new jump occurs. A graphic representation of a PDMP is shown in figure 4.7. In a more formal way, the state space $\mathbf{X}$ is an open subset of $\mathbb{R}^d$, $d \in \mathbb{N}$; $\partial\mathbf{X}$ corresponds to its boundary. The flow associated with the process $\phi(x, t) : \mathbb{R}^d \times \mathbb{R} \mapsto \mathbb{R}^d$ is a Lipschitz continuous vector field in $\mathbb{R}^d$, meaning that $\phi(x, t + s) = \phi(\phi(x, s), t)$ for all $x \in \mathbb{R}^d$ and $(t, s) \in \mathbb{R}^2$. The active boundary is defined as $\mathbf{\Xi} = \{x \in \partial\mathbf{X} : x = \phi(y, t)\}$ for some $y \in \mathbf{X}$ and $t \in \mathbb{R}_+^*$. We will then denote $\overline{\mathbf{X}} = X \cup \mathbf{\Xi}$ and for $x \in \overline{\mathbf{X}}$, we can define $t^*(x) = \inf\{t \in \mathbb{R}_+ : \phi(x, t) \in \mathbf{\Xi}\}$. The definition of the flow ϕ outside the space $\overline{\mathbf{X}}$ can be arbitrary, as it has no impact on the definition of the process. The controlled jump intensity η is a $\mathbb{R}_+$-valued measurable function and determines the law of the stochastic jumps. The transition measure $Q(.|x)$ corresponds to the distribution of the state after a natural jump occurring at $x \in \mathbf{X}$.

A PDMP can be constructed using the flow, the jump intensity rate and the transition

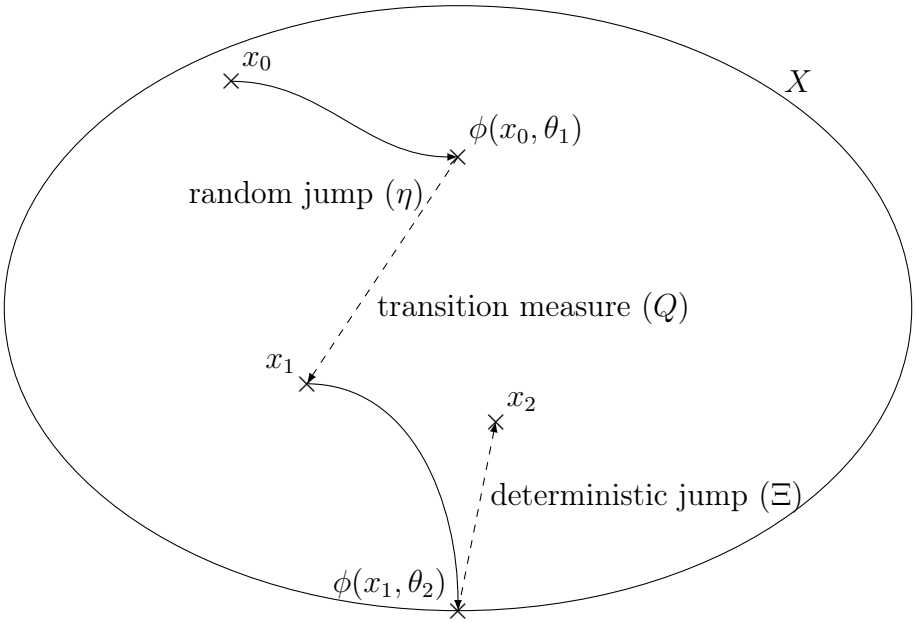


Figure 4.7 – Graphic representation of a PDMP, starting from x_0, with a random jump at time $\tau_1 = \theta_1$, then starting again from x_1 and undergoing a deterministic jump at time $\tau_2 = \theta_1 + \theta_2$ when reaching the boundary of state space X, before starting again from x_2.

measure. We will not focus on technical details here, but roughly speaking, a canonical space of trajectories Ω can be defined. A trajectory in this space can be written as $\boldsymbol{\omega} = (x_0, \theta_1, x_1, \theta_2, x_2, ..) \in \Omega$, with initial state $x_0 \in \mathbf{X}$. Written this way, θ_i corresponds to the time interval between two consecutive jumps and x_i the value of the process right after the jump. Note that the θs depend on the jump intensity rate as well as the flow and the boundary of the state space, and the xs are determined by the transition measure. We note $h_n = (x_0, \theta_1, x_1, \theta_2, x_2, .., \theta_n, x_n)$ the path up to $n \in \mathbb{N}$ and $\mathbf{H}_n$ the set of all such paths. The process has only n jumps when $\theta_n < \infty$ and $\theta_{n+1} = \infty$. In that case, we can fix $\theta_m = \infty$ and $x_m = x_\infty$ for every $m \geq n + 1$, with x_∞ corresponding the an isolated artificial point after which no jump occurs. We note $\mathbf{X}_\infty = \mathbf{X} \cup \{x_\infty\}$ and then we can define mappings $X_n : \Omega \to \mathbf{X}_\infty$ such that $X_n(\omega) = x_n$, corresponding to the post-jump locations and $\Theta_n : \Omega \to \overline{\mathbb{R}}_+^*$ by $\Theta_0(\omega) = 0$, $\Theta_n(\omega) = \theta_n$ corresponding to the sojourn time (in a deterministic regime). We can also define the sequence of $\overline{\mathbb{R}}_+^*$-valued mappings $(T_n)_{n\in\mathbb{N}^*}$ corresponding to the time at which jump occurs, with $T_n(\omega) = \sum_{i=1}^n \Theta_i(\omega)$ and $T_\infty(\omega) = lim_{n\to\infty} T_n(\omega)$. From X_n and Θ_n, we can determine $H_n = (X_0, \Theta_1, X_1.., X_n)$. Finally we can define the PDMP $\{\xi_t\}_{t\in\mathbb{R}_+}$ such that:

$$\xi_t(\omega) = \begin{cases} \phi(X_n, t - T_n) & \text{if } T_n \leq t < T_{n+1} \text{ for } n \in \mathbb{N}, \\ x_\infty & \text{if } T_\infty \leq t. \end{cases} \tag{54}$$

4.2.2.2 Impulse control of PDMPs

In this section, we will remind the theoretical results presented in Costa et al. [2016] and adapt them to a particular context which will constitutes our framework in the case of protocols of IL-7 injections: here, we will be interested in an impulse control problem where the decision maker can modify the measure of transition Q, but only on the boundary of the system. The process is not controlled inside of the state space. It means that punctual decisions can be taken to modify from where the process starts again when a boundary is reached.

The action space is denoted $\mathbf{A}$. For a particular point $z \in \mathbf{\Xi}$ at the boundary of the system, the set of possible actions at this point is written $\mathbf{A}(z)$. We can then define the state $\mathbf{K} = \{(z, a) \in \mathbf{\Xi} \times A : a \in \mathbf{A}(z)\}$ of pairs containing all the elements of the active boundary and their corresponding possible actions. The stochastic kernel Q on $\mathbf{X}$ given $\mathbf{K}$ determines the state of the process after any jump. For any $(z, a) \in \mathbf{K}$, $Q(.|z, a)$ describes the distribution of the state after the jump induced by the impulsive action a from point z on the active boundary. If a natural jump occurs in state $x \in \mathbf{X}$, the measure of transition does not depend on any action, as the process is uncontrolled inside of the state space. We can write $Q(.|x, a) = Q(.|x)$ the distribution of the state after the jump as well. In the case of a stochastic jump, the signed kernel computes the difference between the state before and after the jump:

$$q(dy|x) = \eta(x)\big[Q(dy|x) - \delta_x(dy)\big]. \tag{55}$$

For any function $V : \mathbf{X} \to \mathbb{R}$, we can define for $(z, a) \in \mathbf{K}$:

$$QV(z, a) = \int V(y)Q(dy|z, a), \tag{56}$$

and for $x \in \mathbf{X}$:

$$qV(x) = \int V(y)q(dy|x). \tag{57}$$

The set of all actions realized by the decision-maker constitutes a control strategy. Formally, an admissible control strategy is a sequence $u = (\chi_n)_{n\in\mathbb{N}}$ with, for any $n \in \mathbb{N}$, χ_n a stochastic kernel on $\mathbf{A}$ given $\mathbf{H}_n$. When the process hits the boundary from x_n, the decision maker chooses randomly an action according to the distribution $\chi_n(.|\mathbf{H}_n)$, satisfying $\chi_n\big(\mathbf{A}(\phi(x_n), t^*(x_n))|h_n\big) = 1$, for $h_n = (x_0, \theta_1, x_1, \theta_2, x_2, .., \theta_n, x_n) \in H_n$, with $x_n \neq x_\infty$ and $t^*(x_n) < \infty$. The set of all admissible control strategies is denoted by $\mathcal{U}$.

The optimization problem corresponds to determining a strategy in $\mathcal{U}$ which minimizes a performance criterion. This criterion is defined from a cost function, itself divided in two parts: the gradual cost C^g, penalizing continuously the trajectory of the process, and

the cost C^i associated to an impulsive action. The associated performance criterion for strategy $u \in \mathcal{U}$ starting from state $x_0 \in \mathbf{X}$ is:

$$\begin{aligned}\mathcal{V}(u, x_0) = \; & \mathbb{E}^u_{x_0}\left[\int_{]0,+\infty[} e^{-\alpha s} C^g(\xi_s) ds\right] \\ & + \mathbb{E}^u_{x_0}\left[\sum_{n=1}^{+\infty} e^{-\alpha T_n} \mathbb{1}_{\left\{\xi_{T_n-} \in \Xi\right\}} \int_{A\left(\xi_{T_n-}\right)} C^i(\xi_{T_n-}, a) \chi_n(da|H_n)\right],\end{aligned} \tag{58}$$

where α is a discount factor ensuring the convergence of the integral when the process is defined on an infinite horizon. However, when working on a finite horizon T_h, it is possible to define a boundary of the process when $t \geq T_h$ at which the process enters a so-called absorbing state Δ, where nothing happens, $\phi(\Delta, t) = \Delta$ and $C^g(\Delta) = C^i(\Delta) = 0$. In that case, equation (58) can be computed with integrals on $]0, T_h]$ instead of $]0, +\infty[$. The optimization problem aims at finding an optimal strategy $\hat{u} \in \mathcal{U}$ such that:

$$\mathcal{V}(\hat{u}, x_0) = \inf_{u \in \mathcal{U}} \mathcal{V}(u, x_0). \tag{59}$$

Below are a list of assumptions that should be verified to solve the optimization problem, as defined in section 3.2 of Costa et al. [2016]. They mainly ensure the ability to compute values of the performance criterion and the existence of an optimal solution to the problem.

Assumption A. There are constants $K \geq 0, \varepsilon_1 > 0$ and $\varepsilon_2 \in [0, 1[$ such that

(A1) For any $x \in \mathbf{X}$, $\eta(x) \leq K$.

(A2) For any $(z, a) \in \mathbf{K}$, $Q(A_{\varepsilon_1}|z, a) \geq 1 - \varepsilon_2$, where

$$A_{\varepsilon_1} = \{x \in \mathbf{X} : t^*(x) > \varepsilon_1\}.$$

Assumption B.

(B1) The set $\mathbf{A}(y)$ is compact for every $y \in \overline{\mathbf{X}}$.

(B2) The kernel Q is weakly continuous.

(B3) The function η is continuous on $\mathbf{X}$.

(B4) The flow ϕ is continuous on $\mathbb{R}^d \times \mathbb{R}_+$.

(B5) The function t^* is continuous on $\overline{\mathbf{X}}$.

Assumption C.

(C1) The multifunction Ψ from Ξ to $\mathbf{A}$ defined by $\Psi(z) = \mathbf{A}(z)$ is upper semicontinous.

(C2) The cost function C^g (respectively, C^i) is bounded and lower semicontinuous on $\mathbf{X}$ (respectively, $\mathbf{K}$).

When assumptions (A1) and (A2) hold, constants K_A and K_B can be defined as in section 5 from Costa et al. [2016]:

$$\begin{array}{rcl} K_B & \geq & \dfrac{K}{1-\varepsilon_2}, \\ K_A & \geq & \dfrac{K(1+K_B)(1-e^{-(K+\alpha)\varepsilon_1})+(K+\alpha)(K+K_B\varepsilon_2)e^{-(K+\alpha)\varepsilon_1}}{\alpha(1-e^{-(K+\alpha)\varepsilon_1})}. \end{array} \tag{60}$$

The theorem allowing to determine the optimal cost and providing an optimal strategy is adapted from theorem 5.5 in Costa et al. [2016]. It is stated as follows:

Theorem 1. Suppose assumptions A, B and C are verified. We define the sequence of functions $\{W_m\}_{m\in\mathbb{N}}$ for any $x \in \overline{\mathbf{X}}$ as follows:

$$\left\{ \begin{array}{rcl} W_{m+1}(x) & = & \mathcal{B}W_m(x) \text{ for m} \in \mathbb{N}, \\ W_0(x) & = & -K_A\mathbf{1}_{A_{\varepsilon_1}}(x) - (K_A+K_B)\mathbf{1}_{A^c_{\varepsilon_1}}(x), \end{array} \right. \tag{61}$$

with constants K_A and K_B defined as in equation (60), $A_{\varepsilon_1} = \{x \in \mathbf{X} : t^*(x) > \varepsilon_1\}$ and

$$\mathcal{B}V(y) = \int_{[0,t^*(y)[} e^{-(K+\alpha)t}\mathcal{R}V(\phi(y,t))dt + e^{-(K+\alpha)t^*(y)}\mathcal{T}V(\phi(y,t^*(y))), \tag{62}$$

with real-value functions $\mathcal{R}V$ and $\mathcal{T}V$ defined for any V respectively on $\mathbf{X}$ and $\mathbf{\Xi}$:

$$\mathcal{R}V(x) = C^g(x) + qV(x) + \eta V(x),$$
$$\mathcal{T}V(z) = \inf_{a\in\mathbf{A}(z)} \Big\{ C^i(z,d) + QV(z,a) \Big\},$$

with qV and QV defined as in equations (57) and (56) respectively.
The sequence of functions $\{W_m\}_{m\in\mathbb{N}}$ converges to a function W defined on the state space and such that :
i) $W(x_0) = \inf_{u\in\mathcal{U}} \mathcal{V}(u, x_0)$
ii) there is a measurable mapping $\widehat{\varphi} : \mathbf{\Xi} \to \mathbf{A}$ such that $\widehat{\varphi}(z) \in \mathbf{A}(z)$ for any $z \in \mathbf{\Xi}$ and satisfying

$$C^i(z,\widehat{\varphi}(z)) + QW(z,\widehat{\varphi}(z)) = \inf_{a\in\mathbf{A}(z)} \Big\{ C^i(z,a) + QW(z,a) \Big\}. \tag{63}$$

The value function W as previously defined verifies $W(x) = \mathcal{B}W(x)$ for any $x \in \overline{\mathbf{X}}$, which is an integral form of the Hamilton-Jacobi-Bellman equation. This theorem shows that the optimal cost can be obtained by computing the value function W in x_0. Note that, due to the definition of W as the convergence of sequence $\{W_m\}$, its value should be computed on the state space $\mathbf{X}$ in order to obtain the specific value $W(x_0)$. Moreover, the optimal strategy is obtained by choosing the optimal action $\widehat{\varphi}(z)$ for every point $z \in \mathbf{\Xi}$ reached on the trajectory of the process. The optimal strategy is simulated as follows: we start with the trajectory from x_0, then when a boundary is reached, the chosen action corresponds to the one minimizing the criterion $C^i(z,a) + QW(z,a)$, as given by equation (63).

4.2.2.3 Numerical aspects

The theory of impulse control on PDMPs, as developed in previous sections 4.2.2.1 and 4.2.2.2, can be applied to biological or physical questions by using a two-steps method: first, the PDMP associated to the process should be developed, then the impulse control problem can be solved using theorem 1.

The defined PDMP can be simulated using a numerical software. Simulations constitute a way to check if the PDMP is able to reproduce the biological process. Moreover, simulations of the PDMP allow to generate and compare the effect of several injections strategies on the process. It helps getting a first evaluation of the ability of the cost function to distinguish between the different strategies. It also gives a basis of comparison to the optimal strategy, after it is determined.

Then, the impulse control problem should be solved numerically, but at the moment there is no general computation method [de Saporta et al., 2015]. Some methods have been developed in de Saporta and Dufour [2012], and they have also shown that some tools are still under development. Results from Costa et al. [2016], and in particular theorem 1 gives a natural method to solve the problem. Indeed, the value function W is obtained by iteration of a sequence $\{W_m\}_{m\in\mathbb{N}}$. This sequence is defined by equation (61), with $W_{m+1}(x) = \mathcal{B}W_m(x)$. An approximation of the operator $\mathcal{B}$ defined in equation (62) can be realized on a grid of the state space and it will allow to compute an approximation of all functions W_m on the grid. This grid, written Γ, must be chosen to be stable by transformation with $\mathcal{B}W_m$, for all $m \in \mathbb{N}$: it can be ensured by approximating $\mathcal{B}W_m(x)$ on Γ, for every $x \in \Gamma$. Finally, it will give an approximation of the function W on Γ, and in particular of $W(x_0)$, corresponding to the optimal cost. Our main contribution in this work was to determine the PDMP associated to the context of IL-7 and to develop a numerical method able to solve the optimal control problem for some given patients. The challenge comes with finding a right way to organize the grid Γ to ensure good computation of the sequence $\{W_m\}_{m\in\mathbb{N}}$ on every point of the grid. Moreover, a computational challenge is induced by the size of the grid, as a large number of elements is easily reached, which can dramatically increases the time of computation.

In the next section, we include the article corresponding to the application of the theory of optimal control on PDMPs to the IL-7 question. Details of the two steps of the method are given: first the PDMP is described and the algorithm for computing the sequence of functions is explained. We decided to focus on a biological process slightly simpler than the one described by the ODE models on INSPIRE data and focused only on the effect of IL-7 on the proliferation of cells. This is partly due to the fact that our work was more a "proof-of-concept" and mainly aimed at showing that the theory of optimal control could

be used in a context of optimization of schedules of injections. It was then natural to choose a simple biological model, but still realistic and good enough to fit the data from INSPIRE studies. The numerical method was applied to 50 pseudo-patients that were generated using the posterior distributions previously estimated on the parameters using INSPIRE data. All codes (both simulation of the PDMP and computation of the iterative sequence for solving the optimal problem) were written in Matlab version R2016b (The MathWorks, Inc., Natick MA, USA, 1984).

4.3 Application of the optimal control to the IL-7 context: "Controlling IL-7 injections in HIV-infected patients"

The following paper was published in Bulletin of Mathematical Biology (2018), volume 80, issue 9, pp 2349-2377.

Bulletin of Mathematical Biology (2018) 80:2349–2377
https://doi.org/10.1007/s11538-018-0465-8

Society for Mathematical Biology

ORIGINAL ARTICLE

Controlling IL-7 Injections in HIV-Infected Patients

Chloé Pasin[1,2,3,4] · François Dufour[1,2,5] · Laura Villain[1,2,3,4] · Huilong Zhang[1,2] · Rodolphe Thiébaut[1,2,3,4]

Received: 23 October 2017 / Accepted: 3 July 2018 / Published online: 2 August 2018

Abstract

Immune interventions consisting in repeated injections are broadly used as they are thought to improve the quantity and the quality of the immune response. However, they also raise several questions that remain unanswered, in particular the number of injections to make or the delay to respect between different injections to achieve this goal. Practical and financial considerations add constraints to these questions, especially in the framework of human studies. We specifically focus here on the use of interleukin-7 (IL-7) injections in HIV-infected patients under antiretroviral treatment, but still unable to restore normal levels of $CD4^+$ T lymphocytes. Clinical trials have already shown that repeated cycles of injections of IL-7 could help maintaining $CD4^+$ T lymphocytes levels over the limit of 500 cells/μL, by affecting proliferation and survival of $CD4^+$ T cells. We then aim at answering the question: how to maintain a patients level of $CD4^+$ T lymphocytes by using a minimum number of injections (i.e., optimizing the strategy of injections)? Based on mechanistic models that were previously developed for the dynamics of $CD4^+$ T lymphocytes in this context, we model the process by a piecewise deterministic Markov model. We then address the question by using some recently established theory on impulse control problem in order to develop a numerical tool determining the optimal strategy. Results are obtained on a reduced model, as a proof of concept: the method allows to define an optimal strategy for a given patient. This method could be applied to optimize injections schedules in clinical trials.

Keywords Optimal control · Immune therapy · Dynamic programming

Rodolphe Thiébaut
rodolphe.thiebaut@u-bordeaux.fr

1 Univ. Bordeaux, Bordeaux, France

2 INRIA Bordeaux Sud Ouest, Talence, France

3 INSERM U1219 Bordeaux Population Health, SISTM Team, Bordeaux, France

4 Vaccine Research Institute-VRI, Hôpital Henri Mordor, Créteil, France

5 Bordeaux INP, IMB, Bordeaux, France

1 Introduction

The infection by the Human Immunodeficiency Virus (HIV) compromises the immune system functions, mainly because of the depletion of $CD4^+$ T lymphocytes. Combined antiretroviral (cART) therapy has led to a spectacular improvement of patients' survival by controlling virus replication and consequently restoring the immune system functions. However, some patients fail at reconstituting their immune system and recovering normal $CD4^+$ T cell levels, especially when they start antiretroviral treatment late (Lange and Lederman 2003). Immune therapy has been considered as a complement to cART to help immune restoration. In particular, interleukin-7 (IL-7), a cytokine produced by non-marrow-derived stromal and epithelial cells, is thought to improve thymic production (Mackall et al. 2001; Okamoto et al. 2002) and cell survival (Tan et al. 2001; Vella et al. 1998; Leone et al. 2010). The safety and beneficial effect of injections of exogenous IL-7 was first shown in phase I trials (Sereti et al. 2009; Levy et al. 2009) and observational studies (Camargo et al. 2009). Then, phase I/II human clinical trials (INSPIRE 1, 2 and 3 studies) have evaluated the effect of repeated cycles of three IL-7 injections and showed that this therapy helped maintaining HIV-infected patients with $CD4^+$ T cells levels above 500 cells/μL (Levy et al. 2012), a level associated with a nearly healthy clinical status (Lewden et al. 2007).

The dynamics of $CD4^+$ T lymphocytes following IL-7 injections can be fitted by mechanistic models based on ordinary differential equations (ODEs). These models contain compartments corresponding to different populations of $CD4^+$ T lymphocytes and biological parameters characterizing these populations. Hence, it was possible to quantify the effect of repeated cycles of IL-7 on $CD4^+$ T lymphocytes on specific parameters. Previous work using data from clinical trials (INSPIRE studies) has shown that IL-7 enhances both proliferation and survival of $CD4^+$ T lymphocytes (Thiebaut et al. 2014). Moreover, a differential effect of the injections within a given cycle has been found, the third injection of a cycle appearing to have a weaker effect on proliferation than the first ones (Jarne et al. 2017).

In addition to providing insight into the most important mechanism of the effect of exogenous IL-7, the models have shown a very good predictive capacity (Thiebaut et al. 2014; Jarne et al. 2017). Hence, the next step was the determination of the best protocol of injections. A first approach, realized in Jarne et al. (2017), consisted in simulating and comparing the regular protocol to three other protocols with different numbers of injections by cycle. In all four protocols, CD4 counts were measured every 3 months, and a new cycle was administered when the CD4 numbers were below 550 cells/μL. Comparison was based on three criteria: number of injections received, mean CD4 count and time spent below 500 cells/μL over a 4-years period. Results showed that cycles of two injections could be sufficient to maintain CD4 levels, while using less injections than in the clinical protocol. These results suggest the possibility to reduce the number of injections in clinical protocols. However, the 3 months delay between visits is independent of the patient and constrains the protocol. While some patients with "not too low" baseline CD4 levels could afford coming back later than 3 months after the last visit, some patients with "low" baseline CD4 levels would need more repeated cycles or more injections by cycle. Individualized protocols could help in achieving the maintenance of the patient's $CD4^+$ T lymphocytes levels over a

given threshold by using different patient-dependent timing of injections and doses. The possibility of conducting the lightest intervention for every patient could be very important for the development of IL-7 in HIV-infected patients especially for further large clinical trials.

Optimization of schedule and doses is an up-to-date question when working on protocol of injections. In their review on mathematical modeling for immunology, Eftimie et al. (2016) emphasize the need for complex optimal control approaches coupled with immunology experiments, in order to improve clinical interventions. Basically, there are two kinds of techniques that can be used to solve optimal control problems: methods involving Pontryagin's maximum principle and dynamic programming approaches. Pontryagin's maximum principle has been applied to a number of biological problems of the form $\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f(x(t), u(t))$, where the solution to the ordinary differential equation depends on the dynamics of the control function $u(t)$. For example, it was applied to the determination of the optimal schedule of dendritic cells vaccine injection in cancer immunotherapy by Castiglione and Piccoli (2006), Cappuccio et al. (2007), Castiglione and Piccoli (2007) and Pappalardo et al. (2010). However, in our case, the model is a piecewise deterministic Markov model (PDMP), where dynamics of IL-7 are unknown and not modeled. Addressing the objective of spending the least time possible under the threshold of 500 cells/μL by using repeated injections of IL-7 corresponds in a more formal way to determining actions (injection or not and choice of dose) at given time points over a horizon of time: this can be treated as a problem of impulse control in the optimal control theory. To the best of our knowledge, there is no maximum principle solving this kind of problem. We will focus on a dynamic programming method, as developed in Costa et al. (2016). In a formal mathematical framework, we addressed the question of optimizing the schedule of IL-7 injections for a given patient by a two-steps method: determining an adapted mathematical model for the process, and developing a numerical method to determine an optimal strategy of IL-7 injections for a given patient.

As described in Davis (1984), most of the continuous-time stochastic problems of applied probability (including those modeling biological processes) consist of some combination of diffusion, deterministic motion and/or random jumps. Ordinary differential equations can be included in the class of deterministic motion with random jumps. In our particular framework of modeling cell dynamics after IL-7 injections, jumps correspond to the change of some parameters value. This can be easily and naturally modeled by the largely studied class of Piecewise Deterministic Markov Processes (PDMPs). A non-controlled version of this model can be described by iteration as follows: from a point in the state space, the process follows a deterministic trajectory determined by the flow, until a jump occurs. This jump happens either spontaneously in a random manner, or when the flow hits the boundary of the state space. After the jump, the system restarts from a new point determined by the transition measure of the process. We will show in this article how to model the dynamics of the $CD4^+$ T cells in HIV-infected patients following IL-7 injections using a PDMP.

According to the problem studied in Costa et al. (2016), impulse control consists in possible actions only when the process reaches its boundary. This will constitute our framework: the decision-maker has the possibility to inject IL-7 when the number of $CD4^+$ T lymphocytes reaches a given level or when a certain amount of time has

passed since the last injection. Each action has a cost, and a strategy is defined as the set of all realized actions over a given horizon. The impulse problem consists in determining a strategy of injections minimizing the optimality criterion induced by the cost function. In our case, the cost function depends on the number of injection realized and the time spent with the $CD4^+$ T lymphocytes levels under the threshold of 500 cells/μL, as both quantities should be minimized.

As emphasized by the authors of Dufour and Zhang (2015), the development of computational methods for the control of PDMPs has been limited, and at the moment, there is no general method allowing the numerical resolution of optimal control on PDMPs (and in particular impulse control). This constitutes a real challenge. We propose in this work a numerical method based on the results developed in Costa et al. (2016). In this paper, the authors studied the existence of a solution of the Bellman–Hamilton–Jacobi equation by showing that the value function is the limit of a sequence of functions given by iteration of an integro-differential operator. This construction leads to a natural method for the computation of the optimal cost and the determination of an optimal strategy of injections. In particular, we have developed an algorithm for the iteration of the operator and applied our numerical tool to the case of the biological model. This provides a proof of concept as it succeeded in determining an optimal strategy for a number of pseudo-patients simulated using previous estimations. The paper is organized as follows: Sect. 2 presents the mathematical modeling of the process, including data and design of INSPIRE studies, as well as mechanistic model and finally the associated PDMP. Section 3 focuses on the optimal control problem, by reminding the main theoretical results from Costa et al. (2016) and adapting them to the IL-7 study. Section 4 presents some numerical aspects of the dynamic programming work, necessary to determine the optimal cost function and strategy for a given patient. Results are presented in Sect. 5, and discussion is done in Sect. 6.

2 Mathematical Modeling

2.1 Material

Our work is based on three phase I/II multicenter studies assessing the effect of a purified glycosylated recombinant human Interleukin 7 (IL-7) treatment for immune restoration in HIV-infected patients under treatment: INSPIRE (Levy et al. 2012), INSPIRE 2 and INSPIRE 3 (Thiébaut et al. 2016). A total of 128 HIV-infected patients under antiretroviral therapy with $CD4^+$ T cell count between 100 and 400 cells/μL and undetectable viral load for at least 6 months were included among the three studies from the time of the first injection. IL-7 was administered in cycles of weekly injections, with a "complete cycle" defined as three weekly injections. In INSPIRE, all 21 patients received complete cycles of IL-7 at different weight-dependent doses: 10, 20 and 30 μg/kg. In INSPIRE 2 and INSPIRE 3, 23 and 84 patients (respectively) received repeated (and sometimes incomplete) cycles of IL-7 at dose 20 μg/kg. Repeated visits and follow-up once every 3 months after the first cycle allowed to measure biomarkers levels in patients, in particular total $CD4^+$ T cell counts and number of proliferating $CD4^+$ T cells through Ki67 marker. At every visit, a new cycle of injections was

administered if the patient's $CD4^+$ T cell level was under 550/μL, in order to globally maintain the levels above 500 cells/μL. The total duration of the studies was 12, 24 and 21 months for INSPIRE, INSPIRE 2 and INSPIRE 3, respectively.

2.2 Mechanistic Model

The dynamics of $CD4^+$ T lymphocytes were largely described in Thiebaut et al. (2014) and Jarne et al. (2017) by using several mechanistic models. We focus here on the following model, described in Fig. 1: it includes two populations of cells, non-proliferating (or resting, R) and proliferating (P). Resting cells are produced by thymic output at rate λ, become proliferating cells at rate π and die at rate μ_R. Proliferating cells die at rate μ_P and can also divide and produce two non-proliferating cells at rate ρ. The system of differential equations is written as follows:

$$\begin{cases} \dfrac{\mathrm{d}R}{\mathrm{d}t} = \lambda - \mu_R R - \pi R + 2\rho P \\ \dfrac{\mathrm{d}P}{\mathrm{d}t} = -\mu_P P - \rho P + \pi R \end{cases} \tag{1}$$

We assume the system is at equilibrium at $t = 0$, before the study begins and any injection is administered. IL-7 injections are realized through cycles containing up to three injections with 7 days elapsed between each injection. Parameters estimation was performed using a population approach. Mixed-effect models including intercept, random and fixed effects, were used on log-transformed parameters, in order to both obtain an estimation across population and account for between-individuals variability. In the controlled framework, the decision-maker can decide to inject IL-7 to a patient at a given dose d, and this will affect the value of the proliferation rate π. Each injection denoted by $n \in \{1, 2, 3\}$ of a cycle has a different effect on the value of π for patient i, defined as follows:

$$\tilde{\pi}^i = \tilde{\pi}_0 + \beta_\pi^{(n)} d_i^{0.25} \mathbf{1}_{\{t \in [t_{\text{inj}}^i, t_{\text{inj}}^i + \tau^i]\}} \tag{2}$$

with $\tilde{\pi} = \log(\pi)$; $\beta_\pi = [\beta_\pi^{(1)}, \beta_\pi^{(2)}, \beta_\pi^{(3)}]$ is the vector of effect of each injection of a single cycle; d is the injected dose; t_{inj} is the time (in days) at which IL-7 is injected,

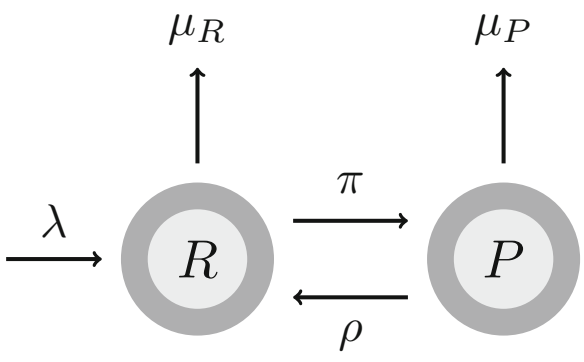


Fig. 1 Mechanistic model for the dynamics of $CD4^+$ T lymphocytes

and τ is the length of effect of the injection (in number of days), considered equal to 7 in previous models (Jarne et al. 2017). Estimation of parameters showed that effect of successive injections on the proliferation rate decreases within a cycle, and in particular, the third injection seems to have a much weaker effect (as $\beta_\pi^{(3)} < \beta_\pi^{(2)} < \beta_\pi^{(1)}$) (Jarne et al. 2017).

2.3 Mathematical Model: Piecewise Deterministic Markov Process

As described in the introduction, ODEs-based mechanistic models can be included into the broader class of PDMPs. A PDMP is characterized by a state space in which it evolves, a flow, a jump intensity and a measure of transition. From a mathematical point of view, we note X the state space, an open subset of $\mathbb{R}^d$, $d \in \mathbb{N}$, and ∂X its boundary. The flow associated with the process is $\phi(x,t) : \mathbb{R}^d \times \mathbb{R} \mapsto \mathbb{R}^d$. The active boundary is defined as $\Xi = \{x \in \partial X : x = \phi(y,t)\}$ for some $y \in X$ and $t \in \mathbb{R}_+^*$. We will then denote $\overline{X} = X \cup \Xi$ and for $x \in \overline{X}$, we can define $t^*(x) = \inf\{t \in \mathbb{R}_+ : \phi(x,t) \in \Xi\}$. The controlled jump intensity η is a $\mathbb{R}_+$-valued measurable function and determines the law of the stochastic jumps. When the process, i.e., the trajectory of $CD4^+$ T lymphocytes, reaches Ξ, the decision-maker can act by injecting IL-7 to the patient. The action varies according to the dose injected. This leads to a jump in some parameters value, and the process restarts from a new point defined by the transition measure $Q(.|\phi(x_0,\tau),d)$, depending on the dose and the position of the state before the jump $\phi(x_0,\tau)$.

In this section, we present the PDMP associated with the biological process described in Sect. 2.2. Here, the PDMP is patient-dependent. As we focus on the control question (and not the estimation one), we suppose that parameters values of the studied patient are known. Previous work has shown that estimation of ODE's parameters based on population approaches can be reliable (Lavielle and Mentré 2007; Prague et al. 2013). Moreover, the model developed in our particular framework for the CD4 dynamics has shown good predictive abilities (Jarne et al. 2017). Therefore, we make the assumption that we determine the strategy for a patient who is already included in a clinical study and for which we had enough observations to estimate its parameters (by running a first cycle of injection for example). As developed in this part, the stochasticity is supposed to be induced by the biological model but not by the uncertainty on the parameters estimation. Sensitivity analysis of the method regarding the estimation uncertainty is provided in "Appendix C".

The PDMP modeling the dynamics of $CD4^+$ T lymphocytes of a given patient is defined using six variables: the state vector is denoted by $x = (\gamma, n, \sigma, \theta, p, r)$. γ determines the value of parameter π when combined with n, the number of injections realized in the ongoing cycle. If $d = [d_0, d_1, .., d_{m_d}]$ is the vector of all possible doses (with $d_0 = 0$), then $\gamma \in \{1..m_d + 1\}$. Injecting dose d_k at the n-th injection of a cycle gives the following: $\gamma(d_k) = k + 1$ and $\pi = \pi_0 + \beta_\pi^{(n)} d(\gamma)^{0.25}$. The two variables σ and θ are time variables, discretized with steps of 1 day. In particular, σ corresponds to the number of days since the last injection and θ to the running time ($\theta = 1$ at the first injection of the first cycle). Finally, variables p and r are values of compartments P and R solutions of system 1 with parameter π defined by γ and n and other parameters

are supposed to have been previously estimated. We suppose the patient is followed until a horizon of time T_h, then the state space is $X = \tilde{X} \cup \Delta$ with

$$\tilde{X} = \{1..m_d + 1\} \times \{1..n_{\text{inj}}\} \times \{0..T_h - 7(n_{\text{inj}} - 1)\} \times \{0..T_h\} \\ \times \{p_{\min}..p_{max}\} \times \{r_{\min}..r_{max}\}$$

and Δ is an absorbing state representing the end of the study, at $t = T_h$: $\Delta = (0, 0, 0, T_h, 0, 0)$. For $x = (\gamma, n, \sigma, \theta, p, r) \in X$, the flow is defined as:

- $\phi(x, t) = (\gamma, n, \sigma + t, \theta + t, p, r)$ if $\theta \leq 1$
- $\phi(x, t) = (\gamma, n, \sigma + t, \theta + t, P(t, \gamma, n), R(t, \gamma, n))$ if $\theta \in [1, T_h - 1]$, with $P(t, \gamma, n)$ and $R(t, \gamma, n)$ solutions of system 1 with initial conditions p and r and π determined with γ and n
- $\phi(\Delta, t) = \Delta$

Moreover, even if the deterministic mechanistic model allowed good fits for the data, we make the hypothesis that the process undergoes some stochasticity: in particular, as the value of parameter π is modified by an injection of IL-7 during some days, we suppose that this modification can last randomly up to 7 days after the injection. Stochastic jumps can then occur with intensity η such that for $x \in \tilde{X}$, $\eta(x) = \eta \mathbf{1}_{\{\gamma > 1\}}$ with η a given value and $\eta(\Delta) = 0$. It means that if we consider the modification of π value after an IL-7 injection through Eq. 2, τ follows there a random exponential law of parameter η. We define the constant $K = \eta$ such that $\eta(x) \leq K$ for every $x \in \tilde{X}$.

IL-7 injections aim at maintaining the $CD4^+$ T cell level over 500 cells/μL. When this value is reached, we consider that the system has reached a deterministic boundary of the state space. A new injection of IL-7 injection is possible at that moment and gives the possibility to increase $CD4^+$ T cell counts. To account for clinical constraints, we assume a minimum time $\sigma_{\min}$ is observed between the beginning of two consecutive cycles, even if the number of CD4 falls below the threshold of 500. During cycles, the deterministic boundary corresponds to the 7 days delay between injections. In a more formal way, the boundary can actually be reached in five different situations described in the following:

- for a technical reason due to the mathematical modeling which cannot account for an impulse action at $t = 0$, we define a first artificial boundary when the study begins, at $\theta = 1$: $\Xi_1 = \{x : \theta = 1\}$. This allows a cycle of injections to begin at $\theta = 1$. We suppose the studied patient is already included in the clinical study: it means that its biological parameters are known, and her/his $CD4^+$ T cell count at $t = 0$ as well (either because she/he is at equilibrium, and the values are known from biological parameters, or because some measures have been realized at this time).
- we also define a time corresponding to the end of the study and a boundary when the time reaches the horizon T_h:$\Xi_2 = \{x : \theta \geq T_h\}$
- another boundary is reached when the patient is undergoing a cycle of injections and 7 days have passed since the last injection: $\Xi_3 = \{x : n < n_{\text{inj}}, \sigma = 7, \theta < T_h\}$
- we also consider a boundary when at least one cycle was already achieved and the count of cells is equal to or below the threshold of 500 cells/μL. We also assume a

minimum time $\sigma_{\min}$ is observed between the beginning of two consecutive cycles: $\Xi_4 = \{x : p + r \leq 500, n = n_{\text{inj}}, \sigma \geq \sigma_{\min}, \theta < T_h\}$
– finally, an artificial boundary is created when π has not returned to its baseline value 7 days after the last injection of a cycle: $\Xi_5 = \{x : \gamma > 1, n = n_{\text{inj}}, \sigma = 7, \theta < T_h\}$

We define the active boundary as $\Xi = \Xi_1 \cup \Xi_2 \cup \Xi_3 \cup \Xi_4 \cup \Xi_5$. In this process, actions (IL-7 injections) can only be realized when the process hits the active boundary. We model the possibility of not doing an injection in a given cycle by using a fictive dose d_0 equal to zero. When beginning a new cycle of injections, the first injection needs to be positive though. The possible action made by the decision-maker depends on the boundary reached. Therefore, for every $x \in \Xi$:

$$A(x) = \begin{cases} \{d_1, ..d_{m_d}\} & \text{if } x \in \Xi_1 \cup \Xi_4 \\ \{0, d_1, ..d_{m_d}\} & \text{if } x \in \Xi_3 \\ \emptyset & \text{if } x \in \Xi_2 \cup \Xi_5 \end{cases}$$

We also define the transition measure (or Kernel): it determines the new point from which the process restarts after a jump. It depends on the injected dose only when the boundary of the process is reached. All possible situations are the following:

– when the flow hits Ξ_1, the study begins with administration of a cycle of injections. γ takes the value corresponding to the chosen dose. $(p, r) = (P_c, R_c)$, known values from either equilibrium or biological measures made on the patient before the beginning of the study
– when the flow hits Ξ_2, the study is over and nothing happens from absorbing state Δ
– when the flow hits Ξ_3, a new injection is administered to the patient. γ takes the value corresponding to the chosen dose $\gamma(d)$, n increases by one, σ goes back to 0
– when the flow hits Ξ_4, a new cycle of injections begins. γ takes the value corresponding to the chosen dose, n goes back to 1, σ goes back to 0
– when the flow hits Ξ_5, there is no injection. γ goes back to 1
– in case of spontaneous jump, there is no injection and γ goes back to 1

In a formal way, the Kernel Q is written:

$$\begin{aligned} Q(\mathrm{d}y|x, d) = {} & \delta_{(\gamma(d),1,0,1,P_c,R_c)}(\mathrm{d}y)\mathbf{1}_{\{x\in\Xi_1\}} + \delta_{\Delta}(\mathrm{d}y)\mathbf{1}_{\{x\in\Xi_2\}} \\ & + \delta_{(\gamma(d),n+1,0,\theta,p,r)}(\mathrm{d}y)\mathbf{1}_{\{x\in\Xi_3\}} + \delta_{(\gamma(d),1,0,\theta,p,r)}(\mathrm{d}y)\mathbf{1}_{\{x\in\Xi_4\}} \\ & + \delta_{(1,n,\sigma,\theta,p,r)}(\mathrm{d}y)\mathbf{1}_{\{x\in\Xi_5\}} + \delta_{(1,n,\sigma,\theta,p,r)}(\mathrm{d}y)\mathbf{1}_{\{x\in\tilde{X}\}} \end{aligned}$$

The impulse control problem consists in determining the optimal scheme of injections and their associated dose according to a given optimality criterion, based on the cost function C: in our case, this cost function depends on the number of injections realized and the time spent with the $CD4^+$ T lymphocytes levels under the threshold of 500 cells/μL. Both quantities need to be minimized, in order to maintain the patient in good health by injecting the least possible. The cost can be divided in two parts. First, the gradual cost penalizes the trajectory of the process through the time spent

under the threshold after the beginning of the first cycle. This time is considered in months, approximately, as it is computed as the number of days divided by 30. For $x = (\gamma, n, \sigma, \theta, p, r) \in \tilde{X}$:

$$C^g(x) = \frac{1}{30}\mathbf{1}_{\{p+r<500\}}\mathbf{1}_{\{\theta\geq 1\}}$$

Then, the cost associated with an impulsive action penalizes the fact of injecting IL-7 to the patient:

$$C^i(x, d) = \mathbf{1}_{\{x\in\Xi_1\cup\Xi_4\}} + \mathbf{1}_{\{d\neq 0\}}\mathbf{1}_{\{x\in\Xi_3\}}$$

After the horizon, the cost is null, as $C^i(\Delta) = C^g(\Delta) = 0$.

3 Optimal Control

In this section, we will first remind the main theoretical results obtained in Costa et al. (2016), then we will transpose these results to our particular context.

3.1 Main Theoretical Results

The objective of this section is to adapt some results obtained in Costa et al. (2016) to our specific context. We follow closely their notation. The set of all realized injections over a given horizon constitutes the strategy of injections. In a more formal way, a strategy u of the decision-maker is a sequence $u = \{u_n\}_{n\in\mathbb{N}}$ of functions $u_n : X \mapsto A$ giving the action to realize at punctual time points $t_n \geq 0$ when the system is in state $x \in X$. The set of all admissible strategies is noted $\mathcal{U}$. According to section 2.2 in Costa et al. (2016), there exists a continuous-time stochastic process ξ defined on probabilistic space using characteristics ϕ, η and Q depending on the action given by u, such that ξ_t, $t \in \mathbb{R}_+$ corresponds to the state of the variables at time t. To each admissible strategy $u \in \mathcal{U}$, we associate a discounted cost optimality criterion depending on the gradual cost on the trajectory of the process ξ, C^g, and the cost related to an injection, C^i, as defined in Sect. 2.3:

$$\begin{aligned}\mathcal{V}(u, x_0) = \mathbb{E}^u_{x_0}\left[\int_{]0,+\infty[} e^{-\alpha s} C^g(s)ds\right] \\ + \mathbb{E}^u_{x_0}\left[\int_{]0,+\infty[} e^{-\alpha s} I_{\{\xi_{s-}\in\Xi\}} \int_{\mathbf{A}(\xi_{s-})} C^i(\xi_{s-}, a)u(da|s)\mu(ds)\right]\end{aligned} \tag{3}$$

with $\alpha > 0$ the discount factor and where μ is a measure that counts the number of jumps in the process. The impulse control problem aims at finding a strategy u minimizing the discounted cost optimality criterion. Here we want to determine the patient-specific schedule of injections and their dose to optimize the patient's $CD4^+$ T lymphocyte numbers by using a minimum number of injections. The theorem allowing to determine the optimal cost and providing an optimal strategy is adapted from Theorem 5.5 in Costa et al. (2016). It is stated as followed:

Theorem 1 *Suppose assumptions A, B and C from section 3.2 in* Costa et al. (2016) *are verified. We define the sequence of functions* $\{W_q\}_{q\in\mathbb{N}}$ *for any* $x \in \overline{\mathbf{X}}$ *as follows:*

$$\begin{cases} W_{q+1}(x) = \mathfrak{B}W_q(x) \text{ for } q \in \mathbb{N} \\ W_0(x) = -K_A \mathbf{1}_{A_{\varepsilon_1}}(x) - (K_A + K_B)\mathbf{1}_{A^c_{\varepsilon_1}}(x) \end{cases} \tag{4}$$

with constants K_A *and* K_B *defined as in section 5 of* Costa et al. (2016), $A_{\varepsilon_1} = \{x \in \mathbf{X} : t^*(x) > \varepsilon_1\}$ *and*

$$\mathfrak{B}V(y) = \int_{[0,t^*(y)[} e^{-(K+\alpha)t}\mathfrak{R}V(\phi(y,t))dt + e^{-(K+\alpha)t^*(y)}\mathfrak{T}V(\phi(y,t^*(y))) \tag{5}$$

with real-value functions $\mathfrak{R}V$ *and* $\mathfrak{T}V$ *defined for any* V*, respectively, on* X *and* Ξ*:*

$$\begin{aligned} \mathfrak{R}V(x) &= C^g(x) + qV(x) + \eta V(x) \\ \mathfrak{T}V(z) &= \inf_{d\in\mathbf{A}(z)} \left\{ C^i(z,d) + QV(z,d) \right\} \end{aligned}$$

q *being the signed kernel, which computes the difference between the states before and after the spontaneous jump. For* $x \in X$*, it is defined with:*

$$q(dy|x) = \eta(x)[Q(dy|x) - \delta_x(dy)]$$

The sequence of functions $\{W_q\}_{q\in\mathbb{N}}$ *converges to a function* W *defined on the state space and such that:*

(i) $W(x_0) = \inf_{u\in\mathcal{U}} \mathcal{V}(u, x_0)$*, with* $\mathcal{V}$ *defined as in Eq. 3*
(ii) *there is a measurable mapping* $\widehat{\varphi} : \Xi \to \mathbf{A}$ *such that* $\widehat{\varphi}(z) \in \mathbf{A}(z)$ *for any* $z \in \Xi$ *and satisfying*

$$C^i(z, \widehat{\varphi}(z)) + QW(z, \widehat{\varphi}(z)) = \inf_{d\in\mathbf{A}(z)} \left\{ C^i(z,d) + QW(z,d) \right\}. \tag{6}$$

This theorem allows to determine the optimal cost and an optimal injection strategy, consisting in choosing the optimal action $\widehat{\varphi}(z)$ for every point $z \in \Xi$ reached on the trajectory of the process. Indeed, the iteration of the sequence $\{W_q\}_{q\in\mathbb{N}}$ defined by Eq. 4 can be realized by numerically approximating the operator $\mathfrak{B}$ defined in Eq. 5. This will give an approximation of the function W, and in particular of $W(x_0)$, corresponding to the optimal cost. Moreover, to obtain an optimal strategy, the process is the following: we simulate a trajectory from x_0, then when a boundary is reached, the chosen action corresponds to the one minimizing the criterion $C^i(z,d) + QW(z,d)$, as given by Eq. 6.

3.2 Application

The process describing the effect of IL-7 on $CD4^+$ T lymphocytes dynamics is now well defined by its characteristics ϕ, η and Q, boundaries and possible actions in

Sect. 2.3. Moreover, both gradual cost on the trajectory and impulse cost were defined in that section. We will quickly describe in this part how to apply the results from Theorem 1 for our specific problem, i.e., determining the function $\mathfrak{B}$ needed for the computation of the optimal strategy. For a more detailed and formal computation, we refer the reader to "Appendix A". We need to compute:

$$\mathfrak{B}V(y) = \int_{[0,t^*(y)[} e^{-(K+\alpha)t}\mathfrak{R}V(\phi(y,t))\mathrm{d}t + e^{-(K+\alpha)t^*(y)}\mathfrak{T}V(\phi(y,t^*(y)))$$

We define

$$G(V,y) = \int_{[0,t^*(y)[} e^{-(K+\alpha)t}\mathfrak{R}V(\phi(y,t))\mathrm{d}t \tag{7}$$

and

$$H(V,y) = e^{-(K+\alpha)t^*(y)}\mathfrak{T}V(\phi(y,t^*(y)) \tag{8}$$

We define a time interval Δt (in practice equal to 1 day) and for every $y = (\gamma, n, \sigma, \theta, p, r) \in \tilde{X}$, we note

$$n^*(y) = \left\lfloor \frac{t^*(y)}{\Delta t} \right\rfloor$$

For every $j \in \{0..n^*(y)-1\}$, we denote $\phi_j(y,t) = \phi(y, j\Delta t)$ and $\phi(y,t^*(y)) = (\gamma, n, \sigma + t^*(y), \theta + t^*(y), p^*(y), r^*(y))$. The integral defined in Eq. 7 is computed by approximation using the classic trapezoidal rule using the $j\Delta t$ nodes. Thus, $G(V,y)$ can be approximated by a linear combination of $\{V(y_j)\}_{j\in\{0..n^*(y)-1\}}$, with y_j depending on $\phi_j(y,t)$. Moreover, $H(V,y)$ is proportional to $V(\overline{y})$, with $\overline{y}$ depending on the boundary reached in $\phi(y,t^*(y))$. Finally, for every point $y \in \tilde{X}$, if we note $\overline{y} = y_{n^*(y)}$, $\mathfrak{B}V(y)$ can be computed as a linear combination of $\{V(y_j)\}_{j\in\{0..n^*(y)\}}$.

4 Numerical Aspects of the Dynamic Programming Method

From Theorem 1, we know that we need to compute the sequence $\{W_q\}_{q\in\mathbb{N}}$ such that for $y \in \overline{\mathbf{X}}$, $W_0(y) = -K_A\mathbf{1}_{A_{\varepsilon_1}}(y) - (K_A + K_B)\mathbf{1}_{A^c_{\varepsilon_1}}(y)$ and $W_{q+1}(y) = \mathfrak{B}W_q(y)$ for $q \in \mathbb{N}$. The sequence converges to a function W defined on $\overline{\mathbf{X}}$ that allows the determination of the optimal cost and the optimal protocol of injections achieving that cost. This computation is realized on a grid of the state space: at each iteration q, a new matrix is computed, each element on line v and column s corresponding to $\mathfrak{B}W_q(x_{vs})$, with x_{vs} element of the grid Γ of the state space. The implementation of our algorithm was realized in Matlab version R2016b (The MathWorks, Inc., Natick MA, USA, 1984). In this section, we give elements to understand how the method is implemented. The structure of the code is detailed in "Appendix B".

4.1 Discretization of the State Space

The grid Γ contains points of the form $(\gamma, n, \sigma, \theta, p, r)$. γ and n are discrete variables with $\gamma \in \{1 \dots m_d + 1\}$, $n \in \{1..n_{\text{inj}}\}$. σ and θ are discretized with a time step of 1 day, with $\sigma \in \{0..\sigma_{\max}\}$ and $\theta \in \{0..T_h\}$. Solutions p and r of the ODE system are continuous and are discretized in a regular grid, with $p \in \{p_{\min}..p_{\max}\}$ with regular step h_p and $r \in \{r_{\min}..r_{\max}\}$ with regular step h_r. We then obtain:

$$n_p = \frac{p_{\max} - p_{\min}}{h_p} + 1$$
$$n_r = \frac{r_{\max} - r_{\min}}{h_r} + 1$$

h_r and h_p are chosen such that both $n_p, n_r \in \mathbb{N}$ count the number of values of p and r on the grid, respectively.

4.2 Organization of the Grid

We arrange all points of the grid Γ in a matrix M of size $N_{\text{sum}} \times N_{pr}$, with N_{sum} corresponding to the number of possible $(\gamma, n, \sigma, \theta)$ combinations and $N_{pr} = n_p n_r$ number of possible (p, r) combinations. Each element $M(v, s)_{\substack{v \in \{1..N_{\text{sum}}\} \\ s \in \{1..N_{pr}\}}}$ corresponds to a given combination $(\gamma, n, \sigma, \theta, p, r)$ of Γ, through the following bijection:

$$\begin{aligned} \chi : \Gamma &\to \{1..N_{\text{sum}}\} \times \{1..N_{pr}\} \\ x_{vs} = (\gamma, n, \sigma, \theta, p, r) &\mapsto (v, s) = \Big(\chi_l(\gamma, n, \sigma, \theta),\, \chi_c(p, r)\Big) \end{aligned}$$

χ_l is defined in the following way: v corresponds to a given value of $(\gamma, n, \sigma, \theta)$. Possible combinations of (σ, θ) depend on the value of (γ, n): for example, during the first cycle, when $n = 1, \sigma = 0$ is associated with $\theta = 1$, while when $n = 2, \sigma = 0$ is associated with $\theta = 8$. We divide the lines of matrix M by defining then $N_{\gamma n} = (m_d+1)n_{\text{inj}}$ blocks, corresponding to the possible combinations of (γ, n). Each block is indexed by $i = f(\gamma, n) = \gamma + (m_d+1)(n-1) \in \{1..N_{\gamma n}\}$ and contains combinations of (σ, θ), indexed by $j = g_i(\sigma, \theta) \in \{1..N_{b_i}\}$ within the i-th block. The total number of lines of matrix M is the sum of the number of lines in each block: $N_{\text{sum}} = \sum_{i=1}^{N_{\gamma n}} N_{b_i}$. We can define a vector $l_{block} = (1, 1 + N_{b_1}, .., 1 + \sum_{i=1}^{k} N_{b_i}, .., \sum_{i=1}^{N_{\gamma n}-1} N_{b_i})$ of length $N_{\gamma n}$, that determines the index of the first line of each block. Finally:

$$v = \chi_l(\gamma, n, \sigma, \theta) = l_{block}(i) + j - 1$$

with $i = f(\gamma, n)$ and $j = g_i(\sigma, \theta)$. χ_c is defined in the following way:

$$s = \chi_c(p, r) = \frac{p - p_{\min}}{h_p} + 1 + n_p \frac{r - r_{\min}}{h_r}$$

such that $s = 1$ when $(p, r) = (p_{\min}, r_{\min})$ and $s = n_p n_r$ when $(p, r) = (p_{\max}, r_{\max})$.

4.3 Iteration of the Algorithm

Each iteration of the algorithm computes then a matrix M_q such that

$$M_q(v,s) = W_q\Big(M(v,s)\Big) = W_q(x_{vs})$$

For every $x = (\gamma, n, \sigma, \theta, p, r) \in \Gamma$, $W_q(x)$ is a linear combination of some $W_q(x_m)$, $m \in \{1..M_x\}$, as shown in Eq. 12 from "Appendix A". Values of $W_q(x_m)$ are given by $M_q(\chi(x_m))$; they are linearly combined and implemented in $M_{q+1}(\chi(x))$.

4.4 Convergence Criterium

We assume that the sequence converges when $\|W_{q+1} - W_q\|_\infty < \epsilon$. In practice, we compute $\max_{v,s} |M_{q+1}(v,s) - M_q(v,s)|$ and we consider that the sequence converges with $\epsilon = 0.001$. It usually occurs after 35 to 45 iterations.

5 Results

We applied the previously described method to the model detailed in Sect. 2.3, with a choice of $m_d = 2$ possible doses: $d = [0, 10, 20]$ (unit = μg/kg), cycles of 3 injections: $n_{\text{inj}} = 3$ and a reduced horizon $T_h = 365$ days. We also assumed a minimum time of $\sigma_{\min} = 30$ days between the end of a cycle and the beginning of a new one. For a given patient with fixed biological parameters, we can approximate the function W in a grid of the state space through convergence of the sequence $\{W_q\}$: this determines the optimal cost over all strategies. Moreover, using Eq. 6 from Theorem 1, we can simulate the strategy choosing the optimal action to realize when reaching the boundary of the state space and compute the cost of the obtained strategy. As some randomness is included in the model by the time of effect of an injection of parameter π, we simulate $N = 5000$ realizations of a protocol on a given patient with a Monte Carlo method and compute the expectation of its cost. From that, we check the numeric performance of our method by first comparing the cost of the optimal strategy to the computation of the optimal cost from the value function W. Moreover, we wish to compare the optimal strategy to other "naive" protocols. For each protocol, including the optimal one, we compute by Monte Carlo the mean cost, the standard deviation and the minimum cost achievable. This is usually reached when the patient responds well to all injections, i.e., the effect of the injection on parameter π lasts 7 days after every injection. In order to compare protocols based on clinical criteria, we also computed by Monte Carlo the mean number of $CD4^+$ T cells count until horizon, the mean time spent under 500 cells/μL (in days) and the mean number of injections over all simulations. These comparisons were realized with 50 pseudo-patients. Parameters values were generated from the posterior law estimated on real data from INSPIRE trials in Thiébaut et al. (2016). Patients are divided in three categories according to their initial levels of $CD4^+$ T cells: "very low" baseline (100−200 cells/μL), "low" baseline (200−300 cells/μL) and "not too low" baseline (300−400 cells/μL). Table 1 sums up the characteristics of the pseudo-patients population.

Table 1 Characteristics of the pseudo-patients population

Parameter	Mean (SD)
λ (cells/day)	2.24 (0.39)
ρ (/day)	1.96 (0.84)
π_0 (/day)	0.0461 (0.0035)
μ_R (/day)	0.0503 (0.0033)
μ_P (/day)	0.0717 (0.014)
β_{π_1}	0.958
β_{π_2}	0.752
β_{π_3}	0.143
Category	**Number of patients (%)**
Very low	4 (8%)
Low	24 (48%)
Not too low	22 (44%)

Table 2 Comparison of cost values from value function and Monte Carlo simulation

	Patient A very low	Patient B low	Patient C not too low
Optimal cost $W(x_0)$	9.47	6.11	2.87
Cost of optimal strategy: mean (SD) (obtained by Monte Carlo)	9.53 (0.85)	6.20 (0.56)	2.90 (0.36)

We first compare the value of the optimal function obtained from the numerical computation of W with the cost of the optimal strategy. For a sake of clarity, we show detailed results in Table 2 only for three chosen patients. Patient A is in category "very low", patient B in category "low" and patient C in category "very low". We note that for these three patients the two cost values are very similar, meaning that we make a good approximation of the value function with our numerical method. We make the same observation on the 47 other patients (data not shown). Also of note is the hierarchy of the cost between the categories of patients. Very low patients have higher optimal costs (between 8.4 and 12) than low (between 3.9 and 9.4) and not too low (between 2.1 and 4.2). This is consistent with the fact that the lower baseline CD4 levels the patient has, the more time will be spent under 500 cells/μL and the more injections are needed, which both increase the cost of the strategy of injections.

We also realized comparisons of several protocols. We simulated five "naive" protocols: P1 with 3-injections cycles, P2 with a first cycle of 3 injections then 2-injections cycles, P3 with 2-injections cycles, P4 with a first cycle of 2 injections then 1-injection cycles and P5 with 1-injection cycles, all protocols with dose 20. Assessing the cost of these protocols is interesting as they imply variable trajectories within the same patient as well as different values for clinical criteria. Moreover, they would be clinically feasible and represent a good basis for comparison for our optimal strategy. For every protocol k, we note $\mathcal{P}_{+k}$ the space of patients such that cost of optimal strategy is lower than cost of protocol k and n_{+k} its size. We have computed the mean relative

Table 3 Computation of the mean relative variation of cost value (MRC) for every protocol allows determining the mean percentage of gain in term of cost function when using the optimal strategy over protocol k

Protocol	P1	P2	P3	P4	P5
n_+	50	50	50	49	50
MRC (%)	43	31	20	5.7	8.8

P1 Cycles of 3 injections. *P2* First cycle of 3 injections then cycles of 2 injections. *P3* Cycles of 2 injections. *P4* First cycle of 2 injections then cycles of 1 injection. *P5* Cycles of 1 injection

positive variation of cost value (MRC), as shown in Table 3. We note C_{opt_i} the mean cost of optimal strategy for patient i and C_{Pk_i} the mean cost of protocol k for patient i. The MRC allows computing the mean percentage of gain in term of cost function when using the optimal strategy over protocol k:

$$\text{MRC}_k = \frac{1}{n_{+k}} \sum_{i \in \mathcal{P}_{+k}} \frac{(C_{Pk_i} - C_{\text{opt}_i})}{C_{Pk_i}} \tag{9}$$

Results show that mean cost of the optimal strategy is always lower than all other simulated strategies (except one patient for protocol 4, but this is due to numerical approximation, as for this patient $W(x_0)$=4.0, $C_{\text{opt}} = 4.1$ and C_{P_4}=4.0). The percentage of cost reduction of the optimal strategy compared to the other protocols in the simulated population of pseudo-patients varies from 5.7 to 43%. It confirms that our numerical method allows optimizing the cost function.

In addition to comparing the cost value of all five protocols to the optimal strategy, we have also compared clinical criteria such as the mean number of $CD4^+$ T cells count until horizon, the mean time spent under 500 cells/μL (in days) and the mean number of injections over all simulations. Results of these comparisons are shown in Fig. 2, where each point corresponds to the value of the criterion for one pseudo-patient, and each color corresponds to the category of the patients ("very low", "low" and "not too low" baseline). We observe that mean cost of the optimal strategy is lower than other simulated protocols and the optimal strategy achieves a good balance between all clinical criteria. Even if $CD4^+$ T cells levels are not as high as for protocols P1, P2 and P3, the optimal strategy allows to spend as much time with levels over 500 cells/μL as these protocols by using less injections. Protocol P5 has the same performance as the optimal strategy for "not too low" patients, as these strategies are very often the same on these patients. The same observation is made on protocol P4 and the "low" patients. Overall, Fig. 2 shows that the determined strategy allows optimization of the cost function through the chosen criteria (time spent under 500 and number of injections).

More detailed results of comparison of cost function and clinical criteria between optimal strategy and protocols Pk are displayed in Table 4 for patients A, B and C. For these three patients, we observe again that mean cost of the optimal strategy is lower than all other simulated strategies. For patient A, the optimal strategy is achieved by two first cycles of two injections then cycles of one injection. For patient B, the optimal strategy consists in a first cycle of 2 injections followed by 1-injection cycles.

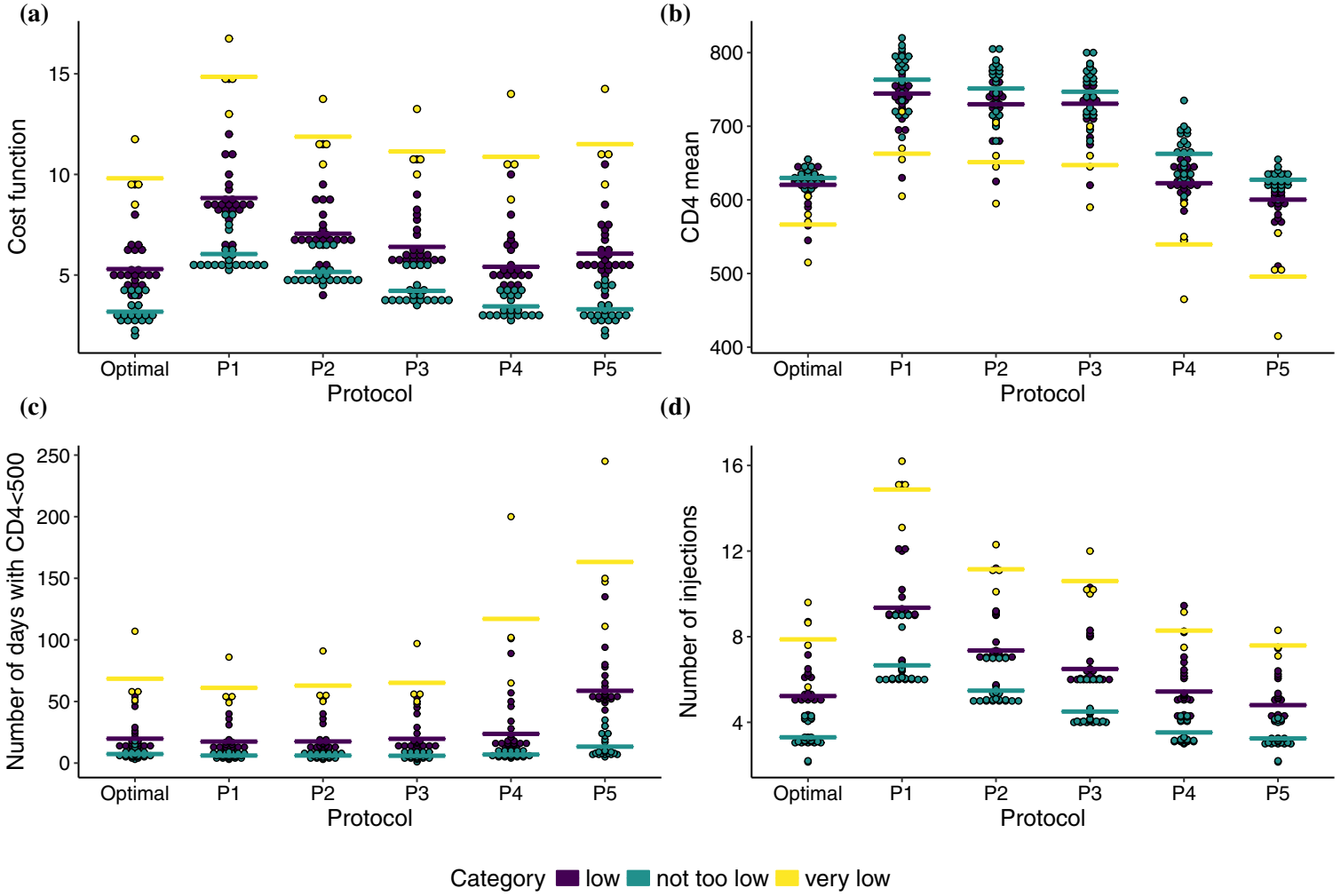


Fig. 2 (Color figure online) Comparison of cost (**a**) and clinical criteria (**b** mean number of $CD4^+$ T cells count until horizon, **c** mean time spent under 500 cells/μL in days, **d** mean number of injections over all simulations) between the determined optimal strategy and the five other protocols. Each point corresponds to the value of a pseudo-patient, with "very low" patients in yellow, "low" in purple and "not too low" in blue. Mean values within each category are represented by horizontal colored lines. P1: cycles of 3 injections. P2: first cycle of 3 injections then cycles of 2 injections. P3: cycles of 2 injections. P4: first cycle of 2 injections then cycles of 1 injection. P5: cycles of 1 injection

Table 4 Comparison of protocols of injections for patients A, B and C

Protocol	P1	P2	P3	P4	P5	Optimal
Patient A						
Mean cost	14.8	11.6	10.7	10.4	11.1	9.53
Std	0.52	0.62	0.68	1.21	1.23	0.85
Min cost	14.6	11.4	10.4	9.56	10.1	8.82
CD4 mean	671	662	659	552	506	578
Days under 500	54.5	55.4	56.5	102	150	58.0
Number of injections	15.1	11.1	10.2	8.25	7.48	8.66
Patient B						
Mean cost	9.35	7.62	6.91	6.26	6.63	6.20
Std	1.14	1.00	0.98	0.70	0.87	0.56
Min cost	8.62	6.91	6.14	5.91	6.12	5.91
CD4 mean	762	742	736	622	598	625
Days under 500	18.4	19.0	24.2	28.1	61.9	24.4
Number of injections	9.83	7.73	6.84	6.05	5.27	6.10
Patient C						
Mean cost	5.54	4.75	3.80	3.08	2.93	2.90
Std	0.29	0.24	0.31	0.37	0.49	0.36
Min cost	5.46	4.70	3.73	2.95	2.79	2.79
CD4 mean	774	762	758	666	631	631
Days under 500	5.56	5.53	5.65	5.90	8.43	5.89
Number of injections	6.02	5.02	4.04	3.13	3.03	3.07

P1 Cycles of 3 injections. *P2* First cycle of 3 injections then cycles of 2 injections. *P3* Cycles of 2 injections. *P4* First cycle of 2 injections then cycles of 1 injection. *P5* Cycles of 1 injection

We can see that the minimum cost is the same for the optimal strategy and protocol P3 (= 5.91): when the patient has a good response to all injections, these strategies are the same. For patient C, the optimal strategy is obtained with 1-injection cycles. Similarly, the minimum cost is the same for the optimal strategy and protocol P5 (= 2.79). For all patients, the optimal strategy is very intuitive: the first complete cycles are needed to raise the number of CD4 over 500 cells/μL; then, 1-injection cycles allow to sustain the levels over 500 cells/μL. For "not too low" patients, CD4 levels are high enough to use only one injection in the first cycle. This helps reducing the number of injections: in patient A, the optimal strategy requires 2–7 less injections than P1, P2 and P3 but allows to spend as much time over 500 cells/ μL. In patient C the optimal strategy requires one less injection as P3 but allows to spend as much time over 500 cells/ μL. It can be noted that a third injection is never used, even for the first cycles of very low patient. It is due to our choice of cost function: it balances the number of injections and the number of months spent under 500 cells/μL. The effect of a third injection is usually too low to allow increasing the time spent over 500 cells/μL by 1 month and is then not chosen as part of the optimal strategy. These results suggest that our numerical method allows to determine an optimal strategy of injections, and the clinical interpretation of the results are consistent with the mathematical method.

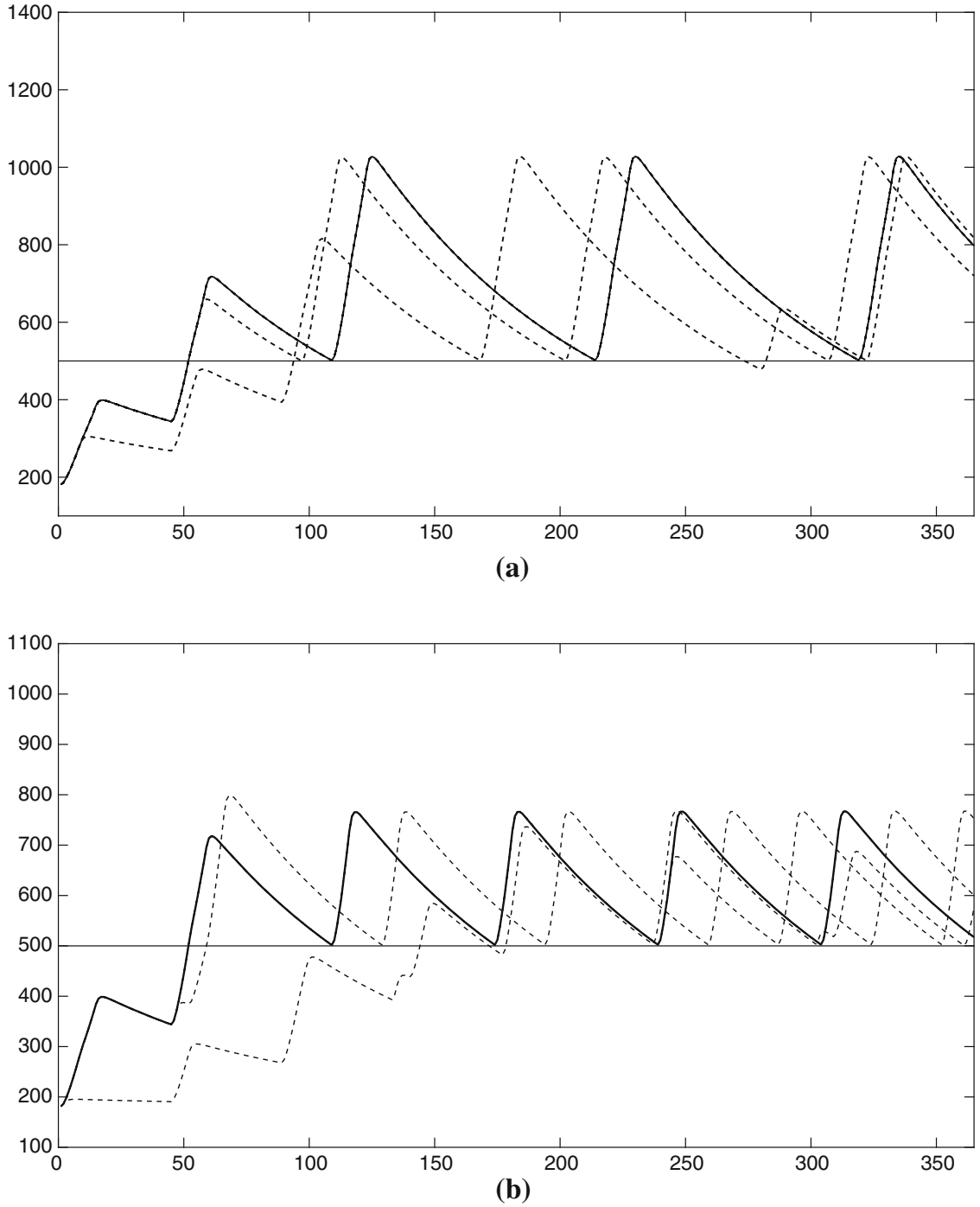


Fig. 3 Dynamics of $CD4^+$ T lymphocytes in patient A. Straight line corresponds to the "best" outcome, i.e., when the effect of all injections lasts 7 days. Dashed line corresponds to other possible trajectories, when this effect can last less than 7 days. **a** Dynamics of $CD4^+$ T lymphocytes in patient A under P3, a 2-injections cycles protocol (dose 20). **b** Dynamics of $CD4^+$ T lymphocytes in patient A under the determined optimal strategy

In terms of trajectories of the process, Fig. 3a, b show some trajectories obtained with, respectively, the 2-injection cycles protocol (P3) and the optimal strategy for patient A. We can note that even if $CD4^+$ levels are globally lower in the optimal strategy compared to the two injections cycles at dose 20 μg/kg, it still allows a maintenance over the threshold of 500 cells/μL by using less injections: indeed, in the best case scenario, the 2-injections cycle strategy implies 5 cycles of 2 injections

which is a total of 10 injections, while the optimal strategy induces 2 cycles of 2 injections and 4 single injections, which is a total of 8 injections. The trajectories for patients B and C are provided in "Appendix D". All together, our results support the interest of determining the optimal strategy based on a criterion combining both the number of injections and the time spent under 500.

6 Discussion

In this work, we have developed a numerical tool allowing to solve an impulse control problem for a PDMP. The specificity of our work is in the development of a dynamic programming method in the context of a specific biological framework. The objective is to determine the optimal strategy of IL-7 injections for a given HIV-infected patient, in order to maintain $CD4^+$ T lymphocytes levels over the threshold of 500 cells/μL. We first modeled the dynamics of $CD4^+$ T lymphocytes during repeated cycles of IL-7 injections by a PDMP. Then, we solved the impulse control problem by iterating a sequence defined by an integro-differential operator. Theoretical results have shown that this sequence converges to the value function, which allows to determine the optimal action that should be realized at every point of the boundary. We proposed a numerical tool approximating the sequence and the value function on a grid of the state space and applied it to our clinical question. As our method relies on numerical approximation, the obtained optimal strategy could be an approximation of the theoretical one. However, the obtained results suggest that we managed to determine optimal strategies for pseudo-patients and that our method allows improving the strategy of injections. Although the horizon of study is only 1 year, these results are also consistent with a clinical interpretation. The optimal strategy determined for different patients is indeed intuitive: the first cycles aim at increasing the $CD4^+$ T lymphocytes levels and should contain as many injections as possible until the levels are acceptable. Then, the following cycles sustain the CD4 levels over the threshold, and punctual injections are sufficient to reach this objective. The optimal strategy, determined with our method, has a lower cost than other possible clinical strategies. Actually, the obtained optimal strategy depends on the cost previously defined, and we could explore other optimal strategies depending on other cost functions. For example, it could be interesting to use different weights on the time spent under 500 cells /μL and the number of injections (depending on the clinician priorities), or to account for the possible negative side effects due to higher doses (this would need additional data on the question). Finally, the model could be extended by studying the patient until a longer horizon (up to 2 years). This rises the issue of the increase in computational time (by increasing the size of the grid of the state space) and constitutes a new challenge in itself. In the end, we hope to use this tool in future possible clinical trial investigating the effect of IL-7 injections with patients-specific schedules of injections, personalized and optimized using this method.

Acknowledgements We would like to thank the main investigators and supervisors of INSPIRE 2 and 3 studies: Jean-Pierre Routy, Irini Sereti, Margaret Fischl, Prudence Ive, Roberto F. Speck, Gianpiero D'Ozi, Salvatore Casari, Sharne Foulkes, Ven Natarajan, Guiseppe Tambussi, Michael M. Lederman, Therese Croughs and Jean-François Delfraissy. This work was supported by the Investissements d'Avenir program managed by the ANR under reference ANR-10-LABX-77.

A Optimal Control: Application

We defined the process describing the effect of IL-7 on CD4$^+$ T lymphocytes dynamics by its characteristics ϕ, η and Q, boundaries and possible actions in Sect. 2.3. We also defined both gradual cost on the trajectory and impulse cost in that section. As we aim at applying the results from Theorem 1 to determine the optimal cost and an optimal strategy by dynamic programming, we need to determine how to compute numerically the function $\mathfrak{B}$ to iterate the sequence $\{W_q\}_{q\in\mathbb{N}}$ defined in Eq. 4. As a reminder, $\mathfrak{B}$ is defined in Costa et al. (2016) by:

$$\mathfrak{B}V(y) = \int_{[0,t^*(y)[} e^{-(K+\alpha)t}\mathfrak{R}V(\phi(y,t))\mathrm{d}t + e^{-(K+\alpha)t^*(y)}\mathfrak{T}V(\phi(y,t^*(y)))$$

We will first detail the computation of $\mathfrak{R}$ then $\mathfrak{T}$, and we will finally show how to compute $\mathfrak{B}$.

Computation of $\mathfrak{R}$

For $x = (\gamma, n, \sigma, \theta, p, r) \in X$, and function $V : \overline{X} \to \mathbb{R}$, $\mathfrak{R}$ is defined as:

$$\mathfrak{R}V(x) = C^g(x) + qV(x) + \eta V(x)$$

with q computing the difference between the states before and after the spontaneous jump. As Q depends on the action only when the process hits the active boundary,

$$\begin{aligned} q(\mathrm{d}y|x,d) &= \eta(x)[Q(\mathrm{d}y|x) - \delta_x(\mathrm{d}y)] \\ &= \mathbf{1}_{\{\gamma>1\}}\eta[\delta_{(1,n,\sigma,\theta,p,r)}(\mathrm{d}y) - \delta_{(\gamma,n,\sigma,\theta,p,r)}(\mathrm{d}y)] \end{aligned}$$

then for every function V, and as $K = \eta$:

$$\begin{aligned} qV(x) &= \int V(y)q(\mathrm{d}y|x) \\ &= \mathbf{1}_{\{\gamma>1\}}K[V(1,n,\sigma,\theta,p,r) - V(\gamma,n,\sigma,\theta,p,r)] \end{aligned}$$

Then

$$\begin{aligned} \mathfrak{R}V(x) &= \frac{1}{30}\mathbf{1}_{\{p+r\leq 500\}} + qV(x) + KV(x) \\ &= \frac{1}{30}\mathbf{1}_{\{p+r\leq 500\}} + KV(1,n,\sigma,\theta,p,r)\mathbf{1}_{\{\gamma>1\}} + KV(x)\mathbf{1}_{\{\gamma=1\}} \end{aligned}$$

Finally,

$$\begin{aligned} \mathfrak{R}V(x) &= \frac{1}{30}\mathbf{1}_{\{p+r\leq 500\}} + KV(1,n,\sigma,\theta,p,r) \\ \mathfrak{R}V(\Delta) &= KV(\Delta) \end{aligned} \tag{10}$$

Computation of $\mathfrak{T}$

For $x \in \Xi$, and function $V : \overline{X} \to \mathbb{R}$, $\mathfrak{T}$ is defined as:

$$\begin{aligned}\mathfrak{T}V(x) &= \inf_{d\in A(x)}\Big\{C^i(x,d)+QV(x,d)\Big\}\\ &= \inf_{d\in A(x)}\Big\{\mathbf{1}_{x\in\Xi_1\cup\Xi_4}+\mathbf{1}_{d\neq 0}1_{x\in\Xi_3}+\int V(y)\Big[\delta_{(\gamma(d),1,0,1,P_c,R_c)}(\mathrm{d}y)\mathbf{1}_{\{x\in\Xi_1\}}\\ &\quad+\delta_{\Delta}(\mathrm{d}y)\mathbf{1}_{\{x\in\Xi_2\}}+\delta_{(\gamma(d),n+1,0,\theta,p,r)}(\mathrm{d}y)\mathbf{1}_{\{x\in\Xi_3\}}\\ &\quad+\delta_{(\gamma(d),1,0,\theta,p,r)}(\mathrm{d}y)\mathbf{1}_{\{x\in\Xi_4\}}\\ &\quad+\delta_{(1,n,\sigma,\theta,p,r)}(\mathrm{d}y)\mathbf{1}_{\{x\in\Xi_5\}}\Big]\Big\}\end{aligned}$$

Finally,

$$\begin{aligned}\mathfrak{T}V(x) &= \inf_{d\in A(x)}\Big\{[1+V(\gamma(d),1,0,1,P_c,R_c)]\mathbf{1}_{x\in\Xi_1}\\ &\quad+[\mathbf{1}_{d\neq 0}+V(\gamma(d),n+1,0,\theta,p,r)]\mathbf{1}_{x\in\Xi_3}\\ &\quad+[1+V(\gamma(d),1,0,\theta,p,r)]\mathbf{1}_{x\in\Xi_4}\Big\}+V(\Delta)\mathbf{1}_{x\in\Xi_2}\\ &\quad+V(1,n,\sigma,\theta,p,r)\mathbf{1}_{x\in\Xi_5}\\ \mathfrak{T}V(\Delta) &= V(\Delta)\end{aligned}\tag{11}$$

Computation of $\mathfrak{B}$

Now, for $Y\in\overline{X}$, and function $V:\overline{X}\to\mathbb{R}$, we need to compute:

$$\mathfrak{B}V(y)=\int_{[0,t^*(y)[}e^{-(K+\alpha)t}\mathfrak{R}V(\phi(y,t))\mathrm{d}t+e^{-(K+\alpha)t^*(y)}\mathfrak{T}V(\phi(y,t^*(y)))$$

As we cannot make an exact computation of $\mathfrak{B}V$ on $\overline{X}$, we need to approximate this computation on a grid of the state space. In order to detail the approximation of the computation, we define

$$G(V,y)=\int_{[0,t^*(y)[}e^{-(K+\alpha)t}\mathfrak{R}V(\phi(y,t))\mathrm{d}t$$

and

$$H(V,y)=e^{-(K+\alpha)t^*(y)}\mathfrak{T}V(\phi(y,t^*(y))$$

as in Eqs. 7 and 8. We define a time interval Δt (in practice equal to 1 day) and for every $y=(\gamma,n,\sigma,\theta,p,r)\in\tilde{X}$, we note

$$n^*(y)=\left\lfloor\frac{t^*(y)}{\Delta t}\right\rfloor$$

For every $j\in\{0..n^*(y)-1\}$, we note $\phi_j(y,t)=\phi(y,j\Delta t)$ and $\phi(y,t^*(y))=(\gamma,n,\sigma+t^*(y),\theta+t^*(y),p^*(y),r^*(y))$. The integral defined in Eq. 7 is computed by approximation using the classic trapezoidal rule using the $j\Delta t$ nodes:

$$\begin{aligned}G(V,y)\simeq\frac{\Delta t}{2}\mathfrak{R}V(y)+\frac{\Delta t}{2}e^{-(K+\alpha)t^*(y)}\mathfrak{R}V(\phi(y,t^*(y)))\\ +\sum_{j=1}^{n^*(y)-2}\Delta t e^{-(K+\alpha)j\Delta t}\mathfrak{R}V(\phi_j(y,t))\end{aligned}$$

with $\mathfrak{R}V(x) = \frac{1}{30}\mathbf{1}_{\{p+r\leq 500\}} + KV(1, n, \sigma, \theta, p, r)$, as computed in Eq. 10. Then we obtain the following for every $y = (\gamma, n, \sigma, \theta, p, r) \in \tilde{X}$:

$$\begin{aligned} G(V, y) = \frac{\Delta t}{2}\left(\frac{1}{30}\mathbf{1}_{\{p+r<500\}} + KV(1, n, \sigma, \theta, p, r)\right) \\ + \frac{\Delta t}{2}e^{-(K+\alpha)t^*(y)}\left(\frac{1}{30}\mathbf{1}_{\{p^*+r^*<500\}}\right. \\ \left. + KV(1, n, \sigma + t^*, \theta + t^*, p^*(y), r^*(y))\right) \\ + \Delta t \sum_{j=1}^{n^*(y)-2} e^{-(K+\alpha)j\Delta t}\left(\frac{1}{30}\mathbf{1}_{\{p_j+r_j<500\}}\right. \\ \left. + KV(1, n, \sigma + j\Delta t, \theta + j\Delta t, p_j, r_j)\right) \end{aligned}$$

Now, we need to compute H as defined in Eq. 8: it depends on $\mathfrak{T}V(\phi(y, t^*(y)))$, which takes different values according to the boundary reached in that point, as written in Eq. 11. Moreover, as we know the flow, we can give conditions on $y = (\gamma, n, \sigma, \theta, p, r)$ to reach a given boundary in $\phi(y, t^*(y))$. Then:

– if $\phi(y, t^*(y)) \in \Xi_1$ $(\theta \leq 1)$ then

$$H(V, y) = \inf_{d\in[d_1,..d_{m_d}]}\left\{e^{-(K+\alpha)t^*(y)}\left[1 + V(\gamma(d), 1, 0, 1, P_c, R_c)\right]\right\}$$

– if $\phi(y, t^*(y)) \in \Xi_2$ $(\theta + t^*(y) \geq T_h)$ then

$$H(V, y) = e^{-(K+\alpha)t^*(y)}V(\Delta)$$

– if $\phi(y, t^*(y)) \in \Xi_3$ $(n < n_{\text{inj}}, \theta + t^*(y) < T_h)$ then

$$\begin{aligned} H(V, y) = \inf_{d\in[0,d_1,..d_{m_d}]}\left\{e^{-(K+\alpha)t^*(y)}\left[\mathbf{1}_{\{d\neq 0\}}\right.\right. \\ \left.\left. + V(\gamma(d), n + 1, 0, \theta + t^*(y), p^*(y), r^*(y))\right]\right\} \end{aligned}$$

– if $\phi(y, t^*(y)) \in \Xi_4$ $(n = n_{\text{inj}}, \gamma = 1, \theta + t^*(y) < T_h)$ then

$$\begin{aligned} H(V, y) = \inf_{d\in[d_1,..d_{m_d}]}\left\{e^{-(K+\alpha)t^*(y)}\left[1 + V(\gamma(d), 1, 0, \theta\right.\right. \\ \left.\left. + t^*(y), p^*(y), r^*(y))\right]\right\} \end{aligned}$$

– if $\phi(y, t^*(y)) \in \Xi_5$ $(n = n_{\text{inj}}, \gamma > 1, \theta + t^*(y) < T_h)$ then

$$H(V, y) = e^{-(K+\alpha)t^*(y)}V(1, n, \sigma + t^*(y), \theta + t^*(y), p^*(y), r^*(y))$$

Finally, for every $y = (\gamma, n, \sigma, \theta, p, r) \in \tilde{X}$:

$$
\begin{aligned}
\mathfrak{B}V(y) = {} & \frac{\Delta t}{2}\left(\frac{1}{30}\mathbf{1}_{\{p+r<500\}} + KV(1,n,\sigma,\theta,p,r)\right) \\
& + \frac{\Delta t}{2}e^{-(K+\alpha)t^*(y)}\Big[\frac{1}{30}\mathbf{1}_{\{p^*+r^*<500\}} + KV(1,n,\sigma \\
& + t^*,\theta+t^*,p^*(y),r^*(y))\Big] \\
& + \Delta t\sum_{j=1}^{n^*(y)-2} e^{-(K+\alpha)j\Delta t}\Big[\frac{1}{30}\mathbf{1}_{\{p_j+r_j<500\}} \\
& + KV(1,n,\sigma+j\Delta t,\theta+j\Delta t,p_j,r_j)\Big] \\
& + \inf_{d\in[d_1,..d_{m_d}]}\Big\{e^{-(K+\alpha)t^*(y)}\Big[1+V(\gamma(d),1,0,1,P_c,R_c)\Big]\Big\}\mathbf{1}_{\{\theta\le 1\}} \\
& + e^{-(K+\alpha)t^*(y)}V(\Delta)\mathbf{1}_{\{\theta+t^*(y)\ge T_h\}} \\
& + \inf_{d\in[0,d_1,..d_{m_d}]}\Big\{e^{-(K+\alpha)t^*(y)}\Big[\mathbf{1}_{\{d\neq 0\}} \\
& + V(\gamma(d),n+1,0,\theta+t^*(y),p^*(y),r^*(y))\Big]\Big\}\mathbf{1}_{\{n<n_{\text{inj}},\theta+t^*(y)<T_h\}} \\
& + \inf_{d\in[d_1,..d_{m_d}]}\Big\{e^{-(K+\alpha)t^*(y)}\Big[1+V(\gamma(d),1,0,\theta+t^*(y),p^*(y),r^*(y))\Big]\Big\} \\
& \times\mathbf{1}_{\{n=n_{\text{inj}},\gamma=1,\theta+t^*(y)<T_h\}} \\
& + e^{-(K+\alpha)t^*(y)}V(1,n,\sigma+t^*(y),\theta+t^*(y),p^*(y),r^*(y)) \\
& \times\mathbf{1}_{\{n=n_{\text{inj}},\gamma>1,\theta+t^*(y)<T_h\}}
\end{aligned}
\tag{12}
$$

and

$$
\mathfrak{B}V(\Delta) = \int_{[0,+\infty)} e^{-(K+\alpha)t}KV(\Delta)\mathrm{d}t = \frac{K}{K+\alpha}V(\Delta)
$$

B Structure of the Code

Structure of the code and its subroutines are shown in Fig. 4. Application in the results section requires the following grid:

- $\gamma \in \{1..3\}$
- n $\in \{1..3\}$
- $\sigma \in \{0..351\}$
- $\theta \in \{1..365\}$
- $p \in \{2..110\}$ depending on the patient
- $r \in \{100..1500\}$ depending on the patient

The grid of the state space created in Matlab contains 67,614 lines and 7755 columns. For a given patient, the computation of 40 iterations of the sequence (convergence is reached between 35 and 45 iterations) requires between 5 and 6 days.

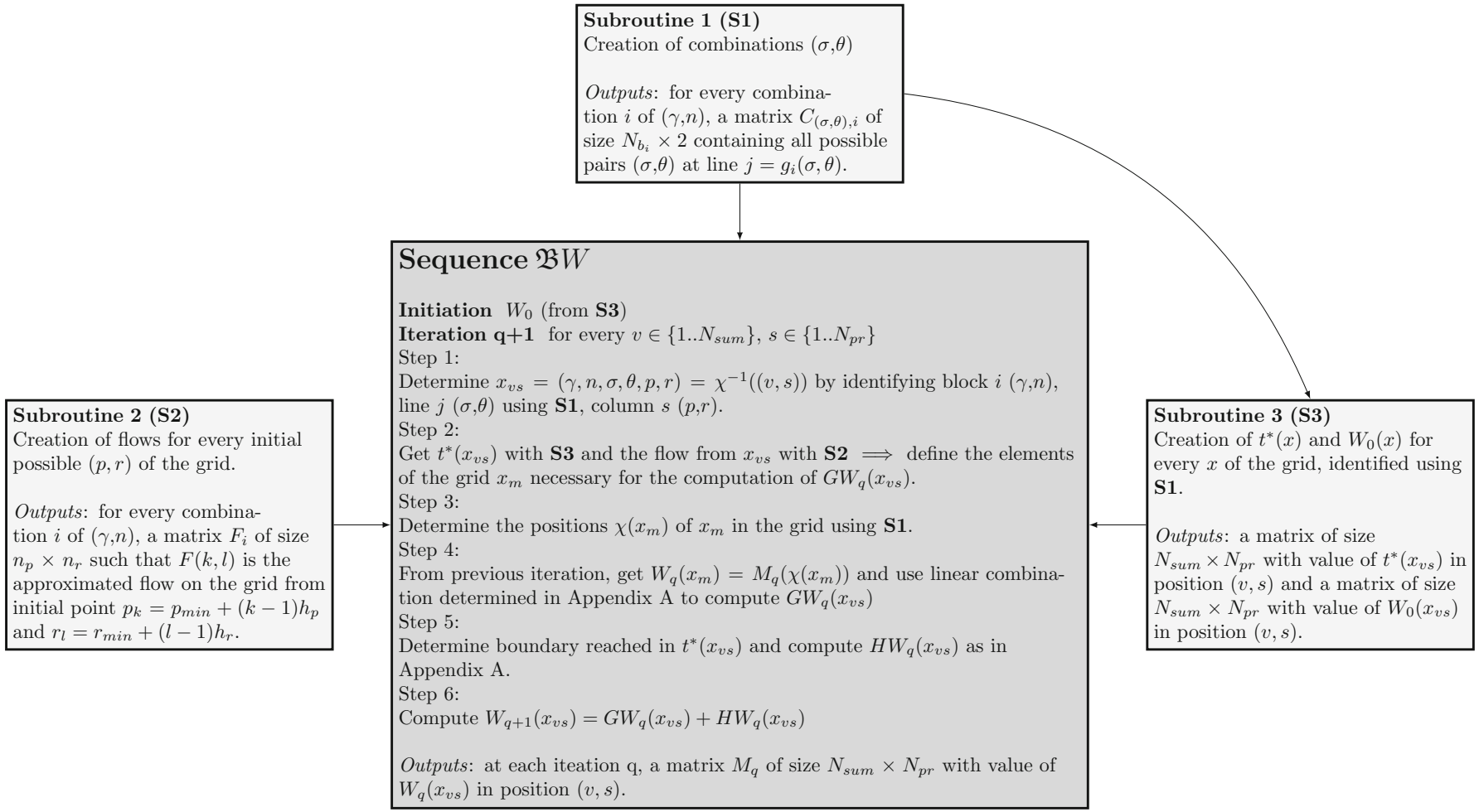


Fig. 4 Structure of the code and its subroutines

C Sensitivity Analysis of the Method

To evaluate how the uncertainty on individual parameters estimation could impact the determination of the optimal strategy, we have realized a sensitivity analysis. For a given patient, we suppose a normal distribution of parameters λ and ρ. We generate $L = 500$ pairs of parameters (λ, ρ) from this joint distribution. Each pair corresponds to an initial value of lymphocytes T $CD4_0$. We determine the empirical quartiles of the distribution of the $CD4_0$ and focus on the pairs inducing values close the first and the third quartiles. Then, for each pair, we simulate the five possible protocols P1 to P5 and compare them to the optimal strategy determined on the mean value of (λ,ρ). In practice, values of pairs and associated values of CD4 are displayed in Table 5. For the mean value of (λ,ρ), we determined the optimal strategy to be a first cycle of 2 injections and then cycles of 1 injection, which corresponds to protocol P4. We show in Table 5 the cost of each protocol for each pair of (λ,ρ), and we put in bold the minimum cost over the five protocols. We can see that protocol P4 achieves the minimum cost for all pairs inducing CD4 values at the first quartile. For pairs inducing CD4 values at the third quartile, the protocol achieving the minimum cost is P5. However, the difference of cost is not huge and P4 actually induces more time spent over the 500 threshold and less than one more injection than P5 on average, which is still acceptable. Overall, this shows that even with some error on the estimation on λ, ρ we would be able to determine a strategy achieving a good balance between clinical criteria.

Table 5 Pairs of (λ,ρ), associated CD4 values and mean cost for protocols P1 to P5

Category	λ	ρ	$CD4_0$	P1	P2	P3	P4	P5
Mean	2.065	2.022	289	6.55	5.54	5.54	**4.16**	4.95
Q1	1.506	2.305	224	8.59	6.94	6.03	**5.20**	5.67
Q1	2.062	1.180	223	11.6	9.19	8.52	**7.28**	8.02
Q1	1.701	1.747	224	8.94	7.31	6.50	**5.60**	6.40
Q1	2.163	1.078	222	11.9	9.50	8.82	**8.04**	8.39
Q1	1.737	1.371	224	9.06	7.39	6.64	**5.71**	6.48
Q1	1.689	1.758	223	8.92	7.31	6.52	**5.62**	6.40
Q1	1.728	1.689	224	9.00	7.35	6.63	**5.71**	6.45
Q1	1.426	2.599	222	8.54	6.90	5.98	**5.14**	5.62
Q1	2.493	0.838	222	13.3	11.0	10.3	**8.85**	9.76
Q1	1.805	1.542	223	9.23	7.59	6.94	**6.36**	6.66
Q3	2.160	2.594	337	5.62	4.81	3.88	3.15	**3.04**
Q3	2.424	1.956	336	5.83	4.99	4.10	3.81	**3.21**
Q3	2.638	1.625	335	6.10	5.22	4.30	3.93	**3.31**
Q3	2.219	2.429	337	5.68	4.85	3.93	3.21	**3.11**
Q3	2.477	1.879	337	5.87	5.04	4.12	3.83	**3.21**
Q3	2.466	1.896	337	5.86	5.01	4.09	3.82	**3.22**

P1 Cycles of 3 injections. *P2* First cycle of 3 injections then cycles of 2 injections. *P3* Cycles of 2 injections. *P4* First cycle of 2 injections then cycles of 1 injection. *P5* Cycles of 1 injection. *P4* Is the optimal protocol for mean value of (λ,ρ)

D Trajectories of Patients B and C

See Figs. 5, 6.

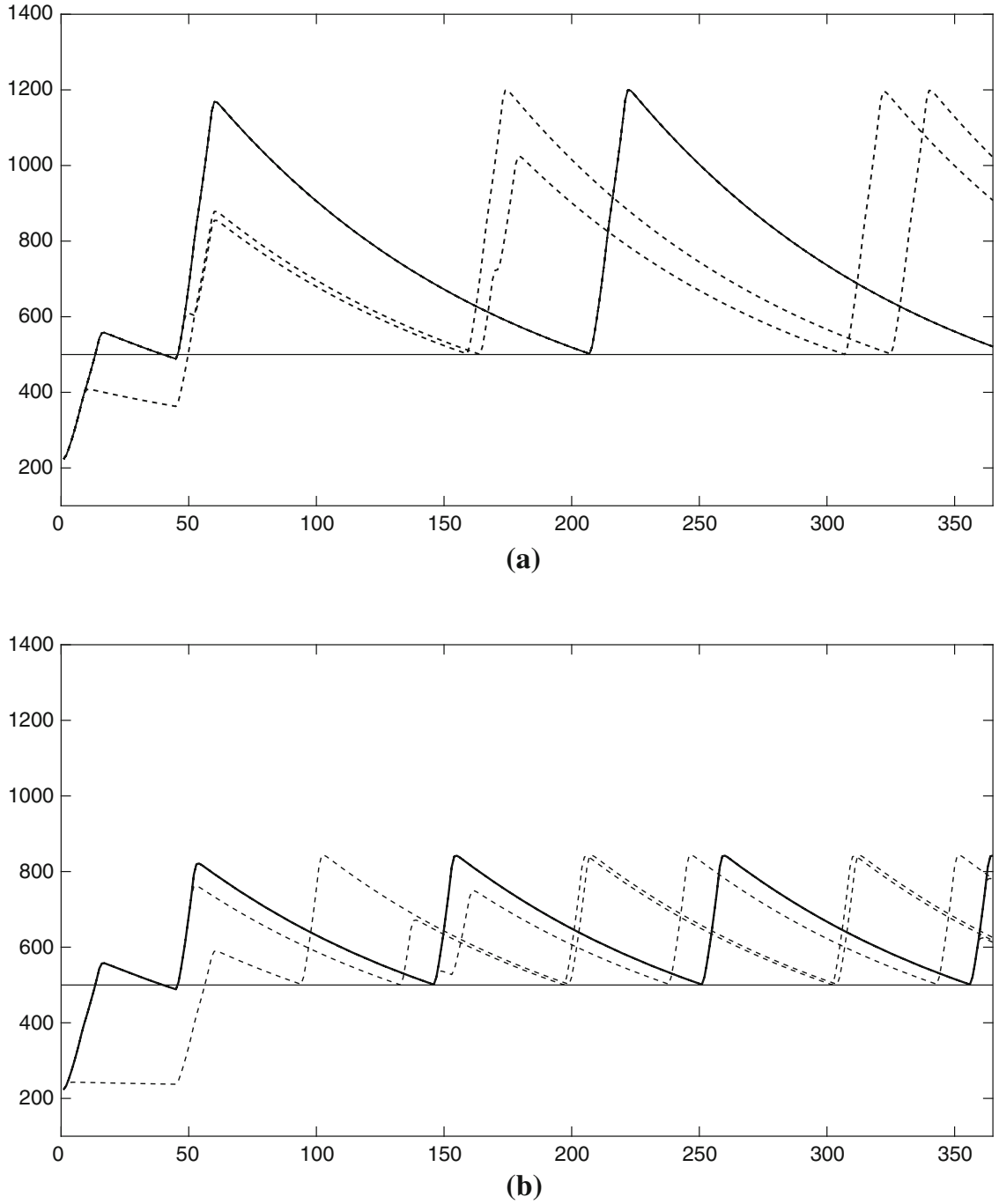


Fig. 5 Dynamics of $CD4^+$ T lymphocytes in patient B. Straight line corresponds to the "best" outcome, i.e., when the effect of all injections lasts 7 days. Dashed line corresponds to other possible trajectories, when this effect can last less than 7 days. **a** Dynamics of $CD4^+$ T lymphocytes in patient B under P3, a 2-injections cycles protocol (dose 20). **b** Dynamics of $CD4^+$ T lymphocytes in patient B under the determined optimal strategy

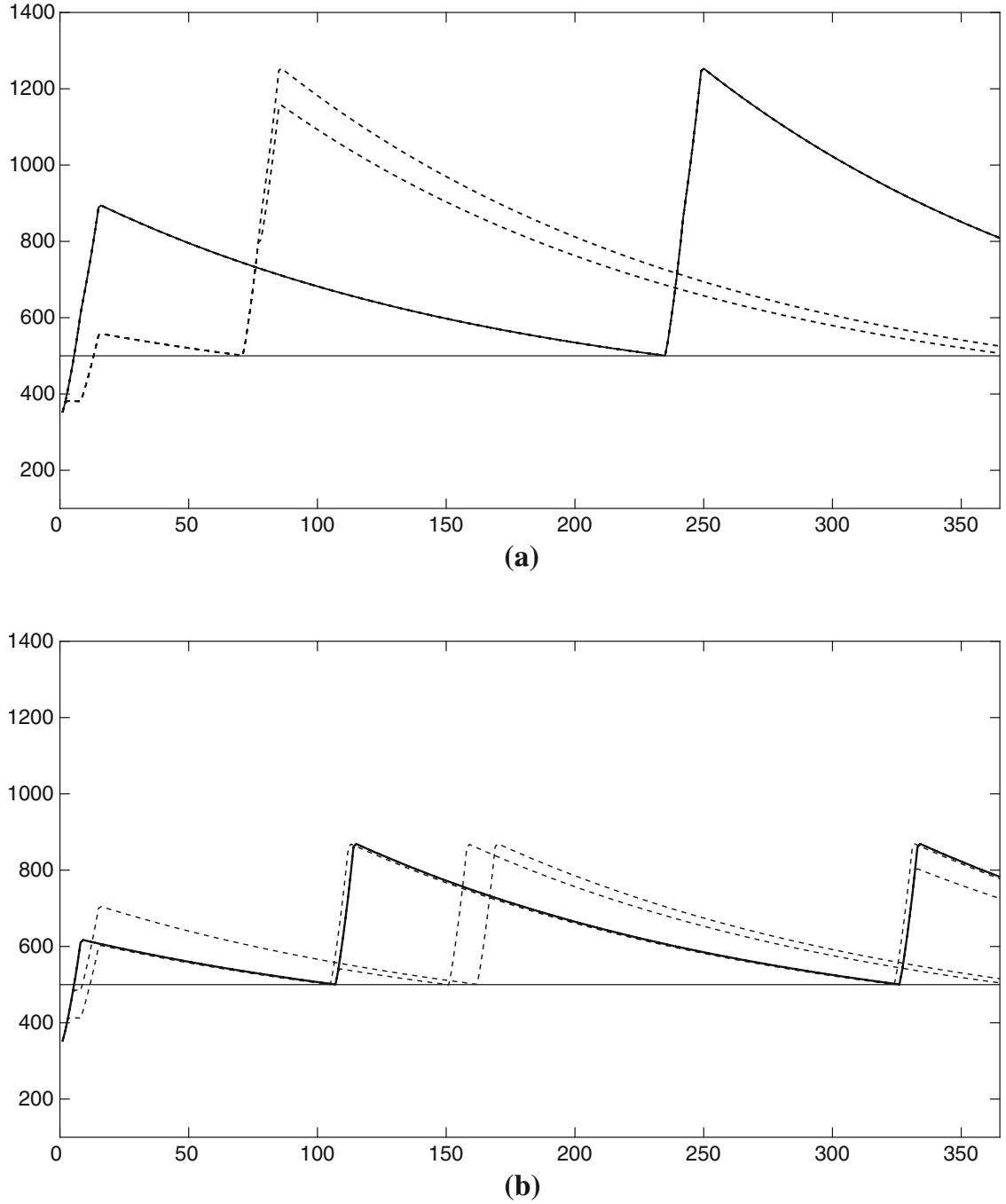


Fig. 6 Dynamics of $CD4^+$ T lymphocytes in patient C. Straight line corresponds to the "best" outcome, i.e., when the effect of all injections lasts 7 days. Dashed line corresponds to other possible trajectories, when this effect can last less than 7 days. **a** Dynamics of $CD4^+$ T lymphocytes in patient C under P3, a 2-injections cycles protocol (dose 20). **b** Dynamics of $CD4^+$ T lymphocytes in patient C under the determined optimal strategy

References

Camargo JF, Kulkarni H, Agan BK, Gaitan AA, Beachy LA, Srinivas S, He W, Anderson S, Marconi VC, Dolan MJ et al (2009) Responsiveness of T cells to interleukin-7 is associated with higher $CD4^+$ T cell counts in HIV-1-positive individuals with highly active antiretroviral therapy-induced viral load suppression. J Infect Dis 199(12):1872–1882

Cappuccio A, Castiglione F, Piccoli B (2007) Determination of the optimal therapeutic protocols in cancer immunotherapy. Math Biosci 209(1):1–13

Castiglione F, Piccoli B (2006) Optimal control in a model of dendritic cell transfection cancer immunotherapy. Bull Math Biol 68(2):255–274

Castiglione F, Piccoli B (2007) Cancer immunotherapy, mathematical modeling and optimal control. J Theor Biol 247(4):723–732

Costa O, Dufour F, Piunovskiy A (2016) Constrained and unconstrained optimal discounted control of piecewise deterministic markov processes. SIAM J Control Optim 54(3):1444–1474

Davis MH (1984) Piecewise-deterministic markov processes: a general class of non-diffusion stochastic models. J R Stat Soc Ser B (Methodol) 46:353–388

Dufour F, Zhang H (2015) Numerical methods for simulation and optimization of piecewise deterministic markov processes. Wiley, New York

Eftimie R, Gillard JJ, Cantrell DA (2016) Mathematical models for immunology: current state of the art and future research directions. Bull Math Biol 78(10):2091–2134

Jarne A, Commenges D, Villain L, Prague M, Lévy Y, Thiébaut R et al (2017) Modeling $CD4^+$ T cells dynamics in HIV-infected patients receiving repeated cycles of exogenous interleukin 7. Ann Appl Stat 11(3):1593–1616

Lange CG, Lederman MM (2003) Immune reconstitution with antiretroviral therapies in chronic HIV-1 infection. J Antimicrob Chemother 51(1):1–4

Lavielle M, Mentré F (2007) Estimation of population pharmacokinetic parameters of saquinavir in HIV patients with the monolix software. J Pharmacokinet Pharmacodyn 34(2):229–249

Leone A, Rohankhedkar M, Okoye A, Legasse A, Axthelm MK, Villinger F, Piatak M, Lifson JD, Assouline B, Morre M et al (2010) Increased $CD4^+$ T cell levels during IL-7 administration of antiretroviral therapy-treated simian immunodeficiency virus-positive macaques are not dependent on strong proliferative responses. J Immunol 185(3):1650–1659

Levy Y, Lacabaratz C, Weiss L, Viard JP, Goujard C, Lelièvre JD, Boué F, Molina JM, Rouzioux C, Avettand-Fénoêl V et al (2009) Enhanced T cell recovery in HIV-infected adults through IL-7 treatment. J Clin Investig 119(4):997

Levy Y, Sereti I, Tambussi G, Routy J, Lelievre J, Delfraissy J, Molina J, Fischl M, Goujard C, Rodriguez B et al (2012) Effects of recombinant human interleukin 7 on T-cell recovery and thymic output in HIV-infected patients receiving antiretroviral therapy: results of a phase I/IIa randomized, placebo-controlled, multicenter study. Clin Infect Dis 55(2):291–300

Lewden C, Chêne G, Morlat P, Raffi F, Dupon M, Dellamonica P, Pellegrin JL, Katlama C, Dabis F, Leport C et al (2007) HIV-infected adults with a CD4 cell count greater than 500 cells/mm3 on long-term combination antiretroviral therapy reach same mortality rates as the general population. JAIDS J Acquir Immune Defic Syndr 46(1):72–77

Mackall CL, Fry TJ, Bare C, Morgan P, Galbraith A, Gress RE (2001) Il-7 increases both thymic-dependent and thymic-independent T-cell regeneration after bone marrow transplantation. Blood 97(5):1491–1497

Okamoto Y, Douek DC, McFarland RD, Koup RA (2002) Effects of exogenous interleukin-7 on human thymus function. Blood 99(8):2851–2858

Pappalardo F, Pennisi M, Castiglione F, Motta S (2010) Vaccine protocols optimization: in silico experiences. Biotechnol Adv 28(1):82–93

Prague M, Commenges D, Guedj J, Drylewicz J, Thiébaut R (2013) NIMROD: a program for inference via a normal approximation of the posterior in models with random effects based on ordinary differential equations. Comput Methods Programs Biomed 111(2):447–458

Sereti I, Dunham RM, Spritzler J, Aga E, Proschan MA, Medvik K, Battaglia CA, Landay AL, Pahwa S, Fischl MA et al (2009) Il-7 administration drives T cell-cycle entry and expansion in HIV-1 infection. Blood 113(25):6304–6314

Tan JT, Dudl E, LeRoy E, Murray R, Sprent J, Weinberg KI, Surh CD (2001) Il-7 is critical for homeostatic proliferation and survival of naive T cells. Proc Nat Acad Sci 98(15):8732–8737

Thiebaut R, Drylewicz J, Prague M, Lacabaratz C, Beq S, Jarne A, Croughs T, Sekaly RP, Lederman MM, Sereti I (2014) Quantifying and predicting the effect of exogenous interleukin-7 on $CD4^+$ T cells in HIV-1 infection. PLoS Comput Biol 10(5):e1003,630

Thiébaut R, Jarne A, Routy JP, Sereti I, Fischl M, Ive P, Speck RF, D'offizi G, Casari S, Commenges D (2016) Repeated cycles of recombinant human interleukin 7 in HIV-infected patients with low CD4 T-cell reconstitution on antiretroviral therapy: results of 2 phase II multicenter studies. Clin Infect Dis 62(9):1178–1185

Vella AT, Dow S, Potter TA, Kappler J, Marrack P (1998) Cytokine-induced survival of activated T cells in vitro and in vivo. Proc Nat Acad Sci 95(7):3810–3815

4.4 Discussion

4.4.1 Numerical method: consistency with theoretical results, calibration, performance

In order to ensure consistency with the theoretical results obtained in Costa et al. [2016], assumptions A, B and C were verified during our work. However, we will not develop these theoretical considerations, as the focus of our work was mainly on the development of a numerical tool for solving the optimal control problem using dynamic programming.

Part of the proof of theorem 1 in Costa et al. [2016] relies on the fact that the sequence $\{W_m\}_{m\in\mathbb{N}}$ defined in equation (61) is increasing and uniformly bounded, and thus converges. Actually, it can be shown that the sequence of function $\{V_m\}_{m\in\mathbb{N}}$ similarly defined, such that

$$\left\{\begin{array}{rcl} V_{m+1}(x) & = & \mathcal{B}V_m(x) \text{ for m} \in \mathbb{N}, \\ V_0(x) & = & K_A\mathbf{1}_{A_{\varepsilon_1}}(x) + (K_A + K_B)\mathbf{1}_{A^c_{\varepsilon_1}}(x), \end{array}\right. \tag{64}$$

is decreasing and uniformly bounded and converges to the same function as the sequence $\{W_m\}_{m\in\mathbb{N}}$. This property was verified in the numerical method. As reminded in section 4 of the article from section 4.3, the numerical method computes at each iteration a matrix containing the value of W_m at all the points of the grid of the state space. For each of the first pseudo-patients studied, we have checked that $W_{m+1}(x) - W_m(x) \geq 0$ for every point of the grid until convergence. Moreover, we have run the sequence V_m with the same criterion of convergence, namely $\|V_{m+1} - V_m\|_\infty < \epsilon$ with $\epsilon = 0.001$ and checked first that $V_{m+1}(x) - V_m(x) \leq 0$ and that $\|W_\infty - V_\infty\|_\infty < \tilde{\epsilon}$ with a small $\tilde{\epsilon}$. These verifications ensured consistency of the numerical method with theoretical results.

As mentioned in the "Results" section of the article presented in section 4.3, we first evaluated the performance of our optimization algorithm by comparing the approximated value of the optimal function obtained from the numerical computation of W with the cost obtained from numerical approximation of the optimal strategy. The cost is estimated with Monte Carlo method, by simulating the strategy and the corresponding trajectory of the process a large number of times. In the article, results were only presented for 3 pseudo-patients among the 50 generated. We also computed a mean relative variation between these two values on all pseudo-patients by defining:

$$MRV = \frac{1}{50}\sum_{i=1}^{50}\frac{|W(x_0)_i - C_{opt_i}|}{C_{opt_i}}, \tag{65}$$

with $W(x_0)_i$ given by the approximation of the value function for patient i and C_{opt_i} the

value obtained by simulating the trajectories associated to the numerical approximation of the optimal strategy for patient i. We find a mean relative variation of 1.5%, showing the consistency of our numerical results. However, the performance of the method could depend on the precision of the grid of the state space. To evaluate the sensitivity of the numerical approximation to the grid of the state space, we ran the sequencing algorithm on the 3 patients A, B and C presented in the article with 3 other grids of the state space. As mentioned in Appendix B of the article presented in section 4.3, the grid contains vectors of six elements: $(\gamma, n, \sigma, \theta, p, r)$. Parameters γ and n only takes a few number of values, and it makes sense to discretize the time in days, so we evaluated variations only on the choice of step for the discretization on values of p and r. For all grids, we consider $p_{min} = 2$, $p_{max} = 142$, $r_{min} = 100$, $r_{max} = 1500$ to include most of the patients' values. Note that the intervals $[p_{min}; p_{max}]$ and $[r_{min}; r_{max}]$ can be reduced and adapted to patients if necessary. The program was run on each grid for the same patients A, B and C as those presented in the results section of the article presented in section 4.3. Table 4.1 below sums up the characteristics of the grids of computation, their size and the computation time requested for each patient. For each patient and each grid, we computed the value

Grid	h_p	h_r	Size of the $p \times r$ grid	Total size of the grid	Computational time ($n_{it}^{\dagger}$) Pat A	Pat B	Pat C
Γ_1^*	2	10	10 011	676 883 754	5.5d (34)	7.2d (40)	7.6d (45)
Γ_2	10	10	2 115	143 003 610	30h (36)	36h (41)	41h (45)
Γ_3	2	50	2 059	139 217 226	28h (35)	36h (43)	37h (44)
Γ_4	10	50	435	29 412 090	6.2h (35)	7.7h (41)	8.3h (45)

Table 4.1 – Characteristics of the different grids used for the numerical approximation. Computational time is in days (d) or hours (h). Γ_1^* corresponds to the grid used in the article presented in section 4.3. † n_{it}: number of iterations of the sequence until convergence.

$W(x_0)$ corresponding to the minimal cost that can be reached and we also simulated and computed the cost of the associated optimal strategy. Table 4.2 shows the values and the MRV computed on these 3 patients. We can see that all grids induce the same optimal strategy (as simulated by choosing the optimal action when reaching the boundary), as the value of the cost stays pretty much the same among the different grids. It should be highlighted here that the strategy is simulated on continuous values of P and R, but the action is determined by projecting their values on the grid of the state space. This can explain the differences of values between the grids when simulations of the optimal strategy are realized. However, we observe larger differences in the determination of the

	Pat A		Pat B		Pat C		
Grid	$W(x_0)$	$C_{opt}(sd)$	$W(x_0)$	$C_{opt}(sd)$	$W(x_0)$	$C_{opt}(sd)$	MRV
Γ_1^*	2.87	2.90 (0.36)	6.11	6.20 (0.56)	9.44	9.54 (0.85)	1.2%
Γ_2	2.91	2.90 (0.36)	6.06	6.21 (0.57)	9.46	9.54 (0.84)	1.2%
Γ_3	2.28	2.91 (0.37)	6.51	6.18 (0.57)	8.97	9.80 (0.80)	12%
Γ_4	2.45	2.91 (0.38)	5.65	6.21 (0.57)	9.45	9.79 (0.75)	10%

Table 4.2 – Comparison of cost value from value function and Monte Carlo simulations (with the standard deviation sd) in patients A, B, C on the 4 different grids.

value function between the grids, meaning that the evaluation of the optimal strategy is more robust that the computation of the value function. In particular, grids Γ_3 and Γ_4 induce a mean variation of 12 and 10% between the value $W(x_0)$ obtained from the value function and cost function computed with Monte Carlo simulation. With grid Γ_3 the computation time is almost the same as for Γ_2, but there is less variation between $W(x_0)$ and the optimal cost obtained with simulations when using Γ_2 (only 1.2%). Actually, grid Γ_2 seems to have similar estimations of the value function on these 3 patients as grid Γ_1, with a much lower computation time. If we had to consider increasing the time of the grid but looking at a longer horizon of time, it could be interesting to investigate further these considerations (on a larger number of patients) and the grid Γ_2 could be a good choice to reduce computation times.

4.4.2 Cost function

The numerical approximation of the optimal strategy is dependent on the definition of the cost function. Indeed, the method allows to determine an approximation of the strategy $\hat{u}$ minimizing $\mathcal{V}(u, x_0)$ as in equation (59) and the criterion $\mathcal{V}$ directly depends on the cost function. In our particular framework, the criterion combines the time spent with CD4 counts below 500 cells/μL and the number of injections realized. The combination could be weighted to balance these two quantities. We chose to allocate the same cost to an injection of IL-7 and to spending 30 days below 500 cells/μL. This choice is not in favor of a third injection in cycle, as we have seen that the third injection has less effect on the cells' proliferation than the two previous ones. Most of the time, the third injection does not allow any patient to increase the sustainability of CD4 levels above 500 cells/μL during more than a month. In order to illustrate the impact of the choice of the cost function on the ranking of the strategies of injections, we have computed for patients A, B and C of the article presented in section 4.3 the mean criterion of several protocols, using either the cost counting the number of months spent below 500 cells/μL

– used in the article (C_m) or the number of days (C_d). This computation was also realized for an additional pseudo-patient (written Patient A', with very low levels of CD4$^+$ T cells at baseline). Table 4.3 sums up the computed values obtained from protocols P1 to P5 as described in the article section 4.3 and the optimal strategy obtained using C_m (as in the article), denoted optimal$_m$. We note that for patient C, with not too low levels of

Patient \ Protocol		P1	P2	P3	P4	P5	Optimal$_m$
A (very low)	C_m	**14.8**	11.6	10.7	10.4	11.0	**9.54**
	C_d	65.7	63.5	63.6	102	**144**	**63.4**
A' (very low)	C_m	**16.8**	13.7	13.2	14.7	14.2	**11.8**
	C_d	**96.3**	98	103	183	**222**	109
B (low)	C_m	**9.41**	7.62	6.91	6.25	6.63	**6.21**
	C_d	27.5	**25.3**	29.7	32.2	**64.2**	29.3
C (not too low)	C_m	**5.53**	4.75	3.80	3.08	2.94	**2.89**
	C_d	10.8	10.0	9.25	8.72	**11.3**	**8.48**

Table 4.3 – Comparison of mean criteria for protocols P1 to P5 and optimal strategy written Optimal$_m$ because it was obtained using C_m, the cost computed from the number of months spent below 500 cells/μL. C_m corresponds to the criterion computed from the number of months spent below 500/μL and C_d from the number of days. For each line, we have written in bold red the worst protocol (highest cost) and in bold green the best one (lowest cost).

CD4$^+$ T cells at baseline, the optimal strategy determined with C_m still has the lowest cost among all other protocols P1 to P5. This is also the case for patient A, although the difference between the optimal$_m$ and protocols P2 and P3 is less clear with cost C_d (63.4, 63.5 and 63.6 respectively). For patients A' and B, the optimal$_m$ strategy does not have the lowest cost when computed with C_d, and in particular for patient A' protocol P1 was the worst using C_m and is the best choice with C_d. Overall, using C_d is not in favor of protocol P5. This is easily explained by the fact that C_d counts all the days spent with CD4$^+$ levels below 500 days/μL, and each day has the same weight as one injection with C_d. This cost function penalizes more the protocols inducing more days with low CD4 levels (even if less injections are realized, as in P5). It is also the reason why C_d is more in favor with protocols with a large number of injections for patients requesting many injections to sustain their CD4 levels. As some subjects tend to spend many more days with low CD4 levels if the number of injections is reduced compared to others, they will be more sensitive to the choice of the balance between number of injections and time

spent with low CD4 levels in the cost function.

We have shown that the choice of the cost function can be crucial for determining the optimal strategy of injections, but the impact of this choice can also differ among all participants of the study. The cost function could account for other parameters than the ones studied. In particular when several doses of a given product are tested, side effects due to higher doses could be included by accounting for a term dependent of the dose in the cost function. Finally, these considerations emphasize the need to define the cost function in agreement with clinicians and real expectations from the optimization of the design.

4.4.3 Improving the biological model

Our main contribution is a "proof-of-concept", showing that the theory of optimal control can be applied successfully to the question of optimizing schedule of injections. As we focused mainly on this proof-of-concept, we chose to work on a simple, although still realistic, mechanistic model of the effect of IL-7 on the $CD4^+$ T cells dynamics. In order to apply this method to actual clinical studies, we would need to consider an improved biological/mechanistic model. This would require first to account for the additional effect of IL-7 on survival, modeled by a modification of parameter μ_R after IL-7 injection and depending on time [Thiébaut et al., 2014]. A cycle effect was also estimated [Jarne et al., 2017]. It could simply be due to the fact that CD4 counts are higher during the repeated cycles than at the beginning of the study and the homeostatic regulation of the immune system limits the number of $CD4^+$ T cells generated; but it could also mean that the immune system is also responding to IL-7 by generating antibodies against it. Both the effect of IL-7 on cell survival and the decreased effect of IL-7 on cells proliferation in the repeated cycles compared to the first one should be accounted in a more biologically plausible model. For example, the cycle effect could be added in the PDMP by adding a state variable c counting the number of cycles realized so far. The value of the parameter π would be different when $c = 1$ and $c > 1$.

In term of modeling, we also chose to introduce stochasticity in the PDMP through the time of effect of an IL-7 injection, τ. Adding stochasticity to the model could be discussed, as the deterministic model had good fitting and predictive abilities. However, it makes biological sense to assume that there is some stochasticity in the mechanism of action of IL-7 on the $CD4^+$ T cells. This stochasticity could have been introduced in other ways in the model: for example, the value of the effect of IL-7 on the proliferation rate, β_π could be random. In term of trajectories, the impact of both hypothesis would be difficult to assess on the available data. As a first way to account for stochasticity in the

process, we decided to add randomness in the time of effect of an IL-7 injection because it could easily be handled by the jump intensity η in the PDMP.

One limitation of the optimal control approach resides in the assumption that the process is continuously observed which allows the action to be adapted when needed. In the clinical context of treatment injections, the patient can only be observed at punctual time points. To account for these constraints, we would need to model the process as partially observable and time of observations should also be optimized in the problem, to determine the best visit times to obtain information on the patient and adapt the decision from this observation. This would lead to a more complex modeling and optimizing problem.

Finally, another inconvenient of the optimal control approach is that the parameters of the studied patient are assumed to be known. Although we are confident in the estimation obtained from the population approach, uncertainty on the parameters could induce some modifications in the determined optimal strategy. It makes this method less applicable in the clinical context of IL-7 than the Bayesian approach mentioned in section 4.2.1.2. This point constitutes one of the reason why we did not pursue in improving the biological model by accounting for the previous points discussed in this section, but we rather focused on understanding the gain from both approaches and trying to determine a framework where the optimal control method would be beneficial. In particular, the optimal control method could be more efficient when the model requires intrinsic stochasticity. This is discussed in the next following sections.

4.4.4 The challenge of both estimation and optimization

The Bayesian approach presented in section 4.2.1.2 and the optimal control approach were developed in parallel and it gave us a unique opportunity to discuss both methods for optimizing and adapting schedule of IL-7 injections in HIV-infected patients. This section does not aim at realizing a formal comparison between the two approaches but constitutes more of a discussion on the interests of both approaches in this framework and in prospects.

First, it should be reminded that both estimation and optimal control problems are difficult to assess simultaneously. Indeed, it would correspond to a partially observed optimal control problem and it represents a real challenge from both theoretical and numerical point of views [Hernández-Lerma, 1989; Bäuerle and Rieder, 2011]. In our particular case, we assumed that the principle of separation was valid, meaning that we are able to separately estimate the parameters of the model and optimize its control. In the optimal control approach presented in section 4.3, we assumed that the stochasticity

was due to the biological process itself and not to the estimation of the parameters. The pseudo-patient parameters are assumed to be determined before the optimization step. It is a strong assumption, especially as the estimation is realized on a deterministic model and not the stochastic model itself. Moreover, it should be underlined that the optimal control method depends on the chosen model, by making the strong assumption that the model is valid throughout the follow-up of the patient. In the previous section, we mentioned possibilities to modify the biological model, by accounting for example for a reduced effect of IL-7 on cell proliferation due to consecutive cycles. This could also modify the final optimal strategy obtained with our method. In our work, we then based the optimization method on two main assumptions: in term of parameters estimation, the population approach was already shown to give good results in ODE-based models [Lavielle and Mentré, 2007; Prague et al., 2013a]. In term of model choice, the previous work on modeling the effect of IL-7 on $CD4^+$ T cells dynamics and developed in section 4.1.2 ensured the model induced good fits of the data and had good prediction abilities.

In the Bayesian approach, the stochasticity was considered to arise from the uncertainty on the estimation of a patient's parameters, but the model for the CD4 dynamics was still deterministic, as previously developed in 4.1.2. The advantage of this method is that the parameters estimation can be updated each time new information is obtained. Moreover, decisions allow to locally optimize some criteria and they are based on the prediction from the model, not directly from the observation of a biological marker. It was also showed that the algorithm only requested a short phase of learning before having a reliable estimation of the parameters and a sequence of right decisions to take.

Overall, we have showed that both approaches could be used to adapt schedules of injections while maintaining patients above 500 $CD4^+$ T cells as long as possible. We should underline here that the IL-7 question provided a specific context of work for several reasons. First, the mechanistic model for the dynamics of the marker of interest ($CD4^+$ T cells) is very simple with linear equations. Moreover, it can be estimated using only two biomarkers (CD4 and Ki67), which makes the clinical context much easier. For other questions, the marker of interest could depend from the dynamics of a large number of markers and its dynamics could be less clear to describe. It would lead to a more complex mathematical model, with more compartments and interactions between the agents of the process and could request more measures for estimation. Moreover, the model is deterministic and induces good fits and predictions, which makes the use of the Bayesian approach easier. If the model needed stochasticity, the computation of the criteria of decision would have much more variance and it would make it more difficult to take any adaptive decision. We believe that in that case, the optimal control approach could be more adapted and robust. This should not be neglected, as we know that most of the

biological and physical processes actually undergo stochasticity, and if we aim at modeling systems with more granularity, it could request accounting for more stochasticity.

Finally, in practice, due to the limitation in term of computation time and not accounting for the uncertainty on the patient's estimation, the optimal control approach is harder to apply directly in clinical studies. The Bayesian approach can be easily implemented and at the moment it offers clinical perspectives, such as the evaluation of the adaptive strategies on other clinical outcomes in larger trials. The optimal control approach offers prospects in other applications where the stochasticity needs to be included and modeling the process with a PDMP is more appropriate; we will discuss the particular case of gene networks in the following section.

4.4.5 Prospects in vaccinology

One of the objectives of developing methods for optimizing schedules of injections is also to be able to optimize prime-boost vaccination regimens, combining several products in distinct immunizations. We have already mentioned in section 3.4.3 that systems vaccinology is a promising approach which consists in integrating data from different sources and measurements techniques into a broader model of the immune response to immunizations. This already represents in itself a challenge: currently, there is no general method for integrating information from the environment, the microbiome and the gene expression into mechanistic models of the biomarker(s) of interest (antibodies, cell populations, or others, possibly acting as a surrogate of protection). In fact, modeling and infering this information is still an open question, especially regarding the gene expression data. We have seen that gene regulatory networks (GRN) have been widely used, but there is no consensus on the best way to model and estimate these processes. In view of the work realized in this thesis, an interesting way to model GRN is actually to use PDMPs: the gene expression is determined by the state of the promoter of the gene, which randomly shifts between active (on) or inactive (off) state. The state of the promoter depends on transcription factors (proteins) that can bind to the DNA. The level of mRNA and proteins can be considered as continuous quantities, and their dynamics are modeled by ODEs. The mRNA is transcribed only during the active state. The PDMP related to this kind of process was described in Zeiser et al. [2010] and Herbach et al. [2017]. In a particular case of this model, the transcription occurs on very short periods which induce a burst of mRNA and of the protein synthesis [Bokes et al., 2013]. In these models, the flow is defined by the ordinary differential equations determining the dynamics of both the mRNA and the protein concentrations; these equations depend on the state $i \in \{0, 1\}$ of the gene promoter, with 0 corresponding to inactive and 1 to active. The transition rates

from state inactive to active and active to inactive can depend on protein concentrations, generated by the gene itself and possibly other genes included in the model. This rate of transition determine the random time at which the jumps occur. Vaccine immunizations can modify the gene expression. This could be modeled in the PDMP by modifying some transition rates following immunizations to induce higher or lower rates of activation, as some gene are observed to be differentially expressed. Further investigations should be realized to assess the possibility to infer complex GRN modeled this way: in Herbach et al. [2017], the theory is not based on a limited number of genes but simulations were conducted for networks with only 2 genes. Moreover, as we have developed in this thesis work a numerical tool for solving impulse control problems, it could be interesting to investigate the possibility to control gene expression levels by vaccine immunizations using a similar framework. This could be realized first on a calibrated GRN if inference is not completely achieved. The decision-maker would have the choice to realize immunizations at given time points, using one or several vaccine platforms. If some genes are related to the dynamics of the biomarker(s) of interest, it could be useful to be able to have developed methodological tools to optimize their expression. In that case, further investigations are also necessary to evaluate how the optimization performs regarding the complexity of the GRN and the number of genes included. The optimization method could help a rapid adaptation of the vaccine immunization regimen depending on the responder/non responder profile of the recipient. A representation of this modeling/optimizing approach in systems vaccinology is given in figure 4.8.

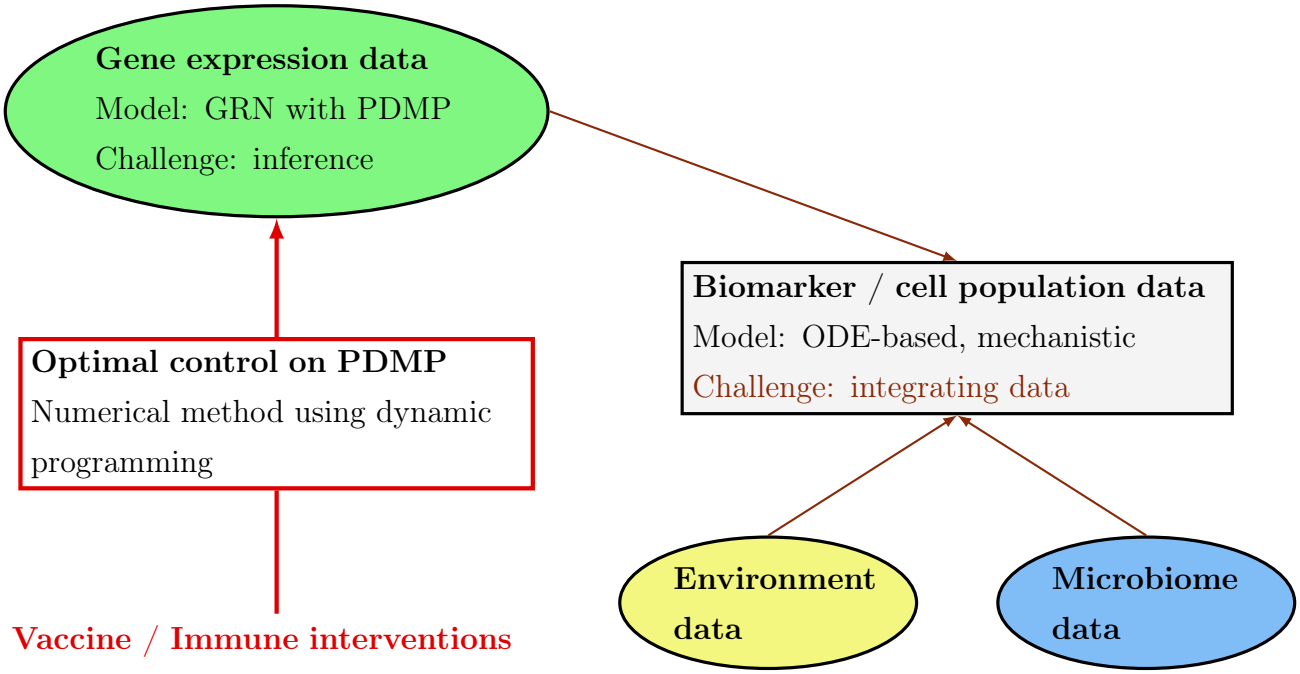


Figure 4.8 – Schematic representation of a modeling/optimizing approach in systems vaccinology

5 General discussion

In this thesis, we have conducted some work combining mathematics, immunology and biostatistics, as part of an effort to use models for understanding, quantifying and optimizing the immune response to preventive and therapeutic interventions against infectious diseases. This work is driven by concrete clinical questions, and sophisticated methods are developed to analyze longitudinal data obtained in human clinical trials: the challenge of parameter estimation in ODE-based models is handled by a population approach based on likelihood maximization, and the optimization of a clinical protocol of repeated injections is realized by developing a dynamic programing numerical method based on recent theoretical results. The originality of this thesis resides in both the analyzed data, as no mechanistic model of the response to Ebola vaccine has been studied on human data so far, and the application of complex optimal control theory on stochastic processes to the specific biological framework of optimal administration of IL-7.

In particular, we have shown that a simple model of the humoral immune response, including the dynamics of two populations of ASCs and of the antibodies, could fit well the binding antibody concentration data measured on subjects receiving Ebola vaccine candidates in phase 1 clinical trials. This model was based on ODEs and the parameters were estimated in order to quantify the dynamics of the humoral immune response and to determine factors involved in its variability. The advantage of the mechanistic model relies in the fact that its structure is biologically justified, and not data-driven; in this way, it can be applied in other contexts and could be estimated on other datasets generated in the EBOVAC consortium or other clinical trials. Moreover, the findings from this modeling work could have clinical implications, in particular in the adaptation of vaccine regimens according to the targeted population. The predictions of the model could be checked using longer follow-up measurements. The model would obviously benefit from additional data measurements, whether it be other biomarker measurements (ASCs for example), genomic, proteomic or microbiome measurements, or measurements from other subjects enrolled in other clinical trials. However, this would raise additional methodological challenges, especially regarding the integration of all necessary compartments in a large model. All this additional information would help: i) refining the estimated durability of the humoral immune response, ii) quantifying the effect of factors impacting on the variability of the immune response, iii) understanding more deeply the mechanisms of the

immune memory, iv) predicting the effect of additional interventions, such as additional immunization or probiotic intervention, aiming at increasing the immunogenicity of the vaccine, v) aiming at optimizing protocols of vaccine immunizations.

Regarding the optimization of vaccine regimens based on several immunizations, we have developed in parallel to the modeling work a numerical tool based on optimal control theory. This was applied to the context of immunotherapeutic interventions in HIV-infected patients, as a proof-of-concept, showing that the theory could indeed be applied in some specific biological framework. We have focused on using PDMPs for modeling the biological process of interest, as this class of model corresponds to stochastic processes where the system undergoes deterministic trajectories that are changed discretely after some random time periods. The deterministic part can indeed be modeled by ODEs, which is our preferential tool in mechanistic modeling. In particular, we have applied some previously developed theory on the optimal control of PDMP to the particular case of impulse control on the boundaries of the system. A numerical method based on dynamic programming was developed and helped determining optimized immunotherapy protocols for HIV-infected patients. This development opens the door to other biological applications, especially in the field of vaccinology.

More generally, the work realized in this thesis highlights the availability of complex methodological tools for analyzing data from clinical trials on preventive and therapeutic interventions against infectious diseases. Of course, these tools should be constantly improved and adapted to the availability and of the data, evolving with the development of new technologies. However, mathematical tools can already be useful to accelerate the clinical development thanks to *in silico* trials. This term corresponds to the use of patient-specific computer simulations [Viceconti et al., 2016] for the development of a product, the diagnosis of the disease or the design of a treatment. For that, all the information regarding the characteristics of the patient and the mechanisms of effect of the studied product should ideally be integrated in a multi-scale model, able to predict the outcome of interest. As described in Viceconti et al. [2016] *in silico* trials could be used in several ways, in particular to reduce the number of subjects enrolled in clinical trials or the length of the study, to refine the efficacy of a product or replace some animal and human studies. In the last few years, there has been growing interest in these computation methods. Some examples of use of this approach cover the feedback control of glucose concentration in diabetic patients [Magni et al., 2009], the development of drugs accounting for variability due to several factors [Rostami-Hodjegan and Tucker, 2007] or the simulation of immunotherapy to design optimized interventions: this is in particular due to the non linearity of the process and the possibility to generate counter-intuitive predictions of the immunotherapy effect with the model [Clermont et al., 2004]. An *in*

silico phase 3 clinical trial was also implemented on 10000 pseudo-patients with Chron's Disease to evaluate the response to new treatments [Abedi et al., 2015]. It was also recently suggested to regulatory agencies to account for modeling and simulations analyses in the development process of biologic products, as explained in Pappalardo et al. [2018]. This review also underlines the potential usefulness of *in silico* trials at all phases of the clinical development of future vaccines. In the team, a classic *in silico* pipeline of analysis has consisted in developing a deterministic, ODE-based mechanistic model and estimating the distribution of the parameters in the population of the study. Two applications can be considered, either at the individual or population level. Indeed, if a new subject is included in the study, its individual parameters can be estimated based on the model and the previous estimations. Simulations of the deterministic model by accounting for the uncertainty around the individual parameters estimation can help adapting the treatment of this subject. This type of approach was used for adapting the schedule of IL-7 injections as already mentioned, and described in appendix B. On the other hand, a large number of pseudo-patients can be generated from the posterior distribution of the parameters. This population is supposed to be representative of the population studied in the clinical trials used for the estimation of the model. In that case, the outcome of several interventions can be simulated on all pseudo-patients. This can help determining the best intervention to use in a target population, or for different classes of profiles of patients. For this type of approach, a work in collaboration with Mélanie Prague aims at developing *in silico* trials to evaluate therapeutic relief in HIV-infected patients [Prague et al., 2018]. In particular, short-cycle therapies consisting of x days under treatment and 7-x days with no treatment, are though to be promising and have been under clinical investigation for a few years. By comparing the results obtained through the simulation approach with existing results from current clinical trials, we have shown that the computer-based approach was a good predictor of the effect of short-cycle therapies on HIV-infected patients, meaning that this approach should be considered in the development of new treatments strategies based on already existing antiretroviral therapies.

Following these results, a similar *in silico* pipeline could be considered in vaccinology and especially in the development of prime-boost regimens. Several designs could be tested, including different orders of immunizations, different intervals between the immunizations or hypothetical additional immunization and its timing. In the light of our findings on Ebola vaccine, it could also be interesting to evaluate optimal vaccine strategies according to the geographical location of the subject. The *in silico* method could help determining optimized vaccine regimens that should be tested in new clinical trials. The good results obtained from the methods used in the IL-7 and the ART frameworks can mostly be attributed to the ability of the deterministic model to fit the data using a

small number of biomarkers and the short phase of learning for estimating the parameters of a new subject. However, we have mentioned the necessity to integrate more than just a few biomarkers' data in a model for the immune response: transcriptomic, proteomic and microbiome data should also be included if possible. A step of selection of the necessary information could be considered to reduce the complexity of such model. But accounting for additional types of data could be crucial and induces other methodological challenges, either in modeling the dynamics of the underlying processes and/or in integrating the information in a large mechanistic model. Indeed, stochasticity cannot be ignored at the single-cell level, but on the other hand, adding stochasticity in the model makes the inference harder. The optimal control work realized in this thesis has highlighted the availability of other tools for modeling and optimizing stochastic processes, in particular by using PDMPs. The optimization is based on theoretical results and not on empirical methods. PDMPs could be used in a modeling strategy when stochasticity cannot be ignored, as we have shown that optimization of this kind of process in a biological framework could be realized. This type of modeling raises other challenges, especially in the inference of the model. However, it could be interesting to investigate the possibility to modify the *in silico* pipeline, by first calibrating the stochastic model on clinical data and then applying the optimal control tools on the calibrated model: an application of interest could be for example the modeling of the gene expression through GRN. Overall, a good balance should be found between the complexity of the model (including the number of compartments, the stochasticity, the time scales...) and the availability of the data. Here, the population approach is very much valuable, as specific measurements are usually realized on sub-cohorts of clinical trials. The population approach allows to pool the data from the principal study with ancillary studies, as having the same amount of data for all subjects is not mandatory in this approach. In any cases of modeling and optimizing methods, the *in silico* methods could help for a more rational development of vaccines, based on the quantification of the immune response generated by the immunizations and the understanding of the underlying mechanisms, instead of the actual empirical methods. It should be kept in mind that these methods are not supposed to replace the experimental methods, as a model is always a simplification of the real biological process: anything that will be predicted from *in silico* trials is related to the knowledge included in the construction of the model and the data used for the inference/calibration of the model. Experimentations should be combined to computation approaches as complementary roles in a collaborative effort between immunologists and mathematicians to further improve the development of new preventive and therapeutic interventions against infectious diseases. This would constitute a loop process, where generated data can be used for the development of models, which themselves can generate new hypotheses and help

designing future studies and determining additional data to be collected. These data can then validate or not the hypotheses and refine the model [Vodovotz et al., 2017].

In conclusion, further work directions can be pursued: in terms of modeling, more complex models of the immune response should be developed, including the immune memory and integrating early signals of the response. Stochasticity should be considered, especially at the single-cell levels, and in this prospect, inference of gene networks should be investigated and possibly adapted depending on the data available. In term of clinical interpretation, additional data (long term follow-up, sub-cohorts) will help refining the quantification of the immune response to vaccine immunizations and the factors associated to its variability. In terms of optimization, the numerical method based on dynamic programming could be improved and applied to other frameworks, especially where the modeling by a PDMP seems to be more adapted and necessary. All together, this will allow for adopting a substantial system vaccinology approach for future clinical developments.

Bibliography

Abbas A. K., Lichtman A. H., and Pillai S. Cellular and molecular immunology. 6th edition. *Philidelphia: Saunders Elsevier*, 2010. 31, 58

Abe M., Okada K., Hayashida K., Matsuo F., Shiosaki K., Miyazaki C., Ueda K., and Kino Y. Duration of neutralizing antibody titer after Japanese encephalitis vaccination. *Microbiology and immunology*, 51(6): 609–616, 2007. 60

Abedi V., Lu P., Hontecillas R., Verma M., Vess G., Philipson C. W., Carbo A., Leber A., Tubau-Juni N., Hoops S., et al. Phase III placebo-controlled, randomized clinical trial with synthetic Crohn's disease patients to evaluate treatment response. *Computational Modeling-Based Discovery of Novel Classes of Anti-Inflammatory Drugs that Target Lanthionine Synthetase C-Like Protein. Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology–Systems & Applications*, 2: 79, 2015. 169

Adams N. M., O'Sullivan T. E., Geary C. D., Karo J. M., Amezquita R. A., Joshi N. S., Kaech S. M., and Sun J. C. NK cell responses redefine immunological memory. *The Journal of Immunology*, 197(8): 2963–2970, 2016. 100

Agnandji S. T., Huttner A., Zinser M. E., Njuguna P., Dahlke C., Fernandes J. F., Yerly S., Dayer J.-A., Kraehling V., Kasonta R., et al. Phase 1 trials of rVSV Ebola vaccine in Africa and Europe. *New England Journal of Medicine*, 374(17): 1647–1660, 2016. 53, 55

Ahmed R. and Gray D. Immunological memory and protective immunity: understanding their relation. *Science*, 272(5258): 54–60, 1996. 36

Amanna I. J. and Slifka M. K. Mechanisms that determine plasma cell lifespan and the duration of humoral immunity. *Immunological reviews*, 236(1): 125–138, 2010. 58, 63

Amanna I. J., Carlson N. E., and Slifka M. K. Duration of humoral immunity to common viral and vaccine antigens. *New England Journal of Medicine*, 357(19): 1903–1915, 2007. 58, 60

Andraud M., Lejeune O., Musoro J. Z., Ogunjimi B., Beutels P., and Hens N. Living on three time scales: the dynamics of plasma cell and antibody populations illustrated for hepatitis A virus. *PLOS Computational Biology*, 8(3): e1002418, 2012. 45, 63

Antia R., Bergstrom C. T., Pilyugin S. S., Kaech S. M., and Ahmed R. Models of CD8+ responses: 1. What is the antigen-independent proliferation program. *Journal of Theoretical Biology*, 221(4): 585–598, 2003. 44, 62

Antia R., Ganusov V. V., and Ahmed R. The role of models in understanding CD8+ T-cell memory. *Nature Reviews Immunology*, 5(2): 101, 2005. 44

Autran B., Carcelain G., Li T. S., Blanc C., Mathez D., Tubiana R., Katlama C., Debre P., and Leibowitch J. Positive effects of combined antiretroviral therapy on CD4+ T cell homeostasis and function in advanced HIV disease. *Science*, 277(5322): 112–116, 1997. 108

Baden L. R., Karita E., Mutua G., Bekker L.-G., Gray G., Page-Shipp L., Walsh S. R., Nyombayire J., Anzala O., Roux S., et al. Assessment of the safety and immunogenicity of 2 novel vaccine platforms for HIV-1 prevention: a randomized trial. *Annals of Internal Medicine*, 164(5): 313–322, 2016. 40, 42

Bailleux F., Coudeville L., Kolenc-Saban A., Bevilacqua J., Barreto L., and André P. Predicted long-term persistence of pertussis antibodies in adolescents after an adolescent and adult formulation combined tetanus, diphtheria, and 5-component acellular pertussis vaccine, based on mathematical modeling and 5-year observed data. *Vaccine*, 26(31): 3903–3908, 2008. 60

Baize S., Pannetier D., Oestereich L., Rieger T., Koivogui L., Magassouba N., Soropogui B., Sow M. S., Keïta S., De Clerck H., et al. Emergence of Zaire Ebola virus disease in Guinea. *New England Journal of Medicine*, 371(15): 1418–1425, 2014. 49

Barouch D. H. A step forward for HIV vaccines. *The Lancet HIV*, 5(7): e338–e339, 2018. 39

Battegay M., Nüesch R., Hirschel B., and Kaufmann G. R. Immunological recovery and antiretroviral therapy in HIV-1 infection. *The Lancet Infectious Diseases*, 6(5): 280–287, 2006. 108

Bäuerle N. and Rieder U. *Markov decision processes with applications to finance*. Universitext. Springer, Heidelberg, 2011. 162

Beal S. L., Sheiner L. B., Boeckmann A., and Bauer R. J. NONMEM users guides. *NONMEM Project Group, University of California, San Francisco*, 1992. 72

Bejon P., Ogada E., Mwangi T., Milligan P., Lang T., Fegan G., Gilbert S. C., Peshu N., Marsh K., and Hill A. V. Extended follow-up following a phase 2b randomized trial of the candidate malaria vaccines FP9 ME-TRAP and MVA ME-TRAP among children in Kenya. *PLOS ONE*, 2(8): e707, 2007. 38

Bekker L.-G., Moodie Z., Grunenberg N., Laher F., Tomaras G. D., Cohen K. W., Allen M., Malahleha M., Mngadi K., Daniels B., et al. Subtype C ALVAC-HIV and bivalent subtype C gp120/MF59 HIV-1 vaccine in low-risk, HIV-uninfected, South African adults: a phase 1/2 trial. *The Lancet HIV*, 2018. 39

Bellman R. *Dynamic programming*. Princeton University Press, 1957. 117

Bensoussan A. Lectures on stochastic control. In *Nonlinear filtering and stochastic control*, pages 1–62. Springer, 1982. 121

Beq S., Delfraissy J.-F., and Theze J. Interleukin-7 (IL-7): immune function, involvement in the pathogenesis of HIV infection and therapeutic potential. *European Cytokine Network*, 15(4): 279–289, 2004. 110

Berkman S. A., Lee M. L., and Gale R. P. Clinical uses of intravenous immunoglobulins. *Annals of Internal Medicine*, 112(4): 278–292, 1990. 58

Berndt E. R., Hall B. H., Hall R. E., and Hausman J. A. Estimation and inference in nonlinear structural models. In *Annals of Economic and Social Measurement, Volume 3, number 4*, pages 653–665. NBER, 1974. 69

Bierman W. F., van Agtmael M. A., Nijhuis M., Danner S. A., and Boucher C. A. HIV monotherapy with ritonavir-boosted protease inhibitors: a systematic review. *AIDS*, 23(3): 279–291, 2009. 108

Black G. F., Weir R. E., Floyd S., Bliss L., Warndorff D. K., Crampin A. C., Ngwira B., Sichali L., Nazareth B., Blackwell J. M., et al. BCG-induced increase in interferon-gamma response to mycobacterial antigens and efficacy of BCG vaccination in Malawi and the UK: two randomised controlled studies. *The Lancet*, 359(9315): 1393–1401, 2002. 42

Bobrowski A., Lipniacki T., Pichór K., and Rudnicki R. Asymptotic behavior of distributions of mRNA and protein levels in a model of stochastic gene expression. *Journal of mathematical analysis and applications*, 333(2): 753–769, 2007. 121

Bokes P., King J. R., Wood A. T., and Loose M. Transcriptional bursting diversifies the behaviour of a toggle switch: hybrid simulation of stochastic gene expression. *Bulletin of Mathematical Biology*, 75(2): 351–371, 2013. 164

Bonabeau E. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99(suppl 3): 7280–7287, 2002. 43

Bouteloup V., Sabin C., Mocroft A., Gras L., Pantazis N., Le Moing V., d'Arminio Monforte A., Mary-Krause M., Roca B., Miro J. M., et al. Reference curves for CD4 T-cell count response to combination antiretroviral therapy in HIV-1-infected treatment-naïve patients. *HIV Medicine*, 18(1): 33–44, 2017. 109

Bovier P., Bock J., Loutan L., Farinelli T., Glueck R., and Herzog C. Long-term immunogenicity of an inactivated virosome hepatitis A vaccine. *Journal of Medical Virology*, 68(4): 489–493, 2002. 60

Brekke O. H. and Sandlie I. Therapeutic antibodies for human diseases at the dawn of the twenty-first century. *Nature Reviews Drug Discovery*, 2(1): 52–62, 2003. 58

Brinkhof M. W., Mayorga O., Bock J., Heininger U., and Herzog C. Kinetics of maternally acquired anti-hepatitis A antibodies: prediction of waning based on maternal or cord blood antibody levels. *Vaccine*, 31(11): 1490–1495, 2013. 58

Brisson M., Bénard É., Drolet M., Bogaards J. A., Baussano I., Vänskä S., Jit M., Boily M.-C., Smith M. A., Berkhof J., et al. Population-level impact, herd immunity, and elimination after human papillomavirus vaccination: a systematic review and meta-analysis of predictions from transmission-dynamic models. *The Lancet Public Health*, 1(1): e8–e17, 2016. 38

Buchbinder S. P., Mehrotra D. V., Duerr A., Fitzgerald D. W., Mogg R., Li D., Gilbert P. B., Lama J. R., Marmor M., del Rio C., et al. Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. *The Lancet*, 372(9653): 1881–1893, 2008. 42

Buckwar E. and Riedler M. G. An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution. *Journal of Mathematical Biology*, 63(6): 1051–1093, 2011. 121

Butler D. What first case of sexually transmitted Ebola means for public health. *Nature News*, 2015. 49

Callard R. E. and Yates A. J. Immunology and mathematics: crossing the divide. *Immunology*, 115(1): 21–33, 2005. 44

Callendret B., Vellinga J., Wunderlich K., Rodriguez A., Steigerwald R., Dirmeier U., Cheminay C., Volkmann A., Brasel T., Carrion R., et al. A prophylactic multivalent vaccine against different filovirus species is immunogenic and provides protection from lethal infections with Ebolavirus and Marburgvirus species in non-human primates. *PLOS ONE*, 13(2): e0192312, 2018. 56

Camargo J. F., Kulkarni H., Agan B. K., Gaitan A. A., Beachy L. A., Srinivas S., He W., Anderson S., Marconi V. C., Dolan M. J., et al. Responsiveness of T Cells to Interleukin-7 Is Associated with Higher CD4+ T Cell Counts in HIV-1–Positive Individuals with Highly Active Antiretroviral Therapy–Induced Viral Load Suppression. *The Journal of Infectious Diseases*, 199(12): 1872–1882, 2009. 110

Capeding M. R., Tran N. H., Hadinegoro S. R. S., Ismail H. I. H. M., Chotpitayasunondh T., Chua M. N., Luong C. Q., Rusmil K., Wirawan D. N., Nallusamy R., et al. *The Lancet*, 384 (9951): 1358–1365, 2014. 41

Cappuccio A., Castiglione F., and Piccoli B. Determination of the optimal therapeutic protocols in cancer immunotherapy. *Mathematical Biosciences*, 209(1): 1–13, 2007. 118

Castiglione F. and Piccoli B. Cancer immunotherapy, mathematical modeling and optimal control. *Journal of Theoretical Biology*, 247(4): 723–732, 2007. 118

Castiglione F. and Piccoli B. Optimal control in a model of dendritic cell transfection cancer immunotherapy. *Bulletin of Mathematical Biology*, 68(2): 255–274, 2006. 118

Castiglione F., Mantile F., De Berardinis P., and Prisco A. How the interval between prime and boost injection affects the immune response in a computational model of the immune system. *Computational and Mathematical Methods in Medicine*, 2012, 2012. 45

Castro M., Lythe G., Molina-París C., and Ribeiro R. M. Mathematics in modern immunology. *Interface focus*, 6(2): 20150093, 2016. 44

CDC. *Bioterrorism Agents/Diseases | Emergency Preparedness*, 2018. `https://emergency.cdc.gov/agent/vhf/index.asp`. 49

Chan P. L., Jacqmin P., Lavielle M., McFadyen L., and Weatherley B. The use of the SAEM algorithm in MONOLIX software for estimation of population pharmacokinetic-pharmacodynamic-viral dynamics parameters of maraviroc in asymptomatic HIV subjects. *Journal of Pharmacokinetics and Pharmacodynamics*, 38(1): 41–61, 2011. 73

Chen S., Zhou Z., Wei F.-X., Huang S.-J., Tan Z., Fang Y., Zhu F.-C., Wu T., Zhang J., and Xia N.-S. Modeling the long-term antibody response of a hepatitis E vaccine. *Vaccine*, 33(33): 4124–4129, 2015. 61

Cheuvart B., Burgess M., Zepp F., Mertsola J., Wolter J., and Schuerman L. Anti-diphtheria antibody seroprotection rates are similar 10 years after vaccination with dTpa or DTPa using a mathematical model. *Vaccine*, 23(3): 336–342, 2004. 60

Clermont G., Bartels J., Kumar R., Constantine G., Vodovotz Y., and Chow C. In silico design of clinical trials: a method coming of age. *Critical care medicine*, 32(10): 2061–2070, 2004. 168

Cloez B., Dessalles R., Genadot A., Malrieu F., Marguet A., and Yvinec R. Probabilistic and piecewise deterministic models in biology. *ESAIM: Proceedings and Surveys*, 60: 225–245, 2017. 119

Commenges D., Jacqmin-Gadda H., Proust C., and Guedj J. A newton-like algorithm for likelihood maximization: The robust-variance scoring algorithm. *arXiv preprint math/0610402*, 2006. 69

Commenges D., Joly P., Gégout-Petit A., and Liquet B. Choice between Semi-parametric Estimators of Markov and Non-Markov Multi-state Models from Coarsened Observations. *Scandinavian Journal of Statistics*, 34(1): 33–52, 2007. 71

Commenges D., Sayyareh A., Letenneur L., Guedj J., Bar-Hen A., et al. Estimating a difference of Kullback–Leibler risks using a normalized difference of AIC. *The Annals of Applied Statistics*, 2(3): 1123–1142, 2008. 72

Commenges D., Proust-Lima C., Samieri C., and Liquet B. A universal approximate cross-validation criterion for regular risk functions. *The International Journal of Biostatistics*, 11 (1): 51–67, 2015. 72

Cook I. F. Sexual dimorphism of humoral immunity with human vaccines. *Vaccine*, 26(29-30): 3551–3555, 2008. 41

Costa O. L. V., Dufour F., and Piunovskiy A. B. Constrained and unconstrained optimal discounted control of piecewise deterministic Markov processes. *SIAM Journal on Control and Optimization*, 54(3): 1444–1474, 2016. 120, 123, 124, 125, 126, 157

Crauste F., Mafille J., Boucinha L., Djebali S., Gandrillon O., Marvel J., and Arpin C. Identification of Nascent Memory CD8 T Cells and Modeling of Their Ontogeny. *Cell systems*, 4(3): 306–317, 2017. 44

Crotty S. A brief history of T cell help to B cells. *Nature Reviews Immunology*, 15(3): 185, 2015. 35

Dans A. L., Dans L. F., Lansang M. A. D., Silvestre M. A. A., and Guyatt G. H. Controversy and debate on dengue vaccine series—paper 1: review of a licensed dengue vaccine: inappropriate subgroup analyses and selective reporting may cause harm in mass vaccination programs. *Journal of Clinical Epidemiology*, 95: 137–139, 2018. 41

David M.-P., Van Herck K., Hardt K., Tibaldi F., Dubin G., Descamps D., and Van Damme P. Long-term persistence of anti-HPV-16 and-18 antibodies induced by vaccination with the AS04-adjuvanted cervical cancer vaccine: modeling of sustained antibody responses. *Gynecologic Oncology*, 115(3): S1–S6, 2009. 61

Davis M. H. Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 353–388, 1984. 121

Da'dara A. A. and Harn D. A. Elimination of helminth infection restores HIV-1C vaccine-specific T cell responses independent of helminth-induced IL-10. *Vaccine*, 28(5): 1310–1317, 2010. 42

de Boer R. J., Oprea M., Antia R., Murali-Krishna K., Ahmed R., and Perelson A. S. Recruitment times, proliferation, and apoptosis rates during the CD8+ T-cell response to lymphocytic choriomeningitis virus. *Journal of Virology*, 75(22): 10663–10669, 2001. 44, 62

de Boer R. J., Homann D., and Perelson A. S. Different dynamics of CD4+ and CD8+ T cell responses during and after acute lymphocytic choriomeningitis virus infection. *The Journal of Immunology*, 171(8): 3928–3935, 2003. 44

de Bruyn G. Cofactors that may influence vaccine responses. *Current Opinion in HIV and AIDS*, 5(5): 404, 2010. 41

de Jong H. Modeling and simulation of genetic regulatory systems: a literature review. *Journal of Computational Biology*, 9(1): 67–103, 2002. 102

De Rosa S. C., Thomas E. P., Bui J., Huang Y., Morgan C., Kalams S. A., Tomaras G. D., Akondy R., Ahmed R., Lau C.-Y., et al. HIV-DNA priming alters T cell responses to HIV-adenovirus vaccine even when responses to DNA are undetectable. *The Journal of Immunology*, 187(6): 3391–3401, 2011. 40

De Santis O., Audran R., Pothin E., Warpelin-Decrausaz L., Vallotton L., Wuerzner G., Cochet C., Estoppey D., Steiner-Monard V., Lonchampt S., et al. Safety and immunogenicity of a chimpanzee adenovirus-vectored Ebola vaccine in healthy adults: a randomised, double-blind, placebo-controlled, dose-finding, phase 1/2a study. *The Lancet Infectious Diseases*, 16 (3): 311–320, 2016. 52, 54

de Saporta B. and Dufour F. Numerical method for impulse control of piecewise deterministic Markov processes. *Automatica*, 48(5): 779–793, 2012. 126

de Saporta B., Dufour F., Zhang H., and Elegbede C. Optimal stopping for the predictive maintenance of a structure subject to corrosion. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 226(2): 169–181, 2012. 121

de Saporta B., Dufour F., and Zhang H. *Numerical methods for simulation and optimization of piecewise deterministic markov processes*. John Wiley & Sons, 2015. 126

de Vrese M., Rautenberg P., Laue C., Koopmans M., Herremans T., and Schrezenmeir J. Probiotic bacteria stimulate virus–specific neutralizing antibodies following a booster polio vaccination. *European journal of nutrition*, 44(7): 406–413, 2005. 42

Delfraissy J.-F., Flandre P., Delaugerre C., Ghosn J., Horban A., Girard P.-M., Norton M., Rouzioux C., Taburet A.-M., Cohen-Codar I., et al. Lopinavir/ritonavir monotherapy or plus zidovudine and lamivudine in antiretroviral-naive HIV-infected patients. *AIDS*, 22(3): 385–393, 2008. 108

Dempster A. P., Laird N. M., and Rubin D. B. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (methodological)*, pages 1–38, 1977. 73

Desai K., Coudeville L., and Bailleux F. Modelling the long-term persistence of neutralizing antibody in adults after one dose of live attenuated Japanese encephalitis chimeric virus vaccine. *Vaccine*, 30(15): 2510–2515, 2012. 60

Dixit N. M. and Perelson A. S. HIV dynamics with multiple infections of target cells. *Proceedings of the National Academy of Sciences*, 102(23): 8198–8203, 2005. 65

Drylewicz J., Commenges D., and Thiébaut R. Maximum a posteriori estimation in dynamical models of primary HIV infection. *Statistical Communications in Infectious Diseases*, 4(1), 2012. 70

Dunning J., Kennedy S. B., Antierens A., Whitehead J., Ciglenecki I., Carson G., Kanapathipillai R., Castle L., Howell-Jones R., Pardinaz-Solis R., et al. Experimental treatment of Ebola virus disease with brincidofovir. *PLOS ONE*, 11(9): e0162199, 2016a. 51

Dunning J., Sahr F., Rojek A., Gannon F., Carson G., Idriss B., Massaquoi T., Gandi R., Joseph S., Osman H. K., et al. Experimental treatment of Ebola virus disease with TKM-130803: a single-arm phase 2 clinical trial. *PLOS Medicine*, 13(4): e1001997, 2016b. 51

Duwal S., Winkelmann S., Schütte C., and von Kleist M. Optimal treatment strategies in the context of 'treatment for prevention'against HIV-1 in resource-poor settings. *PLOS Computational Biology*, 11(4): e1004200, 2015. 119

Eftimie R., Gillard J. J., and Cantrell D. A. Mathematical models for immunology: Current state of the art and future research directions. *Bulletin of Mathematical Biology*, 78(10): 2091–2134, 2016. 119

Eurosurveillance editorial team. First innovative medicines initiative ebola projects launched. *Eurosurveillance*, 20(3): 21014, 2015. 52

Ewer K., Rampling T., Venkatraman N., Bowyer G., Wright D., Lambe T., Imoukhuede E. B., Payne R., Fehling S. K., Strecker T., et al. A monovalent chimpanzee adenovirus Ebola vaccine boosted with MVA. *New England Journal of Medicine*, 374(17): 1635–1646, 2016. 52, 54

Farber D. L., Netea M. G., Radbruch A., Rajewsky K., and Zinkernagel R. M. Immunological memory: lessons from the past and a look to the future. *Nature Reviews Immunology*, 16(2): 124, 2016. 36

Ferreira R. B., Antunes L. C. M., and Finlay B. B. Should the human microbiome be considered when developing vaccines? *PLOS Pathogens*, 6(11): e1001190, 2010. 42

Fink K. Origin and function of circulating plasmablasts during acute viral infections. *Frontiers in Immunology*, 3: 78, 2012. 58

Fischl M. A., Richman D. D., Grieco M. H., Gottlieb M. S., Volberding P. A., Laskin O. L., Leedom J. M., Groopman J. E., Mildvan D., Schooley R. T., et al. The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. *New England Journal of Medicine*, 317(4): 185–191, 1987. 108

Fish E. N. The X-files in immunity: sex-based differences predispose immune responses. *Nature Reviews Immunology*, 8(9): 737, 2008. 41

Folcik V. A., An G. C., and Orosz C. G. The Basic Immune Simulator: an agent-based model to study the interactions between innate and adaptive immunity. *Theoretical Biology and Medical Modelling*, 4(1): 39, 2007. 43

Framstad N. C., Øksendal B., and Sulem A. Sufficient stochastic maximum principle for the optimal control of jump diffusions and applications to finance. *Journal of Optimization Theory and Applications*, 121(1): 77–98, 2004. 121

Fraser C., Tomassini J. E., Xi L., Golm G., Watson M., Giuliano A. R., Barr E., and Ault K. A. Modeling the long-term antibody response of a human papillomavirus (HPV) virus-like particle (VLP) type 16 prophylactic vaccine. *Vaccine*, 25(21): 4324–4333, 2007. 60, 61

Fraser K. A., Schenkel J. M., Jameson S. C., Vezys V., and Masopust D. Preexisting high frequencies of memory CD8+ T cells favor rapid memory differentiation and preservation of proliferative potential upon boosting. *Immunity*, 39(1): 171–183, 2013. 36

Frölich D., Giesecke C., Mei H. E., Reiter K., Daridon C., Lipsky P. E., and Dörner T. Secondary immunization generates clonally related antigen-specific plasma cells and memory B cells. *The Journal of Immunology*, 185(5): 3103–3110, 2010. 58

Gelman A., Lee D., and Guo J. Stan: A probabilistic programming language for Bayesian inference and optimization. *Journal of Educational and Behavioral Statistics*, 40(5): 530–543, 2015. 70

Germain R. N. The art of the probable: system control in the adaptive immune system. *Science*, 293(5528): 240–245, 2001. 43

Germain R. N. Vaccines and the future of human immunology. *Immunity*, 33(4): 441–450, 2010. 38, 39, 99

Germain R. N. Will Systems Biology Deliver Its Promise and Contribute to the Development of New or Improved Vaccines? What Really Constitutes the Study of "Systems Biology" and How Might Such an Approach Facilitate Vaccine Design. *Cold Spring Harbor perspectives in biology*, page a033308, 2017. 43, 99

Golding H., Khurana S., and Zaitseva M. What Is the Predictive Value of Animal Models for Vaccine Efficacy in Humans? The Importance of Bridging Studies and Species-Independent Correlates of Protection. *Cold Spring Harbor perspectives in biology*, 10(4): a028902, 2018. 56

Gómez C. E., Nájera J. L., Perdiguero B., García-Arriaza J., Sorzano C. O. S., Jiménez V., González-Sanz R., Jiménez J. L., Muñoz-Fernández M. A., de Quirós J. C. L. B., et al. The HIV/AIDS vaccine candidate MVA-B administered as a single immunogen in humans triggers robust, polyfunctional and selective effector memory T cell responses to HIV-1 antigens. *Journal of Virology*, pages JVI–05165, 2011. 38

Gong C., Linderman J. J., and Kirschner D. Harnessing the heterogeneity of T cell differentiation fate to fine-tune generation of effector and memory T cells. *Frontiers in Immunology*, 5: 57, 2014. 45

Grassly N. C., Praharaj I., Babji S., Kaliappan S. P., Giri S., Venugopal S., Parker E. P., Abraham A., Muliyil J., Doss S., et al. The effect of azithromycin on the immunogenicity of oral poliovirus vaccine: a double-blind randomised placebo-controlled trial in seronegative Indian infants. *The Lancet Infectious Diseases*, 16(8): 905–914, 2016. 42

Graw F., Weber K. S., Allen P. M., and Perelson A. S. Dynamics of CD4+ T cell responses against Listeria monocytogenes. *The Journal of Immunology*, 189(11): 5250–5256, 2012. 44

Gross L., Lhomme E., Pasin C., Richert L., and Thiebaut R. Ebola vaccine development: Systematic review of pre-clinical and clinical studies, and meta-analysis of determinants of antibody response variability after vaccination. *International Journal of Infectious Diseases*, 74: 83–96, 2018. 51, 57

Guedj J., Thiébaut R., and Commenges D. Maximum likelihood estimation in dynamical models of HIV. *Biometrics*, 63(4): 1198–1206, 2007a. 67

Guedj J., Thiébaut R., and Commenges D. Practical identifiability of HIV dynamics models. *Bulletin of Mathematical Biology*, 69(8): 2493–2513, 2007b. 70

Guedj J., Piorkowski G., Jacquot F., Madelain V., Nguyen T. H. T., Rodallec A., Gunther S., Carbonnelle C., Mentré F., Raoul H., et al. Antiviral efficacy of favipiravir against Ebola virus: A translational study in cynomolgus macaques. *PLOS Medicine*, 15(3): e1002535, 2018. 51

Gurwith M., Lock M., Taylor E. M., Ishioka G., Alexander J., Mayall T., Ervin J. E., Greenberg R. N., Strout C., Treanor J. J., et al. Safety and immunogenicity of an oral, replicating adenovirus serotype 4 vector vaccine for H5N1 influenza: a randomised, double-blind, placebo-controlled, phase 1 study. *The Lancet Infectious Diseases*, 13(3): 238–250, 2013. 38

Hagan T. and Pulendran B. Will systems biology deliver its promise and contribute to the development of new or improved vaccines? From data to understanding through systems biology. *Cold Spring Harbor perspectives in biology*, page a028894, 2017. 103

Hagan T., Nakaya H. I., Subramaniam S., and Pulendran B. Systems vaccinology: enabling rational vaccine design with systems biological approaches. *Vaccine*, 33(40): 5294–5301, 2015. 39, 99

Hammarlund E., Thomas A., Amanna I. J., Holden L. A., Slayden O. D., Park B., Gao L., and Slifka M. K. Plasma cell survival in the absence of B cell memory. *Nature Communications*, 8(1): 1781, 2017. 58

Hammer S. M., Katzenstein D. A., Hughes M. D., Gundacker H., Schooley R. T., Haubrich R. H., Henry W. K., Lederman M. M., Phair J. P., Niu M., et al. A trial comparing nucleoside monotherapy with combination therapy in HIV-infected adults with CD4 cell counts from 200 to 500 per cubic millimeter. *New England Journal of Medicine*, 335(15): 1081–1090, 1996. 108

Hammer S. M., Sobieszczyk M. E., Janes H., Karuna S. T., Mulligan M. J., Grove D., Koblin B. A., Buchbinder S. P., Keefer M. C., Tomaras G. D., et al. Efficacy trial of a DNA/rAd5 HIV-1 preventive vaccine. *New England Journal of Medicine*, 369(22): 2083–2092, 2013. 38

Hanke T., McMichael A. J., Dennis M. J., Sharpe S. A., Powell L. A., McLoughlin L., and Crome S. J. Biodistribution and persistence of an MVA-vectored candidate HIV vaccine in SIV-infected rhesus macaques and SCID mice. *Vaccine*, 23(12): 1507–1514, 2005. 98

Haque R., Snider C., Liu Y., Ma J. Z., Liu L., Nayak U., Mychaleckyj J. C., Korpe P., Mondal D., Kabir M., et al. Oral polio vaccine response in breast fed infants with malnutrition and diarrhea. *Vaccine*, 32(4): 478–482, 2014. 41

Hecker M., Lambeck S., Toepfer S., Van Someren E., and Guthke R. Gene regulatory network inference: data integration in dynamic models—a review. *Biosystems*, 96(1): 86–103, 2009. 102

Henao-Restrepo A. M., Camacho A., Longini I. M., Watson C. H., Edmunds W. J., Egger M., Carroll M. W., Dean N. E., Diatta I., Doumbia M., et al. Efficacy and effectiveness of an rVSV-vectored vaccine in preventing Ebola virus disease: final results from the Guinea ring vaccination, open-label, cluster-randomised trial (Ebola Ça Suffit!). *The Lancet*, 389(10068): 505–518, 2017. 55

Hens N., Ghebretinsae A. H., Hardt K., Van Damme P., and Van Herck K. Model based estimates of long-term persistence of inactivated hepatitis A vaccine-induced antibodies in adults. *Vaccine*, 32(13): 1507–1513, 2014. 61

Heppner Jr D. G., Kemp T. L., Martin B. K., Ramsey W. J., Nichols R., Dasen E. J., Link C. J., Das R., Xu Z. J., Sheldon E. A., et al. Safety and immunogenicity of the rVSVΔ G-ZEBOV-GP Ebola virus vaccine candidate in healthy adults: a phase 1b randomised, multicentre, double-blind, placebo-controlled, dose-response study. *The Lancet Infectious Diseases*, 17(8): 854–866, 2017. 53, 55

Herbach U., Bonnaffoux A., Espinasse T., and Gandrillon O. Inferring gene regulatory networks from single-cell data: a mechanistic approach. *BMC Systems Biology*, 11(1): 105, 2017. 102, 119, 164, 165

Hernández-Lerma O. *Adaptive Markov control processes*, volume 79 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1989. 162

Huang K.-Y. A., Li C. K.-F., Clutterbuck E., Chui C., Wilkinson T., Gilbert A., Oxford J., Lambkin-Williams R., Lin T.-Y., McMichael A. J., et al. Virus-specific antibody secreting cell, memory B-cell, and sero-antibody responses in the human influenza challenge model. *The Journal of Infectious Diseases*, 209(9): 1354–1361, 2014. 58

Huttner A., Dayer J.-A., Yerly S., Combescure C., Auderset F., Desmeules J., Eickmann M., Finckh A., Goncalves A. R., Hooper J. W., et al. The effect of dose on the safety and immunogenicity of the VSV Ebola candidate vaccine: a randomised double-blind, placebo-controlled phase 1/2 trial. *The Lancet Infectious Diseases*, 15(10): 1156–1166, 2015. 53, 55

Inoue T., Moran I., Shinnakasu R., Phan T. G., and Kurosaki T. Generation of memory B cells and their reactivation. *Immunological reviews*, 283(1): 138–149, 2018. 58

Ishizuka A. S., Lyke K. E., DeZure A., Berry A. A., Richie T. L., Mendoza F. H., Enama M. E., Gordon I. J., Chang L.-J., Sarwar U. N., et al. Protection against malaria at 1 year and immune correlates following PfSPZ vaccination. *Nature Medicine*, 22(6): 614, 2016. 102

Isolauri E., Joensuu J., Suomalainen H., Luomala M., and Vesikari T. Improved immunogenicity of oral D x RRV reassortant rotavirus vaccine by Lactobacillus casei GG. *Vaccine*, 13(3): 310–312, 1995. 42

Janeway C. A., Travers P., Walport M., and Shlomchik M. Immunobiology: the immune system in health and disease. 2001. *New York: Garland Science*, 5. 37

Jarne A., Commenges D., Villain L., Prague M., Lévy Y., Thiébaut R., et al. Modeling $CD4^+$ T cells dynamics in HIV-infected patients receiving repeated cycles of exogenous Interleukin 7. *The Annals of Applied Statistics*, 11(3): 1593–1616, 2017. 113, 114, 116, 120, 161

Jarne Munoz A. *Modeling the effect of exogenous Interleukin 7 in HIV patients under antiretroviral therapy with low immune reconstitution.* PhD thesis, Bordeaux, 2015. 113

Jilg W., Schmidt M., Deinhardt F., and Zachoval R. Hepatitis B vaccination: how long does protection last? *The Lancet*, 324(8400): 458, 1984. 58

Johnson K., Webb P., Lange J., Murphy F., et al. Isolation and partial characterisation of a new virus causing acute haemorrhagic fever in Zaire. *The Lancet*, pages 569–571, 1977. 47

Joyce A. R. and Palsson B. Ø. The model organism as a system: integrating 'omics' data sets. *Nature Reviews Molecular Cell Biology*, 7(3): 198, 2006. 102

Julg B. and Walker B. D. The Paradox of Incomplete CD4+ Cell Count Restoration Despite Successful Antiretroviral Treatment and the Need to Start Highly Active Antiretroviral Therapy Early. *Clinical Infectious Diseases*, 48(6): 795–797, 2009. 109

Kass R. E. and Steffey D. Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models). *Journal of the American Statistical Association*, 84(407): 717–726, 1989. 71

Kelley C. F., Kitchen C. M. R., Hunt P. W., Rodriguez B., Hecht F. M., Kitahata M., Crane H. M., Willig J., Mugavero M., Saag M., Martin J. N., and Deeks S. G. Incomplete Peripheral CD4+ Cell Count Restoration in HIV-Infected Patients Receiving Long-Term Antiretroviral Treatment. *Clinical Infectious Diseases*, 48(6): 787–794, 2009. 109

Kennedy S. B., Bolay F., Kieh M., Grandits G., Badio M., Ballou R., Eckes R., Feinberg M., Follmann D., Grund B., et al. Phase 2 placebo-controlled trial of two vaccines to prevent Ebola in Liberia. *New England Journal of Medicine*, 377(15): 1438–1447, 2017. 52, 54

Keshwara R., Johnson R. F., and Schnell M. J. Toward an effective Ebola virus vaccine. *Annual Review of Medicine*, 68: 371–386, 2017. 51

Kibuuka H., Berkowitz N. M., Millard M., Enama M. E., Tindikahwa A., Sekiziyivu A. B., Costner P., Sitar S., Glover D., Hu Z., et al. Safety and immunogenicity of Ebola virus and Marburg virus glycoprotein DNA vaccines assessed separately and concomitantly in healthy Ugandan adults: a phase 1b, randomised, double-blind, placebo-controlled clinical trial. *The Lancet*, 385(9977): 1545–1554, 2015. 52

Kim P. S., Lee P. P., and Levy D. Modeling regulation mechanisms in the immune system. *Journal of Theoretical Biology*, 246(1): 33–69, 2007. 44

Kitahata M. M., Gange S. J., Abraham A. G., Merriman B., Saag M. S., Justice A. C., Hogg R. S., Deeks S. G., Eron J. J., Brooks J. T., et al. Effect of early versus deferred antiretroviral therapy for HIV on survival. *New England Journal of Medicine*, 360(18): 1815–1826, 2009. 109

Kohler B. Mathematically modeling dynamics of T cell responses: predictions concerning the generation of memory cells. *Journal of Theoretical Biology*, 245(4): 669–676, 2007. 44

Kozal M. J., Kroodsma K., Winters M. A., Shafer R. W., Efron B., Katzenstein D. A., and Merigan T. C. Didanosine resistance in HIV-infected patients switched from zidovudine to didanosine monotherapy. *Annals of Internal Medicine*, 121(4): 263–268, 1994. 108

Krause P. R., Bryant P. R., Clark T., Dempsey W., Henchal E., Michael N. L., Regules J. A., and Gruber M. F. Immunology of protection from Ebola virus infection. *Science translational medicine*, 7(286): 286ps11–286ps11, 2015. 56

Kuhn E. and Lavielle M. Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics & Data Analysis*, 49(4): 1020–1038, 2005. 73

Kuhn J. H., Bao Y., Bavari S., Becker S., Bradfute S., Brister J. R., Bukreyev A. A., Chandran K., Davey R. A., Dolnik O., et al. Virus nomenclature below the species level: a standardized nomenclature for natural variants of viruses assigned to the family Filoviridae. *Archives of Virology*, 158(1): 301–311, 2013. 48

Lalor M. K., Ben-Smith A., Gorak-Stolinska P., Weir R. E., Floyd S., Blitz R., Mvula H., Newport M. J., Branson K., McGrath N., et al. Population differences in immune responses to Bacille Calmette-Guerin vaccination in infancy. *The Journal of Infectious Diseases*, 199(6): 795–800, 2009. 42

Lambe T., Bowyer G., and Ewer K. J. A review of Phase I trials of Ebola virus vaccines: what can we learn from the race to develop novel vaccines? *Phil. Trans. R. Soc. B*, 372(1721): 20160295, 2017. 51

Lange C. G. and Lederman M. M. Immune reconstitution with antiretroviral therapies in chronic HIV-1 infection. *Journal of Antimicrobial Chemotherapy*, 51(1): 1–4, 2003. 109

Lanzavecchia A. and Sallusto F. Human B cell memory. *Current Opinion in Immunology*, 21 (3): 298–304, 2009. 58

Larder B. A., Darby G., and Richman D. D. HIV with reduced sensitivity to zidovudine (AZT) isolated during prolonged therapy. *Science*, 243(4899): 1731–1734, 1989. 108

Lavielle M. and Mentré F. Estimation of population pharmacokinetic parameters of saquinavir in HIV patients with the MONOLIX software. *Journal of Pharmacokinetics and Pharmacodynamics*, 34(2): 229–249, 2007. 73, 163

Le D., Miller J. D., and Ganusov V. V. Mathematical modeling provides kinetic details of the human immune response to vaccination. *Frontiers in Cellular and Infection Microbiology*, 4: 177, 2015. 45, 62, 65

Ledgerwood J., Costner P., Desai N., Holman L., Enama M., Yamshchikov G., Mulangu S., Hu Z., Andrews C., Sheets R., et al. A replication defective recombinant Ad5 vaccine expressing Ebola virus GP is safe and immunogenic in healthy adults. *Vaccine*, 29(2): 304–313, 2010. 52, 54

Ledgerwood J. E., Sullivan N. J., and Graham B. S. Chimpanzee Adenovirus Vector Ebola Vaccine–Preliminary Report. *New England journal of medicine*, 373(8): 776, 2015. 52, 54

Ledgerwood J. E., DeZure A. D., Stanley D. A., Coates E. E., Novik L., Enama M. E., Berkowitz N. M., Hu Z., Joshi G., Ploquin A., et al. Chimpanzee adenovirus vector Ebola vaccine. *New England Journal of Medicine*, 376(10): 928–938, 2017. 52

Lee H. Y., Topham D. J., Park S. Y., Hollenbaugh J., Treanor J., Mosmann T. R., Jin X., Ward B. M., Miao H., Holden-Wiltse J., et al. Simulation and prediction of the adaptive immune response to influenza A virus infection. *Journal of Virology*, 83(14): 7151–7165, 2009. 44

Lee Y., Newport M., Goetghebuer T., Siegrist C.-A., Weiss H., Pollard A. J., Marchant A., Group M. T. S., et al. Influence of genetic and environmental factors on the immunogenicity of Hib vaccine in Gambian twins. *Vaccine*, 24(25): 5335–5340, 2006. 41

Leino T., Auranen K., Mäkelä P., Käyhty H., and Takala A. Dynamics of natural immunity caused by subclinical infections, case study on Haemophilus influenzae type b (Hib). *Epidemiology & Infection*, 125(3): 583–591, 2000. 60

Lenz D. C., Kurz S. K., Lemmens E., Schoenberger S. P., Sprent J., Oldstone M. B., and Homann D. IL-7 regulates basal homeostatic proliferation of antiviral CD4+ T cell memory. *Proceedings of the National Academy of Sciences*, 101(25): 9357–9362, 2004. 110

Leone A., Rohankhedkar M., Okoye A., Legasse A., Axthelm M. K., Villinger F., Piatak M., Lifson J. D., Assouline B., Morre M., et al. Increased CD4+ T cell levels during IL-7 administration of antiretroviral therapy-treated simian immunodeficiency virus-positive macaques are not dependent on strong proliferative responses. *The Journal of Immunology*, 185(3): 1650–1659, 2010. 110

Lesaffre E. and Spiessens B. On the effect of the number of quadrature points in a logistic random-effects model: An example. *Journal of the Royal Statistical Society. Series C*, 50: 325–335, 2001. 67

Leuridan E., Hens N., Hutse V., Aerts M., and Van Damme P. Kinetics of maternal antibodies against rubella and varicella in infants. *Vaccine*, 29(11): 2222–2226, 2011. 58

Levy Y., Sereti I., Tambussi G., Routy J., Lelievre J., Delfraissy J., Molina J., Fischl M., Goujard C., Rodriguez B., et al. Effects of recombinant human interleukin 7 on T-cell recovery and thymic output in HIV-infected patients receiving antiretroviral therapy: results of a phase I/IIa randomized, placebo-controlled, multicenter study. *Clinical Infectious Diseases*, 55(2): 291–300, 2012. 110

Levy Y., Lacabaratz C., Weiss L., Viard J.-P., Goujard C., Lelièvre J.-D., Boué F., Molina J.-M., Rouzioux C., Avettand-Fénoêl V., et al. Enhanced T cell recovery in HIV-1–infected adults through IL-7 treatment. *The Journal of Clinical Investigation*, 119(4): 997–1007, 2009. 110

Lévy Y., Lane C., Piot P., Beavogui A. H., Kieh M., Leigh B., Doumbia S., D'Ortenzio E., Lévy-Marchal C., Pierson J., et al. Prevention of Ebola virus disease through vaccination: where we are in 2018. *The Lancet*, 2018. 53

Lewden C., Chêne G., Morlat P., Raffi F., Dupon M., Dellamonica P., Pellegrin J.-L., Katlama C., Dabis F., Leport C., et al. HIV-infected adults with a CD4 cell count greater than 500 cells/mm3 on long-term combination antiretroviral therapy reach same mortality rates as the general population. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 46(1): 72–77, 2007. 109

Li J.-X., Hou L.-H., Meng F.-Y., Wu S.-P., Hu Y.-M., Liang Q., Chu K., Zhang Z., Xu J.-J., Tang R., et al. Immunity duration of a recombinant adenovirus type-5 vector-based Ebola vaccine and a homologous prime-boost immunisation in healthy adults in China: final report of a randomised, double-blind, placebo-controlled, phase 1 trial. *The Lancet Global Health*, 5 (3): e324–e334, 2017. 52, 54

Liu G., Wong G., Su S., Bi Y., Plummer F., Gao G. F., Kobinger G., and Qiu X. Clinical Evaluation of Ebola Virus Disease Therapeutics. *Trends in Molecular Medicine*, 23(9): 820–830, 2017. 51

Liu W. B., Li Z. X., Du Y., and Cao G. W. Ebola virus disease: from epidemiology to prophylaxis. *Military Medical Research*, 2(1): 7, 2015. 49

López E. L., Contrini M. M., Mistchenko A., Kieffer A., Baggaley R. F., Di Tanna G. L., Desai K., Rasuli A., and Armoni J. Modeling the long-term persistence of hepatitis A antibody after a two-dose vaccination schedule in Argentinean children. *The Pediatric infectious disease journal*, 34(4): 417–425, 2015. 60

Louis T. A. Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 226–233, 1982. 67

Mackall C. L., Fry T. J., Bare C., Morgan P., Galbraith A., and Gress R. E. IL-7 increases both thymic-dependent and thymic-independent T-cell regeneration after bone marrow transplantation. *Blood*, 97(5): 1491–1497, 2001. 110

MacLeod M. K., Kappler J. W., and Marrack P. Memory CD4 T cells: generation, reactivation and re-assignment. *Immunology*, 130(1): 10–15, 2010. 36

Magni L., Raimondo D. M., Dalla Man C., De Nicolao G., Kovatchev B., and Cobelli C. Model predictive control of glucose concentration in type I diabetic patients: An in silico trial. *Biomedical Signal Processing and Control*, 4(4): 338–346, 2009. 168

Magwira C. A. and Taylor M. B. Composition of gut microbiota and its influence on the immunogenicity of oral rotavirus vaccines. *Vaccine*, 2018. 42

Marbach D., Prill R. J., Schaffter T., Mattiussi C., Floreano D., and Stolovitzky G. Revealing strengths and weaknesses of methods for gene network inference. *Proceedings of the National Academy of Sciences*, 2010. 102

Marquardt D. W. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2): 431–441, 1963. 69

Martin J. E., Sullivan N. J., Enama M. E., Gordon I. J., Roederer M., Koup R. A., Bailer R. T., Chakrabarti B. K., Bailey M. A., Gomez P. L., et al. A DNA vaccine for Ebola virus is safe and immunogenic in a phase I clinical trial. *Clinical and vaccine immunology*, 13(11): 1267–1277, 2006. 52

Martin M. D. and Badovinac V. P. Influence of time and number of antigen encounters on memory CD8 T cell development. *Immunologic research*, 59(1-3): 35–44, 2014. 36

Marzi A., Robertson S. J., Haddock E., Feldmann F., Hanley P. W., Scott D. P., Strong J. E., Kobinger G., Best S. M., and Feldmann H. VSV-EBOV rapidly protects macaques against infection with the 2014/15 Ebola virus outbreak strain. *Science*, 349(6249): 739–742, 2015. 56

Mast T. C., Kierstead L., Gupta S. B., Nikas A. A., Kallas E. G., Novitsky V., Mbewe B., Pitisuttithum P., Schechter M., Vardas E., et al. International epidemiology of human pre-existing adenovirus (Ad) type-5, type-6, type-26 and type-36 neutralizing antibodies: correlates of high Ad5 titers and implications for potential HIV vaccine trials. *Vaccine*, 28(4): 950–957, 2010. 38

Metzger W. G. and Vivas-Martínez S. Questionable efficacy of the rVSV-ZEBOV Ebola vaccine. *The Lancet*, 391(10125): 1021, 2018. 53

Milligan I. D., Gibani M. M., Sewell R., Clutterbuck E. A., Campbell D., Plested E., Nuthall E., Voysey M., Silva-Reyes L., McElrath M. J., et al. Safety and immunogenicity of novel adenovirus type 26–and modified vaccinia Ankara–vectored Ebola vaccines: a randomized clinical trial. *JAMA*, 315(15): 1610–1623, 2016. 40, 52

Montefiori D. C., Metch B., McElrath M. J., Self S., Weinhold K. J., Corey L., and Network H. V. T. Demographic factors that influence the neutralizing antibody response in recipients of recombinant HIV-1 gp120 vaccines. *The Journal of Infectious Diseases*, 190(11): 1962–1969, 2004. 41

Morlat P. Prise en charge médicale des personnes vivant avec le VIH. Recommandations du groupe d'experts. Rapport 2017. Sous la direction du Pr. Philippe Morlat et sous l'égide du CNS et de l'ANRS, 2017. 108

Morris C. N. Parametric empirical Bayes inference: theory and applications. *Journal of the American Statistical Association*, 78(381): 47–55, 1983. 71

Muyanja E., Ssemaganda A., Ngauv P., Cubas R., Perrin H., Srinivasan D., Canderan G., Lawson B., Kopycinski J., Graham A. S., et al. Immune activation alters cellular and humoral responses to yellow fever 17D vaccine. *The Journal of Clinical Investigation*, 124(7): 3147–3158, 2014. 42

Nakaya H. I., Wrammert J., Lee E. K., Racioppi L., Marie-Kunze S., Haining W. N., Means A. R., Kasturi S. P., Khan N., Li G.-M., et al. Systems biology of vaccination for seasonal influenza in humans. *Nature Immunology*, 12(8): 786, 2011. 100

Nakaya H. I., Hagan T., Duraisingham S. S., Lee E. K., Kwissa M., Rouphael N., Frasca D., Gersten M., Mehta A. K., Gaujoux R., et al. Systems analysis of immunity to influenza vaccination across multiple years and in diverse populations reveals shared molecular signatures. *Immunity*, 43(6): 1186–1198, 2015. 41

Newport M., Goetghebuer T., Weiss H. A., Whittle H., Siegrist C.-A., and Marchant A. Genetic regulation of immune responses to vaccines in early life. *Genes and Immunity*, 5(2): 122, 2004. 41

Nielsen C. M., White M. J., Bottomley C., Lusa C., Rodríguez-Galán A., Turner S. E., Goodier M. R., and Riley E. M. Impaired NK cell responses to pertussis and H1N1 influenza vaccine antigens in human cytomegalovirus-infected individuals. *The Journal of Immunology*, 194(10): 4657–4667, 2015. 42

Nommensen F., Go S., and MacLaren D. Half-life of HBs antibody after hepatitis B vaccination: an aid to timing of booster vaccination. *The Lancet*, 334(8667): 847–849, 1989. 58, 59

Nutt S. L., Hodgkin P. D., Tarlinton D. M., and Corcoran L. M. The generation of antibody-secreting plasma cells. *Nature Reviews Immunology*, 15(3): 160, 2015. 36

O'Connor D. and Pollard A. J. Characterizing vaccine responses using host genomic and transcriptomic analysis. *Clinical infectious diseases*, 57(6): 860–869, 2013. 41

O'Connor J. L., Smith C. J., Lampe F. C., Hill T., Gompels M., Hay P., Chadwick D., Fisher M., Ainsworth J., Gilson R., et al. Failure to achieve a CD4+ cell count response on combination antiretroviral therapy despite consistent viral load suppression. *AIDS*, 28(6): 919–924, 2014. 109

O'Dempsey T. J., McArdle T., Ceesay S. J., Secka O., Demba E., Banya W. A., Francis N., and Greenwood B. M. Meningococcal antibody titres in infants of women immunised with meningococcal polysaccharide vaccine during pregnancy. *Archives of Disease in Childhood-Fetal and Neonatal Edition*, 74(1): F43–F46, 1996. 58

Odendahl M., Mei H., Hoyer B. F., Jacobi A. M., Hansen A., Muehlinghaus G., Berek C., Hiepe F., Manz R., Radbruch A., et al. Generation of migratory antigen-specific plasma blasts and mobilization of resident plasma cells in a secondary immune response. *Blood*, 105 (4): 1614–1621, 2005. 58

Okamoto Y., Douek D. C., McFarland R. D., and Koup R. A. Effects of exogenous interleukin-7 on human thymus function. *Blood*, 99(8): 2851–2858, 2002. 110

Opportunistic Infections Project Team Of The Collaboration Of Observational HIV Epidemiological Research In Europe (COHERE) In EuroCoord et al. CD4 cell count and the risk of AIDS or death in HIV-Infected adults on combination antiretroviral therapy with a suppressed viral load: a longitudinal cohort study from COHERE. *PLOS Medicine*, 9(3): e1001194, 2012. 109

Palella Jr F. J., Delaney K. M., Moorman A. C., Loveless M. O., Fuhrer J., Satten G. A., Aschman D. J., Holmberg S. D., and Investigators H. O. S. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *New England Journal of Medicine*, 338(13): 853–860, 1998. 108

Palgen J.-L., Tchitchek N., Elhmouzi-Younes J., Delandre S., Namet I., Rosenbaum P., Dereuddre-Bosquet N., Martinon F., Cosma A., Lévy Y., et al. Prime and boost vaccination elicit a distinct innate myeloid cell immune response. *Scientific reports*, 8(1): 3087, 2018. 100

Pape K. A., Taylor J. J., Maul R. W., Gearhart P. J., and Jenkins M. K. Different B cell populations mediate early and late memory during an endogenous immune response. *Science*, 331(6021): 1203–1207, 2011. 36

Pappalardo F., Pennisi M., Castiglione F., and Motta S. Vaccine protocols optimization: in silico experiences. *Biotechnology advances*, 28(1): 82–93, 2010. 118

Pappalardo F., Russo G., Tshinanu F. M., and Viceconti M. In silico clinical trials: concepts and early adoptions. *Briefings in Bioinformatics*, 2018. 169

Pasin C., Dufour F., Villain L., Zhang H., and Thiébaut R. Controlling IL-7 Injections in HIV-Infected Patients. *Bulletin of Mathematical Biology*, 80(9): 2349–2377, 2018. 119

Patel C., Brotherton J. M., Pillsbury A., Jayasinghe S., Donovan B., Macartney K., and Marshall H. The impact of 10 years of human papillomavirus (HPV) vaccination in Australia: what additional disease burden will a nonavalent vaccine prevent? *Eurosurveillance*, 23(41): 1700737, 2018. 38

Peng S. A general stochastic maximum principle for optimal control problems. *SIAM Journal on control and optimization*, 28(4): 966–979, 1990. 121

Perelson A. S. Modelling viral and immune system dynamics. *Nature Reviews Immunology*, 2 (1): 28, 2002. 44

Perelson A. S. and Ribeiro R. M. Modeling the within-host dynamics of HIV infection. *BMC Biology*, 11(1): 96, 2013. 44

Perelson A. S., Kirschner D. E., and de Boer R. Dynamics of HIV infection of CD4+ T cells. *Mathematical Biosciences*, 114(1): 81–125, 1993. 44

Peters C. and Peters J. An introduction to Ebola: the virus and the disease. *The Journal of Infectious Diseases*, 179(Supplement_1): ix–xvi, 1999. 47

Plotkin S. History of vaccination. *Proceedings of the National Academy of Sciences*, 111(34): 12283–12287, 2014. 37

Plotkin S. A. Correlates of protection induced by vaccination. *Clinical and Vaccine Immunology*, 17(7): 1055–1065, 2010. 40, 57

Pollard A. J., Perrett K. P., and Beverley P. C. Maintaining protection against invasive bacteria with protein–polysaccharide conjugate vaccines. *Nature Reviews Immunology*, 9(3): 213, 2009. 40

Pontryagin L. S. *Mathematical theory of optimal processes*. Routledge, 1987. 117

Pool V., Tomovici A., Johnson D. R., Greenberg D. P., and Decker M. D. Humoral immunity 10 years after booster immunization with an adolescent and adult formulation combined tetanus, diphtheria, and 5-component acellular pertussis vaccine in the USA. *Vaccine*, 36(17): 2282–2287, 2018. 58

Post T. M., Freijer J. I., Ploeger B. A., and Danhof M. Extensions to the visual predictive check to facilitate model performance evaluation. *Journal of Pharmacokinetics and Pharmacodynamics*, 35(2): 185, 2008. 71

Prague M., Commenges D., Drylewicz J., and Thiébaut R. Treatment Monitoring of HIV-Infected Patients based on Mechanistic Models. *Biometrics*, 68(3): 902–911, 2012. 71

Prague M., Commenges D., Guedj J., Drylewicz J., and Thiébaut R. NIMROD : A program for inference via a normal approximation of the posterior in models with random. *Computer Methods and Programs in Biomedicine*, 111(2): 447–458, 2013a. 65, 163

Prague M., Commenges D., and Thiébaut R. Dynamical models of biomarkers and clinical progression for personalized medicine: The HIV context. *Advanced Drug Delivery Reviews*, 65 (7): 954–965, 2013b. 44

Prague M., Pasin C., Wittkop L., Duffau P., Lazaro E., Cazanave C., Vareil M., Bonnet F., and Thiébaut R. In silico clinical trials for evaluation of HIV short-cycles strategies. Poster presented at: Conference on Retroviruses and Opportunistic Infections (CROI), 2018. 169

PREVAIL II Writing Group. A randomized, controlled trial of ZMapp for Ebola virus infection. *New England Journal of Medicine*, 375(15): 1448–1456, 2016. 51

Priddy F. H., Brown D., Kublin J., Monahan K., Wright D. P., Lalezari J., Santiago S., Marmor M., Lally M., Novak R. M., et al. Safety and immunogenicity of a replication-incompetent adenovirus type 5 HIV-1 clade B gag/pol/nef vaccine in healthy adults. *Clinical Infectious Diseases*, 46(11): 1769–1781, 2008. 38

Prokopiou S. A., Barbarroux L., Bernard S., Mafille J., Leverrier Y., Arpin C., Marvel J., Gandrillon O., and Crauste F. Multiscale modeling of the early CD8 T-cell immune response in lymph nodes: an integrative study. *Computation*, 2(4): 159–181, 2014. 43

Pulendran B. Systems vaccinology: probing humanity's diverse immune systems with vaccines. *Proceedings of the National Academy of Sciences*, 111(34): 12300–12306, 2014. 41, 100

Pulendran B. and Ahmed R. Translating innate immunity into immunological memory: implications for vaccine development. *Cell*, 124(4): 849–863, 2006. 100

Pulendran B. and Ahmed R. Immunological mechanisms of vaccination. *Nature Immunology*, 12 (6): 509, 2011. 39

Pulendran B., Li S., and Nakaya H. I. Systems vaccinology. *Immunity*, 33(4): 516–529, 2010. 99

Puszynski K., Gandolfi A., and d'Onofrio A. The role of stochastic gene switching in determining the pharmacodynamics of certain drugs: basic mechanisms. *Journal of Pharmacokinetics and Pharmacodynamics*, 43(4): 395–410, 2016. 119

Querec T. D., Akondy R. S., Lee E. K., Cao W., Nakaya H. I., Teuwen D., Pirani A., Gernert K., Deng J., Marzolf B., et al. Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. *Nature Immunology*, 10(1): 116, 2009. 100

Radbruch A., Muehlinghaus G., Luger E. O., Inamine A., Smith K. G., Dörner T., and Hiepe F. Competence and competition: the challenge of becoming a long-lived plasma cell. *Nature Reviews Immunology*, 6(10): 741, 2006. 58

Rampling T., Ewer K., Bowyer G., Wright D., Imoukhuede E. B., Payne R., Hartnell F., Gibani M., Bliss C., Minhinnick A., et al. A monovalent chimpanzee adenovirus Ebola vaccine — preliminary report. *New England Journal of Medicine*, pages 150202093719007–150202093719007, 2015. 52, 54

Ramshaw I. A. and Ramsay A. J. The prime-boost strategy: exciting prospects for improved vaccination. *Immunology today*, 21(4): 163–165, 2000. 40

Rappuoli R., Mandl C. W., Black S., and De Gregorio E. Vaccines for the twenty-first century society. *Nature Reviews Immunology*, 11(12): 865, 2011. 37

Rappuoli R., Siena E., and Finco O. Will Systems Biology Deliver Its Promise and Contribute to the Development of New or Improved Vaccines? Systems Biology Views of Vaccine Innate and Adaptive Immunity. *Cold Spring Harbor perspectives in biology*, page a029256, 2017. 102

Rechtien A., Richert L., Lorenzo H., Martrus G., Hejblum B., Dahlke C., Kasonta R., Zinser M., Stubbe H., Matschl U., et al. Systems vaccinology identifies an early innate immune signature as a correlate of antibody responses to the Ebola vaccine rVSV-ZEBOV. *Cell Reports*, 20(9): 2251–2261, 2017. 100

Regules J. A., Beigel J. H., Paolino K. M., Voell J., Castellano A. R., Hu Z., Muñoz P., Moon J. E., Ruck R. C., Bennett J. W., et al. A recombinant vesicular stomatitis virus Ebola vaccine. *New England Journal of Medicine*, 376(4): 330–341, 2017. 53, 55

Renard D., Bruckers L., Molenberghs G., Vellinga A., and Van Damme P. Repeated-measures models to evaluate a hepatitis B vaccination programme. *Statistics in Medicine*, 20(6): 951–963, 2001. 60

Rerks-Ngarm S., Pitisuttithum P., Nitayaphan S., Kaewkungwal J., Chiu J., Paris R., Premsri N., Namwat C., de Souza M., Adams E., et al. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *New England Journal of Medicine*, 361(23): 2209–2220, 2009. 40

Ribeiro R. M., Mohri H., Ho D. D., and Perelson A. S. In vivo dynamics of T cell activation, proliferation, and death in HIV-1 infection: why are CD4+ but not CD8+ T cells depleted? *Proceedings of the National Academy of Sciences*, 99(24): 15572–15577, 2002. 65

Rostami-Hodjegan A. and Tucker G. T. Simulation and prediction of in vivo drug metabolism in human populations from in vitro data. *Nature Reviews Drug Discovery*, 6(2): 140, 2007. 168

Rougeron V., Feldmann H., Grard G., Becker S., and Leroy E. Ebola and Marburg haemorrhagic fever. *Journal of Clinical Virology*, 64: 111–119, 2015. 48

Sallusto F., Lanzavecchia A., Araki K., and Ahmed R. From vaccines to memory and back. *Immunity*, 33(4): 451–463, 2010. 40

Samson A., Lavielle M., and Mentré F. The SAEM algorithm for group comparison tests in longitudinal data analysis based on non-linear mixed-effects model. *Statistics in Medicine*, 26 (27): 4860–4875, 2007. 73

Sarwar U. N., Costner P., Enama M. E., Berkowitz N., Hu Z., Hendel C. S., Sitar S., Plummer S., Mulangu S., Bailer R. T., et al. Safety and immunogenicity of DNA vaccines encoding Ebolavirus and Marburgvirus wild-type glycoproteins in a phase I clinical trial. *The Journal of Infectious Diseases*, 211(4): 549–557, 2014. 52

Savy M., Edmond K., Fine P. E., Hall A., Hennig B. J., Moore S. E., Mulholland K., Schaible U., and Prentice A. M. Landscape analysis of interactions between nutrition and vaccine responses in children. *The Journal of Nutrition*, 139(11): 2154S–2218S, 2009. 41

Seddon B., Tomlinson P., and Zamoyska R. Interleukin 7 and T cell receptor signals regulate homeostasis of CD4 memory cells. *Nature Immunology*, 4(7): 680, 2003. 34

Seder R. A. and Ahmed R. Similarities and differences in CD4+ and CD8+ effector and memory T cell generation. *Nature Immunology*, 4(9): 835, 2003. 36

Seiden P. E. and Celada F. A model for simulating cognate recognition and response in the immune system. *Journal of Theoretical Biology*, 158(3): 329–357, 1992. 43

Seidman J. C., Richard S. A., Viboud C., and Miller M. A. Quantitative review of antibody response to inactivated seasonal influenza vaccines. *Influenza and other respiratory viruses*, 6 (1): 52–62, 2012. 41

Sereti I., Dunham R. M., Spritzler J., Aga E., Proschan M. A., Medvik K., Battaglia C. A., Landay A. L., Pahwa S., Fischl M. A., et al. IL-7 administration drives T cell–cycle entry and expansion in HIV-1 infection. *Blood*, 113(25): 6304–6314, 2009. 110

Sheets R. L., Stein J., Bailer R. T., Koup R. A., Andrews C., Nason M., He B., Koo E., Trotter H., Duffy C., et al. Biodistribution and toxicological safety of adenovirus type 5 and type 35 vectored vaccines against human immunodeficiency virus-1 (HIV-1), Ebola, or Marburg are similar despite differing adenovirus serotype vector, manufacturer's construct, or gene inserts. *Journal of Immunotoxicology*, 5(3): 315–335, 2008. 98

Shen Y., Meng Q., and Shi P. Maximum principle for mean-field jump–diffusion stochastic delay differential equations and its application to finance. *Automatica*, 50(6): 1565–1579, 2014. 121

Shen-Orr S. S. and Furman D. Variability in the immune system: of vaccine responses and immune states. *Current Opinion in Immunology*, 25(4): 542–547, 2013. 41

Sissoko D., Laouenan C., Folkesson E., M'lebing A.-B., Beavogui A.-H., Baize S., Camara A.-M., Maes P., Shepherd S., Danel C., et al. Experimental treatment with favipiravir for Ebola virus disease (the JIKI Trial): a historically controlled, single-arm proof-of-concept trial in Guinea. *PLOS Medicine*, 13(3): e1001967, 2016. 51

Slifka M. K., Antia R., Whitmire J. K., and Ahmed R. Humoral immunity due to long-lived plasma cells. *Immunity*, 8(3): 363–372, 1998. 58

Sridhar S., Luedtke A., Langevin E., Zhu M., Bonaparte M., Machabert T., Savarino S., Zambrano B., Moureau A., Khromava A., et al. Effect of dengue serostatus on dengue vaccine safety and efficacy. *New England Journal of Medicine*, 2018. 41

Stanley D. A., Honko A. N., Asiedu C., Trefry J. C., Lau-Kilby A. W., Johnson J. C., Hensley L., Ammendola V., Abbate A., Grazioli F., et al. Chimpanzee adenovirus vaccine generates acute and durable protective immunity against ebolavirus challenge. *Nature Medicine*, 20(10): 1126, 2014. 56

Stengel R. F. Mutation and control of the human immunodeficiency virus. *Mathematical Biosciences*, 213(2): 93–102, 2008. 118

Stephenson K. E., T D'Couto H., and Barouch D. H. New concepts in HIV-1 vaccine development. *Current Opinion in Immunology*, 41: 39–46, 2016. 39

Sullivan N. J., Martin J. E., Graham B. S., and Nabel G. J. Correlates of protective immunity for Ebola vaccines: implications for regulatory approval by the animal rule. *Nature Reviews Microbiology*, 7(5): 393, 2009. 56

Sullivan N. J., Hensley L., Asiedu C., Geisbert T. W., Stanley D., Johnson J., Honko A., Olinger G., Bailey M., Geisbert J. B., et al. CD8+ cellular immunity mediates rAd5 vaccine protection against Ebola virus infection of nonhuman primates. *Nature Medicine*, 17(9): 1128, 2011. 56

Tan J. T., Dudl E., LeRoy E., Murray R., Sprent J., Weinberg K. I., and Surh C. D. IL-7 is critical for homeostatic proliferation and survival of naive T cells. *Proceedings of the National Academy of Sciences*, 98(15): 8732–8737, 2001. 110

Tapia M. D., Sow S. O., Lyke K. E., Haidara F. C., Diallo F., Doumbia M., Traore A., Coulibaly F., Kodio M., Onwuchekwa U., et al. Use of ChAd3-EBO-Z Ebola virus vaccine in Malian and US adults, and boosting of Malian adults with MVA-BN-Filo: a phase 1, single-blind, randomised trial, a phase 1b, open-label and double-blind, dose-escalation trial, and a nested, randomised, double-blind, placebo-controlled trial. *The Lancet infectious diseases*, 16 (1): 31–42, 2016. 40, 52, 54

Tarlinton D. and Good-Jacobson K. Diversity among memory B cells: origin, consequences, and utility. *Science*, 341(6151): 1205–1211, 2013. 58

Taylor J. J., Jenkins M. K., and Pape K. A. Heterogeneity in the differentiation and function of memory B cells. *Trends in Immunology*, 33(12): 590–597, 2012. 37

Theeten H., Van Herck K., Van Der Meeren O., Crasta P., Van Damme P., and Hens N. Long-term antibody persistence after vaccination with a 2-dose Havrix™(inactivated hepatitis A vaccine): 20 years of observed data, and long-term model-based predictions. *Vaccine*, 33(42): 5723–5727, 2015. 60

Thiébaut R., Lewden C., and Chêne G. Epidemiology of discordant virologic and immunologic responses in HIV-1 infected patients. *Medecine et maladies infectieuses*, 35: S31–2, 2005. 109

Thiébaut R. and Walker S. Commentary When it is better to estimate a slope with only one point. *QJM Quaterly Journal of Medicine*, 101(10): 821–4, 2008. 65

Thiébaut R., Drylewicz J., Prague M., Lacabaratz C., Beq S., Jarne A., Croughs T., Sekaly R.-P., Lederman M. M., Sereti I., et al. Quantifying and predicting the effect of exogenous interleukin-7 on CD4+ T cells in HIV-1 infection. *PLOS Computational Biology*, 10(5): e1003630, 2014. 110, 111, 120, 161

Thiébaut R., Jarne A., Routy J.-P., Sereti I., Fischl M., Ive P., Speck R. F., D'offizi G., Casari S., Commenges D., et al. Repeated cycles of recombinant human interleukin 7 in HIV-infected patients with low CD4 T-cell reconstitution on antiretroviral therapy: results of 2 phase II multicenter studies. *Clinical Infectious Diseases*, 62(9): 1178–1185, 2016. 113, 114, 120

Thorson A., Formenty P., Lofthouse C., and Broutet N. Systematic review of the literature on viral persistence and sexual transmission from recovered ebola survivors: evidence and recommendations. *BMJ open*, 6(1): e008859, 2016. 49

UNAIDS. Unaids data 2017, 2017. 105

Ura T., Okuda K., and Shimada M. Developments in viral vector-based vaccines. *Vaccines*, 2 (3): 624–641, 2014. 38

Van Damme P., Thoelen S., Cramm M., De Groote K., Safary A., and Meheus A. Inactivated hepatitis A vaccine: reactogenicity, immunogenicity, and long-term antibody persistence. *Journal of Medical Virology*, 44(4): 446–451, 1994. 59

Van der Vaart A. W. Asymptotic statistics (Cambridge series in statistical and probabilistic mathematics). 2000. 70

van Leeuwen E. M., Sprent J., and Surh C. D. Generation and maintenance of memory CD4+ T cells. *Current Opinion in Immunology*, 21(2): 167–172, 2009. 36

Van Loveren H., Van Amsterdam J., Vandebriel R. J., Kimman T. G., Rümke H. C., Steerenberg P. S., and Vos J. G. Vaccine-induced antibody responses as parameters of the influence of endogenous and environmental factors. *Environmental health perspectives*, 109(8): 757, 2001. 41

Vella A. T., Dow S., Potter T. A., Kappler J., and Marrack P. Cytokine-induced survival of activated T cells in vitro and in vivo. *Proceedings of the National Academy of Sciences*, 95(7): 3810–3815, 1998. 110

Venkatraman N., Silman D., Folegatti P. M., and Hill A. V. Vaccines against Ebola virus. *Vaccine*, 36(36): 5454–5459, 2018. 51

Viceconti M., Henney A., and Morley-Fletcher E. In silico clinical trials: how computer simulation will transform the biomedical industry. *International Journal of Clinical Trials*, 3(2): 37–46, 2016. 168

Vidor E. Evaluation of the persistence of vaccine-induced protection with human vaccines. *Journal of comparative pathology*, 142: S96–S101, 2010. 60

Vilajeliu A., Ferrer L., Munrós J., Goncé A., López M., Costa J., Bayas J. M., Group P. W., et al. Pertussis vaccination during pregnancy: Antibody persistence in infants. *Vaccine*, 34 (33): 3719–3722, 2016. 58

Vodovotz Y., Xia A., Read E. L., Bassaganya-Riera J., Hafler D. A., Sontag E., Wang J., Tsang J. S., Day J. D., Kleinstein S. H., et al. Solving immunology? *Trends in Immunology*, 38(2): 116–127, 2017. 43, 171

Volberding P. A., Lagakos S. W., Koch M. A., Pettinelli C., Myers M. W., Booth D. K., Balfour Jr H. H., Reichman R. C., Bartlett J. A., Hirsch M. S., et al. Zidovudine in asymptomatic human immunodeficiency virus infection: a controlled trial in persons with fewer than 500 CD4-positive cells per cubic millimeter. *New England Journal of Medicine*, 322(14): 941–949, 1990. 108

Voysey M., Pollard A. J., Sadarangani M., and Fanshawe T. R. Prevalence and decay of maternal pneumococcal and meningococcal antibodies: A meta-analysis of type-specific decay rates. *Vaccine*, 35(43): 5850–5857, 2017. 58

Wang Y., Li J., Hu Y., Liang Q., Wei M., and Zhu F. Ebola vaccines in clinical trial: the promising candidates. *Human vaccines & immunotherapeutics*, 13(1): 153–168, 2017. 51

Wherry E. J. and Ahmed R. Memory CD8 T-cell differentiation during viral infection. *Journal of Virology*, 78(11): 5535–5545, 2004. 40

White M. T., Verity R., Griffin J. T., Asante K. P., Owusu-Agyei S., Greenwood B., Drakeley C., Gesase S., Lusingu J., Ansong D., et al. Immunogenicity of the RTS, S/AS01 malaria vaccine and implications for duration of vaccine efficacy: secondary analysis of data from a phase 3 randomised controlled trial. *The Lancet infectious diseases*, 15(12): 1450–1458, 2015. 60

WHO. Ebola virus disease, 2018a. `http://www.who.int/en/news-room/fact-sheets/detail/ebola-virus-disease`. 49, 50

WHO. Ebola virus disease. Democratic Republic of Congo. External situation report 17, 2018b. `http://apps.who.int/iris/bitstream/handle/10665/273348/SITREP_EVD_DRC_20180725-eng.pdf`. 50

WHO. Ebola virus disease. Democratic Republic of Congo. External situation report 3, 2018c. `http://apps.who.int/iris/bitstream/handle/10665/274258/SITREP_EVD_DRC_20180820-eng.pdf`. 50

WHO. Categorization and prioritization of drugs for consideration for testing or use in patients infected with Ebola, 2015. http://www.who.int/medicines/ebola-treatment/2015_0703TablesofEbolaDrugs.pdf?ua=1. 51

WHO. 10 facts on immunization, 2018d. http://www.who.int/features/factfiles/immunization/en/. 37

WHO. Measles: key facts, 2018e. http://www.who.int/en/news-room/fact-sheets/detail/measles. 38

WHO. Correlates of vaccine-induced protection: methods and implications, 2013. http://apps.who.int/iris/bitstream/handle/10665/84288/WHO_IVB_13.01_eng.pdf?sequence=1. 56

WHO. Infection prevention and control guidance for care of patients in health-care settings, with focus on Ebola, interim guidance, 2016a. http://apps.who.int/iris/bitstream/handle/10665/131828/WHO_EVD_Guidance_IPC_14.1_eng.pdf?sequence=1. 50

WHO. Situation report. Ebola virus disease. 10 June 2016. Geneva, Switzerland, 2016b. http://apps.who.int/iris/bitstream/handle/10665/208883/ebolasitrep_10Jun2016_eng.pdf?sequence=1. 49

WHO Ebola Response Team. Ebola virus disease in West Africa — the first 9 months of the epidemic and forward projections. *New England Journal of Medicine*, 371(16): 1481–1495, 2014. 50

Wiedermann G., Kundi M., Ambrosch F., Safary A., D'hondt E., and Delem A. Inactivated hepatitis A vaccine: long-term antibody persistence. *Vaccine*, 15(6-7): 612–615, 1997. 60

Wiens B. L., Bohidar N. R., Pigeon J. G., Egan J., Hurni W., Brown L., Kuter B. J., and Nalin D. R. Duration of protection from clinical hepatitis A disease after vaccination with VAQTA®. *Journal of Medical Virology*, 49(3): 235–241, 1996. 59

Williams W. B., Han Q., and Haynes B. F. Cross-reactivity of HIV vaccine responses and the microbiome. *Current Opinion in HIV and AIDS*, 13(1): 9–14, 2018. 42

Wilson J. N., Nokes D. J., Medley G. F., and Shouval D. Mathematical model of the antibody response to hepatitis B vaccines: implications for reduced schedules. *Vaccine*, 25(18): 3705–3712, 2007. 61

Winkelmann S., Schütte C., and von Kleist M. Markov control processes with rare state observation: Theory and application to treatment scheduling in HIV-1. *Communications in Mathematical Sciences*, 12(5): 859–877, 2014. 119

Winslow R. L., Milligan I. D., Voysey M., Luhn K., Shukarev G., Douoguih M., and Snape M. D. Immune Responses to Novel Adenovirus Type 26 and Modified Vaccinia Virus Ankara–Vectored Ebola Vaccines at 1 Year. *JAMA*, 317(10): 1075–1077, 2017. 52, 55

Wong G., Richardson J. S., Pillet S., Patel A., Qiu X., Alimonti J., Hogan J., Zhang Y., Takada A., Feldmann H., et al. Immune parameters correlate with protection against Ebola virus infection in rodents and nonhuman primates. *Science translational medicine*, 4(158): 158ra146–158ra146, 2012. 56

Wu S., Liu Z.-P., Qiu X., and Wu H. Modeling genome-wide dynamic regulatory network in mouse lungs with influenza infection using high-dimensional ordinary differential equations. *PLOS ONE*, 9(5): e95276, 2014. 102

Yang Y., Xiao Y., and Wu J. Pulse HIV vaccination: feasibility for virus eradication and optimal vaccination schedule. *Bulletin of Mathematical Biology*, 75(5): 725–751, 2013. 118

Yates A., Chan C. C., Callard R. E., George A. J., and Stark J. An approach to modelling in immunology. *Briefings in Bioinformatics*, 2(3): 245–257, 2001. 44

Yong J. and Zhou X. Y. *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43. Springer Science & Business Media, 1999. 119

Yotebieng M., Maskew M., and Van Rie A. CD4 gain percentile curves for monitoring response to antiretroviral therapy in HIV-infected adults. *AIDS*, 29(9): 1067, 2015. 109

Young J., Psichogiou M., Meyer L., Ayayi S., Grabar S., Raffi F., Reiss P., Gazzard B., Sharland M., Gutierrez F., et al. CD4 cell count and the risk of AIDS or death in HIV-Infected adults on combination antiretroviral therapy with a suppressed viral load: a longitudinal cohort study from COHERE. *PLOS Medicine*, page e1001194, 2012. 108

Yu J., Smith V. A., Wang P. P., Hartemink A. J., and Jarvis E. D. Advances to bayesian network inference for generating causal networks from observational biological data. *Bioinformatics*, 20(18): 3594–3603, 2004. 102

Zeiser S., Franz U., and Liebscher V. Autocatalytic genetic networks modeled by piecewise-deterministic Markov processes. *Journal of Mathematical Biology*, 60(2): 207, 2010. 164

Zhu F.-C., Hou L.-H., Li J.-X., Wu S.-P., Liu P., Zhang G.-R., Hu Y.-M., Meng F.-Y., Xu J.-J., Tang R., et al. Safety and immunogenicity of a novel recombinant adenovirus type-5 vector-based Ebola vaccine in healthy adults in China: preliminary report of a randomised, double-blind, placebo-controlled, phase 1 trial. *The Lancet*, 385(9984): 2272–2279, 2015. 52

Zhu F.-C., Wurie A. H., Hou L.-H., Liang Q., Li Y.-H., Russell J. B., Wu S.-P., Li J.-X., Hu Y.-M., Guo Q., et al. Safety and immunogenicity of a recombinant adenovirus type-5 vector-based Ebola vaccine in healthy adults in Sierra Leone: a single-centre, randomised, double-blind, placebo-controlled, phase 2 trial. *The Lancet*, 389(10069): 621–628, 2017. 52, 54

Appendix A:

Article.

"Ebola vaccine development: Systematic review of preclinical and clinical studies, and meta-analysis of determinants of antibody response variability after vaccination"

International Journal of Infectious Diseases 74 (2018) 83–96

Contents lists available at ScienceDirect

International Journal of Infectious Diseases

journal homepage: www.elsevier.com/locate/ijid

Ebola vaccine development: Systematic review of pre-clinical and clinical studies, and meta-analysis of determinants of antibody response variability after vaccination

Lise Gross[b,c], Edouard Lhomme[a,b,c,d], Chloé Pasin[a,b,c], Laura Richert[a,b,c,d], Rodolphe Thiebaut[a,b,c,d,*]

[a] INSERM, Bordeaux Population Health Research Centre, UMR 1219, Univ. Bordeaux, ISPED, F-33000, Bordeaux, France
[b] SISTM Team (Statistics in System Biology and Translational Medicine), INRIA Research Centre, Bordeaux, F-33000, France
[c] Vaccine Research Institute (VRI), Créteil, F-94000, France
[d] Pôle de Santé Publique, CHU de Bordeaux, Bordeaux, F-33000, France

ARTICLE INFO

Article history:
Received 10 April 2018
Received in revised form 20 June 2018
Accepted 28 June 2018
Corresponding Editor: Eskild Petersen, Aarhus, Denmark

Keywords:
Ebola
Vaccine
Review
Meta-analysis
Humans
Nonhuman primates

ABSTRACT

Objectives: For Ebola vaccine development, antibody response is a major endpoint although its determinants are not well known. We aimed to review Ebola vaccine studies and to assess factors associated with antibody response variability in humans.
Methods: We searched PubMed and Scopus for preventive Ebola vaccine studies in humans or non-human primates (NHP), published up to February 2018. For each vaccination group with Ebola Zaire antibody titre measurements after vaccination, data about antibody response and its potential determinants were extracted. A random-effects meta-regression was conducted including human groups with at least 8 individuals.
Results: We reviewed 49 studies (202 vaccination groups including 74 human groups) with various vaccine platforms and antigen inserts. Mean antibody titre was slightly higher in NHP (3.10, 95% confidence interval [293; 327]) than in humans (2.75 [257; 293]). Vaccine platform ($p < 0{\cdot}001$) and viral strain used for antibody detection ($p < 0{\cdot}001$) were associated with antibody response in humans, but adjusted heterogeneity remained at 95%.
Conclusions: Various platforms have been evaluated in humans, including Ad26, Ad5, ChimpAd3, DNA, MVA, and VSV. In addition to platforms, viral strain used for antibody detection influences antibody response. However, variability remained mostly unexplained. Therefore, comparison of vaccine immunogenicity needs randomised controlled trials.

Introduction

Following the deadly 2013-2016 epidemic in West Africa, there has been an accelerated development of several candidates for an Ebola preventive vaccine. Outbreaks of Ebola virus disease (EVD) have occurred recurrently and unpredictably for the past 40 years with a high lethality rate (Liu et al., 2015). The 2013-2015 outbreak was unprecedented in scale, with over 28,000 cases and more than 11,000 deaths (Ebola Situation Report, 2016). Incidental cases are still reported as recently in the Democratic Republic of Congo in May 2017 (Dhama et al., 2015). In the absence of any specific treatment, EVD prevention and control measures are primarily based on case identification and isolation, early non-specific medical care, surveillance of suspect cases, and safe burial practices (Henao-Restrepo et al., 2017). These measures are now sometimes complemented by ring vaccination of contacts of cases, based on the promising results of a phase III cluster-randomized ring vaccination efficacy trial conducted in Guinea in 2015 (Ohimain, 2016). However, the vaccine used for ring vaccination (rVSV ZEBOV vaccine) is not yet licenced and conducting new efficacy trials for licencing is not feasible in the absence of a large outbreak. Nevertheless, preparation for future outbreaks is required and the licensing of one or several preventive vaccines for stockpiling is a priority.

Several candidate vaccines strategies have been investigated since the first reported EVD outbreak in 1976. During and following

* Corresponding author at: INSERM U1219, INRIA SISTM, ISPED, Bordeaux University, 146 Rue Leo Saignat, 33076, Bordeaux Cedex, France.
E-mail address: Rodolphe.Thiebaut@u-bordeaux.fr (R. Thiebaut).

https://doi.org/10.1016/j.ijid.2018.06.022

the 2013-2015 epidemic, the process of vaccine development has been substantially accelerated, and several strategies have been moved into clinical phases. Despite the promising results of the ring vaccination trial in Guinea (Ohimain, 2016), many questions, such as durability of immune responses, and immune responses and protection in specific sub-groups such as young children, remain to be addressed and Ebola vaccine development continues to be very active. Based on their delivery technologies, several candidate vaccine platforms can be distinguished: whole-virus vaccines, DNA vaccines, virus-like particles vaccines, and recombinant vaccines with different viral vectors (vesicular stomatitis virus or VSV, modified vaccinia Ankara or MVA, human adenovirus or Ad, and chimpanzee adenovirus or ChAd) (World Health Organisation, 2013). Each platform may use specific dose levels and Ebola antigen inserts.

Vaccine trials aim to assess vaccine safety and immunogenicity in phase I and II trials in humans prior to testing for a protective effect in phase III. Assessment of vaccine efficacy during pre-clinical and clinical studies is required to go through the vaccine license steps. Clinical protection from EVD in human populations is impossible to observe outside an epidemic period. In the non-epidemic context, Ebola vaccines are thus currently evaluated by using a main immunogenicity endpoint: the antibody response after vaccination. There is no definite evidence that antibody response is the correlate of protection or surrogate endpoint for efficacy in humans, that is a specific immune response to vaccine associated with vaccine-induced protection (Sullivan et al., 2009) and it may vary according to the vaccine platforms (Sullivan et al., 2000a,b). However, we know that antibody response is correlated with survival after challenge in nonhuman primate models, which is the nearest model to humans for EVD and hence the animal gold standard to test candidate Ebola vaccines; this association is found consistently for different Ebola candidate vaccines (Wong et al., 2012; Food and Drug Administration, 2015; Sridhar, 2015).

For these reasons, antibody response is used as the main criterion to assess the Ebola candidate vaccines in phase I/II trials. In the absence of the possibility to conduct additional phase III trials, regulatory pathways not requiring such efficacy results are also under discussion (Food and Drug Administration, 2015). Significant variations in antibody responses are observable across studies, which could be due to the different types of vaccines evaluated, or not. Various factors are suspected to influence the level of antibody response beyond the vaccine features (vaccine platform, Ebola viral insert, dosage, single injection or boost, ...) such as the measurement techniques (time of measurement, antigen used to detect antibody response, ...) or the population type (human or nonhuman primates, age, sex, study site, ...). There is a lack of quantification of the contribution of each factor in the observed variation of the reported antibody responses.

Although previous reviews exist on Ebola vaccines (Ohimain, 2016; Sridhar, 2015; Wu et al., 2015), the specific topic of antibody response determinants has not yet been addressed by a systematic review or meta-analysis. Yet, the identification of factors potentially associated with antibody response after Ebola vaccination could provide relevant information for further vaccine trials and for regulatory decision making.

By conducting this systematic review with a meta-analysis, we aimed to determine whether the reported antibody response variability in Ebola vaccine trials is not only determined by the vaccine platform but also by other characteristics of vaccine and by population and measurement characteristics and to quantify these factors.

Methods

Search strategy and selection criteria

Studies were identified by searching electronic databases PubMed and Scopus. Pubmed was searched using the following terms: (« hemorrhagic fever, ebola » [MeSH Terms] OR « ebola » [All fields] OR « ebolavirus » [MeSH Terms] OR « ebolavirus » [All fields]) AND (« vaccines » [MeSH Terms] OR « vaccines » [All fields] OR « vaccine » [All Fields]). Scopus was searched using the following terms TITLE-ABS-KEY (ebola) AND TITLE-ABS-KEY (vaccine). Additionally, the Clinicaltrials.gov website was searched to identify unpublished and ongoing studies. Several experts in the field were contacted to find papers which could be not indexed in databases. Reference lists of relevant papers and reviews were examined to identify further articles.

The search was performed on March 23, 2016 and updated as of February 24, 2018 with a publication date limit of the same date in order to identify all published studies which met the inclusion criteria and without restriction on language. All preventive Ebola vaccine clinical trials conducted in humans or in nonhuman primates and with a measure of Ebola Zaire antibody titre after vaccination were included in our systematic review. Studies were excluded in case of duplicate study, studies without original data, preclinical studies conducted in animals other than nonhuman primates or in vitro experimentation.

Data extraction

A first step of selection was performed on the title and abstract, and then a second step was performed after reading the full article. Two authors independently assessed each full article to include papers matching the review's inclusion criteria. Disagreements between reviewers were resolved by consensus.

Data were extracted by two independent reviewers, with differences reconciled by consensus. The following variables were extracted: paper identification (title, first author, publication year), study design, inclusion and exclusion criteria, characteristics of the population (number of subjects; human or nonhuman primates; proportion of women, average age and study site for clinical trials; and animal species for pre-clinical studies using nonhuman primates), characteristics of vaccine (vaccine platform in terms of delivery technology used, specific vector for recombinant vaccines, Ebola viral insert, dosage, route of administration, vaccination schedule), characteristics of measurement techniques (time interval between last injection and measure, strain and nature of antigen used to detect antibody response, measurement method), antibody response after vaccination (geometric mean titre and its variance). Regarding the antibody response after vaccination, geometric mean titre was extracted from the text or estimated from figures. If a single vaccination group had more than one measure of antibody response, data from measurement after each injection were extracted. Therefore, if available, measurement post-prime and measurement post-boost from a same vaccination group were both included in our meta-analysis. If several measurements post-prime or if several measurements post-boost were available, for each injection we extracted the one closest to 28 days after injection, which is a standard time point in Ebola vaccine trials. Variance of titre (within-group variance) was extracted directly from the text or calculated from confidence interval or from individual values. The present study was registered in PROSPERO (no. 54303).

Data analysis

For all analyses, the statistical unit used was the vaccination group (one or several groups for a single study), i.e. a protocol-defined group undergoing the same intervention and follow-up procedures (such as a randomized arm of a clinical trial or an animal group in NHP studies).

First descriptive analyses were performed among all groups, separately for nonhuman primates and for humans. Then, a random-effect meta-regression analysis was performed including only human groups with 8 individuals or more. This threshold allowed both to have sufficient inter-individual variability in each group and to avoid excluding too many groups. Thus, it was not possible to perform the regression analysis with NHP studies because of the usual small sample size of the groups. The effect of every potential determinant of antibody response was assessed through fixed effects. A random intercept was allowed to capture between-group variability not explained by the fixed effects. The residual variance (within-group variance) was fixed in the model according to the values resulting from data extraction as described by Van Houwelingen (Van Houwelingen et al., 2002).

Each potential determinant associated in unadjusted analyses with a p-value <0.25 was included in the multivariable model using forward step-wise selection. The heterogeneity was checked visually with forest plots and quantified by using the Q test. The proportion of total variation across groups due to heterogeneity (I^2) and the amount of variability explained by the factors included in the random-effect model (R^2) were estimated. Antibody titres after vaccination were log transformed in the model.

For the meta-regression analysis, the dosage variable was categorized into "low dose" or "high dose" per vaccine platform, since units of measurement for dose level were platform-dependent. For each unit of dose measurement and each vaccine platform, the average dose level among the human groups included in the meta-regression model was used as a classification threshold for this variable; if only one dose level was assessed for a vaccine platform, the dosage variable was defined as undifferentiated. The absence of interaction between vaccine platform and dosage was checked (likelihood ratio test: p = 0.223).

All analyses were performed using the metafor package of R (i386 3.2.2 version, the R Foundation, Vienna, Austria).

Results

Study selection

The selection process of the studies and vaccination groups is described in Figure 1.

The search yielded a total of 2166 studies. Of these, 49 met the inclusion criteria to the research question corresponding to 202 vaccination groups. Unpublished clinical trials and one trial found by contact with an expert were excluded since no results were available. Studies not reporting any antibody measurements were also excluded. This led to the exclusion of the "Ebola ça suffit" ring vaccination trial conducted in Guinea, the only trial that was able to assess clinical efficacy in humans so far (Ohimain et al., 2016). This trial was conducted under emergency conditions and did not collect blood samples for immunogenicity measurements.

Table 1 shows details of all trials included in the systematic review: 32 studies were conducted in NHP, 13 trials in humans were phase 1, two trials phase 1/2, and two phase 2. The number of trials has increased significantly since the last outbreak of EVD. Clinical trials were conducted mostly in Europe and North America (Figure 2).

Description of included vaccination groups

Among the 202 vaccination groups included in our systematic review, 74 were human groups and 128 were non-human primate groups. The distribution of the number of individuals by groups is

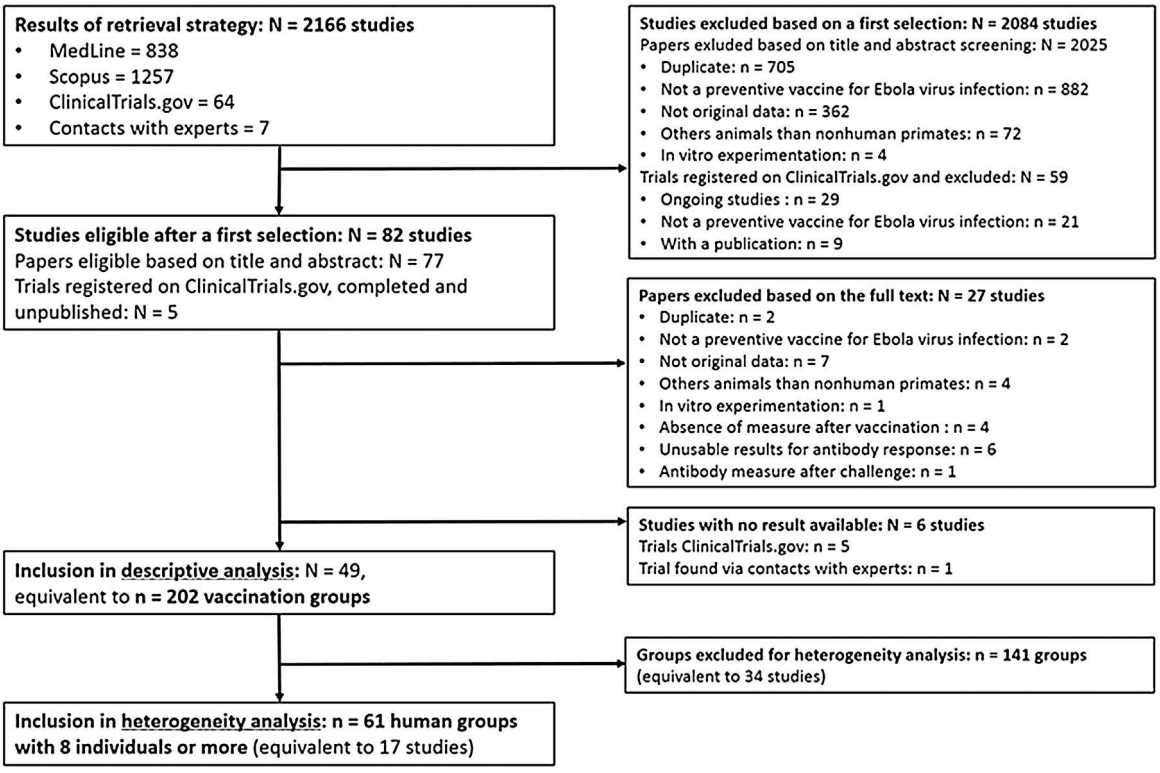


Figure 1. Flow chart for study/vaccination group selection.

Table 1
Main characteristics of the preclinical studies and clinical trials included in the systematic review.

Title	First author	Year of publication	Population and study features	Vaccine(s)	Measurement of antibody response
Phase 1 Trials of rVSV Ebola Vaccine in Africa and Europe (Agnandji et al., 2016)	Agnandji	2016	Humans (Germany, Switzerland, Gabon, Kenya), phase 1, randomization and placebo	Recombinant VSV-GP(Zaire), single injection, IM, 300 000 to 50 million PFU	Antibodies anti GP (Kikwit), D28 or D180
Successful topical respiratory tract immunization of primates against Ebola virus (Bukreyev et al., 2007)	Bukreyev	2007	NHP: rhesus monkeys; placebo	Recombinant HPIV3(+/− modified)-GP +/− NP (Zaire Mayinga), single injection +/− boost D28, IN + IT, 4 to 20 million TCID50	Antibodies anti virion, D28 (or D39 after boost)
Mucosal parainfluenza virus-vectored vaccine against Ebola virus replicates in the respiratory tract of vector-immune monkeys and is immunogenic (Bukreyev et al., 2010)	Bukreyev	2010	NHP: rhesus monkeys (+/− HPIV3 seropositive); placebo	Recombinant HPIV3-GP(Zaire Mayinga), boost D28, IN + IT, 20 million PFU	Antibodies anti virion, D28
Safety and immunogenicity of a chimpanzee adenovirus-vectored Ebola vaccine in healthy adults: a randomised, double-blind, placebo-controlled, dose-finding, phase 1/2a study (De Santis et al., 2016)	De Santis	2016	Humans (Switzerland), phase 1/2, randomization and placebo	Recombinant ChAd3-GP(Zaire Mayinga), single injection, IM, 25 to 50 billion VP	Antibodies anti GP (Mayinga), D28, results in EC90
Respiratory tract immunization of non-human primates with a Newcastle disease virus-vectored vaccine candidate against Ebola virus elicits a neutralizing antibody response (DiNapoli et al., 2010)	DiNapoli	2010	NHP: rhesus monkeys; no placebo	Recombinant NDV-GP(Zaire Mayinga) or HPIV3-GP(Zaire Mayinga), boost D28, IN + IT, 20 million PFU	Antibodies anti virion (Mayinga), D28
A Monovalent Chimpanzee Adenovirus Ebola Vaccine Boosted with MVA (Ewer et al., 2016)	Ewer	2016	Humans (United Kingdom), phase 1, no randomization, no placebo	Recombinant ChAd3-GP(Zaire Mayinga) 10 to 50 billion VP, boost between D7 and D46 with recombinant MVA-GP(Zaire Mayinga + Sudan Gulu)/NP(Taï Forest) 150 to 300 millions PFU, IM	Antibodies anti GP (Mayinga) or anti virion (Makona)
Vesicular stomatitis virus-based vaccines protect nonhuman primates against aerosol challenge with Ebola and Marburg viruses (Geisbert et al., 2008a,b)	Geisbert	2008	NHP: cynomolgus macaques; placebo	Recombinant VSV-GP(Zaire Kikwit), single injection, 20 million PFU	Antibodies anti virion (Kikwit), D14 or D27
Vesicular stomatitis virus-based ebola vaccine is well-tolerated and protects immunocompromised nonhuman primates (Geisbert et al., 2008a,b)	Geisbert	2008	NHP: rhesus monkeys (SHIV infected); placebo	Recombinant VSV-GP(Zaire Mayinga), single injection, IM, 10 million PFU	Antibodies anti virion (Mayinga), D14
Single-injection vaccine protects nonhuman primates against infection with marburg virus and three species of ebola virus (Geisbert et al., 2009)	Geisbert	2009	NHP: cynomolgus macaques and rhesus monkeys; placebo	Recombinant VSV-GP(Zaire Mayinga and/or Sudan Boniface +/− Marburg), single injection +/− boost D14, IM, 10 to 20 million PFU	Antibodies anti virion, between D14 and D28
Recombinant adenovirus serotype 26 (Ad26) and Ad35 vaccine vectors bypass immunity to Ad5 and protect nonhuman primates against ebolavirus challenge (Geisbert et al., 2011)	Geisbert	2011	NHP: cynomolgus macaques (+/− Ad5 seropositive); placebo	Recombinant Ad5, Ad26, or Ad35, or prime Ad26 + boost Ad35 D28 - GP (Zaire + Sudan Gulu), IM, 20 to 200 billion VP	Antibodies anti GP, D21, results in EC90
Codon-optimized filovirus DNA vaccines delivered by intramuscular electroporation protect cynomolgus macaques from lethal Ebola and Marburg virus challenges (Grant-Klein et al., 2015)	Grant-Klein	2015	NHP: cynomolgus macaques; placebo	Vaccin ADN-GP(Zaire +/− Sudan, Reston et Marburg), 3 injections (28 jours apart), electroporation IM, 500 μg to 2 mg	Antibodies anti GP (Mayinga) ΔTM or ΔMuc, D28
Demonstration of cross-protective vaccine immunity against an emerging pathogenic Ebolavirus Species (Hensley et al., 2010)	Hensley	2010	NHP: cynomolgus macaques; placebo	Vaccin ADN-GP(Zaire Mayinga + Sudan Gulu), 4 injections IM, 4 mg (28 to 42 days apart +/− boost D371 recombinant Ad5-GP (Zaire Mayinga) IM 100 billion VP	Antibodies anti GP, D21 or D371, results in EC90
Venezuelan equine encephalitis virus replicon particle vaccine protects nonhuman primates from intramuscular and aerosol challenge with ebolavirus (Herbert et al., 2013)	Herbert	2013	NHP: cynomolgus macaques; placebo	VRP GP(Zaire Kikwit +/− Sudan Boniface), single injection, IM, 10 to 20 billion FFU	Antibodies anti GP, D28
The effect of dose on the safety and immunogenicity of the VSV Ebola candidate vaccine: a randomised double-blind, placebo-controlled phase 1/2 trial (Huttner et al., 2015)	Huttner	2015	Humans (Switzerland), phase 1/2, randomization and placebo	Recombinant VSV-GP(Zaire), single injection, IM, 300 000 PFU	Antibodies anti GP, D28
Live attenuated recombinant vaccine protects nonhuman primates against Ebola and Marburg viruses (Jones et al., 2005)	Jones	2005	NHP: cynomolgus macaques; placebo	Recombinant VSV-GP(Zaire Mayinga), single injection, IM, 10 million PFU	Antibodies anti virion, D28
Phase 2 Placebo-Controlled Trial of Two Vaccines to Prevent Ebola in Liberia (Kennedy et al., 2017)	Kennedy	2017	Humans (Liberia), phase 2, randomization and placebo	Recombinant ChAd3-GP(Saire) 100 billion VP or VSV-GP(Zaire Kikwit) 20 million PFU, single injection, IM	Antibodies anti GP (Kikwit), D28
Safety and immunogenicity of Ebola virus and Marburg virus glycoprotein DNA vaccines assessed separately and concomitantly in healthy Ugandan adults: a phase 1b,	Kibuuka	2015	Humans (Uganda), phase 1b; randomization and placebo	Vaccin ADN GP(Zaire + Sudan +/− Marburg), 3 injections, IM, 4 mg	Antibodies anti GP, D28

Table 1 (*Continued*)

Title	First author	Year of publication	Population and study features	Vaccine(s)	Measurement of antibody response
randomised, double-blind, placebo-controlled clinical trial (Kibuuka et al., 2015)					
A replication defective recombinant Ad5 vaccine expressing Ebola virus GP is safe and immunogenic in healthy adults (Ledgerwood et al., 2010)	Ledgerwood	2010	Humans (USA), phase 1; randomization and placebo	Recombinant Ad5-GP(Zaire Mayinga + Sudan Gulu), single injection, IM, 2 to 20 billion VP	Antibodies anti GP (Mayinga), D28
Chimpanzee Adenovirus Vector Ebola Vaccine – Preliminary Report (Ledgerwood et al., 2015)	Ledgerwood	2015	Humans (USA), phase 1, no randomization and no placebo	Recombinant ChAd3-GP(Zaire Mayinga + Sudan), single injection, IM, 20 or 200 billion VP	Antibodies anti GP (Mayinga or Zaire-Guinea), D28, results in EC90
Immunity duration of a recombinant adenovirus type-5 vector-based Ebola vaccine and a homologous prime-boost immunisation in healthy adults in China: final report of a randomised, double-blind, placebo-controlled, phase 1 trial (Li et al., 2017)	Li	2017	Humans (China), phase 1; randomization and placebo	Recombinant Ad5-GP(Zaire Makona), 2 injections (168 days apart), IM, 40 or 160 billion VP	Antibodies anti GP, D28, results in EC90
A DNA vaccine for Ebola virus is safe and immunogenic in a phase I clinical trial (Martin et al., 2006)	Martin	2006	Humans (USA), phase 1, randomization and placebo	Vaccin ADN GP/NP(Zaire Mayinga) + GP (Sudan Gulu), 3 injections (28 days apart), IM, 2 to 8 mg	Antibodies anti GP or NP (Mayinga), D28
Antibodies are necessary for rVSV/ZEBOV-GP-mediated protection against lethal Ebola virus challenge in nonhuman primates (Marzi et al., 2013)	Marzi	2013	NHP: cynomolgus macaques (with depletion CD4+ or CD8 + or CD20+); placebo	Recombinant VSV-GP(Zaire Mayinga), single injection, IM, 10 million PFU	Antibodies anti GP, D28
Vesicular stomatitis virus-based vaccines against Lassa and Ebola viruses (Marzi et al., 2015a,b,c)	Marzi	2015	NHP: cynomolgus macaques (vaccinated with VSV-Lassa); placebo	Recombinant VSV-GP(Zaire Mayinga), single injection, IM, 10 million PFU	Antibodies anti GP, day of measurement non specified
Vaccines. An Ebola whole-virus vaccine is protective in nonhuman primates (Marzi et al., 2015a,b,c)	Marzi	2015	NHP: cynomolgus macaques; placebo	Attenuated whole-virus Zaire Mayinga, single injection, IM, 10 to 20 million FFU	Antibodies anti GP, D28
EBOLA VACCINE. VSV-EBOV rapidly protects macaques against infection with the 2014/15 Ebola virus outbreak strain (Marzi et al., 2015a,b,c)	Marzi	2015	NHP: cynomolgus macaques; placebo	Recombinant VSV-GP(Zaire Kikwit), single injection, unique, 50 million PFU	Antibodies anti GP, between D3 and D28
Cytomegalovirus-based vaccine expressing Ebola virus glycoprotein protects nonhuman primates from Ebola virus infection (Marzi et al., 2016)	Marzi	2016	NHP: rhesus monkeys (CMV seropositive); placebo	Recombinant RhCMV-GP(Zaire Mayinga), boost D84, SC, 10 million PFU	Antibodies anti GP, D28
Vaccination With a Highly Attenuated Recombinant Vesicular Stomatitis Virus Vector Protects Against Challenge With a Lethal Dose of Ebola Virus (Matassov et al., 2015)	Matassov	2015	NHP: rhesus monkeys; placebo	Recombinant VSV-GP(Zaire Mayinga), single injection, IM, 10 million PFU	Antibodies anti GP, D21
Aerosolized Ebola vaccine protects primates and elicits lung-resident T cell responses (Meyer et al., 2015)	Meyer	2015	NHP: rhesus monkeys; placebo	Recombinant HPIV3-GP(Zaire Mayinga) 40 to 400 million PFU or VRP(Zaire Mayinga) 10 billion PFU, boost D28, IM or aerosol or IN + IT	Antibodies anti virion (Mayinga), D23 or D28
Safety and immunogenicity of novel adenovirus type 26–and modified vaccinia ankara–vectored ebola vaccines: A randomized clinical trial (Milligan et al., 2016)	Milligan	2016	Humans (United Kingdom), phase 1, randomization and placebo	Recombinant Ad26-GP(Zaire Mayinga) 50 billion VP or recombinant MVA-GP(Zaire Mayinga + Sudan Gulu)/NP(Taï Forest) 100 millions TCID50, boost between D15 and D56, IM	Antibodies anti GP Kikwit), D28 after prime and D21 after boost
Vesicular stomatitis virus-based vaccines protect nonhuman primates against Bundibugyo ebolavirus (Mire et al., 2013)	Mire	2013	NHP: cynomolgus macaques; placebo	Recombinant VSV-GP(Zaire Mayinga and/or Sudan Boniface or Bundibugyo) +/− boost VSV-GP(Zaire Mayinga) D14, IM, 20 million PFU	Antibodies anti GP, between D22 and D29
Single-dose attenuated Vesiculovax vaccines protect primates against Ebola Makona virus (Mire et al., 2015)	Mire	2015	NHP: cynomolgus macaques; placebo	Recombinant VSV-GP(Zaire Mayinga), single injection, IM, 20 million PFU	Antibodies anti GP, D28
Protection of nonhuman primates against two species of Ebola virus infection with a single complex adenovirus vector (Pratt et al., 2010)	Pratt	2010	NHP: cynomolgus macaques or rhesus monkeys (+/− Ad5 seropositive); placebo	Recombinant CAdVax-GP(Zaire Kikwit + Sudan Boniface +/− Marburg), boost between D65 and D238, IM, 100 million to 20 billion PFU	Antibodies anti virion, between D7 and D49
A Kunjin Replicon Virus-like Particle Vaccine Provides Protection Against Ebola Virus Infection in Nonhuman Primates (Pyankov et al., 2015)	Pyankov	2015	NHP: African green monkeys; placebo	Recombinant VLP Kunjin-GP(Zaire Mayinga), boost D28, SC, 1 billion VLP	Antibodies anti virion, D21 or D28
A Monovalent Chimpanzee Adenovirus Ebola Vaccine - Preliminary Report (Rampling et al., 2015)	Rampling	2015	Humans (United Kingdom), phase 1, no randomization, no placebo	Recombinant ChAd3-GP(Zaire), single injection, IM, 10 to 50 billions VP	Antibodies anti GP, D28, results in EC90
A Recombinant Vesicular Stomatitis Virus Ebola Vaccine - Preliminary Report (Regules et al., 2015)	Regules	2015	Humans (USA), phase 1, randomization and placebo	Recombinant VSV-GP(Zaire Kikwit), single injection, IM, 3 to 20 million PFU	Antibodies anti GP (Kikwit or Mayinga), D28

Table 1 (*Continued*)

Title	First author	Year of publication	Population and study features	Vaccine(s)	Measurement of antibody response
Safety and immunogenicity of DNA vaccines encoding Ebolavirus and Marburgvirus wild-type glycoproteins in a phase I clinical trial (Sarwar et al., 2015)	Sarwar	2015	Humans (USA), phase 1, no randomization and no placebo	Vaccin ADN GP(Zaire + Sudan), 3 injections (28 days apart) + boost D168, IM, 4 mg	Antibodies anti GP, D28
Chimpanzee adenovirus vaccine generates acute and durable protective immunity against ebolavirus challenge (Stanley et al., 2014)	Stanley	2014	NHP: cynomolgus macaques; placebo	Recombinant ChAd3-GP(Zaire + Sudan), 1 to 10 billion VP or recombinant ChAd3-GP (Zaire + Sudan) or recombinant MVA-GP (Zaire + Sudan) 100 million VP, single injection, IM	Antibodies anti GP, D21, results in EC90
Development of a preventive vaccine for Ebola virus infection in primates (Sullivan et al., 2000a,b)	Sullivan	2000	NHP: cynomolgus macaques; placebo	DNA vaccine GP/NP(Zaire) + GP(Sudan + Taï Forest), 3 injections, 4 mg (28 days apart), boost D84 recombinant Ad5-GP(Z) 10 billion PFU, IM	Nature of viral antigen non specified, D28
Accelerated vaccination for Ebola virus haemorrhagic fever in non-human primates (Sullivan et al., 2003)	Sullivan	2003	NHP: cynomolgus macaques; placebo	Recombinant Ad5-GP/NP(Zaire) +/− boost D63, IM, 2000 billion VP	Antibodies anti virion, between D7 and D63
CD8+ cellular immunity mediates rAd5 vaccine protection against Ebola virus infection of nonhuman primates (Sullivan et al., 2011)	Sullivan	2011	NHP: cynomolgus macaques; placebo	Recombinant Ad5-GP(Zaire), single injection, IM, 10 billion VP	Antibodies anti GP, day of measurement non specified, results in EC90
Vaccine to confer to nonhuman primates complete protection against multistrain Ebola and Marburg virus infections (Swenson et al., 2008)	Swenson	2008	NHP: cynomolgus macaques; placebo	Recombinant Ad5-GP/NP(Zaire) + GP(Sudan Boniface), boost D63, IM, 40 billion PFU	Antibodies anti virion, D14 after prime and D21 after boost
Use of ChAd3-EBO-Z Ebola virus vaccine in Malian and US adults, and boosting of Malian adults with MVA-BN-Filo: a phase 1, single-blind, randomised trial, a phase 1b, open-label and double-blind, dose-escalation trial, and a nested, randomised, double-blind, placebo-controlled trial (Tapia et al., 2016)	Tapia	2016	Humans (Mali), phase 1, randomization and placebo	Recombinant ChAd3-GP(Zaire), 10 to 100 billion VP, boost D97 recombinant MVA-GP (Zaire + Sudan + Marburg) + NP (Taï Forest) 200 millions PFU, IM	Antibodies anti GP, D28
Ebola virus-like particle-based vaccine protects nonhuman primates against lethal Ebola virus challenge (Warfield et al., 2007)	Warfield	2007	NHP: cynomolgus macaques; placebo	VLP GP/VP40/NP(Zaire), 2 injections (42 days apart), boost D42, IM, 250 μg	Antibodies anti virion, D42
Vaccinating captive chimpanzees to save wild chimpanzees (Warfield et al., 2014)	Warfield	2014	NHP: chimpanzee; no placebo	VLP (with adjuvant: IDC-1001 ou CpG) GP/VP40/NP(Zaire), 2 injections (29 days apart), boost D27, IM, 3 mg	Antibodies anti GPΔTM or VP40, between D27 and D29, results in EC50
Homologous and heterologous protection of nonhuman primates by Ebola and Sudan virus-like particles (Warfield et al., 2015)	Warfield	2015	NHP: cynomolgus macaques; placebo	VLP GP/VP40/NP(Zaire et/ou Sudan), boost D42, IM, 3 mg	Antibodies anti GPΔTM or VP40, between D14 and D28
Immune parameters correlate with protection against ebola virus infection in rodents and nonhuman primates (Wong et al., 2012)	Wong	2012	NHP: cynomolgus macaques; placebo	Recombinant VSV-GP(Zaire Mayinga), single injection, IM or IT or PO, 20 millions PFU	Antibodies anti GP, D28
An Adenovirus Vaccine Expressing Ebola Virus Variant Makona Glycoprotein Is Efficacious in Guinea Pigs and Nonhuman Primates (Wu et al., 2016)	Wu	2016	NHP: cynomolgus macaques; placebo	Recombinant Ad5-GP(Zaire Makona), single injection, IM, 40 or 200 billion VP	Antibodies anti GP, D28
Safety and immunogenicity of a novel recombinant adenovirus type-5 vector-based Ebola vaccine in healthy adults in China: preliminary report of a randomised, double-blind, placebo-controlled, phase 1 trial (Zhu et al., 2015)	Zhu	2015	Humans (China), phase 1, randomization and placebo	Recombinant Ad5-GP(Zaire Makona), single injection, IM, 40 to 160 billions VP	Antibodies anti GP (Makona), D28
Safety and immunogenicity of a recombinant adenovirus type-5 vector-based Ebola vaccine in healthy adults in Sierra Leone: a single-centre, randomised, double-blind, placebo-controlled, phase 2 trial (Zhu et al., 2017)	Zhu	2017	Humans (Sierra Leone), phase 2; randomization and placebo	Recombinant Ad5-GP(Zaire Makona), single injection, IM, 40 or 160 billion VP	Antibodies anti GP, D28

presented in Figure 3. The vast majority (82.4%) of human groups included 8 or more individuals, while only 6 for non-human primate groups (range 2; 22 with an average of 4.1 individuals by group).

Characteristics of nonhuman primate and human groups are described in Tables 2 and 3, respectively.

There is a wide heterogeneity of features among studies included in the systematic review. Vaccine platforms varied between studies, especially in NHP (18 different vaccine platforms in NHP groups versus 8 in human groups). The strain of Ebola virus used as vaccine insert or for the antibody detection after vaccination was also variable. For almost a third of the human

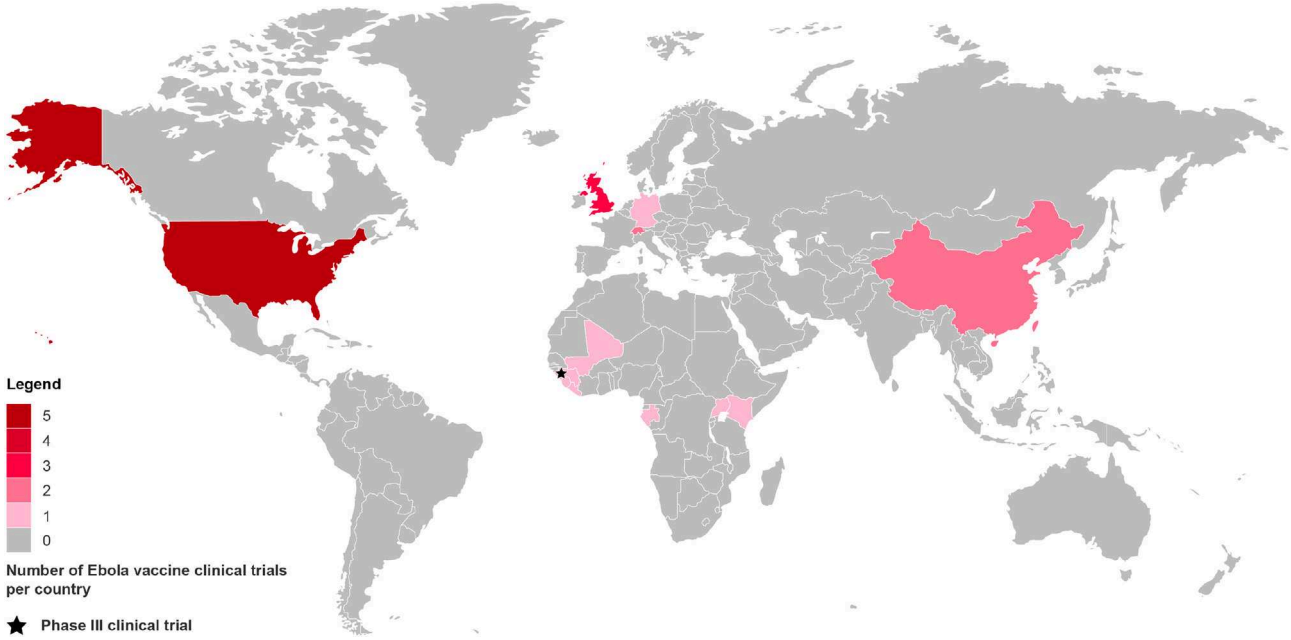


Figure 2. Description of the number of vaccine clinical trials against Ebola per country.
The Ring trial, single phase 3 trial (Guinea), has been excluded from the systematic review.
Several other vaccine clinical trials against Ebola are currently ongoing worldwide but only published trials are reported in the figure.

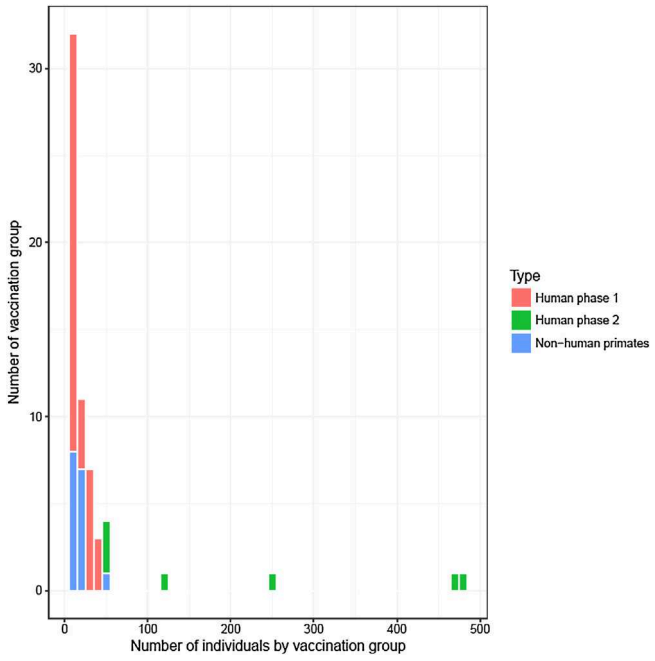


Figure 3. Number of vaccination groups of humans and of nonhuman primates, according to the number of individuals by group.

groups, the detection of antibody response was done with a heterologous strain. The time interval between the last vaccination and the antibody detection was also remarkably variable (range 3; 371 days).

Among all the 202 vaccination groups, the mean antibody titre ranged from 0 (for a group of NHP infected by the simian/human immunodeficiency virus prior to the Ebola vaccination) to 5.81 log10, with an average of 2.97 (95% CI: [2.84; 3.10]).

The NHP groups had a crude antibody response level that was significantly higher than the human groups ($p = 0.006$): in NHP groups the log10 geometric mean titre ranged from 0 to 5.81 with an average of 3.10 (95% CI: [2.93; 3.27]), and in human groups the titre ranged from 0.90 to 4.60 with an average of 2·75 (95% CI: [2.57; 2.93]) Figure S1 (appendix) shows antibody responses in human groups and in NHP groups.

Meta-regression of factors associated with variability in antibody response levels in humans and evaluation of between-groups heterogeneity

Sixty-one human vaccination groups with 8 individuals or more were included in the meta-regression analysis.

Among these, 32 were vaccinated with a low dose of vaccine, 19 with a high dose (for 10 groups, the dose category was undeterminable as only one dose level was assessed for the given vaccine platform).

The distribution of the antibody titres after Ebola vaccination per vaccination group is shown by vaccine platform in Figure 4. The antibody response seems to be higher in groups with a prime-boost strategy (Ad26/MVA or ChAd3/MVA) than in the other groups. The distribution of the antibody titres by viral strain used for antibody detection is presented in Figure S2 (appendix).

In univariate meta-regression analyses (appendix: Table S1), the antibody response after Ebola vaccination was significantly associated with the vaccine platform ($p < 0.001$), the viral strain used to detect the antibody response after vaccination ($p < 0.001$), the year of publication (for publication in 2014 and after versus before 2014: +1.15, $p < 0.001$), the mean age of vaccinated population (for ≥39 years versus <32 years: +0.90; $p < 0.001$), the vaccine dosage (for high dose versus low dose: +0.57, $p = 0.006$), the use of a vaccine boost (for boost versus no boost: +0.63, $p = 0.009$), the similarity between the viral strain used as vaccine insert and the viral strain used to detect the antibody response (for identical strains versus different strains: −0.74, $p = 0.009$), the site of the study ($p = 0.014$), the time interval between the last vaccine injection and the antibody measure (for <28 days versus ≥28 days:

Table 2
Main characteristics of included non-human primates (NHP) groups.

Characteristic	Vaccination schedule				All NHP groups	
	No boost (n = 98)		Boost (n = 30)		n = 128	
Vaccine platform						
DNA vaccine (plasmid)	6	6.1%	0	0.0%	6	4.7%
Adenovirus 26	4	4.1%	0	0.0%	4	3.1%
Adenovirus 26 then adenovirus 35	0	0.0%	1	3.3%	1	0.8%
Adenovirus 35	4	4.1%	0	0.0%	4	3.1%
Adenovirus 5	8	8.2%	3	10.0%	11	8.6%
DNA vaccine (plasmid)/adenovirus 5	0	0.0%	2	6.7%	2	1.6%
CAdVax	6	6.1%	2	6.7%	8	6.2%
Chimpanzee adenovirus 3	2	2.0%	0	0.0%	2	1.6%
Chimpanzee adenovirus 63	1	1.0%	0	0.0%	1	0.8%
HPIV3	12	12.2%	7	23.3%	19	14.8%
MVA	1	1.0%	0	0.0%	1	0.8%
NDV	1	1.0%	1	3.3%	2	1.6%
RhCMV	1	1.0%	0	0.0%	1	0.8%
Whole-virus vaccine	4	4.1%	0	0.0%	4	3.1%
VLP	21	21.4%	10	33.3%	31	24.2%
VLP Kunjin	1	1.0%	1	3.3%	2	1.6%
VRP VEEV	3	3.1%	1	3.3%	4	3.1%
VSV	23	23.5%	2	6.7%	25	19.5%
Route of administration						
Intramuscular	78	79.6%	21	70.0%	99	77.3%
Other routes	20	20.4%	9	30.0%	29	22.7%
Vaccine insert: Ebola species						
Monovalent Zaire	63	64.3%	17	56.7%	80	62.5%
Monovalent no Zaire	6	6.1%	2	6.7%	8	6.2%
Monovalent no Zaire + monovalent Zaire	0	0.0%	1	3.3%	1	0.8%
Monovalent no Zaire + multivalent	0	0.0%	1	3.3%	1	0.8%
Multivalent	29	29.6%	7	23.3%	36	28.1%
Multivalent + monovalent Zaire	0	0.0%	2	6.7%	2	1.6%
Vaccine insert: Ebola strain (only for Zaire species)						
Mayinga	34	64.2%	13	76.5%	47	67.1%
Kikwit	16	30.2%	4	23.5%	20	28.6%
Makona	3	5.7%	0	0.0%	3	4.3%
Missing data	*45*	–	*13*	–	*58*	–
Nonhuman primates species						
Cynomolgus macaques	65	66.3%	17	56.7%	82	64.1%
Chimpanzees	10	10.2%	2	6.7%	12	9.4%
Rhesus macaques	22	22.4%	10	33.3%	32	25.0%
African green monkeys	1	1.0%	1	3.3%	2	1.6%
Year of publication						
Publication < 2014	49	50.0%	14	46.7%	63	49.2%
Publication ≥ 2014	49	50.0%	16	53.3%	65	50.8%
Time interval between last injection and antibody measure						
Mean [standard deviation]	29.1	[363]	25.5	[762]	28.3	[319]
Missing data	*3*	–	*0*	–	*3*	–
Antibody measurement method						
Maximal dilution	65	66.3%	18	60.0%	83	64.8%
Effective concentration 90 (EC90)	15	15.3%	2	6.7%	17	13.3%
Effective concentration 50 (EC50)	18	18.4%	10	33.3%	28	21.9%
Antigen used for antibody detection: nature						
Glycoprotein (GP)	55	56.7%	7	24.1%	62	49.2%
Other nature (virion, viral protein 40)	42	43.3%	22	75.9%	64	50.8%
Missing data	*1*	–	*1*	–	*2*	–
Antigen used for antibody detection: Ebola strain						
Mayinga	14	87.5%	8	100.0%	22	91.7%
Kikwit	2	12.5%	0	0.0%	2	8.3%
Missing data	*82*	–	*22*	–	*104*	–
Similarity between strain used as vaccine insert and strain used for antibody detection						
Identical strains	12	100.0%	8	100.0%	20	100.0%
Missing data	*86*	–	*22*	–	*108*	–

CAdVax: complex adenovirus-based vector, DNA: deoxyribonucleic acid, GP: glycoprotein, HPIV3: human parainfluenza virus 3, MVA: modified vaccinia Ankara, NDV: Newcastle disease virus, RhCMV: rhesus cytomegalovirus cytomegalovirus, VLP: virus-like particles, VRP VEEV: Venezuelan equine encephalitis virus replicon particle, VSV: vesicular stomatitis virus.

Table 3
Main characteristics of included human groups.

Characteristics	Vaccination schedule				All human groups	
	No boost (n = 48)		Boost (n = 26)		(n = 74)	
Vaccine platform						
DNA vaccine (plasmid)	9	18.8%	1	3.8%	10	13.5%
Adenovirus 26	3	6.2%	0	0.0%	3	4.1%
Adenovirus 26/MVA or MVA/adenovirus 26	0	0.0%	5	19.2%	5	6.8%
Adenovirus 5	6	12.5%	2	7.7%	8	10.8%
Chimpanzee adenovirus 3	14	29.2%	0	0.0%	14	18.9%
Chimpanzee adenovirus 3/MVA	0	0.0%	18	69.2%	18	24.3%
MVA	2	4.2%	0	0.0%	2	2.7%
VSV	14	29.2%	0	0.0%	14	18.9%
Route of administration						
Intramuscular	48	100.0%	26	100.0%	74	100.0%
Vaccine insert: species						
Monovalent Zaire	31	64.6%	19	73.1%	50	67.4%
Monovalent Zaire + multivalent	0	0.0%	4	15.4%	4	5.4%
Multivalent	17	35.4%	1	3.8%	18	24.3%
Multivalent + monovalent Zaire	0	0.0%	2	7.7%	2	2.7%
Vaccine insert: strain (only for Zaire species)						
Mayinga	22	71.0%	22	91.7%	44	80.0%
Kikwit	5	16.1%	0	0.0%	5	9.1%
Makona	4	12.9%	0	0.0%	6	10.9%
Missing data	*17*	–	*4*	–	*19*	–
Proportion of women						
Mean [standard deviation]	40%	[18%]	52%	[10%]	44%	[17%]
Mean age (years)						
Mean [standard deviation]	34.8	[45]	34.6	[54]	34.7	[48]
Geographic location of the study						
Africa	14	29.2%	1	3.8%	15	20.3%
China	2	4.2%	2	7.7%	4	5.4%
Europe	15	31.2%	22	84.6%	37	50.0%
USA	17	35.4%	1	3.8%	18	24.3%
Year of publication						
Publication < 2014	8	16.7%	0	0.0%	8	10.8%
Publication ≥ 2014	40	83.3%	26	100.0%	66	89.2%
Time interval between last injection and antibody measure (days)						
Mean [standard deviation]	31.2	[220]	26.1	[32]	29.3	[179]
Antibody measurement method						
Maximal dilution	35	72.9%	24	92.3%	59	79.7%
Effective concentration 90 (EC90)	13	27.1%	2	7.7%	15	20.3%
Antigen used for antibody detection: nature						
Glycoprotein (GP)	45	93.8%	25	96.2%	70	94.6%
Other nature (virion, nucleoprotein)	3	6.2%	1	3.8%	4	5.4%
Antigen used for antibody detection: Ebola strain						
Mayinga	14	36.8%	16	72.7%	30	50.0%
Kikwit	18	47.4%	5	22.7%	23	38.3%
Makona	6	15.8%	1	4.5%	7	11.7%
Missing data	*10*	–	*4*	–	*14*	–
Similarity between strain used as vaccine insert and strain used for antibody detection						
Different strains	9	32.1%	6	27.3%	15	30.0%
Identical strains	19	67.9%	16	72.7%	35	70.0%
Missing data	*20*	–	*4*	–	*24*	–

CAdVax: complex adenovirus-based vector, DNA: deoxyribonucleic acid, GP: glycoprotein, HPIV3: human parainfluenza virus 3, MVA: modified vaccinia Ankara, NDV: Newcastle disease virus, RhCMV: rhesus cytomegalovirus cytomegalovirus, VLP: virus-like particles, VRP VEEV: Venezuelan equine encephalitis virus replicon particle, VSV: vesicular stomatitis virus.

+0.70, p = 0.021), and the Ebola species of vaccine insert (for multivalent and other species versus monovalent Zaire: −0.47, p = 0.027).

Alone, the vaccine platform was the factor which explained the largest part of heterogeneity among all the studied factors (R^2 for vaccine platform = 55%). For all the univariate models, the heterogeneity was very high with I^2 ranging from 97% to 99%.

Results of the final multivariate meta-regression model are shown in Table 4. High heterogeneity was found with a I^2 of 95% and a R^2 of 68%, even after adjustment on the factors associated

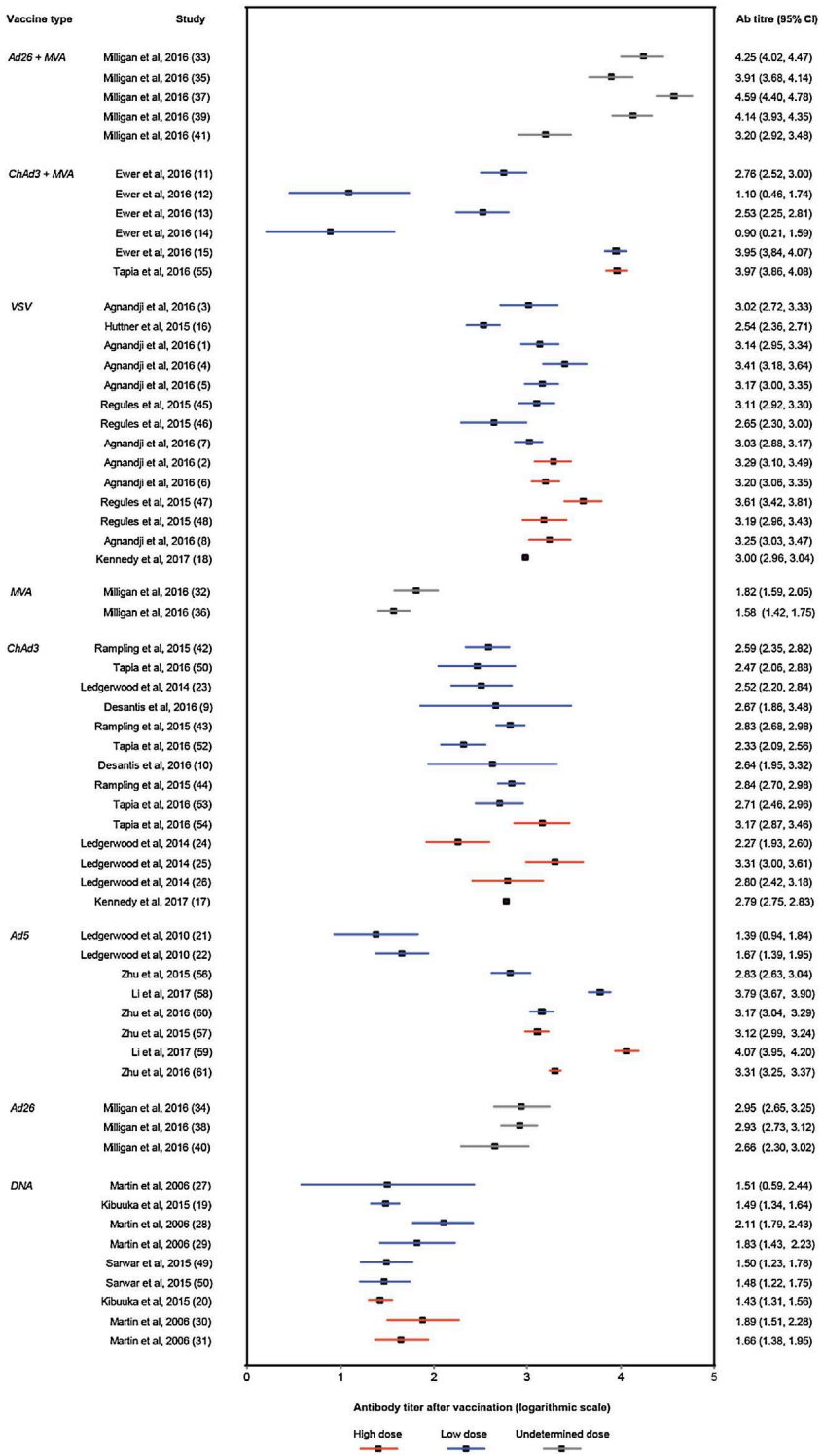


Figure 4. Forest plot of antibody titre after Ebola vaccination for each vaccination group by vaccine platform. Colour codes indicate dose levels within a given platform. GP: glycoprotein. PFU: plaque forming unit. VP: viral particle. TCID: tissue culture infectious dose.

References for Figure 4:

1: Agnandji 2016, VSV vaccine (3.10^6 PFU) with Zaire insert, Germany, detection with Zaire Kikwit GP
2: Agnandji 2016, VSV vaccine (2.10^7 PFU) with Zaire insert, Germany, detection with Zaire Kikwit GP
3: Agnandji 2016, VSV vaccine (3.10^5 PFU) with Zaire insert, Gabon, detection with Zaire Kikwit GP
4: Agnandji 2016, VSV vaccine (3.10^6 PFU) with Zaire insert, Gabon, detection with Zaire Kikwit GP
5: Agnandji 2016, VSV vaccine (3.10^6 PFU) with Zaire insert, Kenya, detection with Zaire Kikwit GP
6: Agnandji 2016, VSV vaccine (2.10^7 PFU) with Zaire insert, Kenya, detection with Zaire Kikwit GP
7: Agnandji 2016, VSV vaccine (1.10^7 PFU) with Zaire insert, Switzerland, detection with Zaire Kikwit GP
8: Agnandji 2016, VSV vaccine (5.10^7 PFU) with Zaire insert, Switzerland, detection with Zaire Kikwit GP
9: De Santis 2016, ChAd3 vaccine ($2.5.10^{10}$ VP) with Zaire Mayinga insert, Switzerland, detection with Zaire Mayinga GP
10: De Santis 2016, ChAd3 vaccine (5.10^{10} VP) with Zaire Mayinga insert, Switzerland, detection with Zaire Mayinga GP

with antibody response in this final model. This emphasises the lack of factors that explained the antibody response among the variables included in the model.

Vaccine platform and viral strain used for detection were the two factors which were independently associated with antibody response after vaccination against Ebola. Compared to the MVA vaccine platform, the recombinant vaccines using DNA or Ad26 (associated or not with an injection of MVA vaccine), ChAd3or VSV vectors were significantly associated with a higher antibody response after vaccination (more than 1.2 log10 units more compared to MVA alone). The statistical association between the vaccine platform and the antibody response was strong and consistent regardless of which other variables were included in the model (sensitivity analyses, data not shown). The antibody response using Makona strain for antibody detection was significantly higher than with use of Mayinga strain (1 log10 unit more compared to the Mayinga strain). By contrast, the antibody response with Kikwit detection strain was not significantly different from the ones with Mayinga strain.

The vaccine dosage, analysed as a binary variable of high versus low dose in the present analyses, was not found to be associated with antibody response variability. Different classifications were tested for this variable (same threshold across the different vaccine platforms corresponding to the mean dose level for categorizing into “low-dose” and “high-dose” groups, classification into three categories, classification of groups with undifferentiated dosages into “low-dose” or into “high-dose” groups), but the dosage was never significant in the multivariate models in these sensitivity analyses (data not shown), nor was the interaction between dose and vaccine platform.

In additional sensitivity analyses, a full model including all variables significantly associated with the antibody response in univariate models (i.e. with no forward selection procedure) did not modify heterogeneity ($I^2 = 92\%$) compared to the model

11: Ewer 2016, ChAd3 vaccine ($2.5.10^{10}$ VP) with Zaire Mayinga insert + boost MVA vaccine at D7 ($1.5.10^{8}$ PFU) with multivalent insert, United Kingdom (UK), detection with Zaire Mayinga GP (Jenner method)
12: Ewer 2016, ChAd3 vaccine ($2.5.10^{10}$ VP) with Zaire Mayinga insert + boost MVA vaccine at D7 ($1.5.10^{8}$ PFU) with multivalent insert, UK, detection with Zaire Mayinga GP (ADI method)
13: Ewer 2016, ChAd3 vaccine ($2.5.10^{10}$ VP) with Zaire Mayinga insert + boost MVA vaccine at D14 ($1.5.10^{8}$ PFU) with multivalent insert, UK, detection with Zaire Mayinga GP (Jenner method)
14: Ewer 2016, ChAd3 vaccine ($2.5.10^{10}$ VP) with Zaire Mayinga insert + boost MVA vaccine at D14 ($1.5.10^{8}$ PFU) with multivalent insert, UK, detection with Zaire Mayinga GP (ADI method)
15: Ewer 2016, ChAd3 vaccine (1 to 5.10^{10} VP) with Zaire Mayinga insert + boost MVA vaccine ($1.5.10^{8}$ PFU) with multivalent insert, UK, detection with Zaire Makona virion
16: Huttner 2015, VSV vaccine (3.10^{5} PFU) with Zaire insert, Switzerland, detection with Zaire Kikwit GP
17: Kennedy 2017, ChAd3 vaccine (2.10^{11} PU) with Zaire insert, Liberia
18: Kennedy 2017, VSV vaccine (2.10^{7} PFU) with Zaire insert, Liberia
19: Kibuuka 2015, 3 injections of DNA vaccine (4 mg) with multivalent insert, Uganda, detection with Zaire GP
20: Kibuuka 2015, 3 injections of DNA vaccine (8 mg) with multivalent insert, Uganda, detection with Zaire GP
21: Ledgerwood 2010, Ad5 vaccine (2.10^{9} VP) with multivalent insert, USA, detection with Zaire GP
22: Ledgerwood 2010, Ad5 vaccine (2.10^{10} VP) with multivalent insert, USA, detection with Zaire GP
23: Ledgerwood 2014, ChAd3 vaccine (2.10^{10} PU) with multivalent insert, USA, detection with Zaire Mayinga GP
24: Ledgerwood 2014, ChAd3 vaccine (2.10^{10} PU) with multivalent insert, USA, detection with Zaire Makona GP
25: Ledgerwood 2014, ChAd3 vaccine (2.10^{11} PU) with multivalent insert, USA, detection with Zaire Mayinga GP
26: Ledgerwood 2014, ChAd3 vaccine (2.10^{11} PU) with multivalent insert, USA, detection with Zaire Makona GP
27: Li 2017, 2 injections of Ad5 vaccine (4.10^{10} VP) with Zaire Makona insert, China
28: Li 2017, 2 injections of Ad5 vaccine ($1.6.10^{11}$ VP) with Zaire Makona insert, China
29: Martin 2006, 3 injections of DNA vaccine (2 mg) with multivalent insert, USA, detection with Zaire NP
30: Martin 2006, 3 injections of DNA vaccine (4 mg) with multivalent insert, USA, detection with Zaire GP
31: Martin 2006, 3 injections of DNA vaccine (4 mg) with multivalent insert, USA, detection with Zaire NP
32: Martin 2006, 3 injections of DNA vaccine (8 mg) with multivalent insert, USA, detection with Zaire GP
33: Martin 2006, 3 injections of DNA vaccine (8 mg) with multivalent insert, USA, detection with Zaire NP
34: Milligan 2016, MVA vaccine (10^{8} TCID50) with multivalent insert, UK, detection with Zaire Kikwit GP
35: Milligan 2016, MVA vaccine (10^{8} TCID50) with multivalent insert + boost Ad26 vaccine at D28 (5.10^{10} VP) with Zaire Mayinga insert, UK, detection with Zaire Kikwit GP
36: Milligan 2016, Ad26 vaccine (5.10^{10} VP) with Zaire Mayinga insert, UK, detection with Zaire Kikwit GP
37: Milligan 2016, Ad26 vaccine (5.10^{10} VP) with Zaire Mayinga insert + boost MVA vaccine at D28 (10^{8} TCID50) with multivalent insert, UK, detection with Zaire Kikwit GP
38: Milligan 2016, MVA vaccine (10^{8} TCID50) with multivalent insert, UK, detection with Zaire Kikwit GP
39: Milligan 2016, MVA vaccine (10^{8} TCID50) with multivalent insert + boost Ad26 vaccine at D56 (5.10^{10} VP) with Zaire Mayinga insert, UK, detection with Zaire Kikwit GP
40: Milligan 2016, Ad26 vaccine (5.10^{10} VP) with Zaire Mayinga insert, UK, detection with Zaire Kikwit GP
41: Milligan 2016, Ad26 vaccine (5.10^{10} VP) with Zaire Mayinga insert + boost MVA vaccine at D56 (10^{8} TCID50) with multivalent insert, UK, detection with Zaire Kikwit GP
42: Milligan 2016, Ad26 vaccine (5.10^{10} VP) with Zaire Mayinga insert, UK, detection with Zaire Kikwit GP
43: Milligan 2016, Ad26 vaccine (5.10^{10} VP) with Zaire Mayinga insert + boost MVA vaccine at D14 (10^{8} TCID50) with multivalent insert, UK, detection with Zaire Kikwit GP
44: Rampling 2015, ChAd3 vaccine (10^{10} VP) with Zaire Mayinga insert, UK, detection with Zaire GP
45: Rampling 2015, ChAd3 vaccine ($2.5.10^{10}$ VP) with Zaire Mayinga insert, UK, detection with Zaire GP
46: Rampling 2015, ChAd3 vaccine (5.10^{10} VP) with Zaire Mayinga insert, UK, detection with Zaire GP
47: Regules 2015, VSV vaccine (3.10^{6} PFU) with Zaire Kikwit insert, USA, detection with Zaire Kikwit GP
48: Regules 2015, VSV vaccine (3.10^{6} PFU) with Zaire Kikwit insert, USA, detection with Zaire Mayinga GP
49: Regules 2015, VSV vaccine (2.10^{7} PFU) with Zaire Kikwit insert, USA, detection with Zaire Kikwit GP
50: Regules 2015, VSV vaccine (2.10^{7} PFU) with Zaire Kikwit insert, USA, detection with Zaire Mayinga GP
51: Sarwar 2015, 3 injections of DNA vaccine (4 mg) with multivalent insert, USA, detection with Zaire GP
52: Sarwar 2015, 4 injections of DNA vaccine (4 mg) with multivalent insert, USA, detection with Zaire GP
53: Tapia 2016, ChAd3 vaccine (10^{10} VP) with Zaire insert, Mali, detection with Zaire GP
54: Tapia 2016, ChAd3 vaccine ($2.5.10^{10}$ VP) with Zaire insert, Mali, detection with Zaire GP
55: Tapia 2016, ChAd3 vaccine (5.10^{11} VP) with Zaire insert, Mali, detection with Zaire GP
56: Tapia 2016, ChAd3 vaccine (10^{12} VP) with Zaire insert, Mali, detection with Zaire GP
57: Tapia 2016, ChAd3 vaccine (10^{10} to 10^{12} VP) with Zaire insert + boost MVA vaccine at D97 (2.10^{8} PFU) with multivalent insert, Mali, detection with Zaire GP
58: Zhu 2015, Ad5 vaccine (4.10^{10} VP) with Zaire Makona insert, China, detection with Zaire Makona GP
59: Zhu 2015, Ad5 vaccine ($1.6.10^{11}$ VP) with Zaire Makona insert, China, detection with Zaire Makona GP
60: Zhu 2016, Ad5 vaccine (4.10^{10} VP) with Zaire Makona insert, Sierra Leone, detection with Zaire Makona GP
61: Zhu 2016, Ad5 vaccine ($1.6.10^{11}$ VP) with Zaire Makona insert, Sierra Leone, detection with Zaire Makona GP

Table 4
Results of a random-effect meta-regression model (with fixed intragroup variance) of determinants of antibody titre (log10) after Ebola vaccination according to characteristics of vaccine, population, and measurement techniques. Multivariate analysis. $I^2 = 95.31\%$, $R^2 = 68.45\%$.

Determinants of antibody response	Estimated β [CI 95%]		p value
Vaccine platform (reference: MVA vaccine)			**<0.001**
DNA	0.43	[−0.52; 1.37]	0.379
Ad26	1.15	[033; 197]	0.006
Ad26/MVA or MVA/Ad26	2.32	[158; 307]	<0.001
Ad5	0.54	[−0.42; 1.50]	0.268
ChAd3	0.97	[010; 183]	0.028
ChAd3/MVA	0.81	[−0.13; 1.76]	0.091
VSV	1.46	[079; 213]	<0.001
Viral strain used for antibody detection (reference: Mayinga strain)			**<0.001**
Kikwit	0.30	[−0.27; 0.86]	0.301
Makona	0.99	[050; 148]	<0.001

presented above. In the full model, the vaccine platform was significantly associated with the antibody response ($p = 0.002$), but the viral strain used to detect the antibody response after vaccination was not ($p = 0.996$). The other variables were not associated with the antibody response.

Discussion

This systematic review on preventive Ebola vaccine trials has found 49 studies conducted in humans or in NHP. The meta-analysis, using a random-effect inverse variance meta-regression including 61 human vaccination groups, showed a major part of antibody response variability in humans that remained unexplained by the factors included in the model. Indeed, the between-group heterogeneity I^2 exceeded 90%, even after adjustment for the factors associated with antibody response. Two significant determinants were independently associated with antibody response after preventive vaccination against EVD: the Ebola vaccine platform and the Ebola strain used for antibody detection.

The use of a systematic review methodology, including solicitation of experts, allowed us to conduct exhaustive descriptive analyses on all Ebola vaccinated groups in NHP or humans published in the literature up to January 2017. Our descriptive results showed an extreme variability of study designs and features, especially in nonhuman primate trials. This variability is related to the recentness of the research topic. The higher variability within nonhuman primate studies compared to human trials is easily explained by the process of vaccine development, which selects for further clinical trials only the subset of candidate vaccines proven to be immunogenic in nonhuman primates. The comparison of antibody response levels between humans and nonhuman primate only had an indicative purpose. It is indeed difficult to compare these very different models, mostly because of potential multiple confounding factors.

Due to the low sample size of each group of nonhuman primates, we decided to restrict heterogeneity analyses to human groups. Human groups with small sample size were excluded, since their between-group variance would have been too low to contribute to the meta-regression model. It was not possible to pool small groups together because of high heterogeneity in the factors likely to influence the antibody response (vaccine and population characteristics, and measure of antibody response). The threshold of at least 8 individuals per group allowed us to include the majority of human groups in the meta-regression. Sensitivity analyses using a threshold of 10 individuals led to the same final results.

The very high heterogeneity between vaccination groups could be explained by various reasons. Firstly, some factors influencing the antibody response may be missing, for instance, genetic factors that are influencing the immunogenicity of the vaccines (Sridhar, 2015). Secondly, the analysis of grouped data, due to unavailability of individual data for the groups included in our meta-regression model, led to a lack of precision in the estimation of influence of factors on antibody response, and also in the evaluation of antibody response heterogeneity across vaccination groups. Thirdly, the enzyme-linked immunosorbent assay (ELISA) measuring relative antibody concentration of immunoglobulin G against EBOV glycoprotein used in the different trials could have a variation of its precision (Logue et al., 2018). Lastly, the extreme variability of study designs certainly explains parts of the high between-group variance for antibody response observed in our results.

Despite the major between-group heterogeneity in our meta-regression model, two factors significantly associated with antibody response variability could be identified. The Ebola strain used for antibody detection seems to influence the results of ELISA tests. This demonstrates the importance of harmonisation for the measurement methods used in vaccines evaluations, and highlights the difficulty in directly comparing published results across several trials. The Ebola vaccine platform was also strongly associated with antibody response.

For the other factors studied in our meta-analysis, no association was found with the antibody response variability. In particular, the vaccine dosage did not have any significant influence on the level of the antibody response in our results. We acknowledge that the use of a binary variable may have limited the ability to detect a dose-effect in the meta-regression. However, the regression result is consistent with the descriptive results that also did not suggest a clear dose-immunogenicity relationship within a given vaccine platform.

No population characteristic was independently associated with the antibody response after Ebola vaccination. It may be possible that the low diversity of the population, which is directly related to the strict criteria for selection of trial participants, prevented the identification of a potential impact of these population characteristics on the antibody response.

Conclusion

Our findings show that there are still significant uncertainties in the determinants of the antibody response after preventive vaccination against Ebola virus disease. This emphasises the interest of harmonizing measurement methods and study designs. Furthermore, it indicates the impossibility to directly compare results from one published study to another or to extrapolate results, due to considerable variations in studies features. Assessment of immunogenicity between Ebola vaccines needs randomised controlled multi-arm trials, as performed in PREVAIL study (NCT02344407) and PREVAC study (NCT02876328).

Conflict of interest

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No. 115861 This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation program and EFPIA.

Laura Richert and Rodolphe Thiebaut are also involved in the ongoing PREVAC (NCT02876328) and Ebovac2 (NCT02564523, NCT02416453) trials. Edouard Lhomme is involved in the PREVAC trial (NCT02876328).

Funding source

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 115861. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA.

The funder of the study had no role in the study design, data collection, data analysis, data interpretation, or writing the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Ethical considerations

According to the National Public Health Code, review and meta-analysis do not require ethical approval.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ijid.2018.06.022.

References

Agnandji ST, Huttner A, Zinser ME, Njuguna P, Dahlke C, Fernandes JF, et al. Phase 1 trials of rVSV Ebola vaccine in Africa and Europe. N Engl J Med 2016;374(17):1647–60.

Bukreyev A, Rollin PE, Tate MK, Yang L, Zaki SR, Shieh W-J, et al. Successful topical respiratory tract immunization of primates against Ebola virus. J Virol 2007;81(June (12)):6379–88.

Bukreyev AA, Dinapoli JM, Yang L, Murphy BR, Collins PL. Mucosal parainfluenza virus-vectored vaccine against Ebola virus replicates in the respiratory tract of vector-immune monkeys and is immunogenic. Virology 2010;399(2):290–8.

De Santis O, Audran R, Pothin E, Warpelin-Decrausaz L, Vallotton L, Wuerzner G, et al. Safety and immunogenicity of a chimpanzee adenovirus-vectored Ebola vaccine in healthy adults: a randomised, double-blind, placebo-controlled, dose-finding, phase 1/2a study. Lancet Infect Dis 2016;16(3):311–20.

Dhama K, Malik YS, Malik SVS, Singh RK. Ebola from emergence to epidemic: the virus and the disease, global preparedness and perspectives. J Infect Dev Ctries 2015;9(5):441–55.

DiNapoli JM, Yang L, Samal SK, Murphy BR, Collins PL, Bukreyev A. Respiratory tract immunization of non-human primates with a Newcastle disease virus-vectored vaccine candidate against Ebola virus elicits a neutralizing antibody response. Vaccine 2010;29(1):17–25.

Ebola Situation Report—30 March 2016 | Ebola [Internet]. [Cited 4 January 2017]. Available from: http://apps.who.int/ebola/current-situation/ebola-situation-report-30-march-2016.

Ewer K, Rampling T, Venkatraman N, Bowyer G, Wright D, Lambe T, et al. A monovalent chimpanzee adenovirus ebola vaccine boosted with MVA. N Engl J Med 2016;374(17):1635–46.

Food and Drug Administration. Licensure of Ebola vaccines: demonstration of effectiveness [Internet]. 2015 Available from: https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/BloodVaccinesandOtherBiologics/VaccinesandRelatedBiologicalProductsAdvisoryCommittee/UCM445819.pdf.

Geisbert TW, Daddario-Dicaprio KM, Geisbert JB, Reed DS, Feldmann F, Grolla A, et al. Vesicular stomatitis virus-based vaccines protect nonhuman primates against aerosol challenge with Ebola and Marburg viruses. Vaccine 2008a;26(52):6894–900.

Geisbert TW, Daddario-Dicaprio KM, Lewis MG, Geisbert JB, Grolla A, Leung A, et al. Vesicular stomatitis virus-based ebola vaccine is well-tolerated and protects immunocompromised nonhuman primates. PLoS Pathog 2008b;4(11)e1000225.

Geisbert TW, Geisbert JB, Leung A, Daddario-DiCaprio KM, Hensley LE, Grolla A, et al. Single-injection vaccine protects nonhuman primates against infection with marburg virus and three species of ebola virus. J Virol 2009;83(14):7296–304.

Geisbert TW, Bailey M, Hensley L, Asiedu C, Geisbert J, Stanley D, et al. Recombinant adenovirus serotype 26 (Ad26) and Ad35 vaccine vectors bypass immunity to Ad5 and protect nonhuman primates against ebolavirus challenge. J Virol 2011;85(9):4222–33.

Grant-Klein RJ, Altamura LA, Badger CV, Bounds CE, Van Deusen NM, Kwilas SA, et al. Codon-optimized filovirus DNA vaccines delivered by intramuscular electroporation protect cynomolgus macaques from lethal Ebola and Marburg virus challenges. Hum Vaccines Immunother 2015;11(8):1991–2004.

Henao-Restrepo AM, Camacho A, Longini IM, Watson CH, Edmunds WJ, Egger M, et al. Efficacy and effectiveness of an rVSV-vectored vaccine in preventing Ebola virus disease: final results from the Guinea ring vaccination, open-label, cluster-randomised trial (Ebola Ça Suffit!). Lancet 2017;389(10068):505–18.

Hensley LE, Mulangu S, Asiedu C, Johnson J, Honko AN, Stanley D, et al. Demonstration of cross-protective vaccine immunity against an emerging pathogenic Ebolavirus species. PLoS Pathog 2010;6(5)e1000904.

Herbert AS, Kuehne AI, Barth JF, Ortiz RA, Nichols DK, Zak SE, et al. Venezuelan equine encephalitis virus replicon particle vaccine protects nonhuman primates from intramuscular and aerosol challenge with ebolavirus. J Virol 2013;87(9):4952–64.

Huttner A, Dayer J-A, Yerly S, Combescure C, Auderset F, Desmeules J, et al. The effect of dose on the safety and immunogenicity of the VSV Ebola candidate vaccine: a randomised double-blind, placebo-controlled phase 1/2 trial. Lancet Infect Dis 2015;15(10):1156–66.

Jones SM, Feldmann H, Ströher U, Geisbert JB, Fernando L, Grolla A, et al. Live attenuated recombinant vaccine protects nonhuman primates against Ebola and Marburg viruses. Nat Med 2005;11(7):786–90.

Kennedy SB, Bolay F, Kieh M, Grandits G, Badio M, Ballou R, et al. Phase 2 placebo-controlled trial of two vaccines to prevent ebola in Liberia. N Engl J Med 2017;377:1438–47.

Kibuuka H, Berkowitz NM, Millard M, Enama ME, Tindikahwa A, Sekiziyivu AB, et al. Safety and immunogenicity of Ebola virus and Marburg virus glycoprotein DNA vaccines assessed separately and concomitantly in healthy Ugandan adults: a phase 1b, randomised, double-blind, placebo-controlled clinical trial. Lancet 2015;385(9977):1545–54.

Ledgerwood JE, Costner P, Desai N, Holman L, Enama ME, Yamshchikov G, et al. A replication defective recombinant Ad5 vaccine expressing Ebola virus GP is safe and immunogenic in healthy adults. Vaccine 2010;29(2):304–13.

Ledgerwood JE, Sullivan NJ, Graham BS. Chimpanzee adenovirus vector ebola vaccine–preliminary report. N Engl J Med 2015;373(8):776.

Li J-X, Hou L-H, Meng F-Y, Wu S-P, Hu Y-M, Liang Q, et al. Immunity duration of a recombinant adenovirus type-5 vector-based Ebola vaccine and a homologous prime-boost immunisation in healthy adults in China: final report of a randomised, double-blind, placebo-controlled, phase 1 trial. Lancet Glob Health 2017;5(3):e324–34.

Liu WB, Li ZX, Du Y, Cao GW. Ebola virus disease: from epidemiology to prophylaxis. Mil Med Res 2015;2:7.

Logue J, Tuznik K, Follmann D, Grandits G, Marchand J, Reilly C, et al. Use of the Filovirus Animal Non-Clinical Group (FANG) Ebola virus Immuno-assay requires fewer study participants to power a study than the Alpha Diagnostic International assay. J Virol Methods 2018;255:84–90.

Martin JE, Sullivan NJ, Enama ME, Gordon IJ, Roederer M, Koup RA, et al. A DNA vaccine for Ebola virus is safe and immunogenic in a phase I clinical trial. Clin Vaccine Immunol 2006;13(11):1267–77.

Marzi A, Engelmann F, Feldmann F, Haberthur K, Shupert WL, Brining D, et al. Antibodies are necessary for rVSV/ZEBOV-GP-mediated protection against lethal Ebola virus challenge in nonhuman primates. Proc Natl Acad Sci U S A 2013;110(5):1893–8.

Marzi A, Feldmann F, Geisbert TW, Feldmann H, Safronetz D. Vesicular stomatitis virus–based vaccines against lassa and ebola viruses. Emerg Infect Dis 2015a;21(2):305–7.

Marzi A, Halfmann P, Hill-Batorski L, Feldmann F, Shupert WL, Neumann G, et al. Vaccines. An Ebola whole-virus vaccine is protective in nonhuman primates. Science 2015b;348(6233):439–42.

Marzi A, Robertson SJ, Haddock E, Feldmann F, Hanley PW, Scott DP, et al. Ebola vaccine. VSV-EBOV rapidly protects macaques against infection with the 2014/15 Ebola virus outbreak strain. Science 2015c;349(6249):739–42.

Marzi A, Murphy AA, Feldmann F, Parkins CJ, Haddock E, Hanley PW, et al. Cytomegalovirus-based vaccine expressing Ebola virus glycoprotein protects nonhuman primates from Ebola virus infection. Sci Rep 2016;6:21674.

Matassov D, Marzi A, Latham T, Xu R, Ota-Setlik A, Feldmann F, et al. Vaccination with a highly attenuated recombinant vesicular stomatitis virus vector protects against challenge with a lethal dose of ebola virus. J Infect Dis 2015;212 Suppl 2:S443–51.

Meyer M, Garron T, Lubaki NM, Mire CE, Fenton KA, Klages C, et al. Aerosolized Ebola vaccine protects primates and elicits lung-resident T cell responses. J Clin Invest 2015;125(8):3241–55.

Milligan ID, Gibani MM, Sewell R, et al. Safety and immunogenicity of novel adenovirus type 26—and modified vaccinia ankara—vectored ebola vaccines: a randomized clinical trial. JAMA 2016;315(15):1610–23.

Mire CE, Geisbert JB, Marzi A, Agans KN, Feldmann H, Geisbert TW. Vesicular stomatitis virus-based vaccines protect nonhuman primates against Bundibugyo ebolavirus. PLoS Negl Trop Dis 2013;7(12)e2600.
Mire CE, Matassov D, Geisbert JB, Latham TE, Agans KN, Xu R, et al. Single-dose attenuated Vesiculovax vaccines protect primates against Ebola Makona virus. Nature 2015;520(7549):688–91.
Ohimain EI. Recent advances in the development of vaccines for Ebola virus disease. Virus Res 2016;211:174–85.
Pratt WD, Wang D, Nichols DK, Luo M, Woraratanadharm J, Dye JM, et al. Protection of nonhuman primates against two species of Ebola virus infection with a single complex adenovirus vector. Clin Vaccine Immunol 2010;17(4):572–81.
Pyankov OV, Bodnev SA, Pyankova OG, Solodkyi VV, Pyankov SA, Setoh YX, et al. A Kunjin Replicon virus-like particle vaccine provides protection against Ebola virus infection in nonhuman primates. J Infect Dis 2015;212 Suppl 2:S368–71.
Rampling T, Ewer K, Bowyer G, Wright D, Imoukhuede EB, Payne R, et al. A monovalent chimpanzee adenovirus Ebola vaccine—preliminary report. N Engl J Med 2015;374(17):1635–46.
Regules JA, Beigel JH, Paolino KM, Voell J, Castellano AR, Muñoz P, et al. A recombinant vesicular stomatitis virus ebola vaccine—preliminary report. N Engl J Med 2015;376(4):330–41.
Sarwar UN, Costner P, Enama ME, Berkowitz N, Hu Z, Hendel CS, et al. Safety and immunogenicity of DNA vaccines encoding Ebolavirus and Marburgvirus wild-type glycoproteins in a phase I clinical trial. J Infect Dis 2015;211(4):549–57.
Sridhar S. Clinical development of Ebola vaccines. Ther Adv Vaccines 2015;3(5–6):125–38.
Stanley DA, Honko AN, Asiedu C, Trefry JC, Lau-Kilby AW, Johnson JC, et al. Chimpanzee adenovirus vaccine generates acute and durable protective immunity against ebolavirus challenge. Nat Med 2014;20(10):1126–9.
Sullivan NJ, Sanchez A, Rollin PE, Yang ZY, Nabel GJ. Development of a preventive vaccine for Ebola virus infection in primates. Nature 2000a;408(6812):605–9.
Sullivan NJ, Sanchez A, Rollin PE, Yang ZY, Nabel GJ. Development of a preventive vaccine for Ebola virus infection in primates. Nature 2000b;408(6812):605–9.
Sullivan NJ, Geisbert TW, Geisbert JB, Xu L, Yang Z-Y, Roederer M, et al. Accelerated vaccination for Ebola virus haemorrhagic fever in non-human primates. Nature 2003;424(6949):681–4.
Sullivan NJ, Martin JE, Graham BS, Nabel GJ. Correlates of protective immunity for Ebola vaccines: implications for regulatory approval by the animal rule. Nat Rev Microbiol 2009;7(5):393–400.
Sullivan NJ, Hensley L, Asiedu C, Geisbert TW, Stanley D, Johnson J, et al. CD8+ cellular immunity mediates rAd5 vaccine protection against Ebola virus infection of nonhuman primates. Nat Med 2011;17(9):1128–31.
Swenson DL, Wang D, Luo M, Warfield KL, Woraratanadharm J, Holman DH, et al. Vaccine to confer to nonhuman primates complete protection against multistrain Ebola and Marburg virus infections. Clin Vaccine Immunol 2008;15(3):460–7.
Tapia MD, Sow SO, Lyke KE, Haidara FC, Diallo F, Doumbia M, et al. Use of ChAd3-EBO-Z Ebola virus vaccine in Malian and US adults, and boosting of Malian adults with MVA-BN-Filo: a phase 1, single-blind, randomised trial, a phase 1b, open-label and double-blind, dose-escalation trial, and a nested, randomised, double-blind, placebo-controlled trial. Lancet Infect Dis 2016;16(1):31–42.
Van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. Stat Med 2002;21(4):589–624.
Warfield KL, Swenson DL, Olinger GG, Kalina WV, Aman MJ, Bavari S. Ebola virus-like particle-based vaccine protects nonhuman primates against lethal Ebola virus challenge. J Infect Dis 2007;196 Suppl 2:S430–7.
Warfield KL, Goetzmann JE, Biggins JE, Kasda MB, Unfer RC, Vu H, et al. Vaccinating captive chimpanzees to save wild chimpanzees. Proc Natl Acad Sci U S A 2014;111(24):8873–6.
Warfield KL, Dye JM, Wells JB, Unfer RC, Holtsberg FW, Shulenin S, et al. Homologous and heterologous protection of nonhuman primates by Ebola and Sudan virus-like particles. PLoS One 2015;10(3)e0118881.
Wong G, Richardson JS, Pillet S, Patel A, Qiu X, Alimonti J, et al. Immune parameters correlate with protection against ebola virus infection in rodents and nonhuman primates. Sci Transl Med 2012;4(158)158ra146.
World Health Organisation. Correlates of vaccine-induced protection: methods and implications. 2013.
Wu X-X, Yao H-P, Wu N-P, Gao H-N, Wu H-B, Jin C-Z, et al. Ebolavirus vaccines: progress in the fight against Ebola virus disease. Cell Physiol Biochem 2015;37(5):1641–58.
Wu S, Kroeker A, Wong G, He S, Hou L, Audet J, et al. An adenovirus vaccine expressing Ebola virus variant makona glycoprotein is efficacious in guinea pigs and nonhuman primates. J Infect Dis 2016;214(Suppl. 3):S326–32.
Zhu F-C, Hou L-H, Li J-X, Wu S-P, Liu P, Zhang G-R, et al. Safety and immunogenicity of a novel recombinant adenovirus type-5 vector-based Ebola vaccine in healthy adults in China: preliminary report of a randomised, double-blind, placebo-controlled, phase 1 trial. Lancet 2015;385(9984):2272–9.
Zhu F-C, Wurie AH, Hou L-H, Liang Q, Li Y-H, Russell JBW, et al. Safety and immunogenicity of a recombinant adenovirus type-5 vector-based Ebola vaccine in healthy adults in Sierra Leone: a single-centre, randomised, double-blind, placebo-controlled, phase 2 trial. Lancet 2017;389(10069):621–8.

Appendix B:

Article.

"Adaptive protocols based on predictions from a mechanistic model of the effect of IL7 on CD4 counts"

Received: 26 January 2018 | Revised: 5 July 2018 | Accepted: 18 August 2018

DOI: 10.1002/sim.7957

RESEARCH ARTICLE

WILEY Statistics in Medicine

Adaptive protocols based on predictions from a mechanistic model of the effect of IL7 on CD4 counts

Laura Villain[1,2,3] | Daniel Commenges[1,2,3] | Chloé Pasin[1,2,3] | Mélanie Prague[1,2,3] | Rodolphe Thiébaut[1,2,3]

[1]University of Bordeaux, Inserm, Bordeaux, Population Health Research Center, Team SISTM, Bordeaux, France

[2]INRIA Bordeaux Sud Ouest, Talence, France

[3]Vaccine Research Institute (VRI), Hôpital Henri Mondor, Créteil, France

Correspondence
Rodolphe Thiébaut, University Bordeaux, Inserm, Bordeaux Population Health Research Center, Team SISTM, 33000 Bordeaux, France; or
INRIA Bordeaux Sud Ouest, 33405 Talence, France; or
Vaccine Research Institute (VRI), Hôpital Henri Mondor, 94010 Créteil, France.
Email: rodolphe.thiebaut@u-bordeaux.fr

Funding information
Ecole des Hautes Études en Santé Publique; École Normale Supérieure

In human immunodeficiency virus–infected patients, antiretroviral therapy suppresses the viral replication, which is followed in most patients by a restoration of CD4+ T cells pool. For patients who fail to do so, repeated injections of exogenous interleukin 7 (IL7) are experimented. The IL7 is a cytokine that is involved in the T cell homeostasis and the INSPIRE study has shown that injections of IL7 induced a proliferation of CD4+ T cells. Phase I/II INSPIRE 2 and 3 studies have evaluated a protocol in which a first cycle of three IL7 injections is followed by a new cycle at each visit when the patient has less than 550 CD4 cells/μL. Restoration of the CD4 concentration has been demonstrated, but the long-term best adaptive protocol is yet to be determined. A mechanistic model of the evolution of CD4 after IL7 injections has been developed, which is based on a system of ordinary differential equations and includes random effects. Based on the estimation of this model, we use a Bayesian approach to forecast the dynamics of CD4 in new patients. We propose four prediction-based adaptive protocols of injections to minimize the time spent under 500 CD4 cells/μL for each patient, without increasing the number of injections received too much. We show that our protocols significantly reduce the time spent under 500 CD4 over a period of two years, without increasing the number of injections. These protocols have the potential to increase the efficiency of this therapy.

KEYWORDS

adaptive protocols, HIV, interleukine 7, mechanistic models

1 | INTRODUCTION

Infection by the human immunodeficiency virus (HIV) leads to a decrease of the concentration of CD4+ T lymphocytes (CD4) associated with a deficiency of the immune system, which increases the risk of opportunistic infections.[1] With an effective combination antiretroviral treatment, the viral load becomes undetectable. The CD4 pool is then reconstituted in most cases.[2] However, this does not happen for some patients[3] who are low immunological responders. These patients, who present CD4 counts below 500 cells/μL of blood, have a lower life expectancy[4] and an increase of nonacquired immunodeficiency syndrome conditions, such as cancer or cardiovascular diseases.[5] To help the reconstitution of the CD4 pool, a treatment based on injections of exogenous interleukin 7 (IL7) has been experimented. The IL7 is a cytokine produced by thymus stromal cells and lymph nodes and is involved in the CD4 homeostasis.[6,7] Several trials

have demonstrated the safety and the beneficial effect of exogenous IL7 on immune markers.[8,9,10] Repeated cycles of three injections of IL7 have been evaluated with the aim of maintaining the CD4 counts above 500 cells/μL because, in this case, HIV infected subjects have about the same life expectancy as the general population.[11] In the phase I/II trials INSPIRE 2 and 3,[12] the repeated cycles of IL7 could indeed maintain CD4 concentration above the limit of 500 cells/μL most of the time, although the best adaptive protocol is yet to be determined.

Mechanistic models, based on ordinary differential equations (ODEs), have been applied to model different infectious diseases, eg, modeling of the HIV,[13-15] the hepatitis C,[16,17] or the human cytomegalovirus,[18] and other health-related processes.[19] They have also been used to understand and predict the effects of IL7. A first work quantified the effect of exogenous IL7 on the proliferation rate of CD4 cells and showed an additional effect on the cells' survival.[20] The model was then extended to fit repeated injections.[21] Thanks to a population approach with random effects, the models were able to predict future individual responses to new injections of IL7 with a very good accuracy. This opened the opportunity to individualize the strategy of IL7 administration.

Dynamical adaptation of the treatment as a function of the response of the patient has been proposed by Murphy[22] and Robins[23] who developed the optimal treatment regime theory. Many papers have followed in this field.[24-27] Methods based on semi parametric models and dynamic treatment regimens exist.[28] However, as Rich et al[28] underlined, these methods are not realistic enough and they often miss some important confounders. This issue can be solved by using mechanistic models.[29] When an ODE-based mechanistic model is available, the modeled treatment can be adapted using this model. This has been described by Rosenberg et al[30] for the supervised treatment interruption strategies, or in the pharmacokinetic-pharmacodynamic field.[17,31] The optimal control theory can be applied for globally optimizing the treatment regime, which has been proposed by Castiglione and Piccoli,[32] and applied to optimizing the treatment of HIV infected patients.[33,34] However, as noted by Chakraborty and Murphy[35] these works do not sufficiently take into account the statistical issues of the problem, ie, model parameters have to be estimated and for efficient estimation and random effects have to be introduced in the statistical model. Such random effect mechanistic model has been applied to tune the dose of an antiretroviral treatment by Prague et al.[36]

In the present paper, we aim to find efficient adaptive protocols for IL7 administration based on predictions from a random effect mechanistic model proposed in the work of Jarne et al.[21] Here, we propose realistic protocols that shorten the time spent under the limit of 500 CD4 cells/μL and limit the number of IL7 injections. Two approaches will be compared, ie, adapting the criterion for a new cycle based on the risk of falling under 500 CD4 cells/μL before the next visit, and adapting the times of control visits. Both of these approaches are based on predictions generated with our random effect mechanistic model at relatively short term to locally optimize the protocols. This is less ambitious than optimal control but is more feasible and the proposed protocols could soon be proposed to real patients. In both approaches, we may or may not adapt also the number of injections per cycle, leading to four possible protocols.

This paper is organized as follows. Section 2 describes the data from INSPIRE 1, 2, and 3, and the design of the protocol for repeated cycles. Section 3 presents the mathematical and statistical features of the model. In Section 4, we describe two prediction-based adaptive protocols and their two variants adapting the number of injections per cycle or not. Section 5 presents the simulation study and its results, while Section 6 shows what the method would have predicted for real data. Section 6 concludes this paper.

2 | DATA

The data used for the mechanistic model are drawn from the INSPIRE 1, 2, and 3 studies.[8,12] These studies evaluated the effect of injections of IL7 on the CD4 concentration in low immunological responders aged 18 years or more. The patients were included in the study if they were under stable combination antiretroviral treatment for at least one year, had CD4 counts between 100 and 400 cells/μL of blood, and had an undetectable viral load for at least 6 months before the beginning of the protocol.

The first study, INSPIRE 1 (initially called simply "INSPIRE"), evaluated the effect of one cycle of injection, which is defined as three injections with one week between each one. Three doses (10, 20, and 30 μg/Kg) were tested, and a placebo was included, for a total of 21 patients. The INSPIRE 2 and 3 studies evaluated the effect of repeated cycles of injections, using only the 20 μg/Kg, which was determined to be the most effective without too many side effects.

Overall, the data from 128 patients are used, with regular measurements of the CD4 counts and the marker of proliferation Ki67. The patients had visits at weeks 1, 2, 3, 4, 5, 6, 9, and 12, and then visited every 3 months. The CD4 counts were measured at each visit, while Ki67 counts were measured only at weeks 1, 2, 3, 5, and 12. For the repeated cycles,

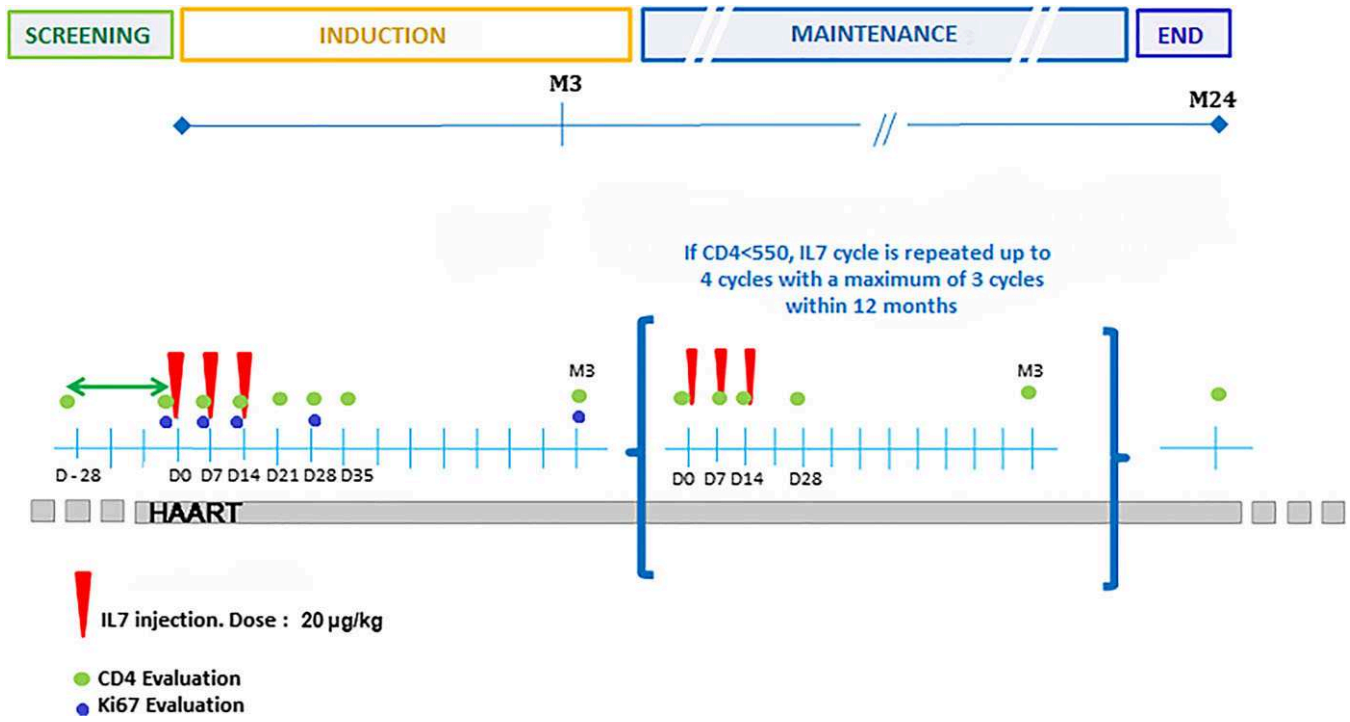


FIGURE 1 Design of INSPIRE 2 study. The screening phase determines the concentrations of CD4 at baseline to include or not the patient. The induction phase starts with a first cycle of three interleukin 7 (IL7) injections at weeks 1, 2, and 3. The maintenance phase then controls the patient every 3 months with a new cycle if the patient is under 550 CD4. The green dots represent the measures of CD4 counts, and blue dots the measures of Ki67 [Colour figure can be viewed at wileyonlinelibrary.com]

the maintenance phase consisted in repeated visits every 3 months. When the CD4 counts were below 550, a new cycle of injections was administered. The durations of the studies were 12, 24 and 21 months for INSPIRE, INSPIRE 2, and 3, respectively. The design of INSPIRE 2 is presented in Figure 1.

The high heterogeneity of the data in terms of number of injections per cycles, number of cycles received and dose, and the availability of regular measurements of different immunological markers makes this an interesting dataset for modeling. Overall, 197 cycles were administered for 128 patients, with 41 incomplete cycles (one or two injections instead of three). Because the patients had different trajectories of CD4 counts, they had different times of injections in the maintenance phase.

3 | MECHANISTIC MODEL FOR IL7 TREATMENT

3.1 | Modeling of the effect of IL7 injections on CD4 concentration

A mechanistic model for the evolution of the CD4 concentration after repeated IL7 injections was proposed in the work of Jarne et al.[21] In this paper, a two compartment model was described, with the compartment P for the proliferating cells and the compartment Q for the quiescent cells. This model and the different effects on the parameters have been selected with the approximate Likelihood Cross Validation criteria (LCVa),[37] as described in the works of Thiebaut et al[20] and Jarne et al.[21] The mathematical structure of the model is written as

$$\begin{cases} \frac{dQ}{dt} = \lambda + 2\rho P - \pi Q - \mu_Q Q \\ \frac{dP}{dt} = \pi Q - \rho P - \mu_P P. \end{cases}$$

We allow the parameters to be different from one subject to another. Thus, we denote by $\boldsymbol{\xi^i}$ the vector of parameters of the ODE system, ie, $\boldsymbol{\xi^i} = [\lambda^i, \rho^i, \pi^i, \mu_Q^i, \mu_P^i]$ for patient i ($i = 1, \dots, n$) and by $X^i = (Q(t, \boldsymbol{\xi^i}), P(t, \boldsymbol{\xi^i}))$ his state vector. All of the parameters are positive because they are rates of proliferation, production, and death for cells. Hence, we use a log transformation denoted by a tilde, ie, $\tilde{\xi}_l^i = \log(\xi_l^i)$. The meaning and units of each parameter are detailed in Table 1.

TABLE 1 Parameters of the model and their estimates from the work of Jarne et al[21] [Correction added on 4 October 2018, after first online publication: the Name and Unit in rows 6-8 and 10 have been corrected]

Parameter	Name	Units	Estimate mean (sd)
λ	Production rate	cells.days^{-1}	5.32 (0.33)
ρ	Reversion rate	days^{-1}	2.44 (0.23)
π	Proliferation rate	days^{-1}	0.06 (0.004)
μ_P	Death rate of P cells	days^{-1}	0.07 (0.005)
μ_Q	Death rate of Q cells	days^{-1}	0.08 (0.02)
β_{π_1}	Effect of IL7 on π (Injection 1)	days^{-1}.μg^{-1}	0.93 (0.04)
β_{π_2}	Effect of IL7 on π (Injection 2)	days^{-1}.μg^{-1}	0.71 (0.04)
β_{π_3}	Effect of IL7 on π (Injection 3)	days^{-1}.μg^{-1}	0.23 (0.04)
β_{μ_Q}	Effect of IL7 on μ_Q	days^{-1}.μg^{-1}	-0.08 (0.006)
β_C	Cycle effect of IL7	days^{-1}	-0.16 (0.02)
σ_λ	Standard deviation of random effect on λ	cells.days^{-1}	0.24 (0.03)
σ_ρ	Standard deviation of random effect on ρ	days^{-1}	0.52 (0.08)
σ_1	Noise parameter on CD4 cells	cells$^{0.25}$	0.29 (0.003)
σ_2	Noise parameter on P cells	cells$^{0.25}$	0.28 (0.02)

Abbreviations: IL7, interleukin 7.

The initial condition for subject i is the equilibrium point, ie, $\frac{dQ}{dt}(0,\boldsymbol{\xi^i}) = 0$, $\frac{dP}{dt}(0,\boldsymbol{\xi^i}) = 0$, which gives the initial points $Q(0,\boldsymbol{\xi^i}) = \frac{\lambda^i(\rho^i+\mu_P^i)}{\pi^i(\mu_P^i-\rho^i)+\mu_Q^i(\rho^i+\mu_P^i)}$, $P(0,\boldsymbol{\xi^i}) = \frac{\lambda^i\pi^i}{\pi^i(\mu_P^i-\rho^i)+\mu_Q^i(\rho^i+\mu_P^i)}$.

A patient-by-patient inference is inefficient if there is not enough information for each subject; hence, we need a statistical model of the variability of the parameters using both explanatory variables with fixed effects and random effects. It can be written as

$$\begin{cases} \tilde{\lambda}^i(t) = \tilde{\lambda}_0 + l_i \\ \tilde{\pi}^i(\mathrm{t}) = \tilde{\pi}_0 + \left[\beta_C \mathbb{1}_{\{C^i(t)>1\}} + \sum_{k=1}^{3} \mathbb{1}_{\{N_t^i=k\}}\ \beta_{\pi_k}\ d_i^{0.25}\right]\ \mathbb{1}_{\{N_t^i - N_{t-7}^i=1\}} \\ \tilde{\rho}^i(t) = \tilde{\rho}_0 + r_i \\ \tilde{\mu_Q^i}(t) = \tilde{\mu}_{Q_0} + \beta_{\mu_Q} f^i(t) d_i^{0.25} \\ \tilde{\mu_P^i}(t) = \tilde{\mu}_{P_0}, \end{cases} \tag{1}$$

where l_i and r_i are normally distributed random effects $l_i \sim \mathcal{N}(0,\sigma_\lambda^2)$ and $r_i \sim \mathcal{N}(0,\sigma_\rho^2)$; $C^i(t)$ counts the number of cycles and β_C represents the cycle effect, meaning that, after one cycle, the effect of IL7 on the proliferation can be lower; β_{π_k} is the effect of the injection on the proliferation, k being 1, 2, or 3 depending on whether the injection is the first, second, or third of the cycle. The effects on π are constant during 7 days after each injection, and they then disappear. N_t^i counts the number of injections that patient i has received until time t; thus, $\mathbb{1}_{\{N_t^i - N_{t-7}^i=1\}}$ is an indicator function taking the value 1 if an injection was administered in the last 7 days. Let T_t^i be the time of the last injection received by the patient i at the time t; the effect on μ_Q is represented by $\beta_{\mu_Q} f(t)$, with f written as

$$f^i(t) = \begin{cases} 1, & \text{if } 2 + T_t^i < t \le 360 + T_t^i \\ 1 - (t-360)/360, & \text{if } 360 + T_t^i < t \le 720 + T_t^i \\ 0, & \text{if } 720 + T_t^i < t. \end{cases}$$

In our data, the variables P and Q are not directly observed. The observations correspond to the total number of CD4 and the number of proliferating cells; hence, we also need an observation model. The observation of M biomarkers are related to the solution of the ODE with the function g, ie, $Y_{mq}^i = g_m(\boldsymbol{\xi}^i, t_q)^{0.25} + \epsilon_{mj}^i$, $m = 1, \dots, M$. Here, $M = 2$, as we observe the CD4 counts and the Ki67 counts. Denoting by Y_{1j}^i the fourth root of the CD4 counts, and by Y_{2k}^i, the Ki67 counts for patient i at times t_{ij} ($j = 1, \dots, J^i$) and t_{ik} ($k = 1, \dots, K^i$), respectively. This gives the following observation model:

$$\begin{cases} \mathrm{Y}_{1j}^i = \left[P\left(t_{ij},\boldsymbol{\xi}^i\right) + Q\left(t_{ij},\boldsymbol{\xi}^i\right)\right]^{0.25} + \epsilon_{1j}^i \\ \mathrm{Y}_{2k}^i = P\left(t_{ik},\boldsymbol{\xi}^i\right)^{0.25} + \epsilon_{2k}^i. \end{cases} \tag{2}$$

We assume that noises variables are normally distributed, ie, $\epsilon^i_{1j} \sim \mathcal{N}(0, \sigma_1^2)$, $\epsilon^i_{2k} \sim \mathcal{N}(0, \sigma_2^2)$. This parameter includes the measurement error and the biological variation not taken into account in the model, eg, the diurnal variation of the CD4 concentration. The fourth root transformation has been applied to make homoscedasticity and normality acceptable.[20]

3.2 | Inference and estimation

To estimate the parameters of the model, Jarne et al[21] used the NIMROD program.[38] The implemented estimation method is based on penalized log-likelihood maximization. This is a quasi-Bayesian approach in the sense that penalization is built from a priori values for the parameters found in the literature. Then, this penalized log-likelihood is maximized using a Newton-Raphson like algorithm, called robust variance scoring.[39] The iterative algorithm stops when the relative distance to maximum criterion is satisfied. The results of this inference are displayed in Table 1. There are enough observations to consider that the posterior distributions are close to normal distribution (in virtue of the Bernstein-von Mises theorem[40]), so that results in Table 1 can be interpreted as summarizing the (marginal) posterior distributions of the parameters by their expectations and their standard deviations.

4 | PREDICTION-BASED ADAPTIVE PROTOCOLS

4.1 | General method

In the original protocol of INSPIRE 2 and 3 studies, which is called the "original" (ORI) protocol, patients start a first cycle of injections, with CD4 counts measurements at the times of injections. There are visits every 3 months. A new cycle is administered if the CD4 counts are below 550 CD4. The aim is to prevent CD4 concentration to fall under 500. Thus, the original protocol is already an adaptive protocol, but because the patients have different CD4 dynamics, this fixed criterion is not always appropriate. If a patient tends to return quickly to his or her baseline concentration, then the margin can be too small and the decision not to administer a new cycle can lead to cross the limit of 500 CD4 shortly after the control. In contrast, for some patients, the CD4 concentration decreases slowly after an injection, and the criterion of 550 CD4 for a new cycle may be too high, which results in unnecessary cycles and visits. Here, we propose protocols that are based on the prediction that can be done with a mechanistic model with the aim of decreasing the time spent under the limit of 500 CD4, while controlling the number of IL7 injections.

We use a mechanistic model that fits the dynamics of CD4 following a cycle of IL7 injections; the fixed effects parameters have been estimated using all observations of the three INSPIRE studies. For the mixed effect parameters, λ and ρ, we have an estimation of the mean of the parameters on the population and the variance of their random effects (σ^2_λ and σ^2_ρ). With these estimates and the information for the patient i available at the time k, called $H_i^{t_k}$, we used an MCMC algorithm to sample the posterior distribution of the individual parameters λ^i and ρ^i. For given values of the parameters of a patient, we can predict the evolution of his or her CD4 concentration.[36] Taking into account the uncertainty on the parameters, we can also compute the distribution of any quantity related to the future CD4 concentration dynamic. This prediction can be used to adapt the treatment. Every time that the patient comes for a control visit, we have access to new data so that $H_i^{t_k} \subset H_i^{t_{k+1}}$; thus, the prediction is more precise as time goes on.

The algorithm used for sampling the random effects of the patient is a Metropolis within Gibbs.[41] At each iteration, λ and ρ are successively sampled. In the Metropolis part of the algorithm, the instrumental function used for the first estimation of the protocol (3-month control) is the posterior law estimated with NIMROD. The standard error of the prior for λ and ρ is the one estimated with NIMROD (respectively, $sd_\lambda = 0.33$ and $sd_\rho = 0.23$; see Table 1) added with the standard error associated with the random effect (respectively, $\sigma_\lambda = 0.24$ and $\sigma_\rho = 0.52$; see Table 1), as the variability comes from both the error of estimation and the interindividuals variability. Then, for each control visit after the first one, the distribution given by the previous MCMC is used as the new prior. The likelihood used in the MCMC procedure is[42]

$$L_i = \prod_{m=1}^{M} \prod_{j=1}^{K_m} \frac{1}{\sigma_m \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{Y^i_{mj} - g_m\left(t_j, \xi^i\right)^{0.25}}{\sigma_m} \right)^2 \right].$$

As presented in the work of van der Vaart,[40] according to Doob's theorem, the distribution of the parameters $\widehat{\xi^i}$ with the data from patient i at time t_k ($H_i^{t_k}$) converges to the Dirac of the true value of the parameters, $\delta(\xi^i)$, when k tends to infinity.

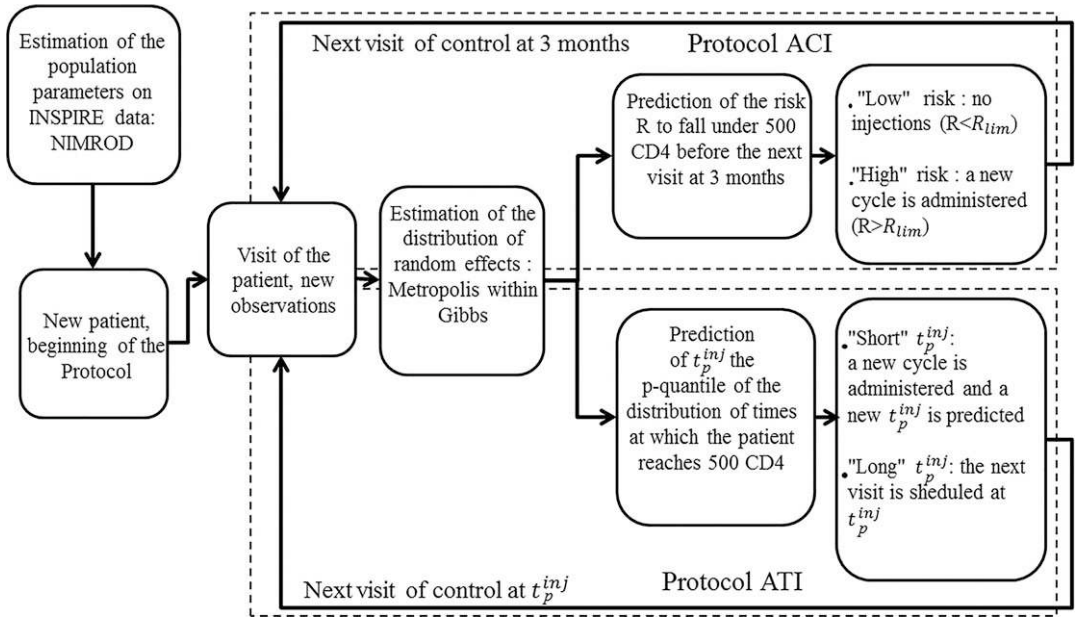


FIGURE 2 Flowchart of the two protocols. Adaptive criterion of injection (ACI) protocol: adaptive criterion protocol, the visits are every 3 months and the decision to administer a new cycle is based on the predicted risk R to fall under 500 CD4 before the next visit. Adaptive time of injection (ATI) protocol: adaptive times protocol, the times of visit are predicted based on the time at which the patient is supposed to reach the 500 CD4 limit, and a new cycle is administered if this predicted time is too short

We propose two prediction-based adaptive approaches; the first is based on an adaptive criterion of injections (ACIs); the second is based on an adaptive time of injections (ATIs). Both approaches have a variant where the number of injections per cycle can be adapted. Figure 2 presents the flowchart of the two approaches.

4.2 | Prediction ability on real data

Before developing and studying prediction-based adaptive protocols, it is essential to study the prediction ability of our model. In this aim, we randomly excluded 10 patients from the data and estimated the parameters of the model on the reduced data set. For each excluded patient, we ran the MCMC algorithm using the estimates from the diminished data set. For an excluded patient i, at each time of control (every 3 months), the MCMC algorithm gives a distribution of CD4 concentration at every time of observation t_j. At each iteration q of the MCMC, we can generate the predicted distribution of observation at each time t_{ij} by computing $\text{CD4}^i_{jq} = P(t_{ij}, \xi^{iq}) + Q(t_{ij}, \xi^{iq})$, and adding a noise variable, ie, $\text{Y}^i_{1jq} = (\text{CD4}^i_{jq})^{0.25} + \epsilon^i_{1jq}$. Figure 3 shows examples of these predicted distributions for four (out of the 10) patients. The black line represents the mean of the future trajectories and the light blue band represents 95% credible intervals; the dark blue band represents 95% predictive interval of observations, and the black dots are the observed values. The 95% credible and predictive intervals were computed by excluding the 2.5% most extreme values of each side of the distribution. If the model is well calibrated, then the black dots should be inside the dark blue band, which is the case here.

To show the good calibration of the predicted distributions, a quantile analysis was done. The process of studying the prediction for 10 excluded patients was repeated 10 times leading to a total of 100 patients, and we analyzed the distributions of the observed data with respect to the predicted distributions by a quantile-quantile plot. The quantiles of the predicted distributions should be the same as the observed quantiles, which was indeed essentially the case, as shown in the quantile-quantile plot presented in the Web Supplementary 1.

4.3 | Protocol with ACIs

The ACI protocol is similar to the original protocol, ie, the patients come every 3 months for a control visit. However, instead of using the fixed criterion of CD4 counts below 550, we predict the risk R that the CD4 concentration will fall

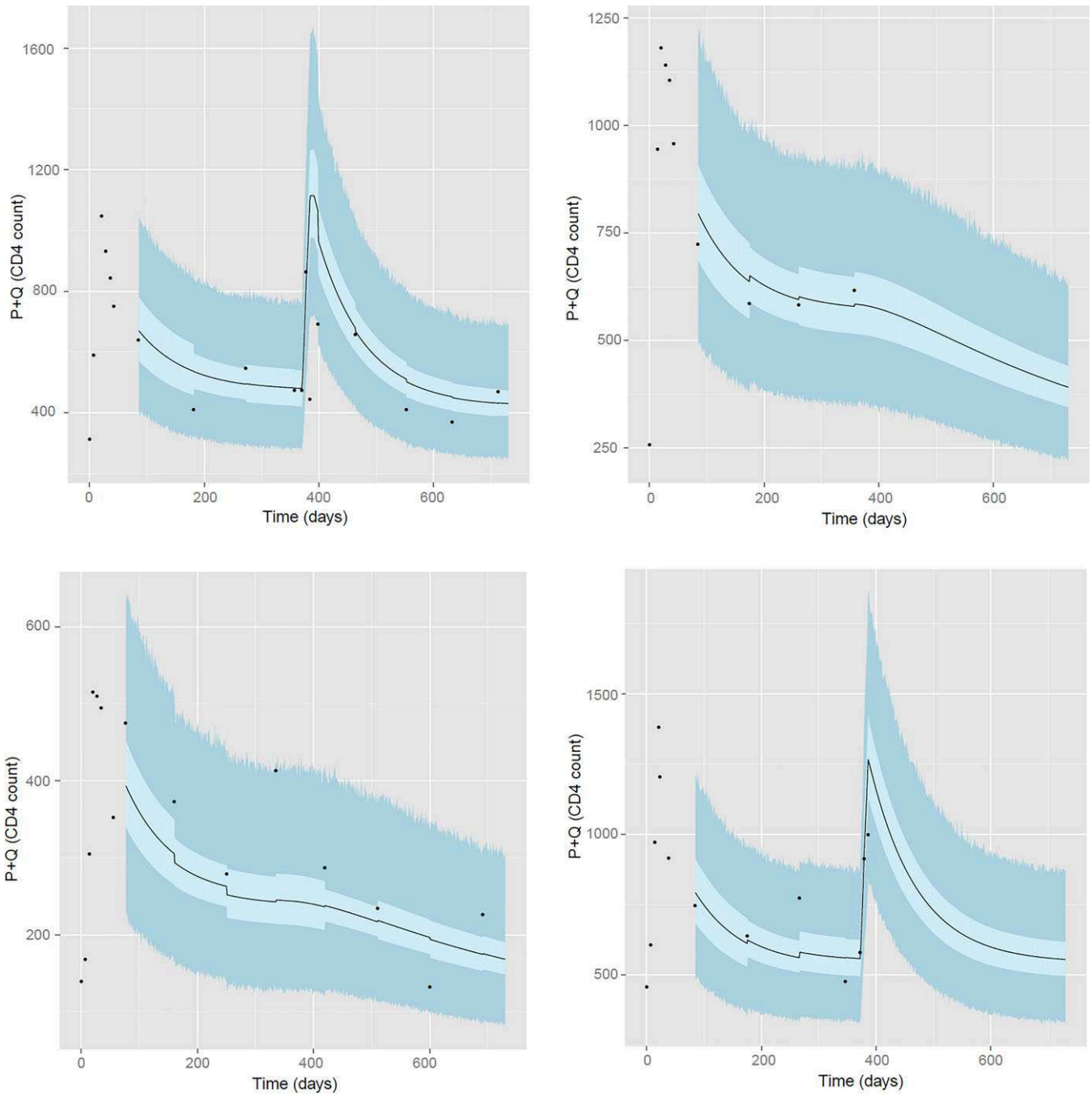


FIGURE 3 Prediction intervals for four random patients. Light blue band: 95% credible interval of trajectories. Dark blue band: 95% predictive interval of observations. Black line: mean of predicted trajectories. Black dots: real-data observations [Colour figure can be viewed at wileyonlinelibrary.com]

below 500 before the next visit. With the distribution of parameters given by the MCMC algorithm presented in Section 4.1, we can directly have the distribution of CD4 concentration at 3 months, which allows us to compute R

$$R\left(\widehat{\boldsymbol{\xi}}^i|H_i^{t_k}\right) = \mathrm{P}\left(\mathrm{g}_1\left(t_k + t_{\text{visit}}, \widehat{\boldsymbol{\xi}}^i\right) > 500|H_i^{t_k}\right),$$

where t_{visit} is the time between two controls (3 months in INSPIRE studies). If R is larger than a limit risk called R_{lim} (for instance, 10%), a new cycle of injections is administered. If $R < R_{\text{lim}}$, then the patients simply comes back 3 months later. At each visit, the MCMC algorithm is done with new data, and the decision is made with the value of the risk R.

4.4 | Protocol with ATIs

In the ATI protocol, the injections times are adaptive. The patient starts his or her first cycle and then comes back 3 months later. We denote t^{inj} such as $g_1(t^{\text{inj}}, \widehat{\xi}^i) = 500$, the time at which the patient will reach the 500 CD4 limit. Then, with the MCMC algorithm, we sample the distribution of t^{inj} and select t^{inj_p} as the p quantile (for instance the 0.1 quantile) of this distribution, ie, $\mathscr{P}(t^{\text{inj}} < t_p^{\text{inj}}) = p$. If t_p^{inj} is larger than a fixed limit (here, 1 month), the patient comes for a new visit at the time t_p^{inj}. If not, then a new cycle is administered immediately and the next time of visit is computed again at the end of the cycle. The time between two cycles is chosen here to be at least 1 month because it is clinically relevant.

4.5 | Adaptive number of injections: ACIC and ATIC protocols

Up to now, the proposed protocols used cycles of three injections. In the work of Jarne et al,[21] the usefulness of the third injection was questioned. The estimation of the effect of this injection, $\beta_{\pi 3}$, showed that the impact of this injection on the proliferation rate is much smaller compared to the first two injections. A comparison of the simulation of protocols showed that the original protocol with only two injections per cycle reduced the number of injections without impacting the time spent under 500 CD4 or the mean number of CD4. We propose a modification of the ACI and ATI protocols for adapting the number of injections per cycle, a C is added in the end of their names to represent the adaptation inside a cycle, leading to the ACIC and ATIC protocols.

- Protocol ACIC: When a decision for a new cycle is taken, the aim is to decide if the decision to diminish the number of injections will change the decision to inject or not at the next time of control. To do this, we compare the risks to fall under 500 CD4 at the next control visit for a cycle of three, two, or one injection, ie, $R(\widehat{\xi}_j^i | H_i^{t_k}) = \mathscr{P}(g_1(\widehat{\xi}_j^i, t_k + 2t_{\text{visit}}) > 500 | H_i^{t_k})$, with $j\,(j = 1, \dots, 3)$ being the number of injections of the cycle, and $\widehat{\xi}_j^i$ depending on this number of injection as presented in Equation 1. If in any case a cycle of injections would be necessary at the next time of visit, meaning that $R(\widehat{\xi}_3^i | H_i^{t_k})$ is superior to the chosen R_{lim}, then we consider that the patient needs a cycle of three injections. If this is not the case, and if both $R(\widehat{\xi}_2^i | H_i^{t_k})$ and $R(\widehat{\xi}_1^i | H_i^{t_k})$ are inferior to R_{lim}, then a cycle of one injection is administered. If $R(\widehat{\xi}_2^i | H_i^{t_k})$ is inferior to R_{lim} but not $R(\widehat{\xi}_1^i | H_i^{t_k})$, a cycle of two injections is administered. Moreover, if only $R(\widehat{\xi}_3^i | H_i^{t_k})$ is inferior to R_{lim}, a cycle of three injections is administered.
- Protocol ATIC: When a decision for a new cycle is taken, the next t_{inj^*} is calculated for one, two or three injections: t^{inj1_p}, t^{inj2_p} and t^{inj3_p}. The relative difference between two and three injections is calculated, ie, $d_3 = \frac{t^{\text{inj2}_\text{p}} - t^{\text{inj3}_\text{p}}}{t^{\text{inj3}_\text{p}}}$. If d_3 is superior to d_{lim}, then three injections are administered. If not, then the same process is repeated to choose between one or two injections, ie, $d_2 = \frac{t^{\text{inj1}_\text{p}} - t^{\text{inj2}_\text{p}}}{t^{\text{inj2}_\text{p}}}$, and if d_2 is superior to d_{lim}, here taken at 10% as it is clinically relevant, then two injections are administered; if not, then one injection is administered.

5 | SIMULATION

5.1 | General description

We simulated the different protocols for 150 "pseudopatients" on a 2-year period. The parameter values of these patients were sampled from the posterior distribution of the parameters estimated with NIMROD over the 138 patients of the INSPIRE studies. Because λ and ρ vary between patients, random effects were generated for all patients for these parameters, using the estimated variance of the random effects on the population. Moreover, we applied the inclusion criterion of the INSPIRE studies, keeping only those patients with baseline CD4 counts between 100 and 400. The simulations of the trajectories were done with R with the DeSolve package.[43] This package numerically solves the ODE for a given set of parameter values ξ^i, using the "lsodes" method, an interface to the FORTRAN ODE solver bearing the same name.[44] Observations were generated at times (0, 7, 14, 21, and 55) before the first time of control (day 90) by adding a noise variable (according to Equation (2)) to the value of the trajectories at these times. Each time that we make a decision, the observations are generated to take into account the decision. If a cycle is administered, then the next observations are at the time of injections and at the next time of control; if not, then the next generated observation is only at the next time of control. The total number of CD4 and the number of proliferating cells are observed each time.

For the ACI and ACIC protocols, the values of the risk limit R_{lim} used were 0.05, 0.1, and 0.2. For the ATI and ATIC protocols, we used the 0.05, 0.1, and 0.2 quantiles to obtain t_p^{inj}. The MCMC algorithm was performed with a total of 5000 iterations each time, with a burn-in phase of 1000 and a thinning of 2. The convergence of the chain was controlled with the Heidel diagnostic; the results of the convergence are presented in Web Supplementary 2.

5.2 | Results for the adaptive protocols

To compare the different protocols, we used eight criteria, ie, the mean number of CD4 over the protocol, the time spent under 500 CD4, the number of cycles administered, the number of visits (including the visits for the injections inside a cycle), the number of injections, and the number of cycles of one, two, and three injections (the number of cycles of three injections includes the first cycle, which is not decided by the protocol because it is the start of the protocol and automatically of three injections). The results are presented in Table 2. We considered the value of R_{lim}(for ACI and ACIC) and the p quantile (for ATI and ATIC) as "threshold" because they both represent the limit value for the decision.

Compared to the original protocol (ORI), the proposed protocols highly reduce the time spent under 500 CD4, while the number of visits is comparable (slightly increased for the ACI protocol, and decreased for the ATI and ATIC protocols). The number of injections is increased for the ACI protocol but this increase is not as important when the number of injections per cycle is also adapted (ACIC) while the time spent under 500 CD4 is similar in both ACI and ACIC protocols. The number of injections is similar between the ATI and ORI protocols, while it is reduced by the ATIC protocol.

The mean of CD4 is higher for the ACI and ACIC protocols than the ATI and ATIC protocols (which are similar to the ORI protocol), because the aim of the ATI and ATIC protocols is to start a cycle just before the patient reaches 500 CD4, while the ACI and ACIC protocols have visits of control only every 3 months, which means that the decision to start a new cycle can be taken while the patient could wait before he or she reaches 500 CD4 but not 3 months. This means that, when a new cycle is started, we expect CD4 concentration to be around 500 in the ATI and ATIC protocols, while these concentrations can be higher in the ACI and ACIC protocols. This induces a higher mean of CD4 for ACI and ACIC protocols. This is consistent with the choice of the criterion of interest, namely, the time spent with CD4 concentrations under 500. This choice is justified by clinical results[11] and consistency with INSPIRE studies. However, our method can be modified by using other criteria. If it is clinically relevant to consider the mean number of CD4, it could be easily implemented in the decision criterion for adapting the protocol.

The sensitivity analysis on the threshold parameter shows that for the ACI and ACIC protocols, the risk R_{lim} does not significantly impact the results in any of the criteria. For the ATI and ATIC protocols, the p-quantile at 0.05 does not increase the frequency of cycles or visits, but it does reduce the time spent under 500 CD4 on average. The p-quantile 0.2 is no better because it increases the time spent under 500 CD4 but does not reduce the number of cycles.

Figure 4 presents the boxplot of the time spent under 500 CD4, the number of visits, the mean of CD4, and the number of injections for each protocol at the threshold 0.05. This shows that the median of time spent under 500 CD4 is at 0 for each of the proposed protocols. The ATI and ATIC protocols have higher third quantiles than the ACI and ACIC protocols, but the outliers are smaller. This means that the ATI and ATIC protocols are more suitable for patients with difficulties to maintain their CD4 concentrations above 500. Indeed, the ATI and ATIC protocols allow more frequent visits for patients with a fast decrease of CD4 after the end of a cycle compared to the ACI and ACIC protocols, where a minimum delay of 3 months between visits has to be respected. Globally, it is clear that all four protocols have better results for the time spent under 500 CD4 than the ORI protocol. We can also see that all protocols have similar distributions for the three other criteria, but the ACI and ACIC induce higher CD4 means, and the ATIC is lower in number of visits and number of injections.

Figure 5 shows the plot of the dynamic of CD4 for three protocols (ORI, ACI, and ATI) with the threshold 0.1 for two pseudopatients. In this figure, the real trajectory is represented by a dark line and the simulated observations are the black dots. The ORI protocol uses those observations for the decision to administer a new cycle with the criterion of 550 CD4. The proposed adaptive protocols predict a distribution of the random effects, which gives a distribution of CD4 at each time point using those observations. The adaptive protocols use those predicted distributions for the decisions while the ORI protocol relies only on the observation at the current time of control. The 95% credible interval of CD4 is represented by the light blue band, and the decision taken by the proposed adaptive protocols are based on this prediction. For each of those distributions of CD4, the distribution of the observations is also predicted by adding a noise variable to the predicted CD4, as it was done in Section 4.2. The 95% predictive interval of observations is represented by the dark blue band. At each time of control, new information is available; it induces adjustment of the prediction of the random effects, the predicted distribution of CD4, and the predicted observations.

TABLE 2 Comparison of the protocols original (ORI), adaptive criterion of injection (ACI), ACIC, adaptive time of injection (ATI), and ATIC. Eight criteria are presented: the mean number of CD4 (CD4), the time spent under 500 CD4 (T500), the number of cycle (Nb Cycles), the number of visits (Nb visits), the number of injections (Nb inj), and the number of cycle of, respectively, one, two, and three injections (C1, C2, and C3). These are calculated per patient over the two year protocols, and their means and quartiles (or standard error for the mean of CD4) are presented. "Threshold" means R_{lim} (for ACI and ACIC) or p-quantile (for ATI and ATIC)

Protocol	Threshold	CD4 mean (sd)	T500 (days) mean [Q1;Q3]	Nb Cycles mean [Q1;Q3]	Nb visits mean [Q1;Q3]	Nb inj mean [Q1;Q3]	C_1 mean [Q1;Q3]	C_2 mean [Q1;Q3]	C_3 mean [Q1;Q3]
ORI	-	722 (112)	107 [6;173]	4.5 [3;6]	18 [15;20]	13 [9;18]	0	0	4.5 [3;6]
ACI	0.05	882 (146)	18[0;7]	5.9 [4;8]	20 [17;24]	18 [12;24]	0	0	5.9 [4;8]
	0.1	866 (143)	19 [0;18]	5.8 [4;8]	20 [17;24]	17 [12;24]	0	0	5.8 [4;8]
	0.2	837(137)	21 [0;25]	5.6 [4;8]	20 [17;24]	17 [12;24]	0	0	5.6 [4;8]
ACIC	0.05	853 (124)	17 [0;7]	6.1 [5;8]	18 [12;24]	15 [8;24]	0.8 [0;2]	1.3 [0;2]	4.0 [1;8]
	0.1	830 (146)	19[0;20]	6.0 [5;8]	18 [12;24]	15 [7;24]	0.9 [0;2]	1.2 [0;2]	3.9 [1;8]
	0.2	804 (130)	23[0;26]	5.8 [4;8]	17 [11;23]	14 [7;23]	0.9 [0;2]	1.1 [0;2]	3.8 [1;7]
ATI	0.05	785 (71)	13[0;18]	5.0 [4;7]	16 [13;20]	15 [12;21]	0	0	5 [4;7]
	0.1	764(67)	19 [0;32]	4.9 [3;6]	15 [12;18]	15 [9;18]	0	0	4.9 [3;6]
	0.2	743 (64)	28[0;48]	4.7 [3;6]	15 [11;18]	14 [9;18]	0	0	4.7 [3;6]
ATIC	0.05	769 (72)	16 [0;24]	5.4 [4;7]	12 [10;14]	12 [8;14]	0.6 [0;1]	3.5 [2;5]	1.4 [1;2]
	0.1	744 (71)	23[0;33]	5.1 [4;7]	12 [9;14]	11 [8;14]	0.6 [0;1]	3.3[2;5]	1.3 [1;2]
	0.2	716 (72)	36 [0;60]	4.9 [3;7]	12 [9;14]	11 [6;14]	0.5 [0;1]	3.0 [1;5]	1.4 [1;2]

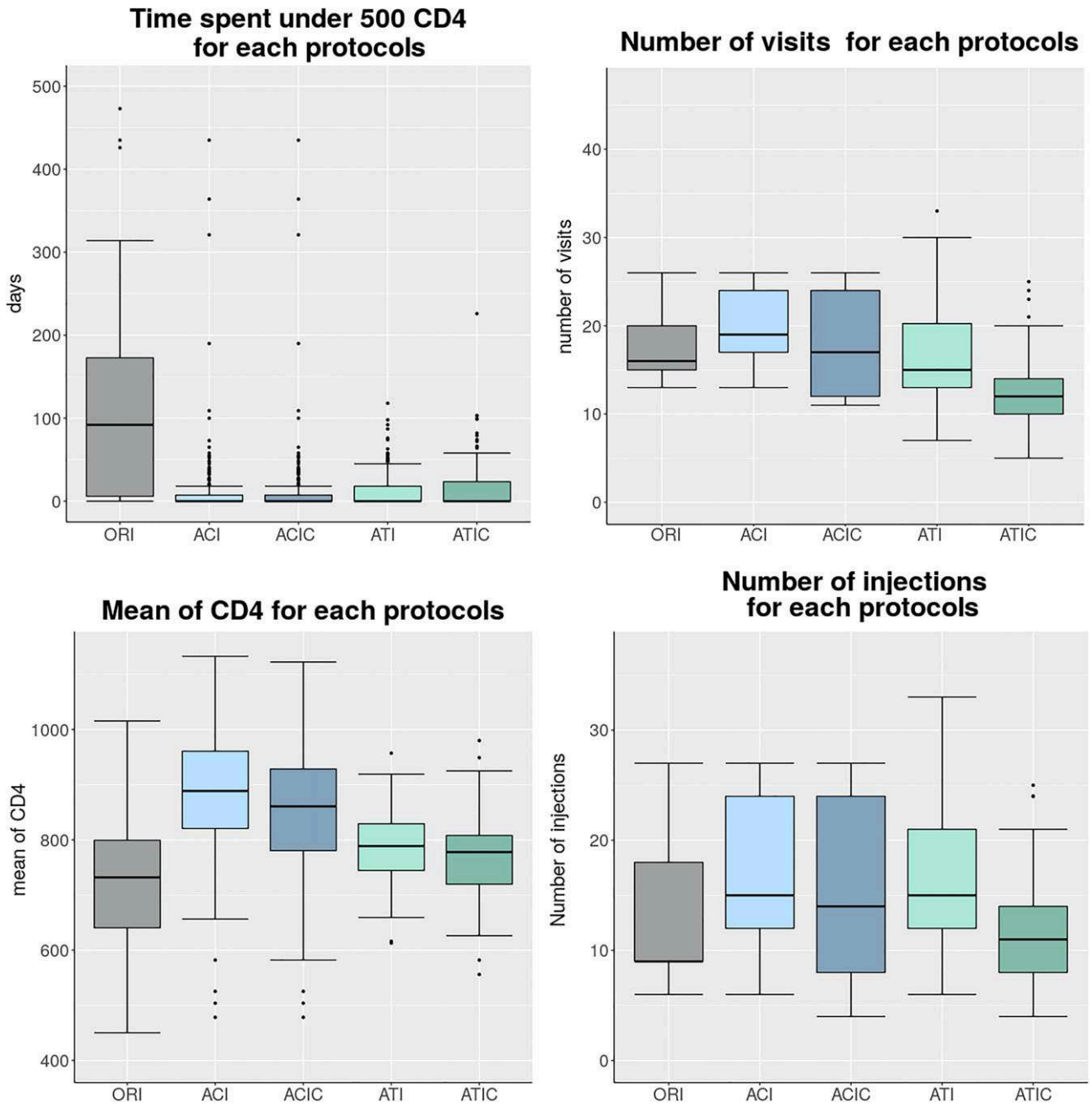


FIGURE 4 Boxplot of time spent under 500 CD4, number of visits, mean number of CD4, and number of injections for each protocol at the threshold 0.05. ACI, adaptive criterion of injection protocol; ATI, adaptive time of injection protocol; ORI, original protocol [Colour figure can be viewed at wileyonlinelibrary.com]

For Patient 1, the ORI protocol predicts a cycle of injection at the second, third, and seventh visits. However, these cycles are unnecessary as the decrease of CD4 in this patient is very slow. The ACI protocol correctly predicts that a new cycle will be necessary only at the fifth visit. The ATI protocol predicts a new cycle at a slightly larger time and has the advantage of having only two visits for a control, instead of eight for the other protocols.

Patient 2 has an opposite problem, as the ORI protocol fails to detect that a new cycle was necessary, eg, at the first visit, resulting in a long time spent under 500 CD4. In contrast, the ACI and ATI protocols correctly predict that a new cycle is necessary. Again, the number of control visits is reduced in the ATI protocol, which yields accurate times of control.

In Web Supplementary 3, the rate of error, defined as the number of times the decision made is not the optimal, is analyzed. Overall, the rate of time when the algorithm made a decision causing the patient to spend some time under 500 CD4 is extremely low for the ACI and ACIC protocols (between 0.1% and 6%). For the ATI and ATIC protocols, it corresponds to the risk taken (5% at the p-quantile of 0.05), while this rate was at 47% for the ORI protocol. The rate of

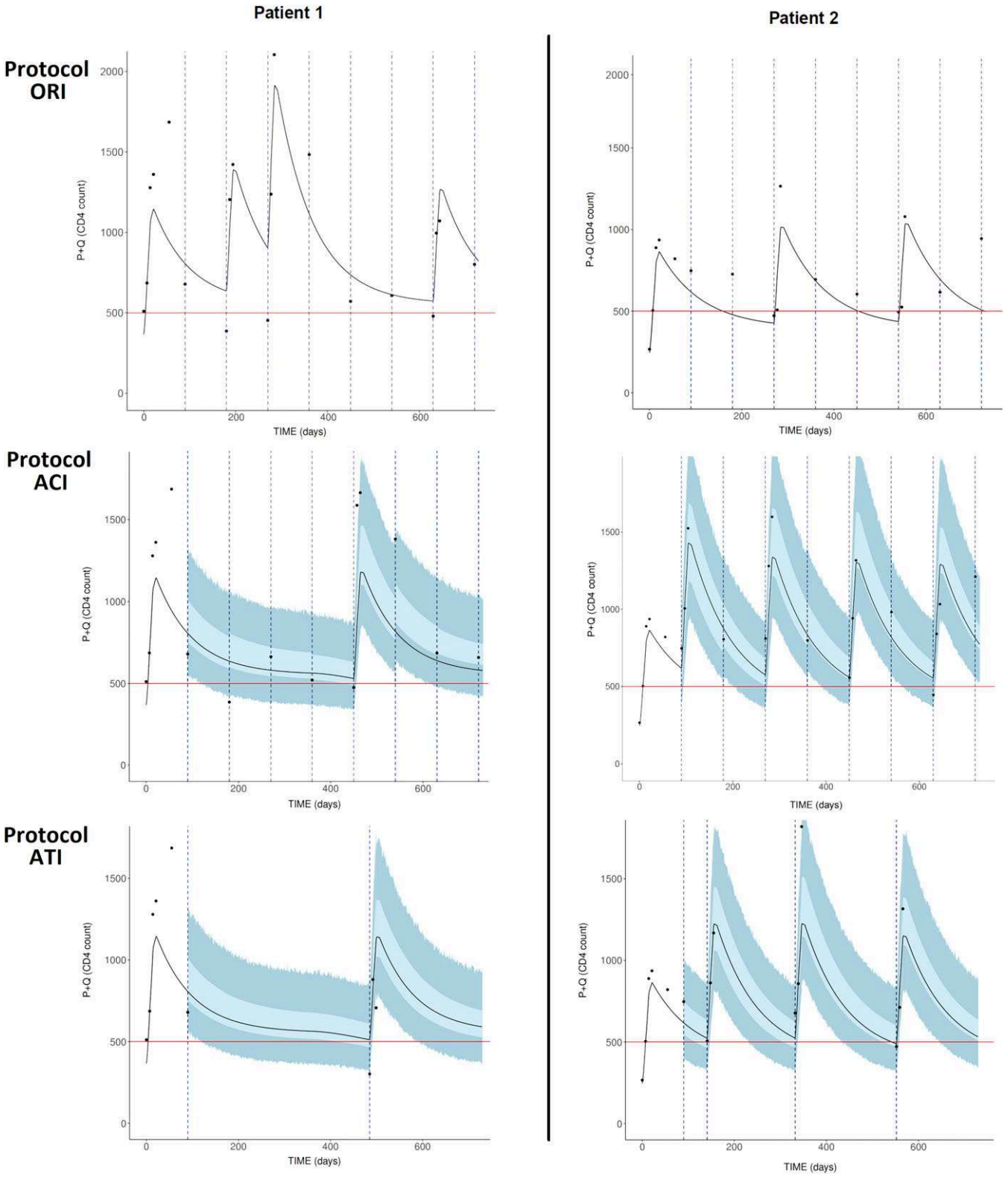


FIGURE 5 Comparison of the three protocols for two random patients. Dots: observations. Black line: simulated trajectory. Horizontal line: limit of 500 CD4. Vertical dashed lines: times of controls. Light band: 95% credible interval of trajectories. Dark band: 95% predictive interval of observations. ACI, adaptive criterion of injection protocol; ATI, adaptive time of injection protocol; ORI, original protocol [Colour figure can be viewed at wileyonlinelibrary.com]

mistake when the algorithm predicts a new cycle for the ACI and ACIC protocols while it could have waited is higher than the ORI protocol; although, this has less impact because this decision can be compensated at the next visit. The rate of time when the ATI and ATIC predicted a time of visit too short is around 10%, but this also has less impact because it simply means that the patient has one visit without injections.

Overall, the ATIC protocol with the p-quantile at 0.05 achieves the best balance between all criteria, with a time spent under 500 CD4 among the lowest, and with the number of injections and visits also among the lowest. Compared with the ORI protocol, the ATIC divided the time spend under 500 CD4 by around six, and spared six visits and one injection over a period of two years. However, an analysis done in Web Supplementary 4 shows that, for the patients who have low CD4 at baseline, the ATI protocol with the p-quantile at 0.05 is the best protocol because, in these patients, we cannot afford to reduce the number of injections per cycle.

6 | CONCLUSION

The very good ability of the mechanistic model to predict CD4 concentrations after a short learning phase to identify individual parameters allows us to embark on optimizing the IL7 administration for each individual patient. Based on this model, we have proposed adaptive protocols with the aim to optimize a criterion, here, the time spent under 500 CD4 cells/μL, using the minimum number of IL7 injections. The simulation of pseudopatients showed that the four proposed protocols succeeded in doing so, reducing the time spent under 500 CD4 cells/μL with a number of injections similar or lower compared to the original protocol.

Compared with classical approaches that are based on optimal control, our approach presents several advantages. First, the parameters are not considered to be known and the adaptation of the IL7 injections is done while the estimations of the individual parameters and the predictions are improved with new observations. This dynamic approach is also referred as dynamic drug monitoring in the work of Murphy et al.[25] Second, the statistical approach of treatment optimization that we propose is less computationally demanding because we are not looking for an optimal strategy over the space of all potential strategies.[45] Rather, we are optimizing the strategy according to that patient's characteristics, by learning the random effects values as information increases. In our application, this was very relevant because we could take into account the diversity of response of the patients.

The success of the proposed approach in this application is due to the validity of the predictions that are obtained after a short learning phase for every patient. However, model misspecification could seriously weaken any optimization of the treatment strategy. Here, the model used was clearly the best model over a series of models tested in this context.[21] The stochasticity was mainly due to inter-individuals variability captured through two parameters (λ and ρ), the other parameters are fixed at the value estimated in Jarne et al[21] and presented in Table 1. Any additional stochasticity, requiring, for instance, to deal with stochastic differential equations, would compromise the feasibility of the approach in a real clinical setting.

The clinical perspective is an evaluation of the adaptive strategy with a standard protocol of injection to confirm the benefit of this intervention on all other immunological markers, such as in the work of Lévy et al,[9] before going to a larger trial to evaluate the impact on clinical outcomes. More generally, this work shows how mechanistic model can help increasing the efficiency of therapies in realistic contexts where patients may respond differently to treatments.

ACKNOWLEDGEMENTS

We thank the investigators involved in the INSPIRE trials (Jean-Pierre Routy, Irini Sereti, Margaret Fischl, Prudence Ive, Roberto F. Speck, Gianpiero D'Offizi, Salvatore Casari, Sharne Foulkes, Ven Natarajan, Thérèse Croughs, Jean-François Delfraissy, Guiseppe Tambussi, Yves Levy, and Michael M. Lederman). We thank the Mésocentre de Calcul Intensif Aquitain (MCIA) of the Université de Bordeaux and of the Université de Pau et des Pays de l'Adour for parallel computing and financial support. Laura Villain gets a PhD grant from Ecole des Hautes Études en Santé Publique (EHESP) and Chloé Pasin from Ecole Nationale Supérieure (ENS).

ORCID

Laura Villain http://orcid.org/0000-0003-4208-900X

REFERENCES

1. Haase AT. Population biology of HIV-1 infection: viral and CD4+ T cell demographics and dynamics in lymphatic tissues. *Annu Rev Immunol*. 1999;17(1):625-656.
2. Hammer SM, Squires KE, Hughes MD, et al. A controlled trial of two nucleoside analogues plus Indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. *N Engl J Med*. 1997;337(11):725-733.
3. Grabar S, Le Moing V, Goujard C, et al. Clinical outcome of patients with HIV-1 infection according to immunologic and virologic response after 6 months of highly active antiretroviral therapy. *Ann Intern Med*. 2000;133(6):401-410.
4. The Antiretroviral Therapy Cohort Collaboration, et al. Life expectancy of individuals on combination antiretroviral therapy in high-income countries: a collaborative analysis of 14 cohort studies. *Lancet*. 2008;372(9635):293-299.
5. CASCADE Collaboration, et al. Effective therapy has altered the spectrum of cause-specific mortality following HIV seroconversion. *AIDS*. 2006;20(5):741-749.
6. Beq S, Delfraissy JF, Theze J. Interleukin-7 (IL-7): immune function, involvement in the pathogenesis of HIV infection and therapeutic potential. *Eur Cytokine Netw*. 2004;15(4):279-289.
7. Vassena L, Proschan M, Fauci AS, Lusso P. Interleukin 7 reduces the levels of spontaneous apoptosis in $CD4^+$ and $CD8^+$ T cells from HIV-1-infected individuals. *Proc Natl Acad Sci USA*. 2007;104(7):2355-2360.
8. Levy Y, Lacabaratz C, Weiss L, et al. Enhanced T cell recovery in HIV-1–infected adults through IL-7 treatment. *J Clin Invest*. 2009;119(4):997-1007.
9. Lévy Y, Sereti I, Tambussi G, et al. Effects of recombinant human interleukin 7 on T-cell recovery and thymic output in HIV-infected patients receiving antiretroviral therapy: results of a phase I/IIa randomized, placebo-controlled, multicenter study. *Clin Infect Dis*. 2012;55(2):291-300.
10. Sereti I, Dunham R, Spritzler J, et al. IL-7 administration drives T cell-cycle entry and expansion in HIV-1 infection. *Blood*. 2009;113(25):6304-6314.
11. Lewden C, Chêne G, Morlat P, et al. HIV-infected adults with a CD4 cell count greater than 500 cells/mm3 on long-term combination antiretroviral therapy reach same mortality rates as the general population. *J Acquir Immune Defic Syndr*. 2007;46(1):72-77.
12. Thiébaut R, Jarne A, Routy J-P, et al. Repeated cycles of recombinant human interleukin 7 in HIV-infected patients with low CD4 T-cell reconstitution on antiretroviral therapy: results of 2 phase II multicenter studies. *Clin Infect Dis*. 2016;62(9):1178-1185.
13. Perelson AS, Ribeiro RM. Modeling the within-host dynamics of HIV infection. *BMC Biol*. 2013;11(1):96.
14. De Boer RJ, Perelson AS. Quantifying T lymphocyte turnover. *J Theor Biol*. 2013;327:45-87.
15. Lavielle M, Samson A, Fermin AK, Mentré F. Maximum likelihood estimation of long-term HIV dynamic models and antiviral response. *Biometrics*. 2011;67(1):250-259.
16. Guedj J, Rong L, Dahari H, Perelson AS. A perspective on modelling hepatitis C virus infection. *J Viral Hepat*. 2010;17(12):825-833.
17. Ke R, Loverdo C, Qi H, Sun R, Lloyd-Smith JO. Rational design and adaptive management of combination therapies for hepatitis C virus infection. *PLoS Comput Biol*. 2015;11(6):e1004040.
18. Kepler GM, Banks HT, Davidian M, Rosenberg ES. A model for HCMV infection in immunosuppressed patients. *Math Comput Model*. 2009;49(7-8):1653-1663.
19. Crauste F, Pujo-Menjouet L, Génieys S, Molina C, Gandrillon O. Adding self-renewal in committed erythroid progenitors improves the biological relevance of a mathematical model of erythropoiesis. *J Theor Biol*. 2008;250(2):322-338.
20. Thiébaut R, Drylewicz J, Prague M, et al. Quantifying and predicting the effect of exogenous interleukin-7 on $CD4^+$ T cells in HIV-1 infection. *PLoS Comput Biol*. 2014;10(5):e1003630.
21. Jarne A, Commenges D, Villain L, Prague M, Lévy Y, Thiébaut R. Modeling $CD4^+$ T cells dynamics in HIV-infected patients receiving repeated cycles of exogenous interleukin 7. *Ann Appl Stat*. 2017;11(3):1593-1616.
22. Murphy SA. Optimal dynamic treatment regimes. *J R Stat Soc Ser B Stat Methodol*. 2003;65(2):331-355.
23. Robins JM. Optimal structural nested models for optimal sequential decisions. In: *Proceedings of the Second Seattle Symposium in Biostatistics: Analysis of Correlated Data*. New York, NY: Springer Science+Business Media Inc; 2004;189-326.
24. Huang X, Ning J, Wahed AS. Optimization of individualized dynamic treatment regimes for recurrent diseases. *Statist Med*. 2014;33(14):2363-2378.
25. Murphy SA, Lynch KG, Oslin D, McKay JR, TenHave T. Developing adaptive treatment strategies in substance abuse research. *Drug Alcohol Depend*. 2007;88:S24-S30.
26. Petersen ML, Deeks SG, van der Laan MJ. Individualized treatment rules: generating candidate clinical trials. *Statist Med*. 2007;26(25):4578-4601.
27. Moodie E, Richardson T, Stephens DA. Demystifying optimal dynamic treatment regimes. *Biometrics*. 2007;63(2):447-455.
28. Rich B, Moodie E, Stephens DA. Optimal individualized dosing strategies: a pharmacologic approach to developing dynamic treatment regimens for continuous-valued treatments. *Biom J*. 2016;58(3):502-517.
29. Prague M, Commenges D, Gran JM, et al. Dynamic models for estimating the effect of HAART on CD4 in observational studies: application to the Aquitaine cohort and the Swiss HIV cohort Study. *Biometrics*. 2017;73(1):294-304.
30. Rosenberg ES, Davidian M, Banks HT. Using mathematical modeling and control to develop structured treatment interruption strategies for HIV infection. *Drug Alcohol Depend*. 2007;88:S41-S51.
31. Toffoli G, Corona G, Basso B, Boiocchi M. Pharmacokinetic optimisation of treatment with oral etoposide. *Clin Pharmacokinet*. 2004;43(7):441-466.

32. Castiglione F, Piccoli B. Optimal control in a model of dendritic cell transfection cancer immunotherapy. *Bull Math Biol*. 2006;68(2):255-274.
33. Banks HT, Jang T, Kwon HD. Feedback control of HIV antiviral therapy with long measurement time. *Int J Pure Appl Math*. 2011;66(4):461-485.
34. Croicu AM. Short- and long-term optimal control of a mathematical model for HIV infection of $CD4^+$ T cells. *Bull Math Biol*. 2015;77(11):2035-2071.
35. Chakraborty B, Murphy SA. Dynamic treatment regimes. *Annu Rev Stat Its Appl*. 2014;1:447-464.
36. Prague M, Commenges D, Drylewicz J, Thiébaut R. Treatment monitoring of HIV-infected patients based on mechanistic models. *Biometrics*. 2012;68(3):902-911.
37. Commenges D, Joly P, Gégout-Petit A, Liquet B. Choice between semi-parametric estimators of Markov and non-Markov multi-state models from coarsened observations. *Scand J Stat*. 2007;34(1):33-52.
38. Prague M, Commenges D, Guedj J, Drylewicz J, Thiébaut R. NIMROD: a program for inference via a normal approximation of the posterior in models with random effects based on ordinary differential equations. *Comput Methods Programs Biomed*. 2013;111(2):447-458.
39. Commenges D, Jacqmin-Gadda H, Proust C, Guedj JA. Newton-like algorithm for likelihood maximization: the robust-variance scoring algorithm. arXiv preprint math/0610402; 2006.
40. van der Vaart AW. *Asymptotic Statistics*. Cambrigde, UK: Cambridge University Press; 2000.
41. Gilks WR, Best NG, Tan KKC. Adaptive rejection metropolis sampling within Gibbs sampling. *Appl Stat*. 1995;45(4):455-472.
42. Guedj J, Thiébaut R, Commenges D. Maximum likelihood estimation in dynamical models of HIV. *Biometrics*. 2007;63(4):1198-1206.
43. Soetaert K, Petzoldt T, Setzer RW. Solving differential equations in R: package deSolve. *J Stat Softw*. 2010;33(9).
44. Hindmarsh AC. *Ordinary Differential Equation System Solver*. Livermore, CA: Lawrence Livermore National Laboratory; 1992.
45. Pasin C, Dufour F, Villain L, Zhang H, Thiébaut R. Controlling IL-7 injections in HIV-infected patients. arXiv preprint arXiv:1801.06227; 2018.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Villain L, Commenges D, Pasin C, Prague M, Thiébaut R. Adaptive protocols based on predictions from a mechanistic model of the effect of IL7 on CD4 counts. *Statistics in Medicine*. 2018;1–15. https://doi.org/10.1002/sim.7957

Modélisation et optimisation de la réponse à des vaccins et à des interventions immunothérapeutiques. Application au virus Ebola et au VIH.

Résumé : Les vaccins ont été une grande réussite en matière de santé publique au cours des dernières années. Cependant, le développement de vaccins efficaces contre les maladies infectieuses telles que le VIH ou le virus Ebola reste un défi majeur. Cela peut être attribué à notre manque de connaissances approfondies en immunologie et sur le mode d'action de la mémoire immunitaire. Les modèles mathématiques peuvent aider à comprendre les mécanismes de la réponse immunitaire, à quantifier les processus biologiques sous-jacents et à développer des vaccins fondés sur un rationnel scientifique. Nous présentons un modèle mécaniste de la dynamique de la réponse immunitaire humorale après injection d'un vaccin Ebola basé sur des équations différentielles ordinaires. Les paramètres du modèle sont estimés par maximum de vraisemblance dans une approche populationnelle qui permet de quantifier le processus de la réponse immunitaire et ses facteurs de variabilité. Le schéma vaccinal n'a d'impact que sur la réponse à court terme, alors que des différences significatives entre des sujets de différentes régions géographiques sont observées à plus long terme. Cela pourrait avoir des implications dans la conception des futurs essais cliniques. Ensuite, nous développons un outil numérique basé sur la programmation dynamique pour optimiser des schémas d'injections répétées. Nous nous intéressons en particulier à des patients infectés par le VIH sous traitement mais incapables de reconstruire leur système immunitaire. Des injections répétées d'un produit immunothérapeutique (IL-7) sont envisagées pour améliorer la santé de ces patients. Le processus est modélisé par un modèle de Markov déterministe par morceaux et des résultats récents de la théorie du contrôle impulsionnel permettent de résoudre le problème numériquement à l'aide d'une suite itérative. Nous montrons dans une preuve de concept que cette méthode peut être appliquée à un certain nombre de pseudo-patients. Dans l'ensemble, ces résultats s'intègrent dans un effort de développer des méthodes sophistiquées pour analyser les données d'essais cliniques afin de répondre à des questions cliniques concrètes.

Mots clés : Modèles mécanistes ; Equations différentielles ordinaires ; Maximisation de la vraisemblance ; Modèles linéaires mixtes ; Contrôle optimal ; Processus de Markov déterministes par morceaux ; Programmation dynamique ; Vaccin ; Ebola ; Réponse immunitaire ; Durabilité ; Facteurs de variabilité ; VIH ; Immunothérapie ; Injections répétées.

Modeling and optimizing the response to vaccines and immunotherapeutic interventions. Application to Ebola virus and HIV.

Abstract: Vaccines have been one of the most successful developments in public health in the last years. However, a major challenge still resides in developing effective vaccines against infectious diseases such as HIV or Ebola virus. This can be attributed to our lack of deep knowledge in immunology and the mode of action of immune memory. Mathematical models can help understanding the mechanisms of the immune response, quantifying the underlying biological processes and eventually developing vaccines based on a solid rationale. First, we present a mechanistic model for the dynamics of the humoral immune response following Ebola vaccine immunizations based on ordinary differential equations. The parameters of the model are estimated by likelihood maximization in a population approach, which allows to quantify the process of the immune response and its factors of variability. The vaccine regimen is found to impact only the response on a short term, while significant differences between subjects of different geographic regions are found at a longer term. This could have implications in the design of future clinical trials. Then, we develop a numerical tool based on dynamic programming for optimizing schedule of repeated injections. In particular, we focus on HIV-infected patients under treatment but unable to recover their immune system. Repeated injections of an immunotherapeutic product (IL-7) are considered for improving the health of these patients. The process is first modeled by a piecewise deterministic Markov model and recent results of the impulse control theory allow to solve the problem numerically with an iterative sequence. We show in a proof-of-concept that this method can be applied to a number of pseudo-patients. All together, these results are part of an effort to develop sophisticated methods for analyzing data from clinical trials to answer concrete clinical questions.

Key words: Mechanistic modeling; Ordinary differential equations; Likelihood maximization; Linear mixed models; Optimal control; Piecewise deterministic Markov processes; Dynamic programming; Vaccine; Ebola; Immune response; Durability; Variability factors; HIV; Immunotherapy; Repeated injections.

Discipline : Santé publique – option : Biostatistiques
Laboratoire : Unité INSERM U1219, Bordeaux Population Health Center - INRIA - Université de Bordeaux
146 rue Léo Saignat 33000 Bordeaux, FRANCE