



Université de Yaoundé I
Ecole Nationale Supérieure Polytechnique



Université Montpellier I
UFR Médecine

THESE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE DE YAOUNDE 1 & DE L'UNIVERSITE MONTPELLIER 1

Discipline : Mathématiques appliquées-Biologie santé (Biostatistiques)

Section CNU : 46-04 Biostatistiques

Ecoles Doctorales : Ecole Nationale Supérieure Polytechnique, Yaoundé
Cameroun & Sciences Chimiques et Biologiques pour la Santé,
Montpellier, France

Soutenue publiquement

par

Jules Brice TCHATCHUENG MBOUGUA

le 12 juin 2012

**Problématiques statistiques rencontrées dans l'étude du traitement
antirétroviral des adultes infectés par le VIH en Afrique
subsaharienne.**

JURY

M. Tetanye EKOE	Professeur, Yaoundé, Président
M. Maurice TCHUENTE	Professeur, Yaoundé, Rapporteur
M. Roch GIORGI	Professeur, Marseille, Rapporteur
M. Henri GWET	Maitre de conférences, Yaoundé, Co-directeur
M. Christian LAURENT	Chargé de recherche, Montpellier, Co-directeur
M. Nicolas MOLINARI	Maitre de conférences, Montpellier, Co-encadrant

Dédicace

*Je dédie cette thèse à
ma compagne Adrienne
mon fils Patrick Lohan
mes parents Abel et Françoise*

Remerciements

Si la thèse est encore le symbole d'un travail solitaire, les échanges et les collaborations sont indispensables. C'est donc avec plaisir que je remercie ici tous ceux qui m'ont apporté leur concours, leurs conseils et leurs idées.

Ma reconnaissance va tout d'abord à mes encadrants de thèse Henri Gwet, Christian Laurent, Nicolas Molinari pour leurs précieux conseils, leurs encouragements constants et leurs observations critiques, dont l'intérêt et les remarques ont toujours été fort stimulants et sans lesquels ce travail n'aurait pu aboutir. Ce fut un réel bonheur de travailler avec vous.

Toute ma gratitude aux Professeurs Éric Delaporte et Sinata Koulla-Shiro pour avoir accepté de m'accueillir dans leurs laboratoires l'UMI 233 de l'IRD à Montpellier et site de coordination de recherche ANRS du Cameroun, et profiter des compétences qui s'y développent.

Mes remerciements s'adressent ensuite aux rapporteurs de mon travail de thèse les Professeurs Maurice Tchuenté et Roch Giorgi qui ont bien voulu consacrer du temps à la lecture et à l'examen de cette recherche, me permettant ainsi d'améliorer la qualité et d'envisager de nouvelles voies de développement.

Je remercie tout particulièrement le Professeur Tetanye Ekoe d'avoir accepté de présider le jury de cette thèse.

Je remercie mon comité de suivi de thèse (Jean-Christophe Thalabard, Patrick Bastien, Vincent Le Moing) pour le suivi et les conseils sur l'orientation de cette thèse.

Je remercie l'Institut de Recherche pour le Développement pour le soutien financier (36 mois de bourses), sans lequel n'aurait pas eu lieu ce travail. Tout le mérite revient à Christian Laurent pour les efforts consentis à la recherche de ce financement.

Je ne saurais oublier toute l'équipe de l'UMI 233 à Montpellier et du site ANRS Cameroun pour leur accueil, les conditions de travail privilégiées qui m'ont été offertes et tous ces repas que nous avons partagés à la cantine. Je pense à Charles Kouanfack pour avoir toujours répondu efficacement à mes questions. Marlise Dontsop, Jean-Marc Ben, Maguy Ngolle, Odile Ela, Gabrièle laborde-balen pour tous les conseils et les moments passés ensemble. Sabrina pour les astuces d'analyse sous STATA. Suzanne et Alain pour votre grand cœur et les moments de détente que nous avons partagés.

J'exprime ma reconnaissance à l'ensemble des enseignants du projet STAFAY pour m'avoir initiés à la recherche, guidés lors de mon master et confortés dans mon désir de poursuivre en thèse.

Je remercie chaleureusement tous mes camarades de promotion pour leur soutien amical avec une pensée spéciale pour Cyprien Mbogning, Serges Teukam, Colbert Ngouanfouo.

Je remercie mes frères et amis de l'UJEEBay et Dynamite 2008, qui m'accompagnent dans tous mes projets depuis plusieurs années. Que l'aboutissement de cette thèse soit un élément de motivation pour tous ces jeunes qui ont du talent et qui hésitent.

Je ne saurais oublier les familles Pebou et Nembou, mes frères et sœurs Adeline, Carole, Achille, Patrick, Thiery, Raoul, Arlette, Francine pour votre soutien financier et moral indéfectible et tout le reste.

Je remercie ma belle-famille pour la confiance qu'elle m'accorde.

Mes parents Tchatchueng Abel, Magniyimkam Françoise, aucun mot ne peut exprimer ce que j'ai envie de vous dire. L'aboutissement de cette thèse est le fruit de l'éducation exemplaire que j'ai reçue de vous.

Merci Adrienne pour ton soutien quotidien, pour ta patience malgré mes longues absences et surtout pour le beau cadeau que tu m'as donné. Vous donnez un sens à ma vie.

Résumé en français

Partant de problématiques statistiques rencontrées dans l'étude du traitement antirétroviral des adultes infectés par le virus de l'immunodéficience humaine (VIH) en Afrique subsaharienne, cette thèse cherche, d'une part, à favoriser la vulgarisation d'outils méthodologiques relativement récents auprès d'un public d'utilisateurs moins avertis et, d'autre part, à participer au développement de nouveaux outils. Le premier chapitre présente différentes méthodes de modélisation des données longitudinales dont des méthodes d'analyse de l'évolution d'un critère au cours du temps (les modèles linéaires mixtes généralisés et les modèles d'équations d'estimation généralisées) ou de la survenue d'un événement au cours du temps (le modèle semi-paramétrique de Cox et ses extensions à la prise en compte des covariables dépendantes du temps et de la censure informative). Le deuxième chapitre s'intéresse aux tests de non-infériorité et propose deux développements de la procédure classique de ces tests pour les cas où la marge de non-infériorité est relative. Enfin, le troisième chapitre aborde la question des données manquantes et propose une extension de la méthode d'imputation multiple par les distributions conditionnelles univariées qui consiste à prendre en compte des effets non linéaires des covariables dans les modèles d'imputation par des fonctions B-splines. Ces méthodes sont illustrées par des études sur le VIH au Cameroun et au Sénégal.

Mots clés : Méthodes statistiques, données longitudinales, analyse de survie, données manquantes, non-infériorité, VIH/SIDA.

Laboratoires :

Laboratoire Informatique, Modélisation, Simulations et Systèmes (LIMSS), Ecole Nationale Supérieure Polytechnique, Université de Yaoundé 1, BP 8390 Yaoundé Cameroun

IRD UMI 233, Transitions épidémiologiques, recherches translationnelles appliquées au VIH et aux Maladies Infectieuses (TransVIHMI), 911 avenue agropolis BP 64501 4394 Montpellier cedex 5

Résumé en anglais

On the basis of statistical challenges encountered in study of antiretroviral treatment of adults infected with human immunodeficiency virus (HIV) in sub-Saharan Africa, this thesis aims to promote the dissemination of relatively recent methodological tools of less aware audience of users on one hand and to participate to development of new toolson the other hand. The first chapter presents various methods for modeling longitudinal data of which analysis methods for changing of a criterion over time (the generalized linear mixed models and models of generalized estimating equations) or the occurrence of an event over time (the semi-parametric Cox model and its extensions to take into account time-dependent covariates and informative censoring). The second chapter focuses on non-inferiority test and provides two developments of the classical procedure of these tests in cases where the non-inferiority margin is relative. The third chapter addresses the question of missing data and proposes an extension of the multiple imputation method based on fully conditional specification, to take into account nonlinear effects of covariates in the imputation models using B-splines functions. These methods are illustrated by studies on HIV in Cameroon and Senegal.

Keys Word: statistic methods, longitudinal data, survival analysis, missing data, non-inferiority, HIV/AIDS.

Laboratoires :

Laboratoire Informatique, Modélisation, Simulations et Systèmes (LIMSS), Ecole Nationale Supérieure Polytechnique, Université de Yaoundé 1, BP 8390 Yaoundé Cameroun

IRD UMI 233, Transitions épidémiologiques, recherches translationnelles appliquées au VIH et aux Maladies Infectieuses (TransVIHMI), 911 avenue agropolis BP 64501 4394 Montpellier cedex 5

Sommaire

Dédicace	1
Remerciements	2
Résumé en français	5
Résumé en anglais	6
Sommaire	7
Liste des figures	10
Liste des tableaux	13
Introduction	14
1. Infection par le VIH et traitement antirétroviral en Afrique subsaharienne	14
2. Contexte spécifique des travaux de thèse	21
2.1. La situation au Cameroun	22
2.2. La situation au Sénégal	24
3. Thématiques spécifiques et données utilisées pour les travaux de thèse	25
3.1. Co-infection par le VIH et les virus des hépatites B et C.....	25
3.2. Stratégies thérapeutiques de première ligne	28
3.3. Stratégies de suivi des patients sous antirétroviraux dans le contexte de la prise en charge décentralisée	30
4. Présentation des problématiques statistiques abordées dans les travaux de thèse	34
Chapitre 1 : Modélisation des données longitudinales	36
1. Position du problème	36
2. Méthodes d'analyse de l'évolution d'un critère au cours du temps	39
2.1. Modèles mixtes	39
2.1.1. Modèle linéaire mixte	40
2.1.2. Modèle linéaire mixte généralisé (GLMM)	43
2.1.3. Modélisation de la partie fixe et de la partie aléatoire dans les modèles mixtes.....	46
2.2. Modèle marginal : le modèle GEE	57
2.3. Comparaison entre les modèles GLMM et GEE	59

2.3.1.	Sur le plan théorique	60
2.3.2.	Sur le plan pratique	60
3.	Méthodes d'analyse de la survenue d'un évènement au cours du temps.....	62
3.1.	Fonctions associées aux distributions de survie	63
3.2.	Modèle à risque proportionnel.....	65
3.3.	Extensions du modèle semi-paramétrique de Cox.....	67
3.3.1.	Modèle de Cox avec covariables dépendantes du temps	67
3.3.2.	Prise en compte de la censure informative.....	71
4.	Quelques méthodes de sélection de modèles	74
5.	Applications	76
5.1.	Hépatotoxicité et efficacité des traitements antirétroviraux incluant la névirapine chez les patients co-infectés par le VIH et le virus de l'hépatite B ou C au Cameroun	76
5.2.	Efficacité et tolérance à long terme du traitement antirétroviral incluant la didanosine, lalamivudine et l'efavirenz ou la névirapine chez les patients naïfs, infectés par le VIH : étude de cohorte de 9 ans au Sénégal.....	77
	Chapitre 2 : Tests de non-infériorité	82
1.	Principe des essais de non-infériorité	83
1.1.	Intervention de référence	83
1.2.	Marge de non-infériorité	84
1.3.	Choix et analyse du critère de jugement principal.....	84
2.	Démonstration de la non-infériorité.....	85
2.1.	Notions sur les tests d'hypothèse et les méthodes par intervalles de confiance	86
2.1.1.	Les tests d'hypothèse paramétriques	86
2.1.2.	Notions sur les intervalles de confiance	89
2.2.	Relation entre les tests de non-infériorité et les intervalles de confiance..	90
2.2.1.	Cas où la marge de non-infériorité est fixe ou constante	91
2.2.2.	Cas où la marge de non-infériorité est une variable aléatoire	92
3.	Application à l'analyse du critère de jugement principal de l'essai Stratall	95
4.	Discussion	100
	Chapitre 3 : Données manquantes.....	104

1.	Typologie des données manquantes	105
1.1.	Données manquantes complètement au hasard.....	106
1.2.	Données manquantes au hasard	106
1.3.	Données manquantes non au hasard	107
2.	Gestion des données manquantes	108
2.1.	Imputation des données manquantes	109
3.	Imputation multiple par les fonctions B-splines	110
3.1.	Approximation de $fl(.)$ par des fonctions B-splines.....	112
	Conclusion générale et perspectives	117
	Bibliographie.....	119
	Annexe A: article « Hepatotoxicity and effectiveness of a Nevirapine-based antiretroviral therapy in HIV-infected patients with or without viral hepatitis B or C infection in Cameroon »	132
	Annexe B :article « Long-term effectiveness and safety of didanosine combined with lamivudine and efavirenz or nevirapine in antiretroviral-naive patients: a 9-year cohort study in Senegal ».....	144
	Annexe C :article « Monitoring of HIV viral loads, CD4 cell counts and clinical assessments versus clinical monitoring alone for antiretroviral therapy in rural district hospital in Cameroon (Stratall ANRS 12110/ESTHER): a randomised non-inferiority trial. ».....	151
	Annexe D :article « Nonlinear multiple imputation for continuous covariate within semi-parametric Cox model : application to HIV data in Senegal. ».....	161

Liste des figures

Figure 1. Nombre estimé d'adultes et d'enfants vivant avec le VIH en 2010 dans le monde	15
Figure 2. Nombre estimé d'adultes ou d'enfants nouvellement infectés par le VIH en 2010 dans le monde	16
Figure 3. Nombre estimé de décès dus au Sida chez l'adulte et l'enfant en 2010 dans le monde.....	16
Figure 4. Historique et contexte de l'accès au traitement antirétroviral en Afrique ..	18
Figure 5. Nombre de personnes recevant un traitement antirétroviral dans les pays à revenu faible ou intermédiaire, par région, 2002-2010	19
Figure 6. Proportions de personnes recevant un traitement antirétroviral en Afrique subsaharienne parmi celles qui en ont besoin, 2002-2005.....	20
Figure 7. Position du Cameroun et du Sénégal en Afrique.....	22
Figure 8. Nombre de structures sanitaires et de districts de santé impliqués dans la prise en charge des personnes vivant avec le VIH au Cameroun entre 2002 et 2010	24
Figure 9. Profils d'évolution individuels du nombre total de lymphocytes T CD4 pendant les 24 premiers mois de traitement ARV de 20 patients infectés par le VIH, suivis au Cameroun dans les études de cohorte Parvy et Triomune. Données regroupées à M6, M12, M18, M24.....	41
Figure 10. Profils d'évolution individuels du nombre total de lymphocytes CD4 pendant les 24 premiers mois de traitement de 20 patients infectés par le VIH suivis au Cameroun dans les études de cohorte Parvy et Triomune. Sans regroupement des données (données déséquilibrées).....	47
Figure 11. Profil moyen d'évolution du nombre total de lymphocytes CD4 pendant les 24 premiers mois sous traitement de l'ensemble des patients infectés par le VIH suivis au Cameroun dans les études de cohorte Parvy et Triomune. A lorsque les données sont regroupées, moyenne des observations à J0, M6, M12, M18, M24, ± 2 écarts types. B lorsque les données ne sont pas regroupées, courbe d'évolution moyenne obtenue par une méthode de lissage (les moyennes mobiles.	48
Figure 12. Courbes moyennes observées et courbes moyennes prédites (à l'aide du modèle 6 ci-dessous) de l'évolution de la racine carrée du taux de lymphocytes T CD4 chez les patients monoinfectés et chez les patients co-infectés sous traitement	

<i>ARV incluant la névirapine, suivis au Cameroun dans les études Parvy et Triomune.</i>	51
Figure 13. Courbes d'évolution de la corrélation entre les mesures en fonction du temps pour un modèle linéaire mixte à intercepte et pente aléatoires, avec une structure de corrélation identité entre l'intercepte et la pente aléatoire.....	53
Figure 14. Courbes d'évolution de la corrélation entre les mesures en fonction du temps pour un modèle linéaire mixte à intercepte et pente aléatoires, avec une structure de corrélation indépendante entre l'intercepte et la pente.	54
Figure 15. Représentation des profils d'évolution individuels observés et des courbes d'évolution individuelles prédites de la racine carrée du taux de lymphocytes T CD4 (à l'aide du modèle 6 ci-dessus) chez 17 patients sélectionnés au hasard, infectés par le VIH sous traitement antirétroviral, suivis au Cameroun dans les études Parvy et Triomune.	56
Figure 16. Représentation des zones d'acceptation et de rejet pour un test de non-infériorité avec une marge de non-infériorité définie par une variable aléatoire.....	93
Figure 17. Représentation de la procédure de décision du test de non-infériorité basée sur les intervalles de confiance lorsque la marge de non-infériorité est définie par une variable aléatoire.	93
Figure 18. Résultat de l'analyse en intention de traiter du critère principal de l'essai Stratall avec la méthode LOCF (dernière valeur observée). Représentation de la procédure de décision du test de non-infériorité basé sur les intervalles de confiance bilatéraux à 95 %.....	96
Figure 19. Résultat de l'analyse en intention de traiter du critère principal de l'essai Stratall avec la méthode LOCF (dernière valeur observée). Représentation de la procédure de décision du test de non-infériorité basé sur les intervalles de confiance bilatéraux à 90 %.....	97
Figure 20. Résultat de l'analyse en intention de traiter du critère principal de l'essai Stratall avec les données disponibles. Représentation de la procédure de décision du test de non-infériorité basé sur les intervalles de confiance bilatéraux à 95 %.....	97
Figure 21. Résultat de l'analyse en intention de traiter du critère principal de l'essai Stratall avec les données disponibles. Représentation de la procédure de décision du test de non-infériorité basé sur les intervalles de confiance bilatéraux à 90 %.....	98
Figure 22. Résultat de l'analyse en intention de traiter du critère principal de l'essai Stratall avec la méthode LOCF (dernière valeur observée). Représentation de la	

<i>procédure de décision du test de non-infériorité basé sur les intervalles de confiance à 80 %.</i>	99
Figure 23. <i>Résultat de l'analyse en intention de traiter du critère principal de l'essai Stratall avec les données disponibles. Représentation de la procédure de décision du test de non-infériorité basé sur les intervalles de confiance à 80 %.</i>	99
Figure 24. <i>Impact du traitement antirétroviral</i>	101

Liste des tableaux

Tableau 1. Estimation du pourcentage de la couverture du traitement antirétroviral dans les pays à revenu faible ou moyen par région en fonction des lignes directrices de l'OMS de 2010, en décembre 2009 et en décembre 2010	21
Tableau 2. Représentation d'un tableau de données longitudinales en format long	37
Tableau 3. Liste des fonctions de lien canoniques les plus utilisées dans le modèle GLMM.....	45
Tableau 4. Résultat de la modélisation de l'évolution des lymphocytes T CD4 chez les patients co-infectés par le VIH et le VHB ou le VHC sous traitement ARV incluant la névirapine par rapport aux patients mono-infectés par le VIH suivis au Cameroun dans les études de cohorte Parvy et Triomune.	50
Tableau 5. Tableau récapitulatif des principales différences entre les modèles GLMM et GEE.....	62
Tableau 6. Résultat des estimations du seuil et de la puissance des procédures de test proposées en fonction du niveau des intervalles de confiance.....	95
Tableau 7. Proportion des données manquantes de certaines variables d'inclusion dans l'étude comparant le traitement antirétroviral incluant la didanosine et celui incluant la zidovudine.	105

Introduction

Ce travail de thèse se situe à la frontière entre la biostatistique et l'épidémiologie. Partant de problématiques statistiques rencontrées dans l'étude du traitement antirétroviral des adultes infectés par le virus de l'immunodéficience humaine (VIH) en Afrique subsaharienne, il cherche, d'une part, à favoriser la vulgarisation d'outils méthodologiques relativement récents auprès d'un public d'utilisateurs moins avertis et, d'autre part, à participer au développement de nouveaux outils.

Cette thèse a été réalisée sous la cotutelle des Universités Yaoundé 1 et Montpellier 1. Les données utilisées proviennent de projets de recherche menés par l'Unité Mixte Internationale 233 associant l'Institut de Recherche pour le Développement (IRD), l'Université Montpellier 1, l'Université Yaoundé 1 et l'Université Cheikh Anta Diop de Dakar, avec leurs partenaires camerounais et sénégalais.

Ce travail a bénéficié d'une bourse de l'IRD et s'est effectué en alternance à Yaoundé (site de l'Agence Nationale de Recherches sur le Sida et les Hépatites virales – ANRS – à l'Hôpital Central) et à Montpellier (IRD, UMI 233).

1. Infection par le VIH et traitement antirétroviral en Afrique subsaharienne

Trente ans après son émergence, l'épidémie due au VIH reste un problème majeur de santé publique dans le monde. On estimait ainsi à 34 millions [31,6 – 35,2 millions] le nombre de personnes vivant avec le VIH en 2010 (**Figure 1**). La même année, on estimait à 2,7 millions [2,4 – 2,9 millions] le nombre de personnes nouvellement infectées par le VIH (**Figure 2**) et à 1.8 million [1,6 – 1,9 million] le nombre de décès liés au Syndrome d'Immuno-Déficience Acquisée (SIDA ; **Figure 3**)

L'Afrique subsaharienne continue de supporter une part disproportionnée du fardeau mondial du VIH avec plus des deux tiers (**Figure 1**) des personnes infectées (22,9 millions [21,6 – 24,1 millions]) alors que la population africaine représente 12 % de la

population mondiale. Sur ce continent, en 2010, on estimait à 1,9 million [1,7 – 2,1 millions] le nombre de personnes nouvellement infectées (**Figure 2**) et 1,2 million [1,1 – 1,4 million] le nombre de décès liés au SIDA (**Figure 3**). L'impact de cette épidémie sur le plan socio-économique dans la région subsaharienne de l'Afrique est majeur et le SIDA reste l'une des principales causes de mortalité en Afrique [1].

Figure 1. Nombre estimé d'adultes et d'enfants vivant avec le VIH en 2010 dans le monde

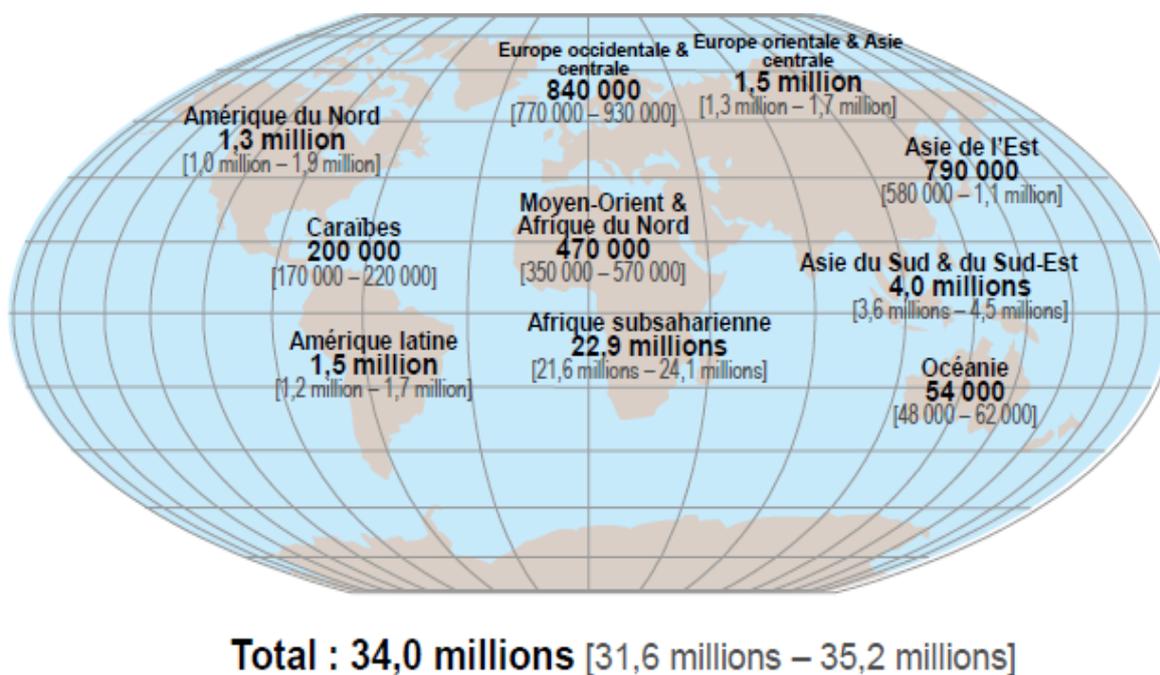


Figure 2. Nombre estimé d'adultes ou d'enfants nouvellement infectés par le VIH en 2010 dans le monde

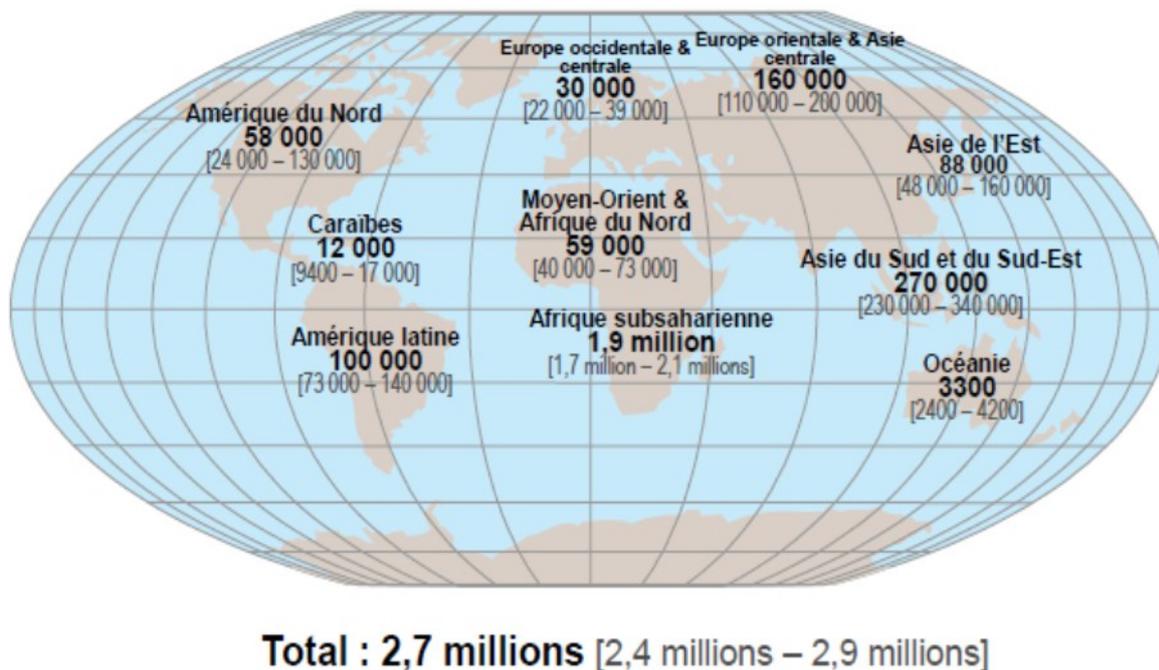
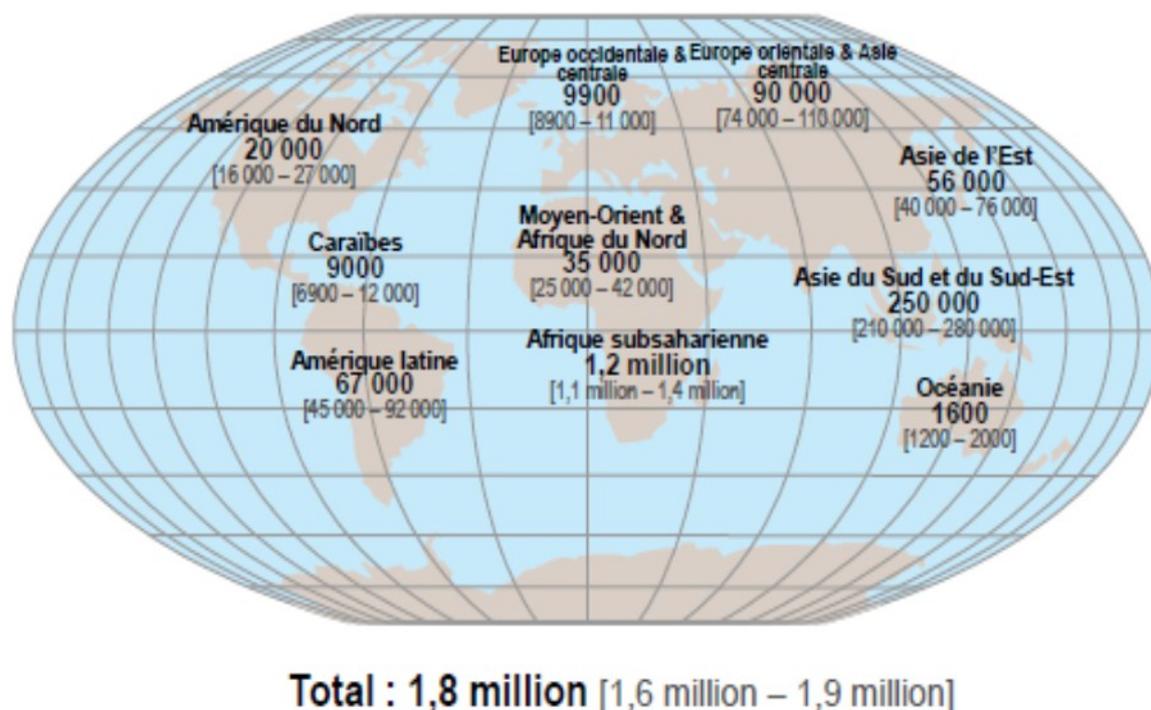


Figure 3. Nombre estimé de décès dus au Sida chez l'adulte et l'enfant en 2010 dans le monde



Source : Journée mondiale sida – Rapport ONUSIDA 2011. Diapositives clés sur les données épidémiologiques.

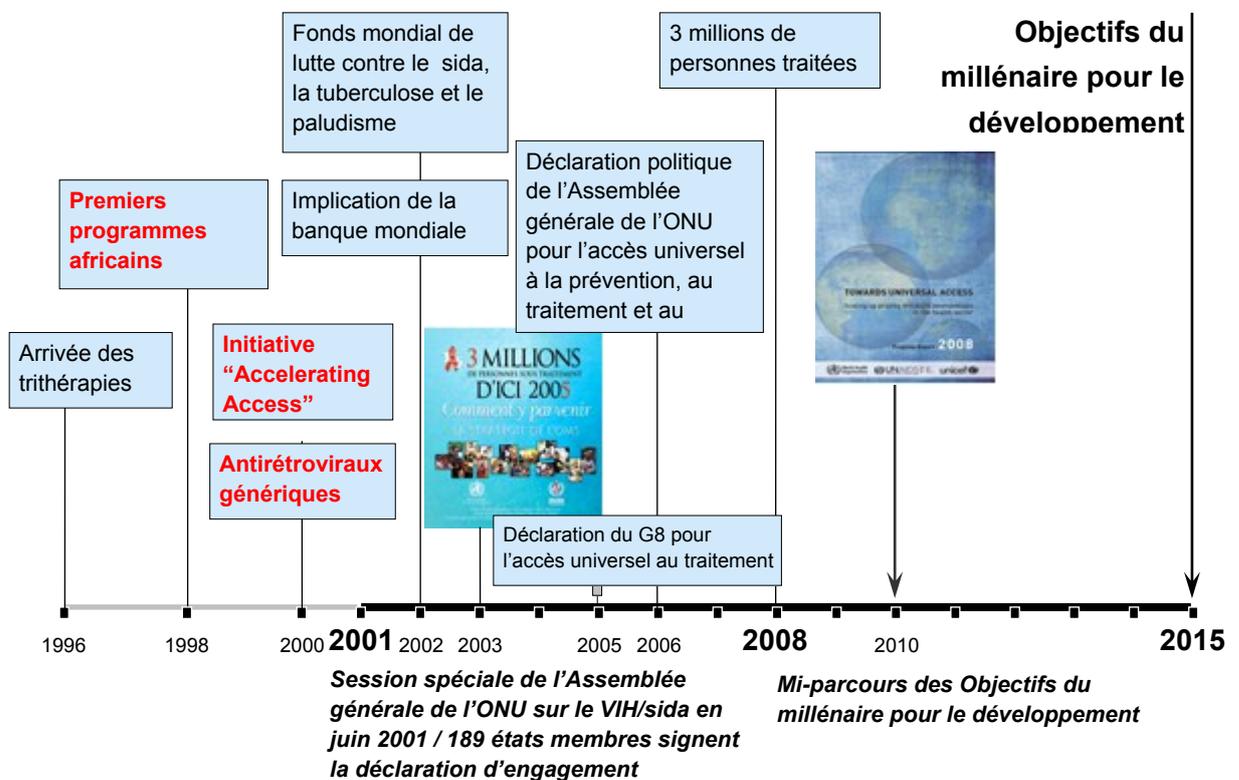
Le traitement antirétroviral (ARV) des personnes infectées par le VIH a débuté dans les pays industrialisés dès 1987, moins de quatre ans après la découverte du virus. Cependant, l'utilisation initiale de monothérapies (à base d'azidothymidine [AZT], également connue sous le nom de zidovudine [ZDV]) puis de bithérapies conduisait à court ou moyen terme à l'échec thérapeutique. Ce n'est qu'en 1996 que sont apparus les traitements dits « hautement actifs » associant trois molécules antirétrovirales [2, 3]. Ces trithérapies, toujours utilisées 15 ans plus tard en première intention, ont permis de réduire considérablement la mortalité et la morbidité liées au VIH et d'améliorer la qualité de vie des personnes infectées [4-6].

En Afrique, les premiers programmes structurés d'accès aux antirétroviraux n'ont vu le jour qu'en 1998 dans de rares pays comme le Sénégal, la Côte d'Ivoire et l'Ouganda (**Figure 4**) [7]. Ces programmes pilotes ont montré la faisabilité et l'efficacité du traitement antirétroviral sur ce continent, avec des résultats comparables à ceux obtenus dans les pays industrialisés en termes de survie, d'efficacité virologique, immunologique et clinique, d'observance, d'émergence des résistances et de toxicité [8-11]. Ces résultats répondaient ainsi aux principales objections d'ordre biomédical quant à l'utilisation des antirétroviraux en Afrique (efficacité sur les souches non-B du VIH-1, tolérance chez des patients fréquemment co-infectés par d'autres pathogènes comme ceux des hépatites...). Depuis, l'accès au traitement s'est étendu à bien d'autres pays africains. Initialement limitée à des services spécialisés dans les capitales et autres grandes villes, la prise en charge des patients sous antirétroviraux s'est progressivement décentralisée vers des structures sanitaires périphériques « généralistes », y compris en milieu rural. L'accès au traitement ARV en Afrique subsaharienne a indubitablement été favorisé par :

- la réduction considérable du prix des ARV depuis 2000 suite aux négociations avec les industries pharmaceutiques dans le cadre du programme ACCESS,
- la disponibilité d'ARV génériques,
- l'engagement de la communauté internationale comme en témoigne la session spéciale de l'Assemblée générale de l'Organisation des Nations Unies (ONU) sur le VIH/sida en juin 2001 (189 états membres signent la déclaration d'engagement),

- l'initiative « 3 X 5 », en 2003, de l'Organisation Mondiale de la Santé (OMS) visant un objectif de 3 millions de personnes sous ARV à la fin de l'année 2005 (l'OMS vise maintenant l'accès universel d'ici à 2015),
- la création du Fonds Mondial de lutte contre le Sida, la Tuberculose et le Paludisme,
- la création du plan américain PEPFAR (*the President's Emergency Plan for AIDS Relief*),
- l'engagement des fondations Clinton et Gates,
- l'engagement des gouvernements africains.

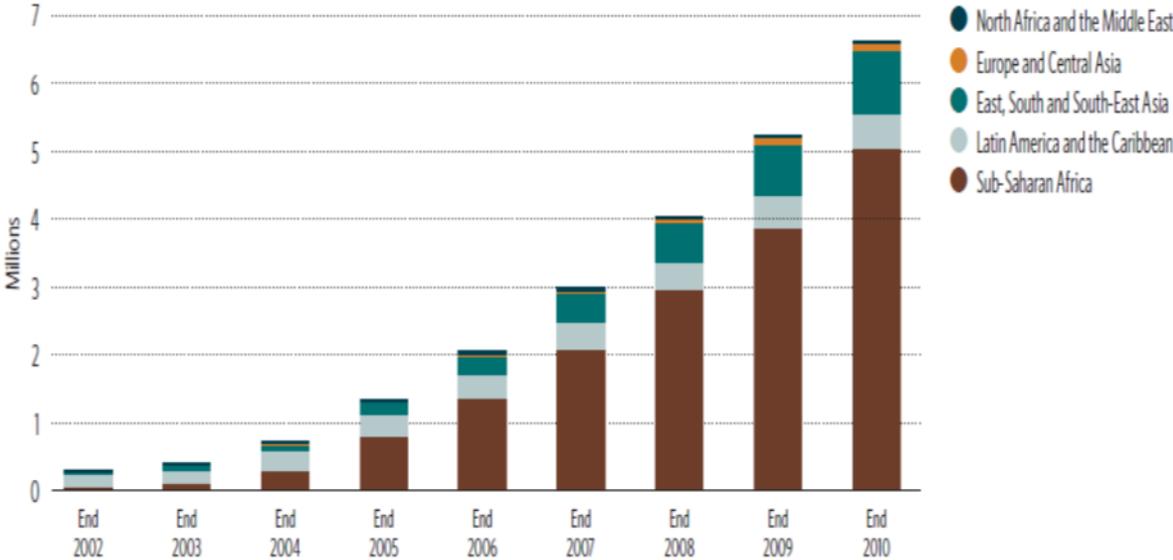
Figure 4. Historique et contexte de l'accès au traitement antirétroviral en Afrique



Source : Adaptée du rapport ONUSIDA. 2008 report on the global AIDS epidemic.

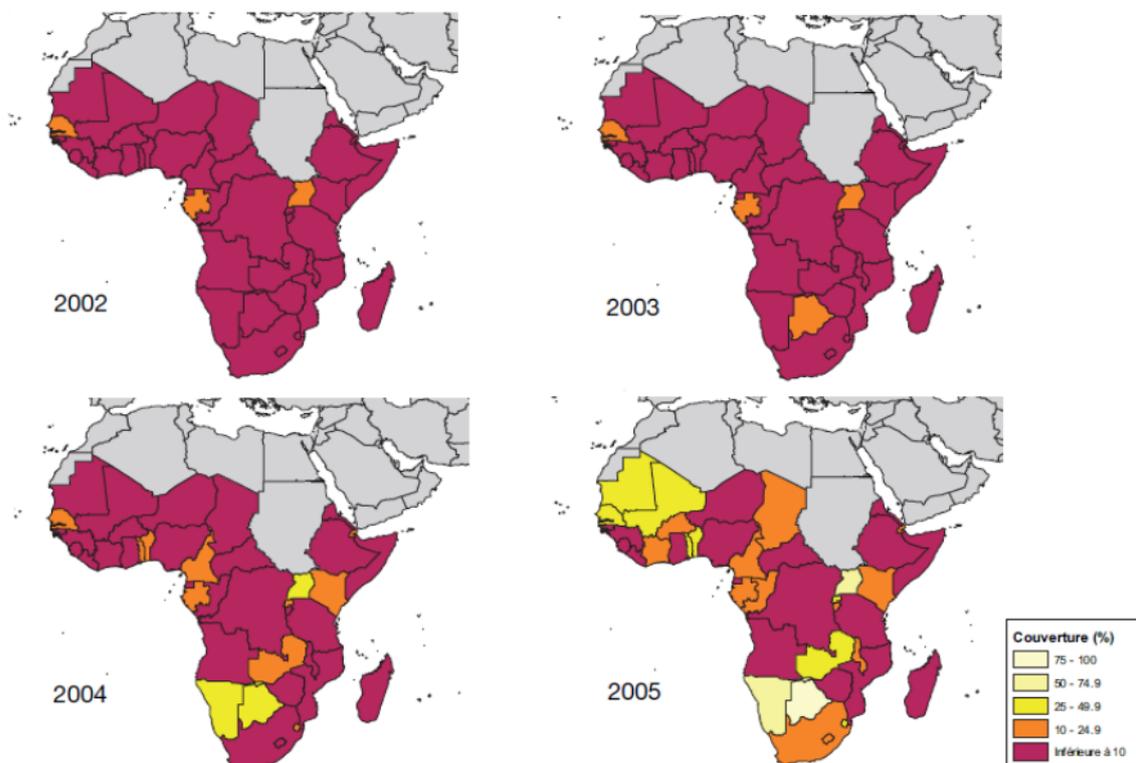
Grâce à ces initiatives, l'accès au traitement ARV a fortement progressé, notamment en Afrique subsaharienne passant de 50000 personnes traitées en 2002 à 5 millions en 2010 (**Figure 5 et Figure 6**)

Figure 5. Nombre de personnes recevant un traitement antirétroviral dans les pays à revenu faible ou intermédiaire, par région, 2002-2010



Source : OMS. Global HIV/AIDS response : epidemic update and health sector progress towards Universal Access. Progress report 2011.

Figure 6. Proportions de personnes recevant un traitement antirétroviral en Afrique subsaharienne parmi celles qui en ont besoin, 2002-2005



Source : OMS. Élargissement de l'accès au traitement antirétroviral dans le monde : rapport sur l'initiative « 3 millions d'ici 2005 » et au-delà, mars 2006.

Néanmoins, l'accès au traitement antirétroviral en Afrique est encore très limité. Ainsi, seulement 49 % [46 – 52 %] du total estimé des personnes infectées par le VIH ayant une indication immédiate de mise sous traitement recevaient ces médicaments en décembre 2010 en Afrique subsaharienne (**Tableau 1**) [12]. Il faut cependant noter que les nouvelles indications de mise sous traitement des recommandations de l'OMS de 2010 (augmentation du seuil minimal de CD4 de 200 à 350/mm³) [13] ont augmenté mécaniquement le nombre de personnes ayant une indication immédiate de traitement. Malgré tous les efforts menés par les pays d'Afrique subsaharienne et la communauté internationale pour l'extension de l'accès au traitement, l'objectif d'accès universel au traitement fixé initialement pour 2010 n'a pas été atteint. Rappelons que l'objectif visé est maintenant d'approcher l'accès universel au traitement pour toutes les personnes infectées par le VIH qui en ont besoin d'ici à l'horizon 2015

Tableau 1. Estimation du pourcentage de la couverture du traitement antirétroviral dans les pays à revenu faible ou moyen par région en fonction des lignes directrices de l'OMS de 2010, en décembre 2009 et en décembre 2010

Geographical region	December 2010			December 2009		
	Number of people receiving antiretroviral therapy	Estimated number of people eligible for antiretroviral therapy [range] ^a	Antiretroviral therapy coverage [range] ^d	Number of people receiving antiretroviral therapy	Estimated number of people eligible for antiretroviral therapy [range] ^a	Antiretroviral therapy coverage [range] ^d
Sub-Saharan Africa	5 064 000	10 400 000 [9 700 000-11 000 000]	49% [46-52%]	3 911 000	9 600 000 [9 000 000-10 200 000]	41% [38-43%]
Eastern and southern Africa	4 221 000	7 600 000 [7 100 000-8 000 000]	56% [53-59%]	3 203 000	7 000 000 [6 600 000-7 400 000]	46% [43-48%]
Western and central Africa	842 000	2 800 000 [2 600 000-3 100 000]	30% [28-33%]	709 000	2 600 000 [2 400 000-2 800 000]	27% [25-30%]
Latin America and the Caribbean	521 000	820 000 [710 000-920 000]	63% [57-73%]	469 000	780 000 [670 000-870 000]	60% [54-70%]
Latin America	461 000	720 000 [620 000-810 000]	64% [57-74%]	416 000	690 000 [590 000-780 000]	60% [53-70%]
Caribbean	60 300	100 000 [91 000-110 000]	60% [53-67%]	52 400	93 000 [84 000-110 000]	56% [50-63%]
East, South and South-East Asia	922 000	2 300 000 [2 100 000-2 500 000]	39% [36-44%]	748 000	2 300 000 [2 000 000-2 400 000]	33% [31-37%]
Europe and Central Asia	129 000	570 000 [500 000-650 000]	23% [20-26%]	114 500	520 000 [450 000-600 000]	22% [19-25%]
North Africa and the Middle East	14 900	150 000 [120 000-190 000]	10% [8-13%]	12 400	140 000 [110 000-180 000]	9% [7-12%]
Total	6 650 000	14 200 000 [13 400 000-15 000 000]	47% [44-50%]	5 255 000	13 300 000 [12 400 000-14 100 000]	39% [37-42%]

Note: some numbers do not add up because of rounding.

a See Box 5.9 for further information on the methods for estimating the need for and coverage of antiretroviral therapy in 2010.

b The 2009 figures may differ from those previously published because countries have submitted newly available data.

c All estimated needs have been developed according to 2010 WHO guidelines and criteria for initiating treatment.

d The coverage estimate is based on the unrounded estimated numbers of people receiving and needing antiretroviral therapy.

Source : OMS. Global HIV/AIDS response : epidemic update and health sector progress towards Universal Access. Progress report 2011.

2. Contexte spécifique des travaux de thèse

Nos travaux de thèse se sont déroulés dans le cadre d'études réalisées au Cameroun (pays d'Afrique centrale) et au Sénégal (pays d'Afrique de l'ouest) (**Figure 7**). Nous présentons ci-dessous la situation de l'infection par le VIH et du traitement antirétroviral dans ces deux pays.

Figure 7. Position du Cameroun et du Sénégal en Afrique.



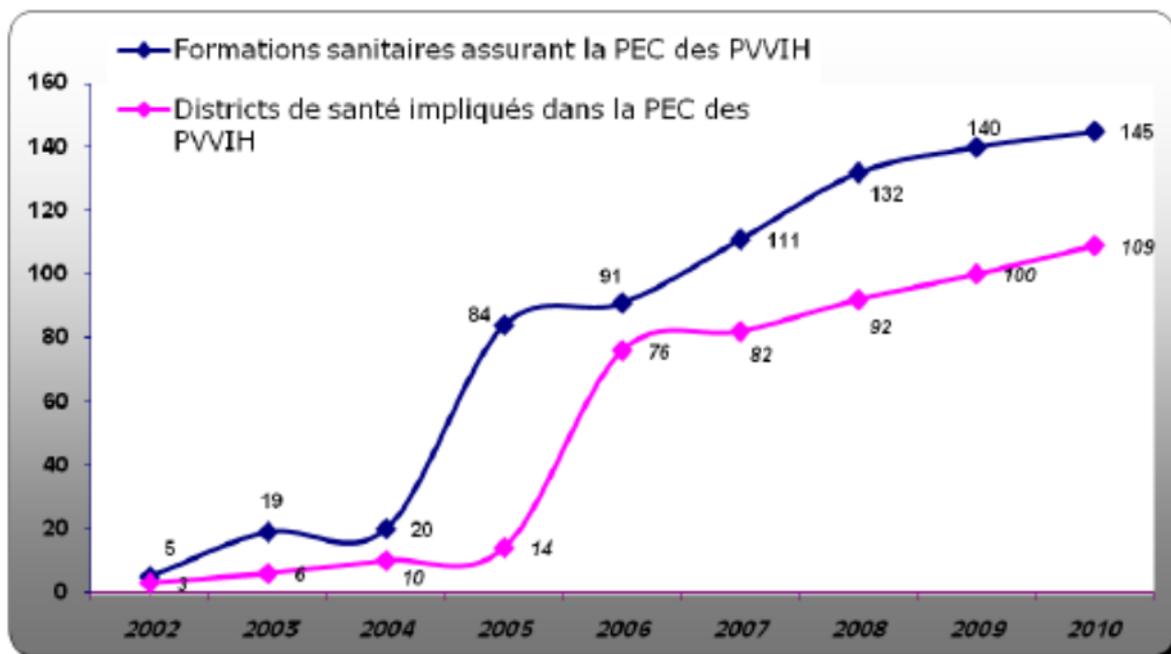
2.1. La situation au Cameroun

Les dernières données d'envergure nationale sur l'épidémie du VIH/Sida au Cameroun datent de 2011. Il s'agit des résultats de l'Enquête Démographique et de Santé et à indicateurs multiples (EDS-MICS 2011). Le rapport préliminaire de cette enquête sur la prévalence du VIH publié en février 2012a confirmé le caractère généralisé de l'épidémie dans le pays (prévalence de 4,3 % dans la tranche d'âge de

15 à 49 ans) et mis en exergue le phénomène de féminisation (prévalence de 5,6 % chez les femmes contre 2,9 % chez les hommes) de l'épidémie [14].

L'introduction des traitements antirétroviraux s'est faite de manière progressive. Avant l'année 2000, peu de patients avaient accès au traitement du fait de son coût élevé et de son indisponibilité au Cameroun. À partir de 2000, de profonds changements sont intervenus avec une organisation au niveau national de la distribution des antirétroviraux accompagnée d'une rationalisation des prescriptions et de la prise en charge. Initialement limitée à quelques hôpitaux publics des deux principales villes, Yaoundé et Douala, la prise en charge des patients a été progressivement étendue au reste du pays grâce à une politique de décentralisation menée par le Ministère de la Santé Publique. En 2001, ont ainsi été créés les premiers Centres de Traitement Agréé (CTA) dans les chefs-lieux de régions. En 2005, ce fut le tour des unités de prise en charge (UPEC) au niveau des districts sanitaires. Entre 2006 et 2010, le nombre de CTA est passé de 23 à 28 et le nombre d'UPEC est passé de 68 à 117. Au total, le nombre de districts de santé ayant au moins un CTA ou une UPEC opérationnels est passé de 76 à 109, couvrant ainsi 61 % des 178 districts de santé (**Figure 8**). Il existe toutefois des disparités interrégionales. Ainsi, la couverture en districts de santé varie de 93,5 % dans la région du Centre à 35,7 % dans la région de l'extrême nord. Parallèlement, le nombre de personnes traitées est passé de 17 156 à la fin de 2005 à 89 455 à la fin de 2010. Ceci correspond à une couverture thérapeutique de 38 % en 2010 [15, 16].

Figure 8. Nombre de structures sanitaires et de districts de santé impliqués dans la prise en charge des personnes vivant avec le VIH au Cameroun entre 2002 et 2010



Source : CNLS. Rapport annuel 2010 des activités du programme national de lutte contre le Sida

Le programme camerounais de lutte contre le sida bénéficie du soutien de la Banque Mondiale, du Fonds Mondial pour le Sida, la Tuberculose et le Paludisme, ainsi que de partenariats hospitaliers dans le cadre du programme « Ensemble pour une Solidarité Thérapeutique Hospitalière en Réseau » (ESTHER) initié par le Ministère de la Santé français. Le programme ESTHER associe notamment l'Hôpital Central de Yaoundé (impliqué dans plusieurs études de nos travaux de thèse) au CHU de Montpellier.

2.2. La situation au Sénégal

Les données les plus récentes d'envergure nationale sur l'épidémie du VIH/Sida au Sénégal datent de 2011. Il s'agit des résultats de la cinquième Enquête Démographique et de Santé à indicateurs multiples (EDSV) qui ont montré une faible prévalence en population générale (0,7 %) et une féminisation de l'épidémie (0,7 % chez les femmes contre 0,5 % chez les hommes [17]).

Le traitement ARV des personnes infectées par le VIH a débuté au Sénégal en 1998 avec l'Initiative Sénégalaise d'Accès aux ARV (ISAARV), l'une des toutes premières initiatives en Afrique, avec les étapes suivantes :

- De 1998 à 2002 : prise en charge centralisée au niveau du centre hospitalier national universitaire (CHNU) de Fann et de l'Hôpital Principal de Dakar
- De 2002 à 2004 : décentralisation au niveau des hôpitaux régionaux
- A partir 2005 : décentralisation au niveau des hôpitaux de district, unités opérationnelles du système de santé

C'est ainsi que le nombre de sites de prise en charge est passé de 2 en 2001 à 101 en 2009, et que le nombre de patients traités est passé de 4 200 en 2005 à 12 249 en 2009, la couverture thérapeutique en ARV chez les personnes vivants avec le VIH était ainsi estimée à 75,6 % [18] (sur la base des recommandations OMS de 2006 c'est-à-dire avec une initiation du traitement ARV au seuil de 200 CD4/mm³). Depuis 2008, l'essentiel des nouvelles inclusions (60,3%) se fait au niveau périphérique.

3. Thématiques spécifiques et données utilisées pour les travaux de thèse

Nos travaux de thèse ont été réalisés dans le cadre d'études sur les thématiques suivantes :

- la co-infection par le VIH et les virus des hépatites B ou C,
- les stratégies thérapeutiques de première ligne,
- les stratégies de suivi des patients sous ARV dans le contexte de la prise en charge décentralisée.

3.1. Co-infection par le VIH et les virus des hépatites B et C

Les hépatites virales chroniques B et C sont endémiques dans de nombreux pays africains, mais leur impact sur le traitement antirétroviral est peu connu en raison du manque d'étude dans le contexte spécifique de l'Afrique. Les estimations rapportent des séroprévalences, selon les pays, de 8 à 20 % pour l'hépatite B (antigènes HBs positifs) et de 1 à 18 % pour l'hépatite C (anticorps anti-VHC positifs) [19]. Dans la

pratique clinique, la recherche des antigènes HBs et des anticorps anti-VHC n'est le plus souvent pas réalisée lors du bilan pré thérapeutique des patients initiant un traitement antirétroviral en Afrique en raison des contraintes financières des patients et des limites structurelles des laboratoires (tests non réalisables dans certaines structures sanitaires, notamment en milieu rural). Le traitement antirétroviral est alors initié sans connaître le statut hépatique des patients.

Or, le traitement antirétroviral de première ligne recommandé par l'OMS pour les pays à ressources limitées est composé d'un inhibiteur non nucléosidique de la reverse transcriptase (INNRT) et de deux inhibiteurs nucléosidiques de la reverse transcriptase (INRT). Les INNRT les plus utilisés sont l'éfavirenz et la névirapine. La névirapine est particulièrement connue pour être associée à des réactions d'hypersensibilité dans les trois premiers mois de mise sous traitement et à une hépatotoxicité à long terme, avec un risque plus élevé chez les patients co-infectés par le VIH et le VHB ou le VHC [20-23]. Ce surrisque d'hépatotoxicité lié à la névirapine pourrait ainsi compromettre l'efficacité du traitement antirétroviral chez des patients souffrant déjà d'une atteinte hépatique.

Une étude menée en Afrique du Sud chez des patients recevant un traitement antirétroviral incluant l'éfavirenz avait montré un risque d'hépatotoxicité 3 fois plus élevé chez les patients co-infectés par le VIH et le VHB que chez les patients mono-infectés par le VIH. En revanche, l'efficacité virologique et immunologique ainsi que la mortalité n'étaient pas liées au statut du VHB [24].

Bien que la toxicité hépatique de la névirapine soit plus élevée que celle de l'éfavirenz, aucune étude spécifique chez des patients africains sous névirapine n'avait été réalisée. Nous avons donc comparé l'hépatotoxicité et l'efficacité virologique, immunologique et clinique du traitement antirétroviral incluant la névirapine entre des patients mono-infectés par le VIH et des patients co-infectés par le VIH et le VHB ou le VHC au Cameroun (cf **annexe A**).

Cette analyse a été réalisée à partir des données de deux études de cohorte menées à Yaoundé, la capitale politique du Cameroun. L'étude Parvy, réalisée entre 2001 et 2004 chez 109 patients à l'hôpital militaire de Yaoundé, avait pour objectif d'évaluer la faisabilité du traitement antirétroviral dans le contexte africain avec l'utilisation de

médicaments génériques [25]. De son côté, l'étude Triomune, réalisée entre 2002 et 2006 chez 60 patients à l'hôpital militaire et à l'hôpital Central de Yaoundé, avait pour objectif d'évaluer l'efficacité et la tolérance de la combinaison antirétrovirale à doses fixes la plus utilisée en Afrique [26]. Ces deux études, coordonnées par les mêmes investigateurs, avaient été réalisées avec des méthodes communes, dont les critères d'inclusion, la stratégie de suivi, les équipes médicales et sociales. Brièvement, les critères d'inclusion étaient les suivants :

- Être infecté par le VIH-1,
- Être âgé de plus de 18 ans,
- Être naïf de traitement (sauf pour les femmes ayant pris de la névirapine dans le cadre de la prévention de la transmission du VIH de la mère à l'enfant),
- Être éligible au traitement antirétroviral suivant les recommandations de l'OMS de 2000,
- Ne pas avoir de contre-indication au traitement antirétroviral,
- Avoir un taux de lymphocytes T CD4 inférieur à 350 cellules par millimètre cube.

Les données recueillies au cours de ces études étaient constituées :

- des caractéristiques d'inclusion : le sexe, l'âge, la durée d'infection, le poids corporel, la taille, l'indice de Karnofsky, le stade clinique, la charge virale du VIH, le nombre de lymphocytes CD4, et les bilans hématologiques et biochimiques standard,
- des données de suivi : données cliniques (poids corporel, indice de Karnofsky, stade clinique, événements indésirables) bimensuelles dans les deux premiers mois (J15, M1, J45, M2), puis mensuelles (M3, M4, M5, M6), et ensuite trimestrielle (M9, M12, M15, M18, M24), charge virale du VIH et lymphocytes CD4 tous les 6 mois, bilans hématologiques et biochimiques mensuels dans les trois premiers mois (M1, M2, M3) puis trimestriels (M6, M9, M12), ensuite semestriels (M18, M24) . Des données supplémentaires des transaminases étaient recueillies à J15 et J45.

Pour l'analyse de l'impact de la co-infection sur la réponse thérapeutique, les patients ont été classés en deux groupes en fonction de la présence d'une hépatite active (détection de l'ADN du VHB ou de l'ARN du VHC) : le groupe des patients

monoinfectés par le VIH et le groupe des patients co-infectés par le VIH et le VHB ou le VHC.

Les marqueurs d'hépatites B et C avaient été recherchés rétrospectivement sur les échantillons de plasma sanguin collectés à l'inclusion et congelés à -80°C [27]. Un test ELISA avait été utilisé pour détecter l'antigène HBs (Monolisa Ag HBs Plus, Bio-Rad, Marnes la Coquette, France). L'ADN du VHB avait ensuite été recherché chez les patients ayant des AgHBs positifs ou indéterminés (Cobas Ampliprep/Cobas Taq-Man quantitative assay, Roche Diagnostics GmbH, Mannheim, Germany ; limite de détection de 12 UI/mL). Le dépistage des anticorps anti-VHC avait été réalisé avec un test ELISA de troisième génération (Ortho HCV EIA 3.0, Ortho-clinical Diagnostics, Riratan, New Jersey, USA) ; les échantillons positifs ou indéterminés avaient été confirmés par un test RIBA (Chiron RIBA HCV 3.0 SIA, Chiron Corporation, Emeryville, California, USA). L'ARN du VHC avait été recherché (Cobas Ampliprep/Cobas Taq-Man quantitative assay, Roche Diagnostics GmbH, Mannheim, Germany ; limite de détection de 15 UI/mL) chez les patients ayant au moins un test sérologique anti-VHC positif ou indéterminé.

3.2. Stratégies thérapeutiques de première ligne

Dans les pays à ressources limitées, le choix des combinaisons antirétrovirales de première ligne est limité. Comme indiqué précédemment, le traitement antirétroviral de première ligne recommandé par l'OMS est composé d'un inhibiteur non nucléosidique de la reverse transcriptase (INNRT ; éfavirenz ou névirapine) et de deux inhibiteurs nucléosidiques de la reverse transcriptase (INRT). Les combinaisons d'INRT recommandées sont maintenant la zidovudine plus la lamivudine, ou le ténofovir plus soit la lamivudine soit l'emtricitabine[13]. La stavudine, longtemps utilisée grâce à son efficacité et sa disponibilité dans des combinaisons à doses fixes [26], n'est plus recommandée du fait de sa toxicité à long terme. Bien que fréquente, l'utilisation de la zidovudine est entravée par sa toxicité hématologique [28, 29]. De son côté, le ténofovir est encore peu disponible en Afrique en raison de son coût relativement élevé et des inquiétudes liées à sa toxicité rénale [30]. Dans ce contexte, l'utilisation de la didanosine en première ligne a de

nouveau été proposée particulièrement pour les pays à ressources limitées, en raison de son efficacité en association avec soit la lamivudine soit l'emtricitabine, de son faible coût et de sa disponibilité [31]. Cependant, cette proposition ne s'appuyait que sur des données scientifiques partielles de courte durée (14 mois) et incluant peu de patients des pays à ressources limitées (12 %).

Nous avons donc comparé la réponse thérapeutique à long terme entre les patients recevant un traitement incluant la didanosine et ceux recevant un traitement incluant la zidovudine (cf **annexe B**). Cette étude a été réalisée dans le cadre de la cohorte ANRS 1215 au Sénégal.

Comme indiqué précédemment, le gouvernement sénégalais a décidé de mettre en place dès 1998 l'initiative sénégalaise d'accès aux ARV (ISAARV). Cette initiative originale pour l'époque s'accompagnait d'une étude de cohorte observationnelle (la cohorte ANRS 1215). Celle-ci s'inscrivait dans le contexte du plaidoyer international pour l'accès aux ARV dans les pays à ressources limitées. L'objectif était d'évaluer la faisabilité, l'efficacité, la tolérance, l'acceptabilité et l'accessibilité du traitement ARV au Sénégal [8, 9, 32-34]. Cette cohorte a progressivement inclus les 404 premiers patients de l'ISAARV et s'est poursuivie jusqu'en 2010.

A l'inclusion, environ la moitié des patients a reçu une combinaison antirétrovirale contenant un inhibiteur de protéase (nelfinavir ou indinavir), l'autre moitié une combinaison contenant l'efavirenz ou la névirapine. Parmi ces derniers, 80 recevaient un traitement incluant la zidovudine et 71 recevaient un traitement incluant la didanosine. Un groupe de 18 patients recevaient une bithérapie (deux INRT).

Parmi les 404 patients inclus dans la cohorte ANRS 1215, 80 provenaient de deux essais cliniques (ANRS 1204/IMEA 0011 et ANRS 1206/IMEA 012). Les principaux critères d'inclusion étaient :

- être infecté par le VIH-1,
- être âgé de plus de 15 ans,
- être naïf de traitement (sauf pour les femmes ayant pris de la névirapine dans le cadre de la prévention de la transmission du VIH de la mère à l'enfant),

- score de Karnofsky ≥ 70 %
- pour les patients provenant des essais cliniques ANRS 1204/IMEA 0011 et ANRS 1206/IMEA 012
 - taux de lymphocytes T CD4 < 350 cellules/mm³,
 - charge virale plasmatique $> 30\ 000$ copies/mL
- pour les patients hors essais
 - stade clinique A et
 - taux de lymphocytes T CD4 < 350 cellules/mm³,
 - charge virale plasmatique $> 10^5$ copies/mL
 - stade clinique B et
 - taux de lymphocytes T CD4 < 350 cellules/mm³,
 - charge virale plasmatique $> 10^4$ copies/mL
 - stade clinique C
 - taux de lymphocytes T CD4 < 200 cellules/mm³ pour les patients inclus à partir d'octobre 2000

Les données recueillies au cours de ces études étaient constituées :

- des caractéristiques d'inclusion : le sexe, l'âge, la durée d'infection, le poids corporel, la taille, l'indice de Karnofsky, le stade clinique, la charge virale du VIH, le nombre de lymphocytes CD4, et les bilans hématologiques et biochimiques standard,
- des données de suivi : données cliniques (poids corporel, indice de Karnofsky, stade clinique, événements indésirables) au minimum tous les deux mois, charge virale du VIH et lymphocytes CD4 et biologie courante tous les 6 mois,

3.3. Stratégies de suivi des patients sous antirétroviraux dans le contexte de la prise en charge décentralisée

Afin d'étendre l'accès au traitement antirétroviral dans les pays à ressources limitées, l'OMS propose depuis 2003 une approche de santé publique tenant compte des limites des infrastructures sanitaires et du manque de personnels hautement qualifiés [35, 36]. Cette approche de santé publique repose notamment sur une standardisation :

- Des critères d'initiation du traitement antirétroviral,
- Des schémas thérapeutiques et une simplification des formulations,
- Des critères d'échec thérapeutique et de changement de traitement,
- Du suivi biologique,
- De la surveillance de la pharmaco-résistance du VIH au niveau de la population.

Pour le suivi biologique, l'approche de santé publique recommande d'effectuer des mesures des lymphocytes CD4 pour l'optimisation de la mise sous traitement et de l'évaluation de son efficacité. Cependant, le manque de mesures de lymphocytes CD4 ne doit pas être un obstacle à la mise sous traitement (les personnes à un stade avancé doivent être mises sous traitement même si les mesures de lymphocytes CD4 ne sont pas disponibles). Elle recommande aussi maintenant d'utiliser les mesures de charge virale (si elles sont disponibles) pour détecter ou confirmer un échec thérapeutique.

Cette stratégie de passage à l'échelle de l'OMS est basée sur la décentralisation de l'accès aux soins des centres spécialisés vers les structures périphériques (hôpitaux régionaux, hôpitaux de district, centres de santé) dans les zones urbaines et les zones rurales. Cependant, les équipements de laboratoire notamment dans ces structures périphériques sont souvent très limités. De ce fait, les mesures de charge virale et, à un moindre degré, celles des lymphocytes CD4 sont peu accessibles pour les patients ruraux. L'initiation du traitement et la détection des échecs thérapeutiques sont alors basées sur des critères cliniques seulement.

Cette stratégie pragmatique a initialement été proposée par l'OMS malgré l'absence d'évaluation rigoureuse. Une étude par modélisation mathématique a ensuite suggéré qu'une prise en charge sans surveillance immunologique ou virologique avait un effet limité sur la survie des patients et le développement des résistances [37]. Toutefois, les essais randomisés Home-Based AIDS Care (HBAC) [38] et Development of AntiRetroviral Therapy in Africa (DART) [39] ont permis de montrer respectivement que

- le décès et les évènements classant au stade Sida survenaient plus vite chez les patients suivis sans surveillance de la charge virale et des lymphocytes CD4 que chez ceux pour lesquels ces marqueurs étaient disponibles.

- L'utilisation en routine des lymphocytes CD4 lors du suivi des patients sous traitement ARV apportait un bénéfice en termes de progression clinique et de survie.

L'essai HBAC était mené en Ouganda, et comparait trois stratégies de prise en charge : la première stratégie était constituée du suivi clinique et du monitoring de la charge virale plasmatique et des lymphocytes T CD4, la deuxième stratégie était constituée du suivi clinique et du monitoring des lymphocytes T CD4, et la troisième stratégie était constituée du suivi clinique exclusif. L'essai DART était mené en Ouganda et au Zimbabwe et comparait deux stratégies de prise en charge : la première stratégie était constituée du suivi clinique et du monitoring des lymphocytes T CD4, la deuxième stratégie était constituée du suivi clinique exclusif. Dans ces deux essais, l'effet limité du suivi biologique ne permettait pas de tirer des conclusions définitives.

Un troisième essai, l'essai Stratall, a été réalisé au Cameroun pour comparer un suivi clinique exclusif à un suivi clinico-biologique (incluant la charge virale et les lymphocytes CD4) chez des patients recevant un traitement antirétroviral dans des hôpitaux de district ruraux au Cameroun (cf **annexe C**).

L'essai Stratall était un essai d'intervention randomisé, multicentrique, sans insu sur l'intervention, de non-infériorité dans 9 hôpitaux de district de la région du Centre au Cameroun (Ayos, Bafia, Mbalmayo, Mfou, Monatélé, NangaEboko, Ndikiniméki, Obala, Sa'a). Au total, 459 patients infectés par le VIH ont été inclus entre mai 2006 et janvier 2008, randomisés individuellement dans les groupes de suivi avec un ratio 1:1 et suivis pendant 24 mois.

Dans le groupe « suivi clinico-biologique », le suivi des patients était identique à celui pratiqué dans les CTA au Cameroun conformément aux recommandations nationales (décision ministérielle du 13 août 2003). Plus précisément, l'évaluation de l'efficacité thérapeutique se basait sur les résultats semestriels de charge virale et de lymphocytes T CD4 en complément des critères cliniques. Dans le groupe « suivi clinique », l'évaluation de l'efficacité thérapeutique se basait exclusivement sur les critères cliniques.

L'objectif principal de cette étude était de comparer le gain en lymphocytes T CD4 entre des patients adultes bénéficiant du suivi clinique exclusif et ceux bénéficiant d'un suivi clinico-biologique. Les objectifs secondaires étaient de comparer, entre les deux groupes, l'efficacité virologique du traitement, la survie, la progression clinique, les arrêts de traitement, la proportion de patients perdus de vue, la tolérance clinique et biologique, l'observance, l'émergence de résistances virales, l'impact du traitement sur la vie quotidienne des patients, l'acceptabilité des deux stratégies par les patients et les soignants, et leurs performances coût-efficacité.

Les critères d'inclusion étaient :

- Homme ou femme, âgé(e) d'au moins 18 ans,
- Ayant un accès relativement facile à l'un des hôpitaux de district de l'essai,
- Infection à VIH-1 groupe M confirmée,
- Répondant aux critères de mise sous ARV dans les hôpitaux de district tels que définis par le protocole camerounais et recommandés par l'OMS
 - Stade III ou IV (classification OMS) quel que soit le taux de lymphocytes totaux
 - Stade II (OMS) et taux de lymphocytes totaux $\leq 1200/\text{mm}^3$
- Patient acceptant d'être suivi mensuellement et traité pendant au moins 24 mois, selon les modalités du protocole
- Consentement éclairé signé

Les patients recevaient les traitements antirétroviraux recommandés et disponibles dans le programme national. Le traitement de première intention était composé de stavudine ou zidovudine, lamivudine, et névirapine ou efavirenz. La fréquence des consultations cliniques systématiques était bimensuelle pendant le premier mois (J14 et M1) puis trimestrielle (M3, M6, M9, M12, M15, M18, M21 et M24). Les mesures de charge virale et de lymphocytes CD4 étaient réalisées à l'inclusion puis tous les 6 mois. Les résultats de celles-ci étaient rendus aux cliniciens au fur et à mesure pour les patients du groupe « suivi clinico-biologique », mais à la fin de l'étude seulement pour les patients du groupe « suivi clinique ».

4. Présentation des problématiques statistiques abordées dans les travaux de thèse

Dans cette thèse, nous nous focalisons sur certaines problématiques statistiques rencontrées lors des études présentées ci-dessus. Au-delà de nos études, ces problématiques statistiques sont fréquentes dans l'étude du traitement antirétroviral des adultes infectés par le VIH en Afrique subsaharienne (notamment).

Nous abordons ainsi dans le premier chapitre la modélisation des données longitudinales. En effet, dans le cadre du traitement antirétroviral des personnes infectées par le VIH, la surveillance de la charge virale, des lymphocytes T CD4, de la progression clinique et de la toxicité permet d'évaluer la réponse au traitement. Cette surveillance se fait par un recueil régulier des données clinico-biologiques (charge virale plasmatique, taux de lymphocytes T CD4, bilans hématologiques et biochimiques, évènements indésirables, etc.). Cette répétition des mesures dans le temps pour chaque patient entraîne une structure de données longitudinales qui nécessite des méthodes particulières d'analyse statistique. Ainsi, dans ce chapitre, nous faisons une revue de la littérature sur les méthodes d'analyse de l'évolution d'un critère au cours du temps (première partie) et les méthodes d'analyse de la survenue d'un événement au cours du temps (deuxième partie). Plus particulièrement, dans la première partie de ce chapitre, nous présentons deux familles de modèles (les modèles mixtes et les modèles d'équations d'estimation généralisées), nous étudions la structure de la matrice de corrélation des données en fonction des hypothèses faites dans la partie aléatoire du modèle mixte dans le cas simple des modèles à intercepte ou pente aléatoires, et enfin nous comparons le modèle linéaire mixte généralisé au modèle d'équations d'estimation généralisées sur le plan théorique et pratique. Dans la deuxième partie de ce chapitre, nous présentons le modèle semi-paramétrique de Cox et ses extensions à la prise en compte des covariables dépendantes du temps et de la censure non-informative.

Le deuxième chapitre est consacré aux méthodes d'analyse des données d'essais de non-infériorité. Dans ce chapitre, nous nous sommes particulièrement intéressés à la démonstration de la non-infériorité à l'aide des tests d'hypothèses. Nous présentons d'abord le test classique de non-infériorité par intervalle de confiance lorsque la marge de non-infériorité est fixe puis nous proposons deux extensions du

test classique lorsque la marge de non-infériorité est relative (de la forme d'un pourcentage). Nous présentons également les résultats de l'étude par simulations que nous avons menée, pour déterminer la relation entre le seuil des procédures de test que nous avons proposées et le niveau des intervalles de confiance à considérer dans ces procédures de test de non-infériorité.

Enfin, le troisième chapitre aborde la question de la gestion des données manquantes. En effet, cette question est au cœur de l'analyse statistique des données partiellement observées depuis que les limites de l'analyse des cas complets ou l'analyse des cas disponibles ont été montrées. Nous faisons en premier lieu une revue de la littérature des méthodes de gestion des données manquantes en portant un accent particulier sur les méthodes d'imputation multiple par les distributions conditionnelles univariées. Nous posons ensuite le problème de la prise en compte des effets non linéaires des covariables continues et proposons une solution en modélisant la relation entre la variable à imputer et les covariables continues à l'aide des fonctions B-splines.

Chaque chapitre contient une ou plusieurs applications à des données réelles des méthodes présentées. Ces applications ont fait l'objet des articles dont les tirés à part (pour les articles publiés) et le texte intégral (pour l'article en cours de soumission) sont joints en annexe. Les deux premiers chapitres contiennent en outre une discussion sur les méthodes statistiques utilisées, ces questions n'ayant pas été abordées dans les discussions des articles.

Chapitre 1 : *Modélisation des données longitudinales*

Dans ce chapitre, nous faisons une revue des méthodes d'analyse de données longitudinales utilisées dans les deux premières études que nous avons menées. Rappelons que l'objectif de la première étude était de comparer l'hépatotoxicité et la réponse au traitement antirétroviral incluant la névirapine entre des patients monoinfectés par le VIH et des patients co-infectés par le VIH et le VHB ou le VHC. L'objectif de la seconde étude était de comparer la réponse thérapeutique à long terme entre des patients recevant un traitement incluant la didanosine et ceux recevant un traitement incluant la zidovudine. Les principaux critères d'évaluation dans ces deux études étaient la réponse virologique (c'est-à-dire la proportion de patients ayant une charge virale du VIH indétectable), la réponse immunologique (c'est-à-dire l'évolution du nombre de lymphocytes CD4), la progression clinique (c'est-à-dire la survenue d'un décès ou d'un événement classant au stade SIDA) et la toxicité du traitement (c'est-à-dire la survenue d'un événement indésirable).

Dans le cadre du traitement antirétroviral, la surveillance de la charge virale et des lymphocytes CD4 est particulièrement importante. En effet, le VIH se réplique massivement dès sa pénétration dans l'organisme (1 à 10 milliards de particules virales produites par jour) et détruit un nombre à peu près équivalent de lymphocytes CD4 (des cellules de l'immunité). Or, le risque de transmission du VIH croît avec la charge virale et le risque d'infection opportuniste croît avec le déficit immunitaire. Le traitement antirétroviral permet de contrôler la réplication virale et de favoriser la restauration immunitaire.

1. Position du problème

Lorsqu'on parle de données longitudinales, on fait référence à un ensemble de données obtenues en effectuant des mesures répétées à des temps successifs sur

les mêmes individus. Le **Tableau 2** est une représentation d'un tableau de données longitudinales d'un critère Y dont les mesures sont répétées (quatre fois) à des temps successifs chez tous les individus. Les temps de mesures peuvent être distincts entre les individus (on parle dans ce cas de données déséquilibrées) et certaines valeurs de la variable réponse Y peuvent être manquantes.

Tableau 2. Représentation d'un tableau de données longitudinales en format long

Identifiant patient	Temps de mesures	Critère mesuré
1	t_{10}	Y_{10}
1	t_{11}	Y_{11}
1	t_{12}	Y_{12}
1	t_{13}	Y_{13}
2	t_{20}	Y_{20}
2	t_{21}	Y_{21}
2	t_{22}	Y_{22}
⋮	⋮	⋮

Les données longitudinales permettent d'étudier l'évolution au cours du temps du critère mesuré (par exemple les lymphocytes T CD4 ou le grade de toxicité hépatique). Cependant, lorsque le critère étudié est une indicatrice de survenue d'un évènement qui ne peut survenir qu'une fois au cours de l'étude (le décès, le passage au stade SIDA, le premier passage en dessous du seuil de détectabilité de la charge virale de VIH), on s'intéresse plutôt au délai de survenue de cet évènement.

Les techniques classiques d'analyse statistique fréquemment utilisées (modèle de régression linéaire et extension, tests d'hypothèses, etc.) reposent sur l'hypothèse fondamentale d'indépendance des observations (ou des mesures) des critères étudiés. Cependant, lorsque les observations du critère étudié sont répétées à des temps successifs sur les mêmes individus, cette hypothèse d'indépendance des observations n'est pas systématiquement vérifiée. En effet, les mesures répétées à des temps successifs chez le même individu peuvent être corrélées entre elles, suivant le critère mesuré. Si on considère par exemple les mesures successives

(semestrielles) de lymphocytes T CD4 chez des personnes infectées par le VIH sous traitement ARV, la mesure de CD4 effectuée à une date donnée chez un individu est corrélée aux mesures précédentes effectuées sur le même individu. Ainsi, l'analyse des données longitudinales par des techniques classiques d'analyse statistique peut entraîner des inférences incorrectes, notamment en ce qui concerne les écart-types des estimations et les p-valeurs des tests associées aux estimateurs [40]. En effet, la répétition des mesures à des temps successifs chez les mêmes individus induit différentes sources de variabilité du critère mesuré : une variabilité entre les individus dite *interindividuelle*, une variabilité entre les mesures répétées effectuées sur le même individu à des temps successifs dite *intra-individuelle*, et enfin une variabilité résiduelle représentant l'erreur de mesure. En général, les variabilités intra-individuelles et résiduelles sont confondues. L'analyse de l'évolution d'un critère d'intérêt au cours du temps nécessite donc des méthodes spécifiques prenant en compte la corrélation entre les mesures répétées et les deux niveaux de variabilité du critère mesuré.

Comme indiqué plus haut, l'intérêt du suivi longitudinal peut porter sur l'étude du délai de survenue d'un évènement lorsque le critère étudié est une indicatrice de survenue d'un évènement qui ne peut survenir qu'une fois au cours du suivi. Au cours de notre travail de thèse, nous nous sommes particulièrement intéressés au délai entre l'initiation du traitement ARV et le décès, le premier passage en dessous du seuil de détectabilité de la charge virale du VIH ou encore la première survenue d'un évènement indésirable classant au stade SIDA. Lorsqu'on s'intéresse au délai de survenue d'un évènement, l'information concernant la survenue de cet évènement n'est pas forcément connue chez tous les individus à la fin du suivi, soit parce que l'évènement n'est pas survenu au cours du suivi, soit parce que certains individus sont sortis de l'étude avant l'observation de l'évènement, soit encore parce que certains individus ne sont pas à risque de l'évènement. Dans les deux premiers cas, on dit que la connaissance de l'occurrence de l'évènement est censurée chez ces individus par leur durée de suivi. Les méthodes d'analyse de survie permettent d'étudier le délai de survenue de l'évènement en prenant en compte la censure dans ces deux cas [41, 42]. Au contraire, la troisième situation fait intervenir une variable latente d'identification des individus non à risque. L'introduction de cette variable latente dans les modèles d'analyse de survie conduit aux modèles de fragilité [43-46].

Ci-dessous, nous présentons les méthodes d'analyse de l'évolution d'un critère au cours du temps, plus précisément les modèles mixtes et les modèles d'équations d'estimation généralisées. Nous présenterons les méthodes d'analyse de la survenue d'un évènement au cours du temps dans la section suivante.

2. Méthodes d'analyse de l'évolution d'un critère au cours du temps

Un cadre conceptuel est nécessaire pour pouvoir analyser la dynamique d'un critère à partir des données longitudinales. Il doit en particulier définir de la façon la plus complète possible le phénomène étudié, allant de la modélisation de la forme de l'évolution moyenne à la recherche des facteurs pouvant expliquer la variabilité interindividuelle [47]. L'évolution des moyens de calcul au cours des trois dernières décennies a favorisé l'émergence de méthodes statistiques performantes pour la modélisation des données longitudinales [47-52]. Deux approches de modélisations prédominent dans la littérature. La première est une approche de modélisation de l'évolution individuelle qui consiste à modéliser l'évolution moyenne du critère conditionnellement à la connaissance des effets individuels. Les modèles utilisés dans cette approche sont les modèles mixtes. La deuxième est une approche marginale qui consiste à modéliser l'évolution moyenne du critère dans la population d'étude en faisant une hypothèse de travail sur la structure de corrélation des mesures répétées. Les modèles utilisés dans cette approche sont les modèles d'équations d'estimation généralisées.

2.1. Modèles mixtes

Les modèles mixtes sont de plus en plus utilisés pour étudier l'évolution des marqueurs biologiques des personnes atteintes de maladies chroniques, notamment la dynamique des infections virales. Ainsi, l'évolution immuno-virologique des patients infectés par le VIH sous traitement antirétroviral a été largement décrite par les modèles mixtes [53-55].

De façon générale, on peut considérer au moins deux niveaux dans l'analyse de l'évolution d'un critère au cours du temps à partir des données longitudinales : le premier niveau réfère au changement du critère chez un individu au cours du temps tandis que le deuxième niveau réfère aux différences d'évolution entre les individus. Pour prendre en compte ces niveaux, on introduit dans le modèle des effets aléatoires spécifiques de chaque individu en considérant que l'effet de certaines covariables varie aléatoirement d'un individu à l'autre. On a ainsi un modèle qui est constitué d'un groupe de covariables dont les effets sont fixes et d'un groupe de covariables dont les effets sont aléatoires. Le modèle mixte doit ainsi sa dénomination à sa composition en deux parties dont l'une est fixe et l'autre aléatoire. À partir de ce modèle, on peut prédire les courbes d'évolution individuelle.

En fonction du type de critère mesuré (quantitatif continu ou non), on distingue deux grandes classes de modèles mixtes : le modèle linéaire mixte pour les critères quantitatifs continus et le modèle linéaire mixte généralisé (en référence au modèle linéaire généralisé) qui est une extension du modèle linéaire mixte aux critères qui ne sont pas quantitatifs continus. Parallèlement, il existe des modèles non linéaires mixtes [56-58], très utilisés en pharmacocinétique et pharmacodynamique, que nous n'avons pas utilisés dans nos analyses et que nous ne présenterons donc pas.

2.1.1. Modèle linéaire mixte

Le modèle linéaire mixte a été introduit par Laird et Ware en 1982 [52]. Il est l'un des modèles les plus utilisés pour étudier l'évolution d'un critère quantitatif continu au cours du temps. Ce modèle est spécifié comme suit :

Dans une population de N sujets, considérons $Y_i = (Y_{i_1}, \dots, Y_{i_{n_i}})$ le vecteur de réponses correspondant aux n_i mesures répétées du critère effectuées sur le sujet i . Le modèle linéaire mixte s'écrit de la façon suivante :

$$Y_i = X_i\beta + Z_i b_i + \varepsilon_i \quad (\text{modèle 1})$$

où X_i est une matrice de taille $n_i \times p$ contenant les variables explicatives (notamment le temps, ainsi qu'une fonction du temps) associée au vecteur d'effets fixes β . Z_i est une sous matrice de X_i de taille $n_i \times q$ associée au vecteur des effets aléatoires

b_i propre au sujet i , et ε_i est le vecteur d'erreurs. Les hypothèses de ce modèle sont les suivantes :

- Le vecteur des effets aléatoires b_i est distribué suivant une loi normale d'espérance zéro et de matrice de variance-covariance D : $b_i \sim \mathcal{N}(0, D)$.
- Les composantes ε_{ij} du vecteur d'erreurs ε_i suivent chacune une distribution normale centrée et de variance σ^2 . Très souvent, on fait une hypothèse d'identité et d'indépendance des erreurs ε_{ij} , ce qui permet d'écrire la matrice de variance-covariance de ε_i sous la forme $\Sigma_i = \sigma^2 I_{n_i}$. Toutefois, on peut également utiliser des matrices de variance-covariance plus complexes si on pense que la prise en compte de la corrélation entre les mesures répétées chez un sujet n'est pas entièrement captée par l'effet aléatoire introduit.
- Le vecteur d'erreurs ε_i est indépendant du vecteur des effets aléatoires b_i

Considérons par exemple que notre variable réponse est le nombre total de lymphocytes CD4 des personnes sous traitement antirétroviral mesuré à l'initiation du traitement puis après chaque semestre pendant les 24 premiers mois de traitement, la **Figure 9** nous montre les profils d'évolution de quelques patients choisis au hasard.

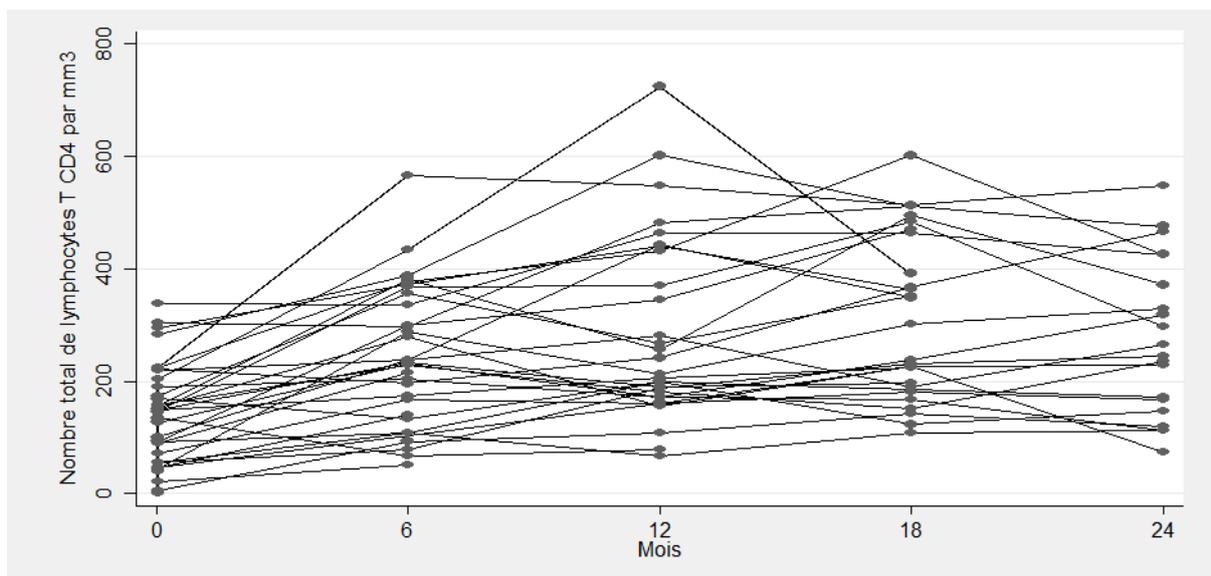


Figure 9. Profils d'évolution individuels du nombre total de lymphocytes T CD4 pendant les 24 premiers mois de traitement ARV de 20 patients infectés par le VIH, suivis au Cameroun dans les études de cohorte Parvy et Triomune. Données regroupées à M6, M12, M18, M24.

La **Figure 9** montre une grande variabilité des mesures de lymphocytes CD4 à l'initiation du traitement entre les individus. Cette variabilité est trop grande pour être attribuée aux seules erreurs de mesure. On fait donc l'hypothèse que les vraies mesures de lymphocytes CD4 à l'initiation du traitement diffèrent entre les patients et cela se traduit dans le modèle par un intercepte aléatoire (intercepte individuel) α_i de moyenne α_0 et de variance σ_{0i}^2 , $\alpha_i = \alpha_0 + b_{0i}$ avec $b_{0i} \sim \mathcal{N}(0, \sigma_{0i}^2)$. Les profils d'évolution individuels sont également différents les uns des autres au cours du temps, donc chaque patient a une pente d'évolution (dans un premier temps on considère que l'effet du temps est linéaire) différente de celle des autres patients au cours du suivi et cela se traduit dans le modèle par une pente aléatoire (individuelle) associée au temps γ_i de moyenne b_0 et de variance σ_{1i}^2 , $\gamma_i = b_0 + b_{1i}$ avec $b_{1i} \sim \mathcal{N}(0, \sigma_{1i}^2)$. Ensuite, pour diminuer la variabilité interindividuelle, on peut introduire d'autres covariables dans le modèle telles que les caractéristiques des patients à l'initiation du traitement (mesures anthropométriques, état clinique, mesures biologiques) ou des variables dépendantes du temps. On met toutes ces covariables dans une matrice X_i et on peut écrire le modèle de la façon suivante :

$$\sqrt{CD4}_i = X_i \beta + \alpha_i + \gamma_i * t_i + \varepsilon_i = \underbrace{X_i \beta + \alpha_0 + b_0 * t_i}_{\text{partie fixe}} + \underbrace{b_{0i} + b_{1i} * t_i}_{\text{partie aléatoire}} + \varepsilon_i \quad (\text{modèle 2})$$

où $CD4_i$ est le vecteur des mesures du nombre total de lymphocytes CD4 effectuées sur le patient i et t_i est le vecteur des temps successifs de mesure pour le patient i . Dans ce cas, nous avons utilisé une transformation par la racine carrée du nombre total de lymphocytes CD4 d'une part pour réduire la variabilité de cette variable, d'autre part pour pouvoir approximer la distribution de $\sqrt{CD4}_i$ par une distribution gaussienne.

Nous avons supposé pour le modèle présenté ci-dessus que les profils individuels sont linéaires, ce qui en fait n'est pas vrai. On voit bien en effet sur la **Figure 9** que les profils individuels sont linéaires par morceaux. Nous présenterons dans la **section 2.1.3** comment prendre cela en compte dans le modèle.

Jusqu'ici, nous n'avons pas présenté la façon dont la corrélation entre les mesures est prise en compte dans le modèle. Mais, comme nous l'avons dit plus haut, l'introduction des effets aléatoires dans le modèle permet de prendre implicitement en compte cette corrélation entre les mesures. En effet, à partir des hypothèses faites sur le modèle, on peut écrire le modèle linéaire mixte de la façon suivante :

$$Y_i \sim \mathcal{N}(X_i\beta, V_i = Z_i D Z_i^T + \Sigma_i)$$

où V_i est la matrice de variance-covariance de Y_i . Ainsi, la corrélation entre les mesures est prise en compte à travers V_i . L'expression ci-dessus est aussi utilisée pour estimer les paramètres du modèle, à savoir le vecteur des effets fixes β et les éléments des matrices de variance-covariance D et Σ_i .

Le principe de l'estimation est de déterminer les estimateurs des paramètres du modèle qui rendent le modèle le plus vraisemblable possible pour les données observées. Pour cela, on définit la vraisemblance du modèle qui est la probabilité que les données observées soient issues du modèle

$$P(Y|\beta, D, \Sigma) = L(\beta, D, \Sigma) = \prod_{i=1}^N \left(\frac{1}{2\pi} \right)^{\frac{n_i}{2}} |V_i|^{-\frac{1}{2}} \exp \left(-\frac{1}{2} (Y_i - X_i\beta)^T V_i^{-1} (Y_i - X_i\beta) \right)$$

Les estimateurs des paramètres qui rendent le modèle le plus vraisemblable sont ceux qui maximisent la vraisemblance : on les appelle les estimateurs du maximum de vraisemblance. Les estimateurs du maximum de vraisemblance des paramètres β, D, Σ sont obtenus en maximisant le logarithme de la vraisemblance. On utilise généralement des algorithmes itératifs pour déterminer ces estimateurs car les équations de maximisation ne sont pas linéaires. Deux algorithmes itératifs sont très utilisés dans la littérature : l'algorithme EM (Expectation Maximization) [52] et l'algorithme de Newton-Raphson [59]. Cependant, l'estimateur du maximum de vraisemblance des composantes des matrices de variance-covariance D, Σ est en général biaisé. Un tel biais ne peut disparaître même avec une grande taille d'échantillon si le nombre d'effets fixes est proportionnel à la taille de l'échantillon. En effet, dans ce cas, les estimateurs du maximum de vraisemblance ne sont pas consistants [60]. Une alternative à l'estimateur du maximum de vraisemblance pour les composantes des matrices de variance-covariance est l'estimateur obtenu par la méthode du maximum de vraisemblance restreinte proposée par *Thompson* [61]. Cet estimateur est sans biais et consistant.

2.1.2. Modèle linéaire mixte généralisé (GLMM)

Le modèle linéaire généralisé (GLM en anglais) pour les données indépendantes, a été proposé par *McCullagh* et *Nelder* [62] comme extension du modèle linéaire pour les variables réponses quantitatives discrètes ou qualitatives. Notons que les hypothèses clés du modèle linéaire classique sont :

- (i) Les observations sont indépendantes
- (ii) L'espérance de la variable d'intérêt est une fonction linéaire des variables explicatives
- (iii) La variance des observations est indépendante de la moyenne.

L'extension proposée par *McCullagh* et *Nelder*, consiste à modifier les hypothèses (ii) et (iii) ci-dessus par :

- (ii)' L'espérance de la variable d'intérêt est reliée à une fonction linéaire des autres covariables par une fonction de lien.
- (iii)' La variance des observations est une fonction de la moyenne.

Le modèle GLM englobe une variété de modèles qui incluent les cas où la variable réponse suit une distribution normale, binomiale, de poisson ou multinomiale. Ce modèle est donc applicable lorsque la variable réponse n'est pas quantitative continue.

Par analogie, on définit le modèle linéaire mixte généralisé (GLMM en anglais) comme une extension du modèle linéaire mixte de la façon suivante :

On suppose que les mesures Y_{i1}, \dots, Y_{in_i} de l'individu i sont conditionnellement indépendantes sachant le vecteur des effets aléatoires b_i tel que $Y_i|b_i \sim \mathcal{N}(X_i\beta + Z_ib_i, \Sigma_i)$, où Σ_i est une matrice diagonale. Les trois éléments clés pour définir le modèle linéaire mixte ci-dessus étaient :

- (i) L'indépendance conditionnelle des observations,
- (ii) La distribution conditionnelle est gaussienne de moyenne $X_i\beta + Z_ib_i$
- (iii) La variance conditionnelle est indépendante de l'espérance conditionnelle.

Ces trois éléments peuvent être généralisés pour définir le modèle GLMM de la façon suivante :

On suppose que la distribution conditionnelle de Y_i sachant b_i est un membre de la famille exponentielle de paramètre ξ_i, ϕ qu'on notera $ED(\xi_i, \phi)$ avec pour densité de probabilité,

$$f_i(\mathbf{y}) = \exp \left\{ \frac{\mathbf{y}\xi_i - \varphi(\xi_i)}{a_i(\phi)} + c_i(\mathbf{y}, \phi) \right\} \quad (\text{modèle 3})$$

où $\varphi(\cdot)$, $a_i(\cdot)$, $c_i(\cdot, \cdot)$ sont des fonctions connues, ϕ est le paramètre de dispersion qui peut être inconnue, ξ_i est une fonction de l'espérance conditionnelle $\mu_i = E(Y_i|b_i)$.

L'espérance conditionnelle $\mu_i = E(Y_i|b_i)$ est associée au prédicteur linéaire $\eta_i = X_i\beta + Z_ib_i$, à travers une fonction de lien $g(\cdot)$ telle que $g(\mu_i) = \eta_i$. D'après les propriétés des familles exponentielles, on a :

$\dot{\phi}(\xi_i) = \mu_i$, $Var(Y_i) = \phi \ddot{\phi}(\xi_i) = \phi \ddot{\phi}(\dot{\phi}^{-1}(\mu_i)) = \phi v(\mu_i)$ où $v(\mu_i)$ est appelée fonction de variance. En particulier, si $g(\mu_i) = \mu_i$, la fonction de lien est l'identité ; c'est le cas de la distribution gaussienne. Si $g(\mu_i) = \xi_i$, la fonction de lien est dite canonique ; c'est le cas avec les fonctions de lien naturelles des distributions binomiales et de Poisson. Une liste de fonctions de lien canoniques est donnée dans le **Tableau 3** ci-dessous.

Tableau 3. Liste des fonctions de lien canoniques les plus utilisées dans le modèle GLMM

Distribution	Fonction de lien $g(\cdot)$	Fonction de variance $v(\mu_i)$	Paramètre de dispersion ϕ
Normale	identité	1	σ^2
Binomiale	logit	$\mu_i(1 - \mu_i)$	1
Poisson	log	μ_i	1
Gamma	log	μ_i^2	1
Binomiale négative	log	$\mu_i(1 + \mu_i)$	1

Comme pour le modèle linéaire mixte, l'estimation des paramètres du modèle GLMM se fait soit par maximum de vraisemblance, soit par maximum de vraisemblance restreinte. Cependant, on ne peut obtenir une forme analytique de la vraisemblance (sauf dans le cas gaussien). En effet, cette vraisemblance fait intervenir des intégrales de grande dimension qui ne peuvent pas être résolues analytiquement et qui sont difficiles à résoudre numériquement. Ainsi, l'inférence basée sur la vraisemblance dans le modèle GLMM est encore un challenge numérique. Différentes approches ont été développées dans la littérature soit pour déterminer les estimateurs du maximum de vraisemblance dans ce contexte, soit pour éviter les difficultés numériques posées par les méthodes d'estimation par maximum de vraisemblance. Les principales approches visant à développer des méthodes

numériques pour l'estimation par le maximum de vraisemblance sont l'algorithme EM par Monte Carlo [50, 63] et une méthode d'estimation par parties [64].

2.1.3. Modélisation de la partie fixe et de la partie aléatoire dans les modèles mixtes

La partie fixe des modèles mixtes permet de modéliser d'une part la forme du changement au cours du temps, d'autre part l'association entre le changement et l'exposition à certains facteurs. La partie aléatoire quant à elle permet de modéliser explicitement la variabilité interindividuelle et implicitement la corrélation entre les mesures. Nous présentons dans la suite quelques approches de modélisation illustrées par des applications.

2.1.3.1. Modélisation de la partie fixe

La première approche de modélisation de la partie fixe consiste à regrouper les différentes mesures des individus à des dates fixes comme sur la **Figure 9 (page 41)**. On peut ainsi modéliser la forme de l'évolution du critère d'intérêt à l'aide d'une fonction linéaire du temps sur chaque intervalle de temps entre deux mesures. Si nous reprenons le modèle d'évolution du nombre total de lymphocytes CD4 présenté dans la section **2.1.1 (page 40)**

$$\sqrt{CD4_i} = \underbrace{X_i\beta + \alpha_0 + b_0 * t_i}_{\text{partie fixe}} + \underbrace{b_{0i} + b_{1i} * t_i}_{\text{partie aléatoire}} + \varepsilon_i$$

Il suffira dans la formulation du modèle de considérer t_i comme une variable catégorielle. Si on choisit comme temps de référence l'initiation au traitement, le coefficient b_{0j} associé au niveau t_{ij} (M6, M12, M18, M24) de la variable t_i sera interprété comme le gain moyen de lymphocytes CD4 entre l'initiation au traitement et le temps t_{ij} (M6, M12, M18, M24).

Ce modèle n'impose aucune contrainte sur l'évolution des moyennes estimées, est facilement interprétable et décrit exactement les données observées. Cependant, il présente quelques inconvénients :

- Le regroupement des données entraîne une perte d'information et peut invalider les résultats.
- Ce modèle utilise de façon inefficace les données puisqu'un nouveau modèle est utilisé pour chaque période sur un nombre restreint de données.

- Le nombre de coefficients de régression estimés dépend du nombre de mesures effectuées.
- Ce modèle ne prend pas en compte l'ordre des temps de mesure

Une alternative dans le même contexte consiste à ne pas regrouper les données. La **Figure 10** ci-dessous nous montre ainsi une représentation des profils d'évolution individuels des lymphocytes T CD4 dans ce cas.

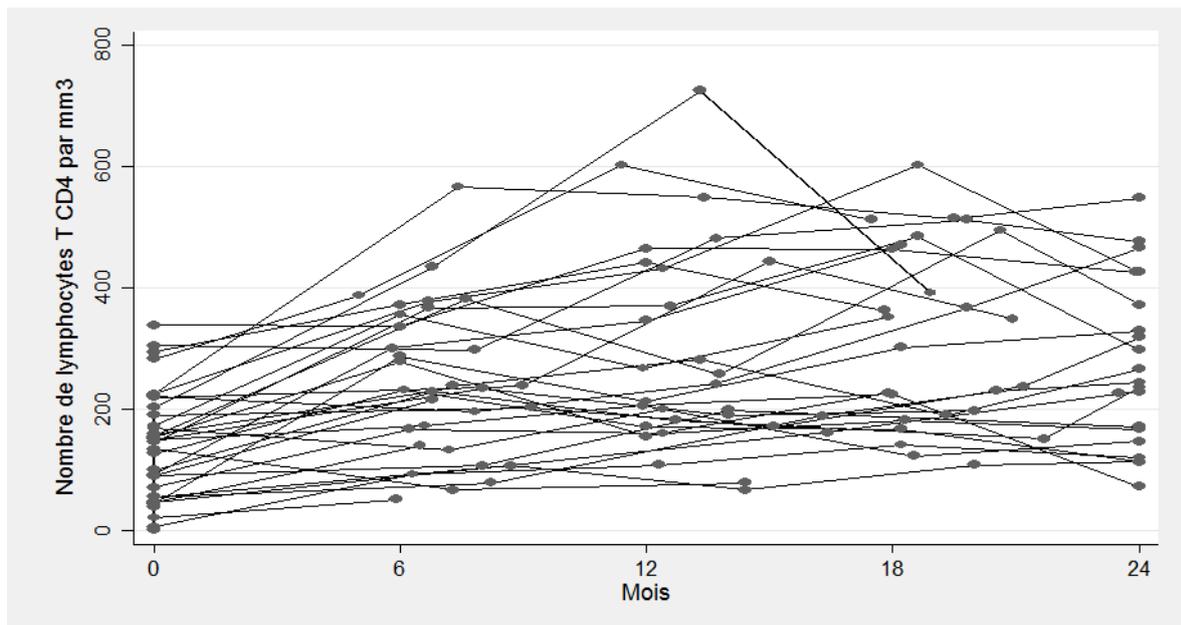


Figure 10. Profils d'évolution individuels du nombre total de lymphocytes CD4 pendant les 24 premiers mois de traitement de 20 patients infectés par le VIH suivis au Cameroun dans les études de cohorte Parvy et Triomune. Sans regroupement des données (données déséquilibrées).

Dans ce cas, on ne peut modéliser les profils d'évolution par des segments de droite comme dans l'approche précédente parce que le vecteur des temps de mesures n'est pas identique pour tous les individus. Généralement, on impose une contrainte sur l'évolution des moyennes estimées (linéaire, linéaire avec rupture de pente, quadratique, splines, etc.), rendant les profils d'évolution individuels estimés les plus proches possible des profils d'évolution observés. Pour choisir la forme de l'évolution moyenne dans nos analyses, nous avons utilisé le profil moyen d'évolution observé lorsque les données étaient regroupées et les moyennes mobiles lorsque les données n'étaient pas regroupées. La **Figure 11** nous présente l'évolution moyenne du nombre total de lymphocytes T CD4 des patients VIH+ sous traitement ARV, suivis au Cameroun selon les deux approches d'analyse.

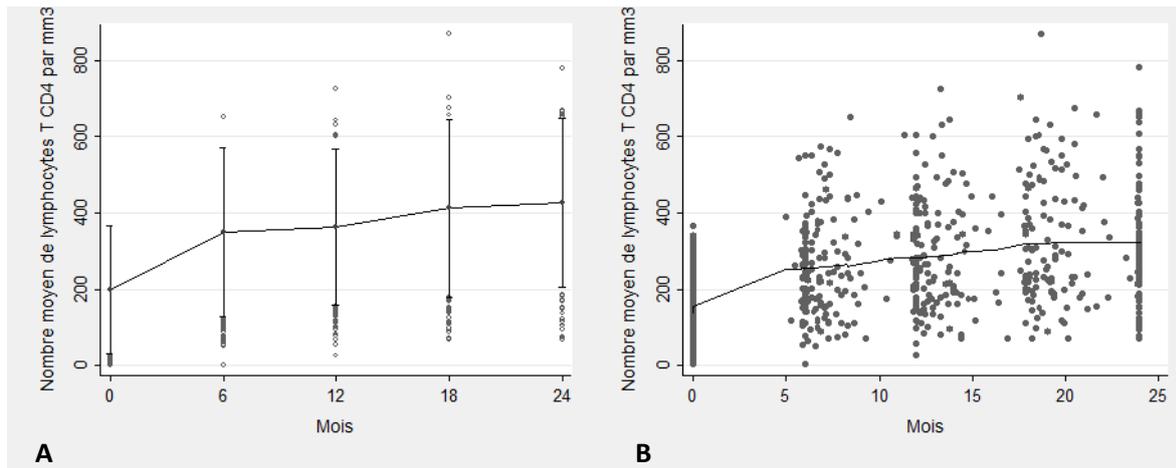


Figure 11. Profil moyen d'évolution du nombre total de lymphocytes CD4 pendant les 24 premiers mois sous traitement de l'ensemble des patients infectés par le VIH suivis au Cameroun dans les études de cohorte Parvy et Triomune. **A** lorsque les données sont regroupées, moyenne des observations à J0, M6, M12, M18, M24, ± 2 écarts types. **B** lorsque les données ne sont pas regroupées, courbe d'évolution moyenne obtenue par une méthode de lissage (les moyennes mobiles).

Ces deux représentations nous montrent que l'évolution des CD4 après la mise sous traitement des patients infectés par le VIH est biphasique. On observe en moyenne une croissance rapide des CD4 dans les 6 premiers mois de traitement et une croissance un peu plus lente au-delà. Cette observation a également été faite sur l'évolution des CD4 des patients infectés par le VIH, sous traitement ARV, suivis au Sénégal (cf. **annexe B**). Nous avons pris cela en compte dans nos analyses en permettant dans nos modèles une rupture de pente à 6 mois. Pour cela, nous avons utilisé un modèle à deux phases dans la partie fixe ; une phase linéaire dans les 6 premiers mois de traitement et une phase quadratique entre 6 et 24 mois (pour prendre en compte la légère courbure que nous avons observée dans la deuxième phase). Cela se traduit de la façon suivante dans l'écriture du modèle.

$$\sqrt{CD4}_i = \underbrace{X_i\beta + \alpha_0 + b_0 * t_i}_{\text{phase linéaire}} + \underbrace{b_1 * t_{(6+)_i} + b_2 * (t_{(6+)_i})^2}_{\text{phase quadratique}} + \underbrace{b_{0i} + b_{1i} * t_i + \varepsilon_i}_{\text{partie aléatoire}} \quad (\text{modèle 4})$$

partie fixe

où $t_{(6+)_i} = \max(0, t_i - 6)$, b_0 est la pente de la courbe dans la première phase, b_1 est la différence de pente entre la première et la deuxième phase et b_2 est la courbure dans la deuxième phase.

L'objectif principal d'une analyse de données longitudinales est très souvent de comparer l'évolution d'un critère entre deux ou plusieurs groupes d'individus. A cet effet, les questions qui se posent souvent au cours des analyses sont les suivantes :

a) Y'a-t-il une différence d'évolution entre les groupes ?

Cette question est souvent la plus importante de l'analyse ; c'est la première question à laquelle on essaye de répondre. Elle concerne l'effet de l'interaction entre les groupes et le temps. Si l'effet de l'interaction entre les groupes et le temps n'est pas significatif, cela se traduit graphiquement par un parallélisme entre les courbes d'évolution moyenne des groupes. Dans ce cas, la seconde question d'intérêt est la suivante.

b) Y'a-t-il une évolution au cours du temps ?

Cette question concerne l'effet associé au temps. Si l'effet du temps n'est pas significatif, cela se traduit graphiquement par des courbes d'évolution moyenne des groupes plates et horizontales, c'est-à-dire que le critère mesuré est constant au cours du temps.

c) Les courbes d'évolution moyenne des groupes sont-elles superposées ?

Cette question concerne l'effet associé au groupe.

Pour répondre à ces questions, il est important d'introduire dans la partie fixe du modèle une variable permettant d'identifier les groupes, une variable permettant de prendre en compte le temps, ainsi que les termes d'interaction entre le groupe et le temps. Dans nos deux premières études (cf. **annexe A et B**), l'objectif principal était de comparer la réponse au traitement antirétroviral entre deux groupes de patients infectés par le VIH, d'une part les mono-infectés vs les co-infectés (cf. **annexe A**), d'autre part les patients recevant un traitement incluant la didanosine vs les patients recevant un traitement incluant la zidovudine (cf. **annexe B**). Au cours de nos analyses, nous avons donc inclus dans le modèle d'évolution des CD4 la variable identifiant les groupes (*grp*), un terme d'interaction entre le groupe et le temps dans la phase linéaire et deux termes d'interaction entre le groupe et le temps dans la phase quadratique. Ceci se traduit dans notre modèle de la façon suivante :

$$\sqrt{CD4_i} = \underbrace{X_i\beta + \alpha_0 + \alpha_1 * grp + b_0 * t_i}_{\text{phaselinéaire}} + \underbrace{\alpha_2 * grp \times t_i}_{\text{interaction}} + \underbrace{b_1 * t_{(6+)_i} + b_2 * (t_{(6+)_i})^2}_{\text{partiefixe}} + \underbrace{\alpha_3 * grp \times t_{(6+)_i} + \alpha_4 * grp \times (t_{(6+)_i})^2}_{\text{interaction}}$$

$$+ \underbrace{b_{0i} + b_{1i} * t_i + \varepsilon_i}_{\text{partieléatoire}} \quad (\text{modèle 5})$$

Dans ces études, l'effet principal du groupe et l'effet de l'interaction entre le groupe et le temps n'étaient pas significatifs. Cependant, l'effet principal du temps était significatif dans les deux phases (**Tableau 4**). On en concluait donc que l'augmentation moyenne du nombre total de lymphocytes CD4 était similaire entre les groupes (**Figure 12**)

Tableau 4. Résultat de la modélisation de l'évolution des lymphocytes T CD4 chez les patients co-infectés par le VIH et le VHB ou le VHC sous traitement ARV incluant la névirapine par rapport aux patients mono-infectés par le VIH suivis au Cameroun dans les études de cohorte Parvy et Triomune.

	Coefficient [IC]	P-value
Effets principaux		
Co-infection	0,61 [-0,86 ; 2,07]	0,8
Pente associée au temps (phase linéaire)	0,65 [0,55 ; 0,75]	<0,001
Pente associée au temps (phase quadratique)	-0,45 [-0,61 ; -0,29]	<0,001
Effet quadratique	-0,00 [-0,01 ; 0,00]	0,06
Effets de l'interaction		
Co-infection*pente (phase linéaire)	0,02 [-0,20 ; 0,24]	0,9
Co-infection*pente (phase quadratique)	-0,04 [-0,38 ; 0,31]	0,8
Co-infection*effet quadratique	0,00 [-0,00 ; 0,01]	0,8

La variable « *Co-infection* » permet d'identifier le groupe des patients (mono-infectés ou co-infectés). Le groupe des patients mono-infectés était le groupe de référence dans cette analyse.

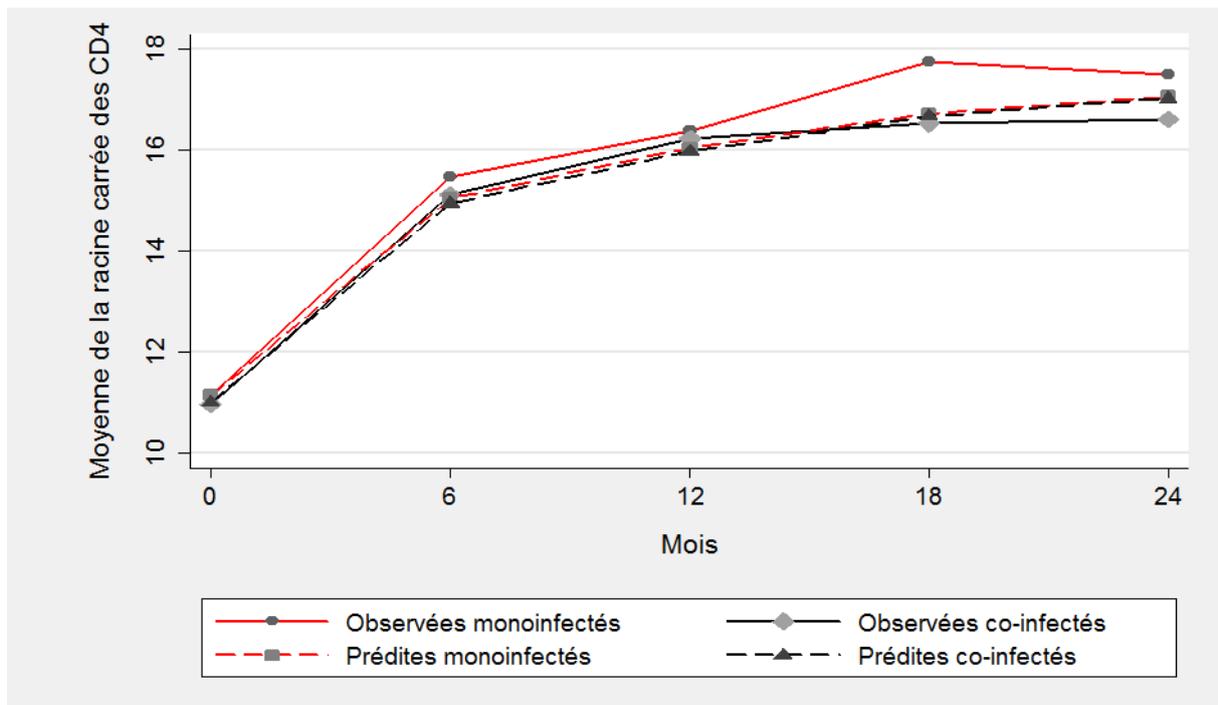


Figure 12. Courbes moyennes observées et courbes moyennes prédites (à l'aide du modèle 6 ci-dessous) de l'évolution de la racine carrée du taux de lymphocytes T CD4 chez les patients mono-infectés et chez les patients co-infectés sous traitement ARV incluant la névirapine, suivis au Cameroun dans les études Parvy et Triomune.

La figure ci-dessus nous permet bien de voir que les courbes des moyennes prédites à l'aide du **modèle 6** dans les deux groupes de patients sont superposées. On constate aussi qu'il y a un biais sur l'estimation des moyennes prédites. Cependant, ce modèle était celui qui était le plus en adéquation avec les données parmi les modèles testés.

2.1.3.2. Modélisation de la partie aléatoire

La modélisation de la variabilité individuelle est une difficulté supplémentaire à l'application des modèles mixtes. En effet, en dehors des choix techniques correspondant aux expériences randomisées et normalisées, le choix *a priori* du caractère fixe ou aléatoire d'un effet dépend du critère étudié. L'objectif de ce paragraphe n'est pas de discuter du choix des effets aléatoires, mais des conséquences de ce choix sur la modélisation lorsque le critère étudié est quantitatif continu. Nous allons nous limiter ici aux modèles à intercepte et pente aléatoires (pente associée au temps).

- L'introduction d'un intercepte aléatoire et d'une pente fixe dans le modèle traduit d'une part la variabilité de la première mesure considérée du critère d'intérêt (variabilité de l'intercepte) et d'autre part une homogénéité des profils d'évolution. Le modèle s'écrit de la façon suivante :

$$Y_i = X_i\beta + b_{0i} + \varepsilon_i$$

avec $b_{0i} \sim \mathcal{N}(0, \sigma_{0i}^2)$, $\varepsilon_i \sim \mathcal{N}(0, \sigma^2 I_{n_i})$. $Cov(Y_{ij}, Y_{ik}) = \sigma_{0i}^2$, $Var(Y_{ij}) = \sigma_{0i}^2 + \sigma^2$. Ce modèle induit implicitement une corrélation constante et positive entre les mesures du même individu et une variance identique des différentes mesures.

- L'utilisation d'un modèle à pente aléatoire sans intercepte aléatoire traduit d'une part une hétérogénéité des profils individuels d'évolution et d'autre part une homogénéité des mesures de la variable réponse à l'inclusion. Le modèle s'écrit de la façon suivante.

$$Y_i = X_i\beta + b_{1i} * t_i + \varepsilon_i$$

avec $b_{1i} \sim \mathcal{N}(0, \sigma_{1i}^2)$, $\varepsilon_i \sim \mathcal{N}(0, \sigma^2 I_{n_i})$. $Cov(Y_{ij}, Y_{ik}) = t_{ij}t_{ik}\sigma_{1i}^2$, $Var(Y_{ij}) = t_{ij}^2\sigma_{1i}^2 + \sigma^2$. Ce modèle induit implicitement une corrélation positive et croissante au cours du temps (tend vers 1 quand t_{ij} et t_{ik} sont très grands, la corrélation entre deux mesures est inversement proportionnelle à l'intervalle de temps entre les deux mesures) entre les mesures du même individu et une variance croissante au cours du temps.

- L'utilisation d'un modèle avec intercepte et pente aléatoires traduit une variabilité de la première mesure considérée du critère d'intérêt et une hétérogénéité des profils individuels d'évolution. Le modèle s'écrit de la façon suivante.

$$Y_i = X_i\beta + b_{0i} + b_{1i} * t_i + \varepsilon_i$$

avec $b_{0i} \sim \mathcal{N}(0, \sigma_{0i}^2)$, $b_{1i} \sim \mathcal{N}(0, \sigma_{1i}^2)$, $\varepsilon_i \sim \mathcal{N}(0, \sigma^2 I_{n_i})$.

$Cov(Y_{ij}, Y_{ik}) = \sigma_{0i}^2 + t_{ij}t_{ik}\sigma_{1i}^2 + (t_{ij} + t_{ik})Cov(b_{0i}, b_{1i})$, $Var(Y_{ij}) = \sigma_{0i}^2 + t_{ij}^2\sigma_{1i}^2 + \sigma^2 + 2t_{ij}Cov(b_{0i}, b_{1i})$. Pour déterminer la matrice de corrélation de Y_i , il est nécessaire de spécifier la structure de la matrice de corrélation de (b_{0i}, b_{1i}) . Cette hypothèse supplémentaire permet en effet de prendre en compte la corrélation entre la première mesure de la variable réponse et la pente de l'évolution de la variable réponse au cours du temps. Plusieurs structures de corrélation entre l'intercepte et la pente aléatoires sont proposées par les logiciels :

- **La structure « identité »** : elle suppose une indépendance entre l'intercepte et la pente aléatoires, une égalité entre les variances de l'intercepte et de la pente aléatoires, $Cov(b_{0i}, b_{1i}) = 0$, $\sigma_{0i}^2 = \sigma_{1i}^2$. Ainsi,

$Cov(Y_{ij}, Y_{ik}) = \sigma_{0i}^2(1 + t_{ij}t_{ik})$, $Var(Y_{ij}) = \sigma_{0i}^2(1 + t_{ij}^2) + \sigma^2$. La **Figure 13** nous montre les courbes d'évolution de la corrélation $Corr(Y_{ij}, Y_{ik})$ entre Y_{ij} et Y_{ik} en fonction de t_{ik} pour des valeurs de t_{ij} et $\frac{\sigma^2}{\sigma_{0i}^2}$ fixées.

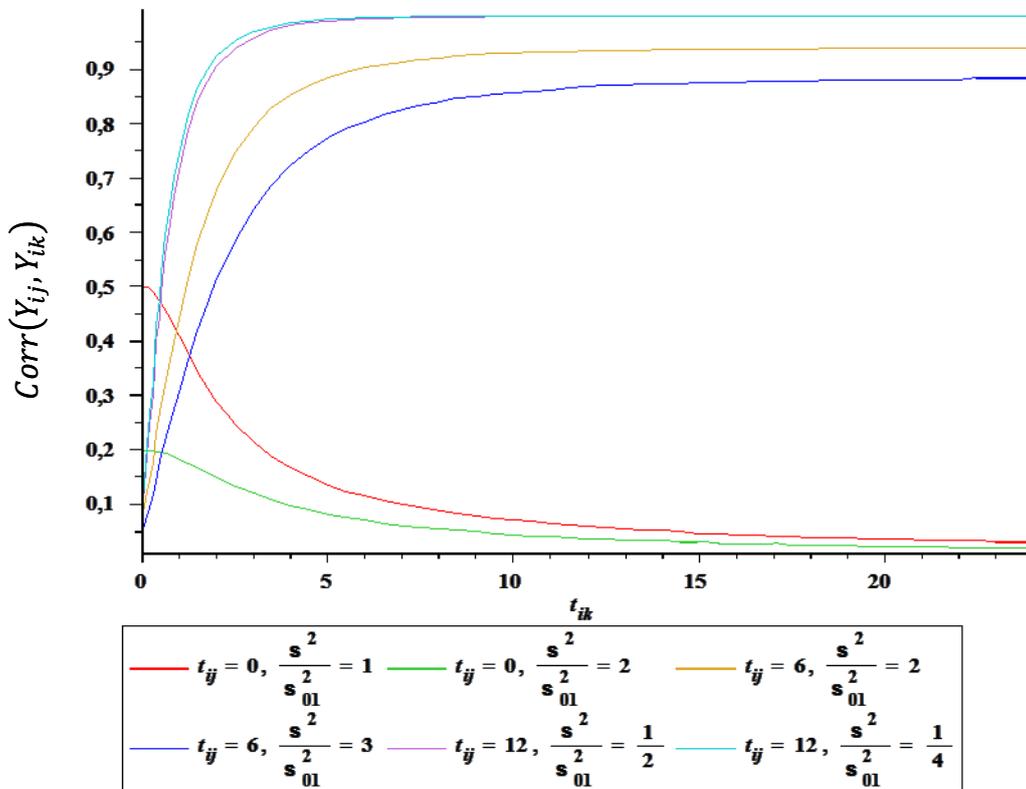


Figure 13. Courbes d'évolution de la corrélation entre les mesures en fonction du temps pour un modèle linéaire mixte à intercepte et pente aléatoires, avec une structure de corrélation identité entre l'intercepte et la pente aléatoire

On déduit de la **Figure 13** que cette hypothèse induit implicitement une corrélation positive et décroissante au cours du temps entre la première mesure Y_{i0} et les autres mesures du temps si $t_{i0} = 0$, une corrélation positive et croissante au cours du temps entre deux mesures Y_{ij} et Y_{ik} de l'individu i (la corrélation $Corr(Y_{ij}, Y_{ik})$ est proportionnelle à l'intervalle de temps $|t_{ik} - t_{ij}|$ entre les deux mesures et au temps de mesure de la première mesure entre Y_{ij} et Y_{ik} , elle tend vers 1 quand t_{ij} et t_{ik} sont grands) entre les mesures Y_{ij} et Y_{ik} si $t_{ij} \neq t_{ik} \neq 0$. La variance des mesures est croissante au cours du temps.

- **La structure « échangeable »** : elle suppose l'existence d'une corrélation et une égalité de variance entre l'intercepte et la pente aléatoires, $Cov(b_{0i}, b_{1i}) = \rho$, $\sigma_{0i}^2 = \sigma_{1i}^2$. Ainsi, $Cov(Y_{ij}, Y_{ik}) = \sigma_{0i}^2(1 + t_{ij}t_{ik}) + \rho(t_{ij} + t_{ik})$, $Var(Y_{ij}) = \sigma_{0i}^2(1 + t_{ij}^2) + 2\rho t_{ij} + \sigma^2$. Cette structure est plus flexible car elle permet des corrélations négatives entre les mesures. Elle fait intervenir un paramètre supplémentaire à estimer, et donc rend le modèle plus complexe.
- **La structure « indépendante »** : elle suppose que l'intercepte et la pente aléatoires sont indépendantes, $Cov(b_{0i}, b_{1i}) = 0$. Cette structure ne fait aucune hypothèse sur les variances de l'intercepte et de la pente aléatoires comme les deux structures précédentes. Ainsi, $Cov(Y_{ij}, Y_{ik}) = \sigma_{0i}^2 + t_{ij}t_{ik}\sigma_{1i}^2$, $Var(Y_{ij}) = \sigma_{0i}^2 + t_{ij}^2\sigma_{1i}^2 + \sigma^2$. Cette structure ne permet pas de corrélation négative entre les mesures. La **Figure 14** nous montre les courbes d'évolution de la corrélation $Corr(Y_{ij}, Y_{ik})$ entre Y_{ij} et Y_{ik} en fonction de t_{ik} pour des valeurs de t_{ij} , $\frac{\sigma_{1i}^2}{\sigma_{0i}^2}$ et $\frac{\sigma^2}{\sigma_{0i}^2}$ fixées.

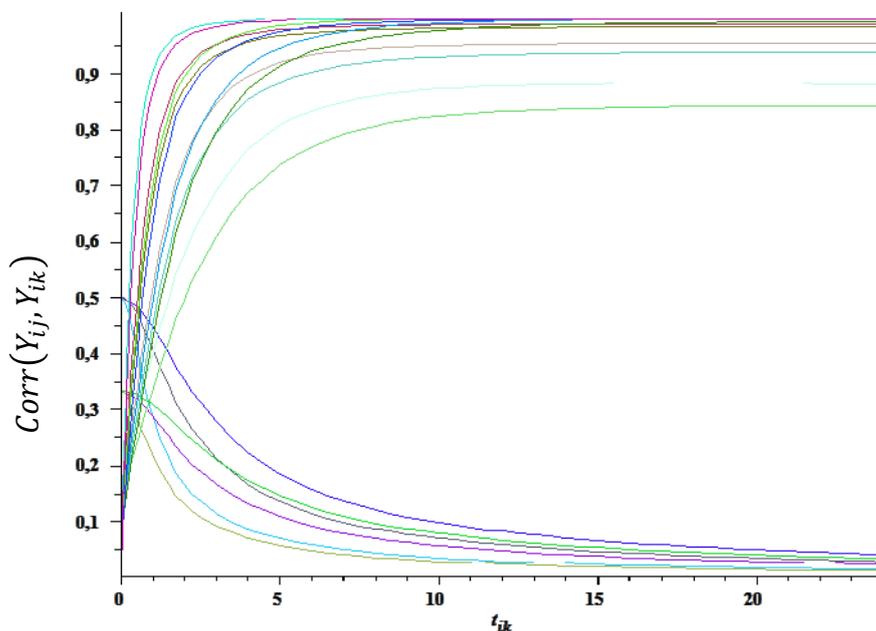


Figure 14. Courbes d'évolution de la corrélation entre les mesures en fonction du temps pour un modèle linéaire mixte à intercepte et pente aléatoires, avec une structure de corrélation indépendante entre l'intercepte et la pente.

Les courbes décroissantes sont celles obtenues pour $t_{ij} = 0$ et les courbes croissantes sont celles obtenues pour $t_{ij} > 0$, pour différentes valeurs de $\frac{\sigma_{1i}^2}{\sigma_{0i}^2}$ et $\frac{\sigma^2}{\sigma_{0i}^2}$ fixées. Ces courbes sont semblables à celles obtenues plus haut avec la structure « identité » : on tire donc les mêmes conclusions.

- **La structure « non-structurée »** : elle ne fait aucune hypothèse sur la corrélation ou la variance de l'intercepte et de la pente aléatoires. Les paramètres de la matrice de variance-covariance de (b_{0i}, b_{1i}) sont estimés à partir des données. Ainsi, $Cov(Y_{ij}, Y_{ik}) = \sigma_{0i}^2 + t_{ij}t_{ik}\sigma_{1i}^2 + (t_{ij} + t_{ik})\rho$, $Var(Y_{ij}) = \sigma_{0i}^2 + t_{ij}^2\sigma_{1i}^2 + \sigma^2 + 2\rho t_{ij}$, avec $\rho = Cov(b_{0i}, b_{1i})$. Cependant, elle fait intervenir un paramètre supplémentaire à estimer par rapport à la structure précédente.

Au cours de l'analyse de l'évolution des CD4 chez les patients sous traitement ARV suivis au Cameroun (cfannexe A), nous avons pris en compte la variabilité des profils individuels (**Figure 9 et Figure 10**) par l'introduction dans le modèle d'un intercepte aléatoire pour prendre en compte la variabilité des mesures à l'inclusion, d'une pente aléatoire dans la phase linéaire et d'une pente aléatoire dans la phase quadratique pour prendre en compte l'hétérogénéité des profils d'évolution au cours du suivi. Nous n'avons fait aucune hypothèse sur la structure de corrélation entre les effets aléatoires. Notre modèle final s'écrivait donc de la façon suivante :

$$\begin{aligned} & \sqrt{CD4_i} \\ &= \underbrace{X_i\beta + \alpha_0 + \alpha_1 * grp + b_0 * t_i}_{\text{phase linéaire}} + \overbrace{\alpha_2 * grp \times t_i}^{\text{interaction}} + \underbrace{b_1 * t_{(6+)_i} + b_2 * (t_{(6+)_i})^2}_{\text{partie fixe}} + \overbrace{\alpha_3 * grp \times t_{(6+)_i} + \alpha_4 * grp \times (t_{(6+)_i})^2}^{\text{interaction}} \\ &+ \underbrace{b_{0i} + b_{1i} * t_i + b_{2i} * t_{(6+)_i}}_{\text{partie aléatoire}} + \varepsilon_i \text{ (modèle 6)} \end{aligned}$$

Nous avons prédit à partir de ce modèle les courbes d'évolution individuelles de la racine carrée du taux de lymphocytes T CD4 que nous avons comparées aux profils individuels observés (**Figure 15**).

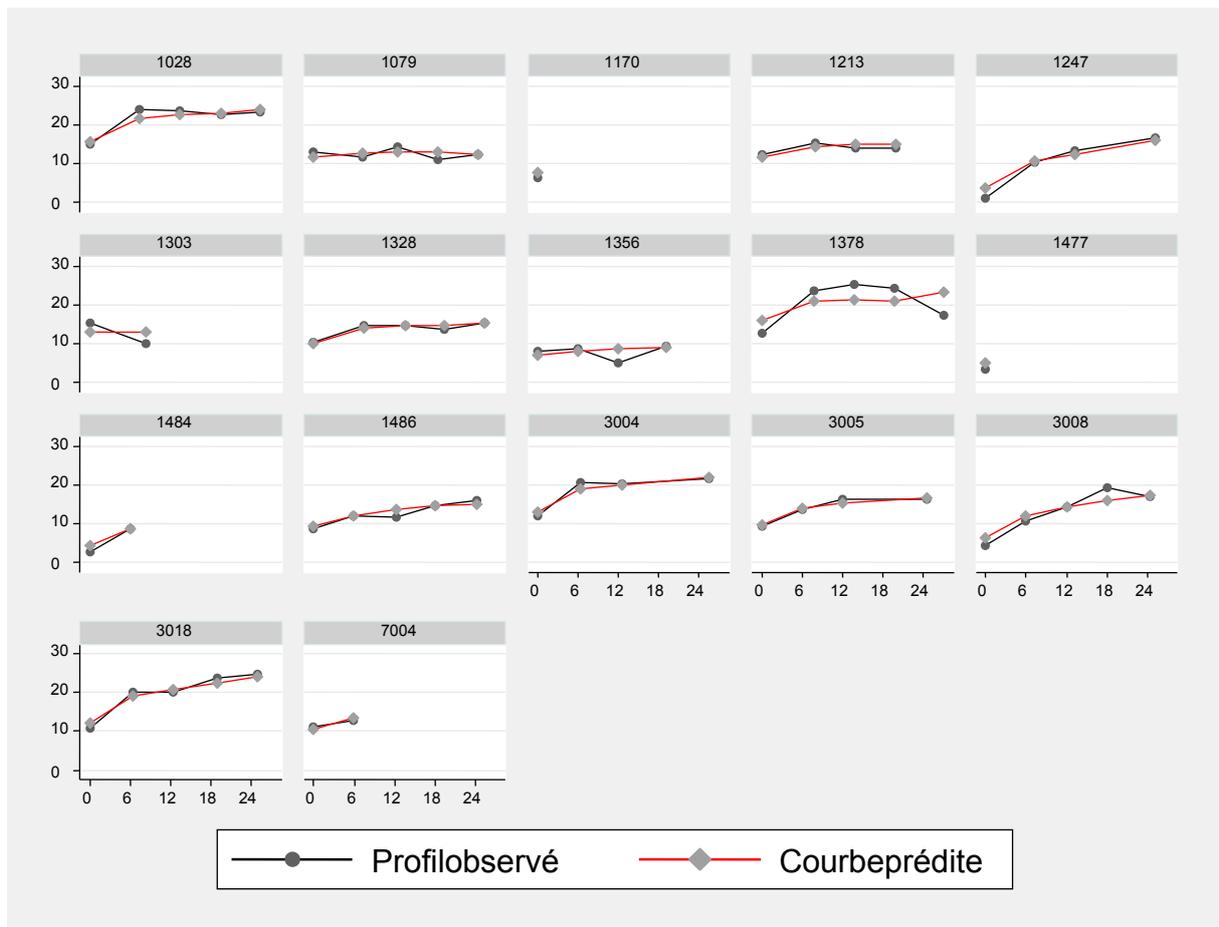


Figure 15. Représentation des profils d'évolution individuels observés et des courbes d'évolution individuelles prédites de la racine carrée du taux de lymphocytes T CD4 (à l'aide du **modèle 6** ci-dessus) chez 17 patients sélectionnés au hasard, infectés par le VIH sous traitement antirétroviral, suivis au Cameroun dans les études Parvy et Triomune.

Comme nous l'avons montré plus haut, l'introduction des effets aléatoires dans le modèle mixte permet de prendre en compte la corrélation entre les mesures successives du critère d'intérêt chez le même individu. Cependant, elle ne permet pas de faire une hypothèse directe sur la structure de la corrélation entre les mesures du critère d'intérêt. Nous présentons dans la suite une famille de modèles qui permettent de faire explicitement une hypothèse sur la structure de corrélation des mesures du critère d'intérêt chez le même individu.

2.2. Modèle marginal : le modèle GEE

Lorsque l'inférence au cours de l'analyse des données longitudinales ne porte que sur les paramètres moyens de la population d'étude, l'intérêt de l'analyse est porté sur l'évolution moyenne du critère d'intérêt. Une approche de modélisation consiste à exprimer la réponse moyenne en fonction des covariables d'intérêt, sans effets aléatoires. Les modèles utilisés dans ce cas sont dits marginaux. Ces modèles permettent de définir implicitement la distribution de la variable réponse par la spécification des deux premiers moments des mesures : la modélisation de la moyenne et la spécification de la structure de la matrice de variance-covariance des mesures. Ainsi, les modèles marginaux fournissent une méthode unique pour analyser divers types de données longitudinales (quantitatives ou qualitatives). L'estimation des paramètres des modèles marginaux ne peut se faire par maximum de vraisemblance car la distribution conjointe de la variable réponse n'est pas définie explicitement. Les méthodes d'estimation utilisées sont basées sur les équations d'estimation généralisées (GEE en anglais) [48, 49]. Dans la suite, on désignera par « modèles GEE » les modèles dont les méthodes d'estimation des paramètres sont basées sur les équations d'estimation généralisées.

Le modèle GEE est une extension naturelle du modèle GLM aux données corrélées. Il requiert une bonne spécification de la distribution marginale uni-variée de la variable mesurée et l'adoption d'une hypothèse de travail sur la structure de corrélation des données. Il combine des équations d'estimations pour les paramètres de régression et une estimation basée sur la méthode des moments pour les paramètres de corrélation, prenant en compte l'hypothèse de travail faite sur la structure de la matrice de variance-covariance des mesures.

La formulation du modèle GEE passe par la spécification des trois composantes suivantes :

1. La distribution marginale des observations : on suppose que chaque mesure de la variable réponse Y_{ij} de l'individu i suit conditionnellement à X_i une distribution uni-variée, membre de la famille des distributions exponentielles

$$Y_{ij}|X_{ij} = x_{ij} \sim ED(\xi_{ij}, \phi)$$

Où ξ_{ij} et ϕ sont des paramètres du modèle, appelés respectivement paramètre naturel et paramètre de dispersion.

2. La fonction de lien : soit $\mu_{ij} = E(Y_{ij}|X_{ij} = x_{ij})$ l'espérance conditionnelle. On suppose que l'espérance conditionnelle μ_{ij} est liée au prédicteur linéaire par la relation suivante :

$$g(\mu_{ij}) = \eta_{ij} = X_{ij}\beta$$

où $g(\cdot)$ est la fonction de lien connue. La matrice X de covariables peut contenir le temps, ainsi qu'une fonction du temps. On a dans le *tableau 2* quelques fonctions de lien canoniques pour les distributions usuelles. L'effet β associé aux covariables X_i est un paramètre moyen. Ainsi, le modèle d'équations d'estimation généralisées permet de faire une inférence sur la réponse moyenne de tous les individus de la population d'étude.

3. La structure de corrélation entre les mesures d'un individu : puisque la distribution de Y_{ij} est une distribution exponentielle, on a $Var(Y_{ij}) = \phi v(\mu_{ij})$. On pose

$$Var(Y_i) = A_i^{1/2}(\phi, \beta) Corr(Y_i) A_i^{1/2}(\phi, \beta), \text{ où } A_i(\phi, \beta) = \begin{pmatrix} \phi v(\mu_{i1}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \phi v(\mu_{in_i}) \end{pmatrix}$$

On fait une hypothèse sur la structure de la matrice de corrélation $Corr(Y_i)$ de Y_i et on la remplace par une matrice paramétrique $R(\rho)$ pour diminuer le nombre de paramètres à estimer. Six types de structures de matrices de corrélation sont fréquemment utilisés dans la pratique : identité, indépendante, échangeable, non-structurée, AR-1 et m-dépendante.

L'estimation du vecteur de coefficients de régression β et du paramètre ρ de la matrice de corrélation se fait par les équations d'estimation généralisées. Une approche de cette méthode d'estimation consiste à procéder comme suit :

- Déterminer l'équation d'estimation pour une distribution conditionnelle uni-variée $Y_{ij}|X_{ij} = x_{ij} \sim ED(\xi_{ij}, \phi)$ en supposant les Y_{ij} indépendants. Cette hypothèse d'indépendance nous place dans un contexte de modèle GLM dont l'équation d'estimation des paramètres est donnée par :

$$\sum_{i=1}^n D_i^T (Var(Y_i))^{-1} (y_i - \mu_i) = 0$$

$$\text{où } D_i^T = \left(x_i \frac{\partial \mu_i}{\partial \eta_i} \right) \text{ et } Var(Y_i) = \begin{pmatrix} Var(Y_{i1}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & Var(Y_{in_i}) \end{pmatrix}$$

- Généraliser l'équation d'estimation ci-dessus aux données corrélées, en faisant une hypothèse de travail sur $Var(Y_i)$. Liang et Zeger [48, 49] proposent de remplacer $Var(Y_i)$ par une matrice de covariance de travail définie par :

$$\Sigma_i = A_i^{1/2}(\beta)R_i(\rho)A_i^{1/2}(\beta)$$

La structure de $R_i(\rho)$ est fixée selon l'hypothèse de travail.

Il en résulte l'équation d'estimation généralisée

$$\sum_{i=1}^n D_i^T(\Sigma_i)^{-1}(y_i - \mu_i) = 0$$

La résolution de cette équation pose une difficulté liée à la présence de trois paramètres (donc trois inconnues) β , ϕ et ρ à estimer. Cette difficulté peut être résolue en remplaçant les paramètres (ϕ, ρ) par leur estimateur $(\hat{\phi}, \hat{\rho})$ obtenu par la méthode des moments. L'estimateur $\hat{\beta}(\hat{\phi}, \hat{\rho})$ est asymptotiquement égal à $\hat{\beta}(\phi, \rho)$.

Bien qu'ils soient tous deux des extensions aux données corrélées du modèle GLM, les modèles GLMM et GEE répondent à des objectifs différents et sous-tendent des méthodes d'estimation distinctes. Cependant en pratique, en fonction des objectifs et des données, on est souvent confronté à une difficulté de choix entre ces deux modèles pour l'analyse des données longitudinales parce que toutes les conditions d'application des deux modèles ne sont pas souvent vérifiées sur les données observées. Nous présentons dans la section suivante les différences entre ces modèles.

2.3. Comparaison entre les modèles GLMM et GEE

Les modèles GLMM et GEE sont les plus utilisés pour l'analyse des données longitudinales discrètes ou catégorielles. Bien que la spécification de ces deux modèles soit semblable, ils conduisent à des estimateurs des paramètres de régression distincts et permettent de répondre à des objectifs différents. Comme pour toute analyse, le choix du modèle dans une analyse de données longitudinales est guidé par l'objectif de l'étude. En général, le modèle GEE est recommandé quand l'objectif est de caractériser la population d'étude. Les modèles GLMM font intervenir des effets aléatoires spécifiques à chaque individu ; les paramètres estimés

permettent de mesurer l'influence des variables explicatives sur la variable réponse pour chaque individu et non pour l'ensemble de la population. Contrairement au modèle GEE, les modèles GLMM permettent d'établir les prédictions individuelles et d'estimer différentes sources de variabilité mais au prix d'hypothèses plus contraignantes (une bonne spécification de la partie fixe et de la partie aléatoire du modèle). Cependant, en fonction des données, les hypothèses sous-tendues par l'un ou l'autre modèle ne sont pas toujours satisfaites. Il est donc nécessaire de présenter les différences qu'impliquent les hypothèses sous-tendues par les deux modèles tant sur le plan théorique que sur le plan pratique.

2.3.1. Sur le plan théorique

La formulation du modèle GLMM nécessite la spécification de la distribution conjointe de $Y_i = (Y_{i1}, \dots, Y_{in_i})$ sachant les effets aléatoires, une hypothèse très contraignante qui impacte directement sur l'efficacité des estimateurs des paramètres du modèle. De son côté, la formulation du modèle GEE nécessite une bonne spécification de la moyenne et une hypothèse sur la structure de la matrice de corrélation. L'estimateur des paramètres de régression obtenus est robuste une fois que la moyenne est bien modélisée. La spécification de la distribution conjointe conditionnelle dans le GLMM permet d'obtenir la vraisemblance utilisée pour l'estimation des paramètres du modèle. Cependant, il est difficile d'obtenir une expression analytique de la vraisemblance, ce qui rend les méthodes numériques d'estimation des paramètres très complexes. A l'inverse, l'estimation des paramètres dans le modèle GEE se ramène à la résolution des équations d'estimation qui présente des difficultés numériques moindres.

2.3.2. Sur le plan pratique

L'objectif du modèle GEE porte sur l'inférence moyenne ce qui suppose une certaine homogénéité des profils d'évolution individuels. Cette hypothèse n'est pas souvent vérifiée sur les données issues d'études épidémiologiques, à l'exemple des données de lymphocytes T CD4 dont l'évolution est très hétérogène entre les individus (**Figure 9 et Figure 10**). Cette situation motive l'approche de modélisation de

l'évolution individuelle par le modèle GLMM qui permet de prendre en compte l'hétérogénéité entre les individus par l'introduction d'effets aléatoires. Cependant, s'il existe une corrélation sérielle entre les mesures répétées, le modèle GLMM ne permettra pas de la prendre complètement en compte. En effet, même en introduisant une auto-corrélation sérielle dans le terme de la variabilité résiduelle, on ne résout que partiellement le problème. A l'inverse, la spécification de la structure de la matrice de variance-covariance dans le modèle GEE permet directement de prendre en compte la corrélation entre les données répétées chez un individu. Cependant, le nombre de composantes de la matrice de variance-covariance augmente avec le nombre de mesures effectuées par individu (si on a n mesures par individu, la matrice de variance-covariance aura $\frac{n(n+1)}{2}$ composantes). Si le nombre de mesures effectuées par individu est important ou si les données sont déséquilibrées, l'utilisation du modèle GLMM permet d'éviter une sur-paramétrisation que peut entraîner l'utilisation du modèle GEE.

Le modèle GLMM est moins contraignant que le modèle GEE en présence de données manquantes car l'inférence est basée sur l'estimation par le maximum de la vraisemblance ou le maximum de la vraisemblance restreinte qui est valide lorsque les données sont manquantes au hasard. Cependant, il existe une extension du modèle GEE, le modèle GEE pondéré (WGEE en anglais) [65-67] dont les inférences sont valides lorsque les données sont manquantes au hasard.

L'interprétation des paramètres de régression diffère dans les deux modèles. Tandis que l'interprétation des paramètres de régression dans le modèle GLMM est rapportée à un individu, celle du modèle GEE est rapportée à la population d'étude. Il a été établi que les paramètres de régression dans le modèle GEE sont généralement inférieurs à ceux des GLMM, à l'exception de quelques cas particuliers. Lorsque le critère d'intérêt suit une distribution gaussienne et que les deux modèles (GLMM et GEE) sont bien définis, il y a égalité entre les coefficients de régression des deux modèles[68].

Le **Tableau 5** ci-dessous récapitule les principales différences entre les modèles GLMM et GEE.

Tableau 5. Tableau récapitulatif des principales différences entre les modèles GLMM et GEE

Modèle GLMM	Modèle GEE
Inférence individuelle	Inférence moyenne
Spécification de la distribution conjointe des données	Spécification de la distribution marginale des données
Estimation des paramètres du modèle par maximum de vraisemblance	Estimation des paramètres du modèle par les équations d'estimation généralisées
Prise en compte de la corrélation entre les mesures à travers l'introduction des effets aléatoires	Prise en compte de la corrélation entre les mesures par la spécification de la structure de la matrice de variance-covariance
Estimation valide lorsque les données sont déséquilibrées	Estimateurs biaisés lorsque les données sont déséquilibrées
Estimation valide en présence de données manquantes au hasard	Estimateurs biaisés en présence des données manquantes

3. Méthodes d'analyse de la survenue d'un évènement au cours du temps

Les données recueillies au cours des études longitudinales se présentent souvent de la façon suivante : les sujets entrent dans l'étude au fur et à mesure qu'elle se déroule. Si on s'intéresse à la survenue d'un évènement, à la date de fin de l'étude ou à la date des dernières nouvelles, le statut de chaque individu par rapport à la survenue de l'évènement d'intérêt et sa durée de participation dans l'étude sont connus. Chacune des situations suivantes peut se produire pour chaque individu. Soit on observe l'évènement d'intérêt au cours de l'étude, et donc le temps de participation est le délai entre la date d'entrée et la date de survenue de l'évènement. Soit l'individu est sorti de l'étude avant la fin de l'étude sans que l'évènement d'intérêt ne soit observé (on dit qu'il est perdu de vue), et donc le temps de participation est le délai entre la date d'entrée dans l'étude et la date de ses dernières nouvelles. Soit l'étude arrive à sa fin sans que l'évènement d'intérêt ne soit

observé, et donc le temps de participation est le délai entre la date d'entrée dans l'étude et la date de fin de l'étude. Pour les deux derniers cas, on dit que le délai de survenue de l'évènement est censuré à droite. Cependant, l'interprétation des résultats est différente suivant les causes de la censure. Les données contenant à la fois le statut par rapport à l'observation de l'évènement d'intérêt à la date de dernières nouvelles ou à la fin de l'étude, et le temps de participation pour chaque individu sont appelées « données de survie ». L'objectif de l'analyse des données de survie consiste à déterminer la distribution du délai de survenue de l'évènement d'intérêt ainsi que les facteurs de risque associés à la survenue de cet évènement. Il existe une abondante littérature statistique qui traite des méthodes d'analyse des données de survie [41, 42, 69-73].

3.1. Fonctions associées aux distributions de survie

Notons T la variable aléatoire positive correspondant au délai de survenue de l'évènement. La loi de probabilité de T peut-être caractérisée par plusieurs fonctions liées entre elles.

La fonction de densité de probabilité, notée $f(t)$:

$$f(t) = \lim_{\Delta t \rightarrow 0^+} \frac{\Pr(t \leq T < t + \Delta t)}{\Delta t}$$

où $f(t)\Delta t$ est la probabilité que l'évènement d'intérêt survienne entre t et $t + \Delta t$,

La fonction de répartition, notée $F(t)$:

$$F(t) = \Pr(T \leq t) = \int_0^t f(u) du$$

définit la probabilité que l'évènement d'intérêt survienne avant t . Cette fonction est monotone et vérifie les propriétés suivantes :

$$F(0) = 0 \text{ et } \lim_{t \rightarrow +\infty} F(t) = 1$$

On déduit la fonction de survie, notée $S(t)$, qui est la probabilité que l'évènement d'intérêt survienne au-delà du temps t .

$$S(t) = \Pr(T > t) = 1 - F(t)$$

C'est une fonction monotone décroissante telle que

$$S(0) = 1 \text{ et } \lim_{t \rightarrow +\infty} S(t) = 0$$

La fonction de risque instantanée, notée $\lambda(t)$:

$$\lambda(t) = \lim_{\Delta t \rightarrow 0^+} \frac{\Pr(t \leq T \leq t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)} = -\frac{d}{dt}(\log(S(t)))$$

$\lambda(t)\Delta t$ est la probabilité d'une nouvelle survenue de l'évènement d'intérêt entre t et $t + \Delta t$ sachant que le sujet était indemne de l'évènement à la date t . En épidémiologie, la fonction $\lambda(t)$ est appelée fonction d'incidence ou de risque instantanée de l'évènement d'intérêt.

La fonction de risque cumulé, notée $\Lambda(t)$:

$$\Lambda(t) = \int_0^t \lambda(u) du$$

Chacune des fonctions définies ci-dessus peut être utilisée pour décrire la distribution du délai de survenue de l'évènement, en raison des relations qui existent entre elles.

$$f(t) = -\frac{dS(t)}{dt}$$

$$S(t) = \exp\left(-\int_0^t \lambda(u) du\right) = \exp(-\Lambda(t))$$

$$f(t) = \lambda(t)S(t)$$

La fonction de risque instantanée $\lambda(t)$ est généralement la plus intéressante. En effet, elle fournit une description probabiliste du futur immédiat d'un sujet encore observé et reflète les différences entre les modèles qui seraient plus difficiles à formuler en termes de fonction de survie, de répartition ou de densité. C'est l'interprétation de cette fonction qui permet souvent de guider le choix du modèle pour le délai de survenue de l'évènement d'intérêt.

Trois approches existent dans la littérature pour décrire la distribution du délai de survenue d'un évènement : une approche paramétrique, une approche semi-paramétrique et une approche non paramétrique.

- L'approche paramétrique consiste à faire l'hypothèse suivant laquelle la distribution du délai de survenue de l'évènement appartient à une famille de distribution $\mathfrak{F}_\theta = (f_\theta(t), \theta \in \Theta)$ paramétrée par θ . Dans ce cas, déterminer la distribution du délai de survenue de l'évènement revient à déterminer θ . Un terme correctif peut être ajouté permettant de prendre en compte les caractéristiques individuelles (les facteurs de risque). Les modèles utilisés dans ce cas sont :

- *Le modèle à temps accélérés* [74-77], $S(t|X) = S_0(t * r(X))$ où S_0 représente la fonction de survie de base retenue, $X = (X_1, \dots, X_p)$ un vecteur de covariables et $r(X)$ une fonction de régression sur les variables X_1, \dots, X_p .
- *Le modèle à risque proportionnel* [41, 78, 79], $\lambda(t|X) = \lambda_0(t)r(X)$ où $\lambda_0(t)$ est la fonction de risque de base retenue. La forme la plus utilisée de $r(X)$ est $r(X) = e^{\beta X}$, où $\beta = (\beta_1, \dots, \beta_p)$ est le vecteur des paramètres de régression.

Dans le modèle à temps accélérés, β mesure l'effet des covariables X_1, \dots, X_p sur le délai de survenue de l'évènement (sur la vitesse de survenue de l'évènement). Au contraire, dans le modèle à risque proportionnel, β mesure l'effet des covariables X_1, \dots, X_p sur le risque de survenue de l'évènement. Au cours de nos analyses, nous nous sommes intéressés aux facteurs de risque de survenue de l'évènement d'intérêt. Nous avons utilisé à cet effet le modèle à risque proportionnel que nous présentons plus en détail dans la suite.

- L'approche semi-paramétrique qui pour des modèles à risque proportionnel cherche à estimer l'influence des facteurs de risque sans hypothèse concernant la distribution de base.
- L'approche non paramétrique qui vise à approximer une ou plusieurs des différentes fonctions caractérisant la distribution observée du délai de survenue de l'évènement d'intérêt sans faire d'hypothèse sur celle-ci. Cette approche est généralement utilisée dans un but descriptif, par exemple avec l'estimateur de Kaplan-Meier de la courbe de survie ou l'estimateur de Nelson Aalen de la fonction de risque cumulative.

3.2. Modèle à risque proportionnel

Ce modèle, introduit par Cox [78], définit la fonction de risque instantanée $\lambda(t)$, conditionnelle à un ensemble de covariables fixes, comme le produit d'une fonction positive déterministe inconnue $\lambda_0(t)$ et une fonction de régression sur l'ensemble des covariables.

$$\lambda(t | X_1, \dots, X_p) = \lambda_0(t) \exp\left(\sum_{i=1}^p \beta_i X_i\right) \quad (\text{modèle 7})$$

$\lambda_0(t)$ correspond à la fonction de risque de survenue de l'évènement d'intérêt de base (i.e lorsque $X_i = 0, i = 1, \dots, p$). C'est à travers cette fonction que le temps intervient dans la fonction de risque. Les $\beta_i, i = 1, \dots, p$, permettent de mesurer l'association entre les facteurs de risque $X_i, i = 1, \dots, p$ et la survenue de l'évènement d'intérêt. On suppose que les $\beta_i, i = 1, \dots, p$ ne dépendent pas du temps ce qui implique que le risque associé à l'exposition au facteur $X_i, i = 1, \dots, p$ est constant au cours du temps. Cette hypothèse appelée hypothèse de risque proportionnelle est l'une des hypothèses les plus importantes du modèle. Plusieurs méthodes ont été proposées dans la littérature pour vérifier cette hypothèse : la méthode graphique qui consiste à vérifier si les courbes du logarithme du risque cumulé sont parallèles pour chaque niveau du facteur de risque d'intérêt et les résidus de Schoenfeld.

Deux approches sont proposées dans la littérature pour l'estimation des paramètres du modèle à risque proportionnel :

- Le modèle paramétrique à risque proportionnel, il consiste à faire une hypothèse sur la fonction de risque de base $\lambda_0(t)$ ou, de façon équivalente, sur la distribution du délai de survenue de l'évènement. Les distributions exponentielle, weibull, log-normal, logistique et gamma sont les plus utilisées. Chacune de ces distributions implique une forme particulière de la fonction de risque au cours du temps. Le choix d'un modèle paramétrique particulier n'est cependant pas aisé en pratique. Il n'est pas non plus forcément utile : si le paramètre d'intérêt est le paramètre de régression, les autres paramètres du modèle apparaissent comme des paramètres de nuisance. Il peut donc être intéressant de ne définir le modèle que partiellement, en ne spécifiant précisément que la partie relative aux paramètres de régression ; c'est l'approche semi-paramétrique.
- Le modèle semi-paramétrique à risque proportionnel le plus utilisé pour analyser le délai de survenue d'un évènement en fonction des covariables est le modèle de Cox [78]. C'est un modèle à risque proportionnel dans lequel la forme de la fonction de risque de base n'est pas précisée. Il consiste ainsi à estimer les paramètres du modèle semi-paramétrique sans hypothèse sur la forme de la fonction de risque de base $\lambda_0(t)$. Cox a proposé de baser l'inférence statistique de son modèle sur une vraisemblance partielle (appelée vraisemblance partielle

de Cox) ne contenant pas $\lambda_0(t)$ qui est considéré comme un paramètre de nuisance [78]. Pour plus de détails sur l'estimation des paramètres du modèle semi-paramétrique de Cox, on peut se référer à [78, 80, 81].

3.3. Extensions du modèle semi-paramétrique de Cox

Le modèle de Cox est l'un des modèles multivariés fréquemment utilisés en épidémiologie pour l'analyse des enquêtes de cohorte. Dans le souci d'adapter les hypothèses sous-jacentes à ce modèle aux données réelles, plusieurs extensions de la formulation de base du modèle de Cox sont proposées dans la littérature statistique. Nous évoquons dans cette section quelques-unes de ces extensions, liées aux problématiques posées par les données analysées au cours de cette thèse ou qui pourront faire l'objet d'études approfondies dans de futurs travaux.

3.3.1. Modèle de Cox avec covariables dépendantes du temps

Dans la section précédente, nous avons introduit le modèle de Cox avec des covariables fixes. Dans cette section, nous présentons une généralisation de ce modèle aux covariables dépendantes du temps. Plusieurs problèmes relatifs aux covariables dépendantes du temps dans le modèle de Cox ont suscité une importante littérature parmi lesquels l'erreur de mesure.

En général, une covariable est mesurée avec erreur lorsqu'elle représente la mesure d'une quantité physique ou biologique qu'il est difficile d'obtenir précisément. Par exemple, la mesure du nombre total de lymphocytes CD4 est souvent entachée d'erreurs qui favorisent une grande variabilité. Une évaluation de la relation entre la survie des patients et les mesures répétées du nombre total de lymphocytes CD4 doit prendre en compte cette erreur. De nombreux travaux ont été proposés dans la littérature dans ce contexte [71, 82-86].

Considérons C le vecteur des délais de censure. On note $T^* = \min(T, C)$ le minimum entre le délai de survenue de l'évènement T et le délai de censure C , en pratique c'est T^* qui est observé. $\Delta = 1_{\{T < C\}}$ est l'indicateur de la survenue de l'évènement,

pour chaque individu, il vaut 1 lorsque l'évènement est observé et 0 si l'individu est censuré. $Y = (Y(t))_{t \geq 0}$ est le processus de covariables, il peut contenir des variables fixes. L'objectif du modèle de Cox est de modéliser la relation entre le délai de survenue de l'évènement T et le processus de covariables $Y = (Y(t))_{t \geq 0}$ par :

$$\lambda_{T|Z}(t) = \lambda_T(t) \exp\{g(Y(s) \ 0 \leq s \leq t; \beta)\}$$

Cependant, on n'observe pas les couples (T, Y) , mais plutôt le triplet $(T^*, \Delta, W(s), 0 \leq s \leq T^*)$ où $W(s)$ est la valeur entachée d'erreur observée de la vraie valeur $Y(s)$ à l'instant s . *Prentice* a montré [87] que l'estimateur du maximum de vraisemblance partielle du paramètre β (qui quantifie l'intensité de la relation entre le processus de covariables Y et le délai de survenue de l'évènement d'intérêt) obtenu à partir de n observations indépendantes du triplet $(T_i^*, \Delta_i, W_i(s), 0 \leq s \leq T_i^*)$ est un estimateur biaisé de l'association entre T et Y .

3.3.1.1. Approche en deux étapes

DeGrutola, Tsiatis et *Wulfsohn* [88] ont proposé une méthode d'estimation de β qui procède en deux étapes.

La première étape consiste à modéliser conjointement $Y(t)$ et $\{W(s), 0 \leq s \leq t\}$ et à estimer l'espérance conditionnelle de $Y(t)$ sachant $\{W(s), 0 \leq s \leq t\}$, $E(Y(t)|W(s), 0 \leq s \leq t)$. Dans la deuxième étape, le paramètre β est estimé en utilisant une estimation de la vraisemblance partielle de Cox dans laquelle les valeurs de $Y(t)$ sont remplacées par les estimations de $E(Y(t)|W(s), 0 \leq s \leq t)$. Plus précisément *DeGrutola, Tsiatis* et *Wulfsohn* supposent que pour chaque individu i ($i = 1, \dots, n$), la covariable Y_i est observée avec erreur aux instants t_j ($j \geq 1$) jusqu'à la survenue de l'évènement ou la censure pour chaque individu i . Les variables observées sont donc pour chaque individu i $(T_i^*, \Delta_i, W_i(t_1), \dots, W_i(t_{m_i}))$, où t_{m_i} est le dernier instant de mesure de Y_i avant T_i^* . Y_i est supposé unidimensionnel.

Les auteurs supposent que pour chaque individu i , $Y(t)$ peut-être modélisée par un modèle linéaire mixte :

$$Y_i(t) = \alpha_{0_i} + \alpha_{1_i} t$$

où $(\alpha_{0_i}, \alpha_{1_i})'$ est le vecteur d'effets aléatoires de loi normale bivariée, d'espérance θ et de matrice de variance-covariance V . Les auteurs supposent ensuite un modèle d'erreur de la forme :

$$W_i(t) = Y_i(t) + \epsilon_i$$

où ϵ_i suit une distribution normale centrée et de variance σ^2 , les ϵ_i sont indépendants, ϵ_i et $(\alpha_{0_i}, \alpha_{1_i})'$ sont indépendants.

La loi conjointe de $Y_i(t)$ et $(W_i(t_1), \dots, W_i(t_{m_i}))$ est une loi normale multivariée et l'espérance conditionnelle $E(Y(t)|W(s), 0 \leq s \leq t)$ peut-être calculée explicitement. Elle dépend des paramètres θ , V et σ^2 que les auteurs estiment par la méthode des moindres carrés, ce qui permet ensuite de calculer l'estimation de $E(Y(t)|W(s), 0 \leq s \leq t)$. Une étude par simulation [82] a permis de montrer que cette approche permet de diminuer le biais de l'estimateur du paramètre β .

Wulfsohn et *Tsiatis* [86] soulignent néanmoins plusieurs inconvénients de cette approche. D'une part, la nécessité d'estimer pour chaque instant d'évènement t les paramètres θ , V et σ^2 pour le calcul de l'espérance $E(Y(t)|W(s), 0 \leq s \leq t)$ pose un problème de parcimonie et d'interprétation de ces paramètres. D'autre part, leur estimateur ne tient pas compte de l'information relative à la durée d'observation des sujets. *Wulfsohn* et *Tsiatis* émettent l'hypothèse que prendre en compte cette information peut permettre de diminuer le biais résiduel de l'estimateur de β observé par *Dafni* et *Tsiatis* [82].

3.3.1.2. Approche par modélisation conjointe

Wulfsohn et *Tsiatis* proposent de modéliser conjointement le délai de survenue de l'évènement censuré et le processus de covariables.

Pour chaque individu i , on dispose des observations $(T_i^*, \Delta_i, W_i(t_1), \dots, W_i(t_{m_i}))$. Un modèle pour (T^*, Δ, Y, W) est obtenu en spécifiant les modèles de Cox pour la distribution conditionnelle du délai de survenue de l'évènement T sachant (Y, W) et la loi marginale pour (Y, W) . *Wulfsohn* et *Tsiatis* supposent que

$$W_i(t_j) = Y_i(t_j) + \epsilon_i(t_j)$$

et

$$Y_i(t_j) = \alpha_{0_i} + \alpha_{1_i}t_j$$

où $(\alpha_{0_i}, \alpha_{1_i})'$ est le vecteur d'effets aléatoires de loi normale bivariée, d'espérance θ et de matrice de variance-covariance V et $\epsilon_i(t_j)$ est l'erreur aléatoire de loi normale centrée et de variance σ^2 .

La fonction de risque $\lambda_{T,Z,W}(t)$ est supposée égale à :

$$\lambda_{T,Z,W}(t) = \lambda_T(t)\exp(\beta(\alpha_{0_i} + \alpha_{1_i}t_j))$$

Plus récemment, *Rizopoulos, Verbeke et Lesaffre* [89] ont étendu cette spécification du modèle conjoint en incluant des caractéristiques d'inclusion comme covariable dans le modèle et à la prise en compte des données déséquilibrées. Ainsi, $Y_i(t_{ij})$ est modélisé par :

$$Y_i(t_{ij}) = X_i(t_{ij})\beta + Z_i(t_{ij})b_i$$

où X_i et Z_i sont les matrices de covariables associées respectivement au vecteur d'effets fixes β et au vecteur d'effets aléatoires b_i . Le vecteur d'effets aléatoires est supposé suivre une distribution multivariée gaussienne, de moyenne zéro et de matrice de variance-covariance V .

La fonction de risque $\lambda_{T|Z,W}(t)$ est supposée égale à

$$\lambda_{T|Z,W}(t) = \lambda_T(t)\exp(\gamma^T X_i + \alpha Y_i(t))$$

Une vraisemblance est calculée pour l'estimation des paramètres du modèle sous les hypothèses suivantes :

- La censure est indépendante des effets aléatoires
- La censure est non-informative par rapport à la variable longitudinale
- La censure est non-informative par rapport à la survenue de l'évènement
- Les effets aléatoires prennent en compte toute la corrélation entre les mesures longitudinales.

La maximisation de cette vraisemblance nécessite une combinaison des méthodes d'intégration numérique et des algorithmes d'optimisation. Des techniques classiques d'intégration numérique telles que la méthode par quadrature de Gauss et la méthode de Monte Carlo ont été appliquées par *Wulfsohn et Tsiatis* [86]. Plus récemment, *Rizopoulos, Verbeke et Lesaffre* [89] ont discuté l'utilisation de la méthode d'approximation de Laplace qui peut être spécialement utile dans le contexte des

effets aléatoires de grande dimension. Ces techniques d'intégration sont généralement combinées à un algorithme EM (Expectation Maximization).

3.3.2. Prise en compte de la censure informative

Les analyses de survie font généralement l'hypothèse d'indépendance du délai de survenue de l'évènement et du délai de survenue de la censure. On parle alors de censure non-informative. Cette hypothèse est respectée dans le cas des censures planifiées. En revanche, elle peut être discutable dans le cas des censures non planifiées. Si on considère par exemple l'étude de la survie des personnes infectées par le VIH sous traitement ARV, comme nous l'avons mentionné dans la section 3, on observe souvent deux types de censure. Le premier type survient lorsqu'il y a des patients encore vivants à la fin de l'étude ; dans ce cas, la censure est planifiée à la date de fin de l'étude. Le deuxième type de censure survient lorsque des patients sont perdus de vue au cours du suivi. Les causes des pertes de vue des patients sont souvent très diverses, les causes rapportées les plus importantes étant les évènements indésirables, les raisons financières, l'arrêt volontaire, etc. On peut supposer dans la première et la dernière situation (et même dans la deuxième situation) que les censures sont informatives, c'est à dire que le fait que la censure se soit produite apporte une information sur le délai de survenue de l'évènement considéré. Cependant, la vérification de cette hypothèse est difficile. Le problème majeur vient du fait qu'il est impossible à partir des seules données observées de savoir si les censures sont informatives ou non-informatives [90]. Une possibilité consiste à recueillir des données supplémentaires sur un sous-échantillon de sujets perdus de vue [91], mais en pratique il peut s'avérer difficile ou impossible d'obtenir ces données.

Le problème de la prise en compte de la censure informative peut être envisagé dans le cadre de l'analyse des risques compétitifs [92] dans lequel on s'intéresse au délai de survenue de plusieurs types d'évènement mais seule la survenue du premier évènement est pris en compte. De ce point de vue, on peut considérer qu'on s'intéresse à deux types d'évènement : le décès et la censure. L'objectif principal d'une telle analyse est de réaliser une inférence sur la distribution marginale obtenue du délai de survenue du décès en tenant compte des censures informatives.

Peterson[93], *Klein et Moeschberger* [94] ont suggéré de déterminer les bornes inférieures et supérieures de la fonction de survie. Plusieurs auteurs ont proposé d'imposer des contraintes non-identifiables pour estimer la fonction de survie sous différentes hypothèses, notamment *Williams et Lagakos* [95], *Slud et Rubinstein* [96], *Link* [97], *Zheng et Klein* [98], *Scharfstein et al.*[99], *Scharfstein et Robins* [100]. Une autre approche consiste à ajuster la censure dépendante en utilisant les facteurs de risque dépendants du temps, le vecteur de délai de survenue de l'évènement d'intérêt et le vecteur de délai de censure [101-104].

3.3.2.1. Approche par analyse de risques compétitifs

La prise en compte du délai de survenue de l'évènement et du délai de censure nécessite la modélisation jointe de ces deux processus. Cette modélisation peut être réalisée dans le cadre des risques compétitifs, en considérant la fonction de survie jointe:

$$S_{T,C}(t, c) = P(T > t, C > c)$$

Dans cette approche, la censure est considérée comme un évènement au même titre que le décès. En pratique, l'inférence porte généralement sur la fonction de survie marginale:

$$S_T(t) = P(T > t) = S_{T,C}(t, 0)$$

Slud et Rubinstein [96] ont développé un estimateur non paramétrique de la fonction de survie marginale $S_T(t)$, basé sur la connaissance de la fonction

$$\rho(t) = \lim_{\Delta \rightarrow 0} \frac{P(t < T < t + \Delta | T > t, C \leq t)}{P(t < T < t + \Delta | T > t, C > t)}$$

Ceci signifie que le risque conditionnel de survenue de l'évènement à l'instant t diffère d'un facteur connu $\rho(t)$ suivant que la survie individuelle jusqu'à l'instant t est censurée avant t ou après t . Ceci peut être justifié dans le cas où la censure est provoquée par la survenue d'un évènement susceptible de modifier le risque de décès. Si $\rho(t) < 1$ pour tout t , alors le risque de décès diminue après la censure tandis qu'il augmente lorsque $\rho(t) > 1$ pour tout t . *Slud et Rubinstein* [96] ont proposé une extension du modèle de Cox dans laquelle la fonction de risque

instantané de survenue de la censure sachant $T \leq C$, $\lambda_C(c|T \leq c)$ est égale à la fonction de risque instantanée conditionnelle sachant $T > C$.

$$\lambda_T(t|C = c) = \begin{cases} \lambda_T(t) & \text{si } t \leq c \\ \lambda_T(t) * \rho(t) & \text{si } t > c \end{cases}$$

La modélisation de la dépendance entre le délai de survenue de l'évènement et le délai de censure via la fonction $\rho(t)$ permet de tenir compte de l'incertitude générée par les censures. Son interprétation en termes de rapport des risques permet d'en déterminer des valeurs raisonnables d'un point de vue clinique. *Slud et Rubinstein* [96] proposent également d'utiliser le rapport de risques de la meilleure variable pronostique pour déterminer un intervalle de valeurs raisonnables de $\rho(t)$. L'utilisation de valeurs raisonnables pour $\rho(t)$ permet alors de déterminer un intervalle d'ignorance [105] contenant l'ensemble des valeurs du paramètre d'intérêt compatibles avec les valeurs raisonnables de $\rho(t)$. Un intervalle d'incertitude tenant compte à la fois de l'ignorance due aux censures et de l'imprécision liée à la taille finie de l'échantillon peut également être déduite.

3.3.2.2. Méthode IPCW

Sous l'hypothèse non-identifiable que toutes les variables dépendantes du temps ou fixes pronostiques de la survenue de l'évènement et de la censure sont disponibles, *Robins*[103] a proposé une méthode d'estimation non-paramétrique appelée méthode IPCW (pour Inverse Probability of Censoring Weighted), pour la prise en compte des censures informatives. L'estimateur IPCW est une version pondérée de l'estimateur de Kaplan-Meier et de l'estimateur de la vraisemblance partielle de Cox dans laquelle les pondérations sont inversement proportionnelles à la probabilité d'être indemne de l'évènement d'intérêt et de ne pas être censuré.

Pour s'assurer de la validité de son estimateur, *Robins* [103] a fait l'hypothèse fondamentale suivante : la fonction de risque instantané de survenue de la censure sachant les variables pronostiques fixes et dépendantes du temps à l'instant t ne dépend pas de la survenue de l'évènement d'intérêt qui n'a pas été observé.

$$\lambda_C(t|Y(s) \ 0 \leq s \leq t, T, T \geq t) = \lambda_C(t|Y(s) \ 0 \leq s \leq t, T \geq t)$$

Pour construire l'estimateur IPCW, basé sur les facteurs pronostiques fixes et dépendants du temps $Y(t)$, la première étape consiste à estimer la fonction de risque instantané de survenue de la censure par le modèle semi-paramétrique de Cox avec des covariables dépendantes du temps.

$$\lambda_c(t|Y(s) \ 0 \leq s \leq t, T \geq t) = \lambda_{0c}(t) * \exp\{\beta * Y(t)\}$$

De cette relation est déduite l'extension suivante de l'estimateur de Kaplan-Meier

$$\widehat{K}_i^Y(t) = \prod_{\{j; T_j^* < t, T_j^* \neq T_j\}} [1 - \hat{\lambda}(T_j^*) * \exp\{\beta * Y(t)\}]$$

où

$$\hat{\lambda}(T_j^*) = \frac{1 - I_{\{T_j^* = T_j\}}}{\left\{ \sum_{i=1}^n \exp(\beta * Y(T_j^*)) I_{\{T_j^* \leq t\}} \right\}}$$

La deuxième étape consiste à construire les poids. Soit $\widehat{K}_i^0(t)$ l'estimateur usuel de Kaplan-Meier, on définit le poids de l'individu i à l'instant t par :

$$\widehat{W}_i(t) = \frac{\widehat{K}_i^0(t)}{\widehat{K}_i^Y(t)}$$

L'estimateur IPCW est ainsi construit en pondérant la contribution de chaque individu i à risque à l'instant T_j^* par $\widehat{W}_i(T_j^*)$.

4. Quelques méthodes de sélection de modèles

Pour résoudre un problème de modélisation, deux étapes se succèdent généralement dans la conception du modèle :

1. En premier lieu, on choisit les variables explicatives pertinentes, c'est-à-dire les variables qui peuvent être associées à la variable réponse.
2. En second lieu, on cherche dans une famille de fonctions (relations), celle qui permet d'estimer « au mieux » l'association entre la variable réponse et les variables explicatives.

Ainsi, dans une procédure de modélisation, on dispose d'un ensemble de variables explicatives à partir desquelles il est possible de construire un ensemble de modèles. L'objectif des méthodes de sélection de modèles est de choisir parmi cet ensemble de modèles, celui qui décrit et explique au mieux le phénomène observé.

La littérature sur la sélection de modèles est abondante. Lorsque le modèle est fixé, la théorie de l'information fournit un cadre rigoureux pour l'élaboration d'estimateurs performants. Mais dans un grand nombre de situations, les connaissances *a priori* sur les données ne permettent pas de déterminer un unique modèle dans lequel se placer pour réaliser une inférence. C'est pourquoi, depuis la fin des années 70, les méthodes de sélection de modèles à partir des données ont été développées. Les exemples classiques d'application de ces méthodes sont les sélections de variables explicatives et de leur forme.

L'une des réponses apportées au problème de sélection de modèles est la minimisation d'un critère pénalisé. Elle consiste à choisir le modèle qui minimise la somme des risques empiriques sur le modèle et une pénalité mesurant la complexité du modèle. Les premiers critères apparus dans la littérature sont l'Akaike Information criterium (AIC) [106], le Bayesian Information Criterium (BIC)[107], le Minimum Description Length (MDL)[108] et le C_p de Mallows[109]. Parmi ces critères, AIC et BIC ont été largement diffusés et appliqués pour les modèles estimés par les méthodes basées sur le maximum de vraisemblance. Beaucoup de travaux ont été réalisés concernant leurs propriétés statistiques et leur adaptation à des modèles bien spécifiques. Il existe une littérature très fournie sur les méthodes de sélection de modèles par critère pénalisé, qui se développe encore actuellement.

L'AIC est un critère qui permet de rechercher le modèle qui fait le meilleur compromis biais-variance pour un échantillon de taille fixe n . Le meilleur modèle au sens de l'AIC dépend donc de la taille de l'échantillon. La formule de calcul de l'AIC est donnée par l'expression suivante :

$$AIC = -2 \times \log(L) + 2p$$

où L est la vraisemblance maximisée et p le nombre de paramètres dans le modèle. Ce critère pénalise la déviance du modèle ($-2 \times \log(L)$) par deux fois le nombre de paramètres.

Le BIC est un critère qui se place dans un contexte bayésien pour rechercher le modèle le plus vraisemblable au vu des données. La formule de calcul du BIC est donnée par l'expression suivante :

$$BIC = -2 \times \log(L) + \log(n) \times p$$

Il est clair que les meilleurs modèles au sens des critères AIC et BIC sont différents. Cependant, ces critères ont souvent fait l'objet de comparaisons empiriques [110, 111]. Dans la pratique, il a été observé que le critère BIC sélectionne les modèles de dimension plus petite que le critère AIC, ce qui n'est pas surprenant puisque le critère BIC pénalise plus que le critère AIC. Toutefois, on ne peut pas comparer au sens absolu ces deux critères de sélection. Seul l'objectif de l'analyse et la connaissance des données peuvent donner un sens à la notion de supériorité d'un critère sur l'autre.

La conception des critères AIC et BIC repose sur une approximation asymptotique qui sous-entend que la liste des modèles est fixée tandis que n tend vers l'infini. *Barron, Birgé et Massart* [112, 113] ont récemment développé une théorie non-asymptotique de la sélection de modèles, basée sur une méthode de sélection automatique des pénalités à partir des données, appelé heuristique des pentes. Cette méthode repose sur le concept de pénalité minimale. *Birgé et Massart* [113] ont déterminé une pénalité minimale et montré que la pénalité optimale est égale à deux fois la pénalité optimale.

Nous nous limitons à cette brève présentation des méthodes de sélection de modèles car cette thématique n'est pas le sujet de notre travail. Toutefois, nous précisons que nous avons adopté un point de vue classique en choisissant d'utiliser au cours de nos analyses des critères de sélection simples tels que l'AIC et le BIC.

5. Applications

5.1. Hépatotoxicité et efficacité des traitements antirétroviraux incluant la névirapine chez les patients co-infectés par le VIH et le virus de l'hépatite B ou C au Cameroun

Dans cette étude, nous avons comparé l'évolution immuno-virologique, la mortalité, la progression clinique de la maladie et l'hépatotoxicité entre des personnes

infectées par le VIH seulement et des personnes co-infectées par le VIH et le VHB ou le VHC sous traitement antirétroviral incluant la névirapine.

L'évolution immunologique a été modélisée par un modèle linéaire mixte avec rupture de pente à 6 mois constitué par une première phase linéaire pour les 6 premiers mois de traitement et par une deuxième phase quadratique pour les 18 mois suivants. Nous avons ensuite utilisé le modèle semi-paramétrique de Cox pour comparer les probabilités de succès virologique, de décès, d'épisodes d'hépatotoxicité et de progression clinique entre les personnes mono-infectées et celles co-infectées. Le modèle de poisson a aussi été utilisé pour comparer le risque d'hépatotoxicité.

L'annexe A présente l'article publié issu de cette étude.

5.2. Efficacité et tolérance à long terme du traitement antirétroviral incluant la didanosine, lamivudine et l'efavirenz ou la névirapine chez les patients naïfs, infectés par le VIH : étude de cohorte de 9 ans au Sénégal

Dans cette étude, nous avons comparé à long terme l'évolution immuno-virologique, la mortalité, la progression clinique de la maladie et la tolérance biologique entre des patients recevant un traitement antirétroviral incluant la didanosine et des patients sous traitement incluant la zidovudine.

L'évolution immunologique a été modélisée via un modèle linéaire mixte avec rupture de pente à 6 mois constitué par une première phase linéaire pour les 6 premiers mois de traitement et par une deuxième phase cubique pour les 18 mois suivants. Le modèle GEE a été utilisé pour comparer l'odds ratio du succès virologique (charge virale plasmatique en dessous du seuil de 500 copies/mL) entre les deux groupes de patients. Nous avons ensuite utilisé le modèle semi-paramétrique de Cox pour comparer le risque ajusté de survenue de la mortalité ou de la progression clinique. Enfin, le modèle binomial négatif a été utilisé pour comparer le risque ajusté d'incidence des événements indésirables biologiques entre les deux groupes de patients.

Par ailleurs, nous avons utilisé la méthode d'imputation multiple pour compléter les variables d'inclusion qui avaient des données manquantes. Nous aborderons plus globalement la question des données manquantes dans le chapitre 3.

L'annexe B présente l'article publié issu de cette étude.

6. Discussion

Dans ce chapitre, nous avons présenté quelques méthodes d'analyse de données longitudinales à savoir le modèle linéaire mixte pour les variables réponses continues, les modèles GLMM et GEE pour les variables réponses quantitatives discrètes ou qualitatives, et enfin le modèle semi-paramétrique de Cox lorsque la variable réponse est une indicatrice de survenue d'un évènement qui ne peut survenir qu'une fois au cours de l'étude.

Nous avons utilisé ces différentes méthodes pour évaluer, d'une part, la réponse au traitement antirétroviral incluant la névirapine dans le contexte particulier de la co-infection par le VIH et le VHB ou le VHC et, d'autre part, la réponse au traitement antirétroviral de première ligne incluant la didanosine.

Le modèle linéaire mixte est largement utilisé pour modéliser la réponse immunologique chez des personnes sous traitement antirétroviral [114-120]. Cependant, la modélisation de la partie fixe et de la partie aléatoire du modèle diffère d'une étude à l'autre. Dans nos deux premières études (cf **annexe A et B**), nous avons utilisé pour modéliser la partie fixe un modèle linéaire avec une rupture de pente à 6 mois pour prendre en compte la différence entre l'augmentation du nombre total de lymphocytes T CD4 dans les 6 premiers mois de traitement et l'augmentation ultérieure. En effet, on observe généralement une reconstitution immunologique après la mise sous traitement des personnes infectées par le VIH. Cette reconstitution immunologique est constituée de deux phases [121-125] ; la première phase caractérisée par une augmentation rapide du nombre total de lymphocytes T

CD4 qui ne dure que quelques mois, correspond à une réactivation des lymphocytes T CD4 quiescents tandis que la seconde phase qui est caractérisée par une augmentation plus lente du nombre total de lymphocytes T CD4, correspond à une augmentation de nouveaux lymphocytes T CD4. Dans nos études, le choix de la durée de 6 mois pour la première phase de reconstitution immunitaire a été imposé par les données (la première mesure des lymphocytes T CD4 sous traitement était faite au bout de 6 mois). Cette même durée pour la première phase de reconstitution a aussi été utilisée par *De Beaudrap et al* [55]. Cependant dans cet article, le modèle utilisé pour modéliser l'évolution immunologique était un modèle mixte non-linéaire. Plusieurs formes de modélisation de la partie fixe du modèle linéaire mixte ont été utilisées dans la littérature pour modéliser la réponse immunologique. *Thibaut et al* ont utilisé un modèle linéaire mixte avec une partie fixe linéaire biphasique [54] tandis que *Nash et al* ont utilisé les polynômes fractionnels d'ordre trois dans la partie fixe du modèle linéaire mixte [118]. Au cours de nos analyses, nous avons testé plusieurs formes de modélisation de la partie fixe du modèle linéaire mixte (linéaire, quadratique, linéaire biphasique, biphasique avec une phase linéaire et une phase quadratique). Un modèle linéaire mixte biphasique avec une première phase linéaire et une seconde phase quadratique a été sélectionné dans notre première étude (cf **annexe A**), et un modèle linéaire mixte biphasique, avec une première phase linéaire et une seconde phase cubique a été sélectionné dans notre deuxième étude (cf **annexe B**). Le critère de sélection des modèles utilisés était le critère d'information d'Akaike (AIC).

Cependant, en présence de données manquantes sur la réponse immunologique, l'estimateur des paramètres du modèle linéaire mixte peut être biaisé. En effet, lorsque la raison pour laquelle les données sont manquantes est par exemple le décès, cela n'a plus de sens de se demander quelle réponse aurait eu le sujet s'il était resté dans l'étude, car le décès rend impossible l'existence d'une réponse. Par conséquent, l'utilisation d'un modèle qui ne prend pas en compte la durée de participation des sujets dans l'étude n'est pas adaptée. De façon générale, lorsque les données manquantes sont informatives, c'est-à-dire lorsque le processus de réponse dépend des données non observées, l'estimateur du maximum de vraisemblance des paramètres du modèle linéaire mixte est biaisé, ce qui peut rendre les inférences incorrectes. Des études récentes [89, 119, 126, 127] ont montré que la modélisation conjointe de la réponse immunologique et du délai de

survenue de l'évènement engendrant la sortie d'étude (**section 3.3.1.2**) permettait, d'une part, de réduire le biais engendré par les sorties d'études informatives et, d'autre part, de prendre en compte les délais de survenue de l'évènement engendrant la sortie d'étude.

Nous avons également présenté dans ce chapitre l'impact de l'inclusion d'effets aléatoires, de la structure de corrélation des effets aléatoires sur la modélisation de la structure de corrélation entre les données, dans le modèle linéaire mixte. Nous avons montré dans les cas simples d'inclusion de deux effets aléatoires au maximum (intercepte aléatoire et/ou pente aléatoire associée au temps) que la corrélation entre les mesures de la variable réponse sous-tendue par le modèle linéaire mixte peut être complexe mais qu'elle reste cependant positive dans la majorité des cas (**section 2.1.3.2**). Ainsi donc, la modélisation de la partie aléatoire du modèle mixte ne doit pas être prise uniquement comme la modélisation de l'hétérogénéité entre les individus mais doit aussi tenir compte de l'implication sur la structure de corrélation entre les mesures de la variable réponse sous-tendue par le modèle.

Toutes les limites relevées pour le modèle linéaire mixte sont identiques pour le modèle GLMM. Le modèle GEE quant à lui, permettant de faire une inférence moyenne reste interprétable même lorsque des individus sont sortis de l'étude pour des raisons qui rendent impossibles l'existence d'une réponse. Cependant, comme nous l'avons précisé dans la **section 2.3.2**, lorsque les données manquantes sont manquantes non au hasard ou informatives, l'estimateur des paramètres du modèle GEE obtenus est biaisé.

Nous avons utilisé le modèle semi-paramétrique de Cox pour estimer le risque ajusté de survenue du décès, d'un évènement classant au stade Sida, d'un épisode d'hépatotoxicité, de succès virologique (premier passage en dessous du seuil de détectabilité). C'est le modèle le plus utilisé dans la littérature pour estimer les risques ajustés de survenue de ce type d'évènement lorsque les dates exactes de survenue de l'évènement d'intérêt sont connues. Cependant, dans plusieurs applications, la date exacte de survenue de l'évènement d'intérêt n'est pas observée. Généralement, dans ces derniers cas, on ne connaît que l'intervalle contenant le délai exact de survenue de l'évènement d'intérêt. Considérons le cas de la survenue d'un évènement indésirable classant au stade Sida. Généralement, les évènements

indésirables sont diagnostiqués au cours des visites cliniques. Ainsi, les dates de début notifiées sont soit la date de manifestation des premiers symptômes de l'évènement signalée par le patient soit la date de diagnostic de l'évènement indésirable. Or, la vraie date de début de l'évènement est comprise entre la date de la dernière visite clinique et la date de manifestation des premiers symptômes signalés par le patient ou la date de diagnostic de l'évènement indésirable. Donc, le fait de considérer la date de manifestation des premiers symptômes signalés par le patient ou la date de diagnostic de l'évènement comme la date de survenue de l'évènement dans une analyse de survie peut entraîner des biais. Il sera important de réfléchir sur une méthode permettant de réduire le biais dans ce cas. Cette réflexion pourrait faire l'objet de travaux futurs.

Les sorties d'étude informatives peuvent aussi engendrer un biais sur l'estimateur des paramètres du modèle semi-paramétrique de Cox. Nous avons présenté dans la **section 3.3.2** des extensions du modèle semi paramétrique de Cox permettant de réduire ce biais. Cependant, les causes de sortie d'étude sont souvent multiples et les modèles présentés ne permettent pas de prendre en compte plusieurs causes de sortie d'étude. En effet, au cours du suivi des personnes sous traitement antirétroviral, plusieurs raisons peuvent entraîner une perte de vue et ces raisons n'ont parfois pas les mêmes valeurs pronostiques sur le risque de décès. Un patient qui sort de l'étude ou qui arrête son traitement parce qu'il se sent mieux n'a pas le même risque de décès qu'un patient qui sort de l'étude ou qui arrête le traitement parce qu'il se sent plus mal. On ne doit donc pas prendre en compte ces deux raisons de la même façon. Très peu d'auteurs ont proposé des méthodes dans ce contexte. Très récemment *Matsuyama*[102] a proposé une extension du modèle semi-paramétrique de Cox en présence de sortie d'étude informative permettant de prendre en compte plusieurs causes de sorties d'étude, utilisant la méthode IPCW. Une de nos perspectives consistera à explorer cette méthode.

Chapitre 2 : *Tests de non-infériorité*

Initialement utilisés en pharmacologie pour le développement de nouvelles thérapeutiques, les essais de non-infériorité sont actuellement fréquents en recherche clinique et épidémiologique, et constituent notamment un outil important d'évaluation des nouvelles interventions de santé publique (nouveau traitement, nouvelle stratégie de prise en charge, etc.). En effet, l'existence d'interventions de référence efficaces pour la majorité des problèmes de santé publique implique que les bénéfiques escomptés en termes d'efficacité avec les nouvelles interventions sont souvent réduits et donc que la démonstration de leur supériorité est plus difficile (les essais de supériorité de la nouvelle intervention *versus* placebo n'étant pas éthiques lorsqu'il existe une intervention de référence). De plus, l'intérêt des nouvelles interventions, outre leur efficacité, peut porter sur la diminution du coût ou des effets indésirables, sur la simplification de la prise en charge ou tous autres aspects opérationnels permettant d'améliorer l'adhérence et la qualité de vie des patients [128].

L'essai Stratall présenté dans l'introduction (**section 3.3, page 30**) s'inscrivait précisément dans ce contexte. Un essai de non-infériorité se justifiait :

- d'une part, par le fait que la stratégie évaluée, le suivi clinique exclusif des patients sous ARV, n'était pas supposée plus efficace que la stratégie de référence, le suivi clinico-biologique des patients
- d'autre part, par l'intérêt du suivi clinique exclusif qui permettait la décentralisation et donc l'extension de l'accès au traitement malgré les limites financières et structurelles, pour de nombreuses années encore, des établissements sanitaires.

Le critère de jugement principal était le gain en lymphocytes T CD4 entre les patients des deux groupes à 24 mois de traitement. Pour cette raison, ce chapitre sera centré sur les procédures de démonstration de la non-infériorité à partir d'un critère de jugement quantitatif continu.

1. Principe des essais de non-infériorité

Contrairement aux essais de supériorité qui consistent à montrer la supériorité d'une intervention par rapport à une autre, les essais de non-infériorité permettent de montrer (ou non) que la nouvelle intervention n'est pas inférieure à l'intervention de référence. En fait, cette démarche ne démontre pas la non-infériorité au sens strict du terme : elle montre que s'il existe une infériorité, celle-ci est contenue dans une certaine limite (appelée marge de non-infériorité) qui n'est pas cliniquement significative. Les essais de non-infériorité supposent donc qu'un certain niveau de perte soit consenti. Cette particularité est au cœur des difficultés posées par ces essais.

La validité et l'interprétation des résultats d'essais de non-infériorité nécessitent le respect de certains principes méthodologiques en ce qui concerne le choix de l'intervention de référence, de la marge de non-infériorité, du critère de jugement principal et de la méthode d'analyse.

1.1. Intervention de référence

L'efficacité de l'intervention de référence doit avoir été établie par un essai de supériorité le comparant soit à un placebo, soit à une autre intervention de référence. Pour que l'intervention de référence d'un essai de non-infériorité soit adéquate, il faut qu'elle soit administrée dans les mêmes conditions que celles du ou des essais ayant démontré son efficacité. Si les conditions des essais ayant permis de démontrer la supériorité de l'intervention de référence ne peuvent être reproduites, les alternatives proposées sont soit d'introduire un bras placebo lorsque cela ne pose aucun problème éthique, pour procéder à une vérification interne de l'efficacité de l'intervention de référence [128-131], soit de comparer l'efficacité observée de l'intervention de référence à celle obtenue dans les études antérieures l'ayant validée [128, 131].

1.2. Marge de non-infériorité

La question de savoir jusqu'à quel point la perte d'efficacité de la nouvelle intervention par rapport à celle de l'intervention de référence peut être considérée comme cliniquement négligeable est une préoccupation majeure lors de la mise en place d'un essai de non-infériorité. C'est dans la fixation et la validation de cette marge que réside la plus grande difficulté des essais de non-infériorité. Plusieurs méthodes sont proposées dans la littérature pour le choix de la marge de non-infériorité en fonction des raisons cliniques et statistiques [128, 132-134], mais il n'y a pas de règle stricte. Il faut en outre noter que la marge de non-infériorité intervient dans les tests statistiques, et donc conditionne les résultats. Dans ce cas, il est crucial dans l'interprétation des résultats d'un essai de non-infériorité de ne pas seulement considérer le résultat en terme de significativité, mais d'évaluer si la marge choisie n'est pas trop tolérante ou trop stricte en fonction du contexte.

1.3. Choix et analyse du critère de jugement principal

Le critère de jugement principal doit être le plus objectif possible et de préférence être celui qui a été utilisé pour l'évaluer l'efficacité de l'intervention de référence.

De nombreux auteurs insistent particulièrement sur la nécessité de réaliser une analyse per protocole pour le critère de jugement principal dans les essais de non-infériorité [135, 136]. En effet, une analyse per protocole permet souvent d'accentuer les différences entre les groupes à comparer, favorisant ainsi l'hypothèse nulle (rendant conservateur le test de non-infériorité). Au contraire, une analyse en intention de traiter rapproche généralement les groupes à comparer, favorisant l'hypothèse de non-infériorité (augmentant l'erreur de première espèce). Cependant, lorsque les raisons des écarts au protocole sont liées aux interventions, l'analyse en intention de traiter peut être non conservatrice. Considérons par exemple un essai de non-infériorité où l'on teste une nouvelle intervention qui est en réalité inférieure à l'intervention de référence. Si, dans le groupe des patients recevant la nouvelle intervention, les écarts au protocole sont liés à la mauvaise réponse à l'intervention, les mauvais répondeurs seront exclus de l'analyse per protocole. En retirant les

mauvais répondeurs de ce groupe, on surestimera l'efficacité de la nouvelle intervention. Dès lors, l'efficacité apparente de la nouvelle intervention se rapprochera de celle de l'intervention de référence, et les deux interventions pourront sembler à tort non-inférieures. Ce phénomène illustre l'attention particulière qui doit être portée à la recherche des raisons des écarts au protocole et à la description des effets secondaires liés aux interventions dans les essais de non-infériorité. *Matsuyama* [137] a montré dans une étude par simulations que l'analyse en intention de traiter favorisait l'hypothèse de non-infériorité et que l'erreur de première espèce augmentait avec la proportion des écarts au protocole et plus encore lorsque les écarts au protocole étaient liés au critère de jugement principal. Au contraire, l'analyse per protocole permettait de maintenir le risque de première espèce lorsque les écarts au protocole ne dépendent pas du critère de jugement principal et était soit conservatrice, soit non conservatrice (en fonction du mécanisme qui engendre les écarts au protocole) lorsque les écarts au protocole dépendent du critère de jugement principal.

Ainsi, d'une façon générale, l'analyse per protocole est recommandée pour les essais de non-infériorité mais elle doit être complétée par une analyse en intention de traiter pour s'assurer de la validité des résultats [138].

2. Démonstration de la non-infériorité

Le principal outil statistique utilisé pour la démonstration de la non-infériorité est le test d'hypothèse. Il s'agit dans ce cas de tester l'hypothèse nulle H_0 : la nouvelle intervention (N) n'est pas non-inférieure à l'intervention de référence (R) *versus* l'hypothèse alternative H_1 : la nouvelle intervention (N) est non-inférieure à l'intervention de référence (R). Plusieurs méthodes statistiques de test ont été développées dans la littérature [128, 139, 140] mais la plus utilisée en pratique est la méthode basée sur les intervalles de confiance. C'est cette méthode, utilisée dans l'essai Stratall, que nous présentons dans la suite.

2.1. Notions sur les tests d'hypothèse et les méthodes par intervalles de confiance

2.1.1. Les tests d'hypothèse paramétriques

Un test d'hypothèse est un procédé d'inférence permettant de décider à partir d'un ou plusieurs échantillons aléatoires, de la validité d'une hypothèse relative à une ou plusieurs populations. On range souvent les tests d'hypothèse en deux catégories : les tests paramétriques et les tests non paramétriques. On parle de tests paramétriques lorsque l'on fait l'hypothèse que les données sont issues d'une distribution paramétrée. Dans ce cas, les caractéristiques des données peuvent être résumées à l'aide de paramètres estimés sur l'échantillon et la procédure de test ne porte alors que sur ces paramètres. A l'inverse, les tests non paramétriques ne font aucune hypothèse sous-jacente sur les données. Le test de non-infériorité que nous avons utilisé au cours de nos analyses était un test par intervalle de confiance, porté sur la moyenne du gain de lymphocytes T CD4 à 24 mois de traitement ARV. C'est la raison pour laquelle nous présentons dans ce qui suit quelques notions sur les tests d'hypothèse paramétriques et les tests par intervalles de confiance.

Soient X une variable aléatoire distribuée selon une loi appartenant à la famille de lois $\mathcal{F} = \{f_\theta, \theta \in \Theta\}$, où θ est le paramètre sur lequel le test est porté, Θ l'ensemble dans lequel le paramètre θ prend ses valeurs et $\mathcal{X} = (X_1, \dots, X_n)$ un échantillon indépendant et identiquement distribué (i.i.d) de la variable d'intérêt X . On note Θ_0 et Θ_1 deux sous-ensembles non vides de Θ tel que $\Theta_0 \cap \Theta_1 = \emptyset$. On appelle statistique, toute variable aléatoire fonction de l'échantillon \mathcal{X} . On dira qu'une statistique est exhaustive pour le paramètre θ si elle concentre toute l'information relative à la connaissance du paramètre θ .

L'objectif d'un test d'hypothèse paramétrique est de décider entre deux hypothèses faites sur la distribution d'une variable X (résumée par θ), à la vue d'un échantillon de cette dernière. Un test paramétrique est formulé de la façon suivante.

$$H_0 : \theta \in \Theta_0 \text{ versus } H_1 : \theta \in \Theta_1 \text{ avec } \Theta_0 \cap \Theta_1 = \emptyset$$

Un test est déterminé par sa région critique (ou région de rejet de H_0) qu'on note \mathcal{W} . C'est l'ensemble des échantillons pour lesquels l'hypothèse nulle H_0 est rejetée. On appelle erreur de première espèce, la probabilité de rejeter H_0 à tort $P_\theta(\mathcal{X} \in \mathcal{W} | \theta \in$

Θ_0) = $P_\theta(\text{rejeter } H_0 | \theta \in \Theta_0)$. On appelle erreur de deuxième espèce, la probabilité de rejeter H_1 à tort $P_\theta(X \notin \mathcal{W} | \theta \in \Theta_1) = P_\theta(\text{accepter } H_0 | \theta \in \Theta_1)$. On appelle puissance du test β , la probabilité de rejeter H_0 lorsque H_1 est vraie, $\beta = P_\theta(\text{rejeter } H_0 | \theta \in \Theta_1)$. Lorsque le cardinal de Θ_0 (qu'on notera dans la suite $\text{card}(\Theta_0)$) est supérieur à 1, on a une erreur de première espèce pour chaque valeur de $\theta \in \Theta_0$. On dira qu'un test est de niveau α , si la borne supérieure de l'ensemble des erreurs de première espèce pour chaque valeur de θ variant dans Θ_0 est égale à α , $\sup_{\theta \in \Theta_0} P_\theta(X \in \mathcal{W}) = \alpha$. On dira qu'un test est de seuil α , si la borne supérieure de l'ensemble des erreurs de première espèce pour chaque valeur de θ variant dans Θ_0 est inférieure ou égale à α , $\sup_{\theta \in \Theta_0} P_\theta(X \in \mathcal{W}) \leq \alpha$. Un test est d'autant meilleur qu'il a un niveau faible et une grande puissance. Pour un seuil de test donné, on peut construire une infinité de régions critiques correspondant à différentes puissances du test. L'identification de la meilleure zone critique (celle qui maximise la puissance du test) est en général difficile. Cependant, il existe des situations dans lesquelles cette tâche est relativement aisée et est complètement résolue par le lemme de Neyman-Pearson et ses extensions [141]. Ainsi, on dira qu'un test de seuil α est Uniformément Plus Puissant (UPP) si sa puissance est maximale pour tout $\theta \in \Theta_1$ parmi tous les tests de seuil α .

En fonction du cardinal de Θ_0 et Θ_1 , on peut formuler différentes catégories d'hypothèses :

- Si $\text{card}(\Theta_0) = \text{card}(\Theta_1) = 1$, on a un test d'hypothèses simples (test de Neyman-Pearson) formulé de la façon suivante :

$$H_0 : \theta = \theta_0 \text{ versus } H_1 : \theta = \theta_1 \text{ avec } \theta_0 \neq \theta_1$$

D'après le lemme de Neyman-Pearson, le test qui rejette H_0 si

$$\frac{L(x, \theta_1)}{L(x, \theta_0)} > c_\alpha \text{ où } P_{\theta_0} \left(\frac{L(x, \theta_1)}{L(x, \theta_0)} > c_\alpha \right) = \alpha$$

est un test UPP parmi tous les tests de seuil α , où $L(x, \theta_i)$, $i = 0, 1$ est la vraisemblance de l'échantillon respectivement sous les hypothèses H_0 et H_1 , $P_{\theta_0}(\cdot)$ la distribution de l'échantillon sous H_0 . Ce lemme permet de construire la meilleure zone critique pour le test d'hypothèses simples.

- Si $\text{card}(\Theta_0) = 1$ et $\text{card}(\Theta_1) > 1$, on a un test entre hypothèses simples et hypothèses composites. L'exemple le plus courant de cette catégorie de test est un test unilatéral (test de supériorité) formulé de la façon suivante.

$$H_0 : \theta = \theta_0 \text{ versus } H_1 : \theta > \theta_0$$

ici $\Theta_0 = \{\theta_0\}$ et $\Theta_1 = \{\theta \in \Theta \mid \theta > \theta_0\}$. On montre que si la région critique du test de Neyman-Pearson pour le test d'hypothèses $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$ avec $\theta_0 < \theta_1$ ne dépend pas de θ_1 , alors le test de Neyman-Pearson de niveau α est le test UPP pour les hypothèses $H_0 : \theta = \theta_0$ versus $H_1 : \theta > \theta_0$ parmi tous les tests de seuil α .

- Si $\text{card}(\Theta_0) > 1$ et $\text{card}(\Theta_1) > 1$, on a un test entre hypothèses composites. Les exemples les plus courants de cette catégorie de tests sont les tests d'équivalence et les tests de non-infériorité, formulés respectivement de la façon suivante :

$$H_0 : |\theta| \leq \Delta_L \text{ versus } H_1 : |\theta| > \Delta_L \text{ et } H_0 : \theta \leq \Delta_L \text{ versus } H_1 : \theta > \Delta_L$$

Une extension du lemme de Neyman-Pearson (théorème de Karlin-Rubin) permet d'avoir une règle de décision pour ce test. Cette règle de décision est stipulée de la façon suivante : si $S(X)$ est une statistique exhaustive pour θ telle que, pour tout $\theta_1 < \theta_2$, la fonction qui à $s \mapsto \frac{f_{S(X)}(s, \theta_2)}{f_{S(X)}(s, \theta_1)}$ est croissante, alors le test qui rejette H_0 si $S(X) > s_\alpha$ est le test UPP parmi tous les tests de seuil $\alpha = P_{\theta_0}(S(X) > s_\alpha)$ où $f_{S(X)}(s, \theta_i)$ est la densité de probabilité de $S(X)$ lorsque $\theta = \theta_i, i = 1, 2$.

Ce dernier test d'hypothèses composites est celui qui nous intéresse pour la démonstration de la non-infériorité. Pour ce test, Δ_L est appelée marge de non-infériorité, $\Theta_0 =]-\infty; \Delta_L]$ et $\Theta_1 =]\Delta_L; +\infty[$.

Dans certains cas, la loi de la statistique $S(X)$ définie dans le test n'est pas connue (par exemple lorsque X est continue et non gaussienne) ce qui rend la procédure de décision présentée ci-dessus irréalisable.

Dans le cas de l'essai Stratall, la marge de non-infériorité était aléatoire et corrélée au gain moyen de lymphocytes T CD4 dans le groupe de référence ce qui rend, là aussi, la procédure de test ci-dessus irréalisable.

Une alternative consiste à utiliser une procédure de décision basée sur les intervalles de confiance ; c'est d'ailleurs la méthode la plus utilisée dans la pratique.

2.1.2. Notions sur les intervalles de confiance

On considère toujours qu'on a un échantillon i.i.d $X = (X_1, \dots, X_n)$ d'une variable aléatoire X distribuée selon une loi appartenant à la famille de lois $\mathcal{F} = \{f_\theta, \theta \in \Theta\}$ paramétrée par θ . Un intervalle de confiance θ est une mesure de précision de l'estimation obtenue à partir d'un échantillon. On le construit de la façon suivante : on considère deux statistiques A_n et B_n ; on dira que $[A_n, B_n]$ est un intervalle de confiance de niveau $1 - \alpha$ pour θ si la probabilité pour que θ appartienne à l'intervalle $[A_n, B_n]$ est égale à $1 - \alpha$, $P_\theta(A_n \leq \theta \leq B_n) = 1 - \alpha$ pour tout $\theta \in \Theta$. On dira que $[A_n, B_n]$ est un intervalle de confiance de niveau asymptotiquement égal à $1 - \alpha$ pour θ si le niveau de l'intervalle de confiance $[A_n, B_n]$ converge vers $1 - \alpha$ lorsque la taille de l'échantillon devient grand, $\lim_{n \rightarrow +\infty} P_\theta(A_n \leq \theta \leq B_n) = 1 - \alpha$ pour tout $\theta \in \Theta$. Pour un niveau de confiance donné, il existe en général une infinité d'intervalles de confiance ayant ce niveau de confiance. Le choix le plus pertinent est celui de l'intervalle le plus court car c'est lui qui localise le mieux la vraie valeur du paramètre θ .

Lorsque la variable X suit une distribution gaussienne de paramètres m et σ^2 , on montre que les intervalles de confiance de l'espérance m les plus courts sont symétriques par rapport à la valeur estimée ce qui justifie l'utilisation systématique des intervalles symétriques. Ainsi, l'intervalle de confiance de niveau $1 - \alpha$ le plus pertinent de m est donné par : $\left[\bar{X}_n - \sigma \frac{z_{1-\alpha/2}}{\sqrt{n}}; \bar{X}_n + \sigma \frac{z_{1-\alpha/2}}{\sqrt{n}} \right]$ si la variance σ^2 est connue et par $\left[\bar{X}_n - \sqrt{S_n} \frac{t_{1-\alpha/2}}{\sqrt{n(n-1)}}; \bar{X}_n + \sqrt{S_n} \frac{t_{1-\alpha/2}}{\sqrt{n(n-1)}} \right]$ si la variance σ^2 est inconnue, où $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ est un estimateur sans biais de m , $z_{1-\alpha/2}$ le quantile d'ordre $1 - \alpha/2$ de la loi normale centrée réduite, $\frac{1}{n-1} S_n = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ est un estimateur sans biais de la variance et $t_{1-\alpha/2}$ le quantile d'ordre $1 - \alpha/2$ de la loi de Student à $n - 1$ degré de liberté.

Lorsque la distribution de la variable X est inconnue, on ne peut pas construire un intervalle de confiance de niveau $1 - \alpha$. Toutefois, même dans ce cas, il est souvent possible de construire un intervalle de confiance de niveau asymptotiquement égal à $1 - \alpha$, c'est-à-dire un intervalle approximatif dont la précision augmente avec la taille de l'échantillon. Pour un grand échantillon, on peut appliquer les lois limites du

théorème central limite pour construire un intervalle de confiance asymptotiquement de niveau $1 - \alpha$ de l'espérance $\theta = E(X)$. Si la variance σ^2 ou un estimateur sans biais de la variance $\widehat{\sigma^2}$ de la variable X est connu, les intervalles $\left[\bar{X}_n - \sigma \frac{z_{1-\alpha/2}}{\sqrt{n}}; \bar{X}_n + \sigma \frac{z_{1-\alpha/2}}{\sqrt{n}}\right]$ et $\left[\bar{X}_n - \sqrt{\widehat{\sigma^2}} \frac{z_{1-\alpha/2}}{\sqrt{n}}; \bar{X}_n + \sqrt{\widehat{\sigma^2}} \frac{z_{1-\alpha/2}}{\sqrt{n}}\right]$ sont respectivement des intervalles de confiance de niveau asymptotiquement égal à $1 - \alpha$. Cependant, il a été montré dans le cas où la distribution de l'espérance θ n'est pas approximativement gaussienne que les intervalles de confiance ci-dessus peuvent être très imprécis [142]. Pour améliorer cette approximation, on transforme souvent l'espérance θ pour approcher le mieux une distribution gaussienne. Une alternative simple qui ne nécessite pas d'analyse préalable, par exemple pour la recherche de la transformation à utiliser, consiste à construire un intervalle de confiance de θ à partir de la distribution bootstrap de θ . Le principe est le suivant :

- A partir de l'échantillon \mathcal{X} , on fait un tirage aléatoire successif et avec remise d'un échantillon de même taille que \mathcal{X} noté $\mathcal{X}^* = (X_1^*, \dots, X_n^*)$ qu'on appelle échantillon bootstrap de \mathcal{X} . On répète cette opération B fois pour obtenir B copies $\mathcal{X}_1^*, \dots, \mathcal{X}_B^*$ de l'échantillon \mathcal{X} .
- A partir de chaque échantillon bootstrap, on estime θ . On obtient ainsi un échantillon $(\theta_1^*, \dots, \theta_B^*)$ de taille B de la distribution bootstrap de θ , qu'on utilise pour construire un intervalle de confiance de θ de niveau asymptotiquement égal à $1 - \alpha$ [142, 143].

Nous avons adopté cette approche de distribution bootstrap pour construire les intervalles de confiance dans la procédure de test de non-infériorité utilisée pour l'analyse du critère de jugement principal dans l'essai Stratall.

2.2. Relation entre les tests de non-infériorité et les intervalles de confiance

La formulation du test de non-infériorité est dépendante du type de critère de jugement principal (quantitatif continu, qualitatif binaire, censuré, etc). Nous nous limiterons à la formulation pour un critère quantitatif continu. Le test statistique sera basé sur la différence des moyennes du critère de jugement principal entre le groupe ayant subi la nouvelle intervention μ_N et celui ayant subi l'intervention de référence

μ_R . Nous supposons qu'une augmentation du critère de jugement principal va dans le sens d'une efficacité accrue.

On note Δ_L la marge de non-infériorité admise, $\Delta = \mu_N - \mu_R$ la différence des moyennes du critère de jugement principal entre le groupe recevant la nouvelle intervention et le groupe recevant l'intervention de référence, α le seuil de significativité du test. Les hypothèses du test sont formulées comme suit :

$$H_0 : \Delta = \mu_N - \mu_R \leq -\Delta_L \text{ contre } H_1 : \Delta = \mu_N - \mu_R > -\Delta_L$$

2.2.1. Cas où la marge de non-infériorité est fixe ou constante

Sachant un intervalle de confiance bilatéral $IC(\mathcal{X})_{2\alpha} = [IC_1(\mathcal{X}); IC_2(\mathcal{X})]$ de Δ de niveau $1 - 2\alpha$ (ou un intervalle de confiance $IC(\mathcal{X})_{2\alpha} = [IC_1(\mathcal{X}); +\infty[$ unilatéral de niveau $1 - \alpha$) ou de niveau asymptotiquement égal à $1 - 2\alpha$, le test qui rejette H_0 si la borne inférieure de l'intervalle de confiance $IC(\mathcal{X})_{2\alpha}$ est supérieure ou égale à $-\Delta_L$ est de seuil α . En effet, $P_{H_0}(IC_1(\mathcal{X}) > -\Delta_L) \leq P_{H_0}(\Delta \notin IC(\mathcal{X})_{2\alpha}) \leq \alpha$. Ce test est d'autant plus puissant que l'amplitude de l'intervalle de confiance bilatéral $IC(\mathcal{X})_{2\alpha}$ est plus petite [144]. Ainsi, le test le plus puissant est celui pour lequel l'intervalle de confiance $IC(\mathcal{X})_{2\alpha}$ est le plus pertinent parmi tous les intervalles de confiance de niveau $1 - 2\alpha$. C'est pour cette raison que dans certains cas (distribution, ou distribution limite de Δ sous H_0 gaussienne), on utilise systématiquement les intervalles de confiance symétriques. On montre dans le cas gaussien que le test basé sur des intervalles de confiance symétriques de niveau $1 - 2\alpha$ est le test UPP parmi tous les tests de seuil α d'après le théorème de Karlin-Rubin énoncé plus haut.

Deux procédures de décision existent en pratique en ce qui concerne le choix du niveau de l'intervalle de confiance lorsque la marge de non-infériorité est fixe :

- La première consiste à utiliser soit un intervalle de confiance bilatéral de niveau $1 - \alpha$ et ne considérer que la borne inférieure, soit un intervalle de confiance unilatéral de niveau $1 - \alpha/2$. Cette approche est recommandée par le Committee for Proprietary Medicinal Products [145].
- La deuxième consiste à utiliser soit un intervalle de confiance bilatéral de niveau $1 - 2\alpha$ et ne considérer que la borne inférieure, soit un intervalle de confiance unilatéral de niveau $1 - \alpha$.

Une étude par simulation menée dans le contexte de tests de non-infériorité a montré que la première approche est conservatrice et qu'elle diminue la puissance de l'étude, tandis que la seconde aboutit à des risques α et β proches de ceux choisis dans le protocole [128].

2.2.2. Cas où la marge de non-infériorité est une variable aléatoire

Dans la procédure de test présentée ci-dessus, nous avons supposé que la marge de non-infériorité était une valeur fixe. Maintenant, nous allons considérer le cas où la marge de non-infériorité Δ_L est une variable aléatoire. C'était le cas dans l'essai Stratall où la marge de non-infériorité relative admise (différence relative maximale du gain moyen de lymphocytes T CD4 après 24 mois de traitement entre les deux groupes) était de 25 %. Nous nous limiterons au cas particulier où la relation entre la marge de non-infériorité Δ_L et la moyenne du critère d'intérêt dans le groupe de référence μ_R est de la forme suivante : $\Delta_L = \frac{\alpha * \mu_R}{100}$, où α est un réel connu. Nous proposons dans la suite deux procédures de décision basées sur des intervalles de confiance. Ces procédures sont des adaptations de la deuxième procédure de décision par intervalle de confiance présentée ci-dessus (**section 2.2.1**)

2.2.2.1. Première procédure de décision

Soient $[I_1; I_2]$ l'intervalle de confiance bilatéral de niveau $1 - \gamma$ de $-\Delta_L$, $\widehat{\Delta}_L$ l'estimation de Δ_L obtenue à partir de l'échantillon. Nous supposons que zéro n'appartient pas à l'intervalle de confiance de $-\Delta_L$ pour rester dans la logique des essais de non-infériorité.

La première procédure consiste à rejeter H_0 si la borne inférieure de l'intervalle de confiance $IC(X)_\gamma$ est supérieure à l'intervalle de confiance de Δ_L (**Figure 16**)

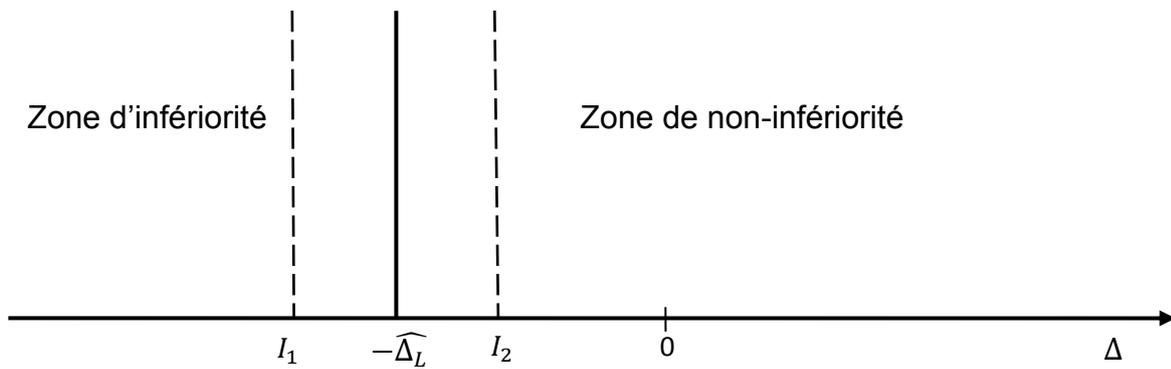


Figure 16. Représentation des zones d'acceptation et de rejet pour un test de non-infériorité avec une marge de non-infériorité définie par une variable aléatoire.

On pourra donc conclure à la non-infériorité si l'intervalle de confiance bilatéral de niveau $1 - \gamma$ de Δ est contenu dans la zone de non-infériorité, i.e si la borne inférieure de l'intervalle de confiance de Δ est supérieure à la borne supérieure de l'intervalle de confiance de $-\Delta_L$. La

Figure 17 nous montre la règle de décision.

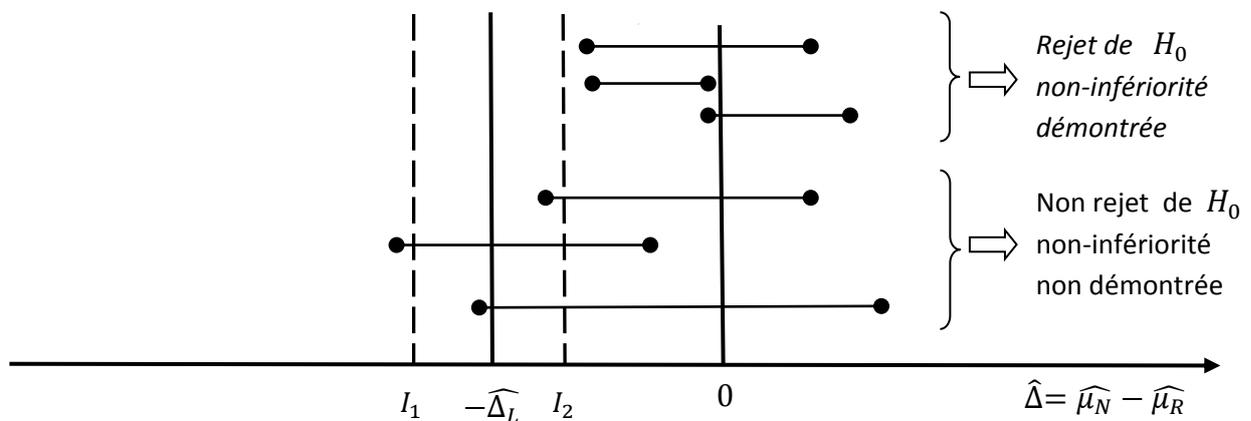


Figure 17. Représentation de la procédure de décision du test de non-infériorité basée sur les intervalles de confiance lorsque la marge de non-infériorité est définie par une variable aléatoire.

Cette procédure de test conduit à un résultat facilement interprétable. Cependant, l'existence d'une corrélation entre Δ et $-\Delta_L$ ($cov(\Delta, -\Delta_L) = -\frac{\alpha}{100} var(\mu_R) \neq 0$) rend le contrôle de l'erreur de première espèce difficile et ne permet pas de déterminer analytiquement la puissance du test.

2.2.2.2. Deuxième procédure de décision

La deuxième procédure de test consiste à reformuler les hypothèses du test (cette reformulation reste équivalente à la première formulation) pour se ramener à une formulation semblable à celle des tests de supériorité. Nous proposons de procéder de la façon suivante :

$$H_0' : \Delta + \Delta_L \leq 0 \text{ contre } H_1' : \Delta + \Delta_L > 0,$$

$$\text{où } \Delta + \Delta_L = \mu_N - \left(1 - \frac{\alpha}{100}\right) * \mu_R$$

Soit $IC(\mathcal{X})$ un intervalle de confiance bilatéral de $\Delta + \Delta_L$ de niveau $1 - \gamma$. Le test qui rejette H_0' si la borne inférieure de l'intervalle de confiance $IC(\mathcal{X})$ est supérieure à zéro, est un test de non-infériorité de H_0 versus H_1 . Ce test est d'autant plus puissant que l'amplitude de l'intervalle de confiance $IC(\mathcal{X})$ est petite.

Nous avons mené une étude par simulation pour déterminer le niveau de l'intervalle de confiance bilatéral à utiliser pour atteindre le seuil α espéré par les deux procédures de test proposées ci-dessus et pour comparer ces deux procédures de test.

Nous avons fixé l'erreur de première espèce $\alpha = 0.05$, la puissance $\beta = 0.8$, le gain moyen de lymphocytes T CD4 dans le groupe de référence $\mu_R = 200$, l'écart type du gain de lymphocytes T CD4 dans le groupe de référence $\sigma_R = 190$ et la marge de non-infériorité $\Delta_L = \frac{25}{100} \mu_R$. A partir de ces valeurs, nous avons calculé la taille de l'échantillon nécessaire n dans chaque groupe. Nous avons ensuite simulé le vecteur de gain de lymphocytes T CD4 dans le groupe de référence suivant une distribution gaussienne de moyenne μ_R et d'écart type σ_R . Pour la simulation du vecteur de gain de lymphocytes T CD4 dans le groupe recevant la nouvelle intervention, nous avons considéré deux cas :

- Le cas où il y a effectivement non-infériorité. Dans ce cas, nous avons simulé le vecteur de gain de lymphocytes T CD4 dans le groupe recevant la nouvelle intervention suivant une distribution gaussienne de moyenne $\mu_N > 0.75 * \mu_R$ et d'écart type $\sigma_N = \sigma_R = 190$. A partir des données simulées et des deux procédures de test proposées ci-dessus (en faisant varier le niveau des intervalles de confiance), nous avons estimé la puissance des procédures de décision proposées.

- Le cas où il n'y a pas non-infériorité. Dans ce cas, nous avons simulé le vecteur de gain de lymphocytes T CD4 dans le groupe recevant la nouvelle intervention suivant une distribution gaussienne de moyenne $\mu_N \leq 0.75 * \mu_R$ et d'écart type $\sigma_N = \sigma_R = 190$. A partir des données simulées et des deux procédures de test proposées ci-dessus (en faisant varier le niveau des intervalles de confiance), nous avons estimé le seuil des procédures de décision proposées.

Le **Tableau 6** présente les résultats obtenus sous ces différentes hypothèses

Tableau 6. *Résultat des estimations du seuil et de la puissance des procédures de test proposées en fonction du niveau des intervalles de confiance.*

	Niveau du test		Puissance du test	
	Première procédure	Deuxième procédure	Première procédure	Deuxième procédure
$\gamma = \alpha$	0.006	0.005	0.99	0.99
$\gamma = 2\alpha$	0.015	0.012	0.92	0.90
$\gamma = 4\alpha$	0.047	0.042	0.84	0.87
$\gamma = 6\alpha$	0.084	0.079	0.86	0.88

Nos résultats étaient les suivants : pour $\gamma = 2\alpha$, le seuil obtenu par les procédures de test proposées était inférieur au seuil $\alpha = 0.05$. Alors que pour $\gamma = 4\alpha$, le seuil et la puissance obtenus approchaient le seuil $\alpha = 0.05$ et la puissance $\beta = 0.8$. La puissance et le seuil des deux procédures de test étaient comparables. Donc, pour que le test de non-infériorité atteigne le seuil $\alpha = 0.05$ par les procédures de test proposées, il faut utiliser des intervalles de confiance à 80 %.

3. Application à l'analyse du critère de jugement principal de l'essai Stratall

Rappelons que le critère de jugement principal dans cet essai était le gain moyen en lymphocytes T CD4 entre l'initiation de traitement et le 24^{ème} mois de suivi. Nous avons utilisé la première procédure de décision proposée ci-dessus (cf **section 2.2.2.1**) pour comparer le gain moyen en lymphocytes T CD4 entre les patients du

groupe « suivi clinique exclusif » et ceux du groupe « suivi clinico-biologique ». Dans un premier temps, nous avons utilisé des intervalles de confiance à 95 % ; ensuite, compte tenu de la marge de non-infériorité relative, nous avons utilisé des intervalles de confiance à 90 % pour approcher un seuil de test de non-infériorité de 5 %. L'annexe C présente l'article publié issu de cette étude.

Deux analyses ont été effectuées. Dans la première analyse incluant tous les patients (238 dans la groupe « suivi clinique exclusif » et 221 dans le groupe « suivi clinico-biologique »), nous avons remplacé les nombres totaux de lymphocytes T CD4 manquants à M24 par les dernières valeurs observées (*last observation carried forward*). Précisons ici que les données de lymphocytes T CD4 à l'initiation de traitement étaient disponibles pour tous les patients. La différence entre le gain moyen de lymphocytes T CD4 du groupe « suivi clinique exclusif » et celui du groupe « suivi clinico-biologique » était de -31 (intervalles de confiance bilatéraux à 95 % [-63; 2] et à 90 % [-58; -3],) et la marge de non-infériorité était de -52 (intervalles de confiance bilatéraux à 95 % [-58; -45] et à 90 % [-57; -46]) **Figure 18.et**

Figure 19

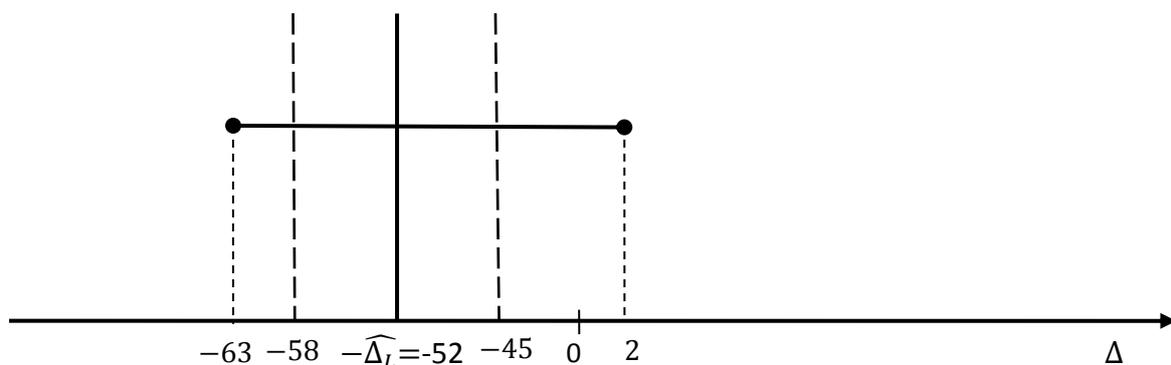


Figure 18. Résultat de l'analyse en intention de traiter du critère principal de l'essai Stratall avec la méthode LOCF (dernière valeur observée). Représentation de la procédure de décision du test de non-infériorité basé sur les intervalles de confiance bilatéraux à 95 %.

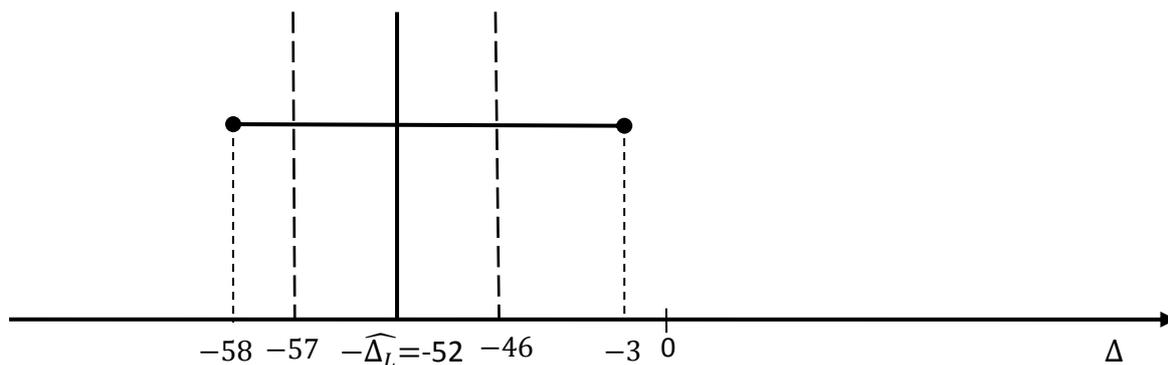


Figure 19. Résultat de l'analyse en intention de traiter du critère principal de l'essai Stratall avec la méthode LOCF (dernière valeur observée). Représentation de la procédure de décision du test de non-infériorité basé sur les intervalles de confiance bilatéraux à 90 %.

La seconde analyse incluait uniquement les patients disposant de données de lymphocytes T CD4 au 24^{ème} mois (165 [69 %] dans le groupe « suivi clinique exclusif » et 169 [76 %] dans le groupe « suivi clinico-biologique »). La différence de gain moyen en lymphocytes T CD4 entre les deux groupes était de -25 (intervalles de confiance bilatéraux à 95 % [-64; 14] et à 90 % [-58; 7]) et la marge de non-infériorité -61 (intervalles de confiance bilatéraux à 95 % [-68; -54] et à 90 % [-67; -56]) **Figure 20** et **Figure 21**.

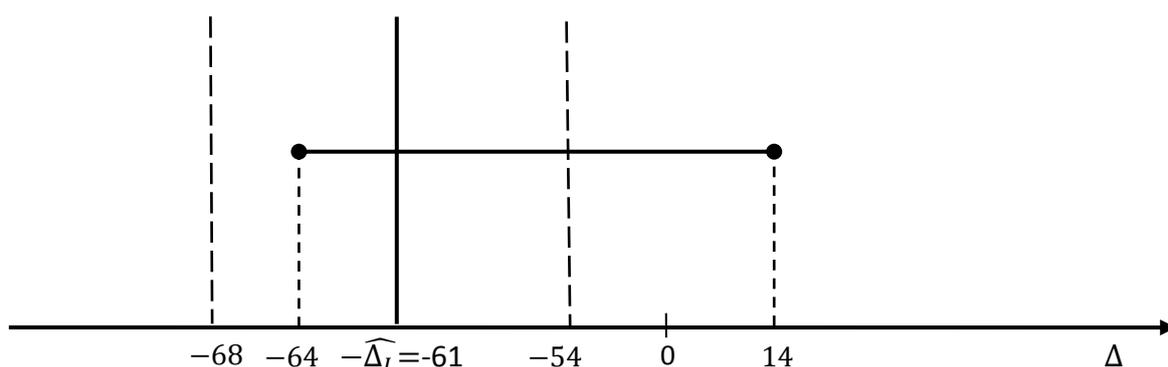


Figure 20. Résultat de l'analyse en intention de traiter du critère principal de l'essai Stratall avec les données disponibles. Représentation de la procédure de décision du test de non-infériorité basé sur les intervalles de confiance bilatéraux à 95 %.

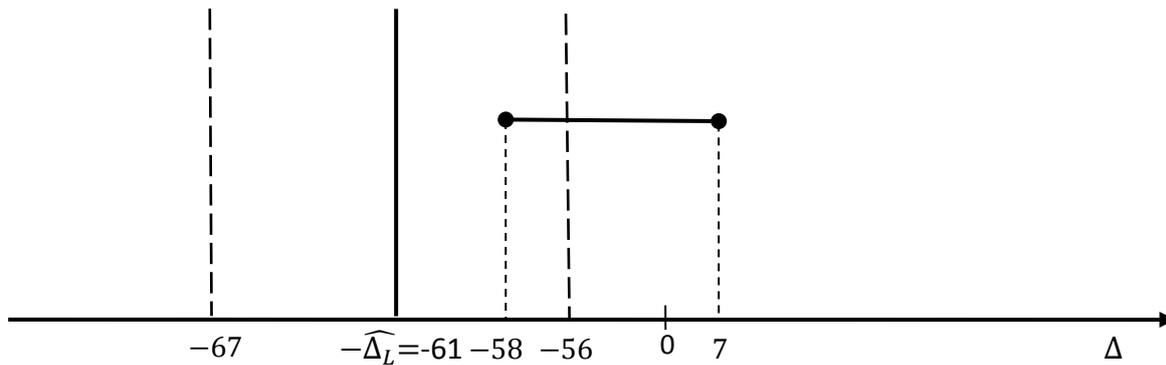


Figure 21. Résultat de l'analyse en intention de traiter du critère principal de l'essai Stratall avec les données disponibles. Représentation de la procédure de décision du test de non-infériorité basé sur les intervalles de confiance bilatéraux à 90 %.

Nous avons conclu conformément à la procédure de décision présentée à la **Figure 17** qu'on ne pouvait pas rejeter l'hypothèse nulle c'est-à-dire que la non-infériorité n'a pas été démontrée.

Cependant, comme indiqué précédemment, l'étude par simulation que nous avons menée pour déterminer la relation entre le seuil des procédures de décision proposées et le niveau des intervalles de confiance utilisés dans les procédures de test nous a montré que le niveau des procédures de test basés sur les intervalles de confiance à 90 % était inférieur à 5 % et qu'il fallait plutôt utiliser des procédures de test basés sur les intervalles de confiance à 80 % pour atteindre un niveau de test proche de 5 % (**Tableau 6**).

Nous avons donc effectué *a posteriori* un test basé sur des intervalles de confiance à 80%. La différence de gain moyen en lymphocytes T CD4 entre le groupe « suivi clinique exclusif » et le groupe « suivi clinico-biologique » à 24 mois de traitement en utilisant la méthode LOCF était alors de -31 (intervalle de confiance bilatéral à 80 % [-52; -9]) et la marge de non-infériorité -52 (intervalle de confiance bilatéral à 80 % [-56; -47]) **Figure 22**.

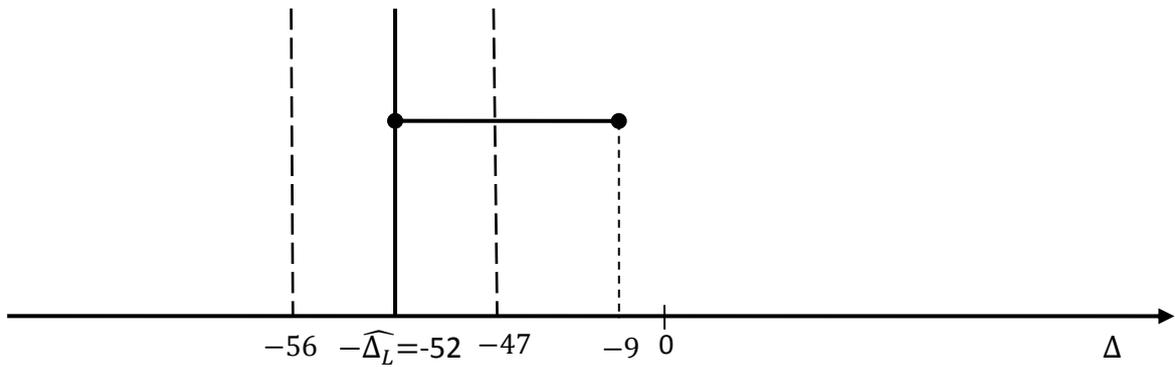


Figure 22. Résultat de l'analyse en intention de traiter du critère principal de l'essai Stratall avec la méthode LOCF (dernière valeur observée). Représentation de la procédure de décision du test de non-infériorité basé sur les intervalles de confiance à 80 %.

En limitant l'analyse aux données disponibles, la différence de gain moyen de lymphocytes T CD4 entre les deux groupes était de -25 (intervalle de confiance bilatéral à 80 % [-50; 0]) et la marge de non-infériorité -61 (intervalle de confiance bilatéral à 80 % [-66; -57])

Figure 23.

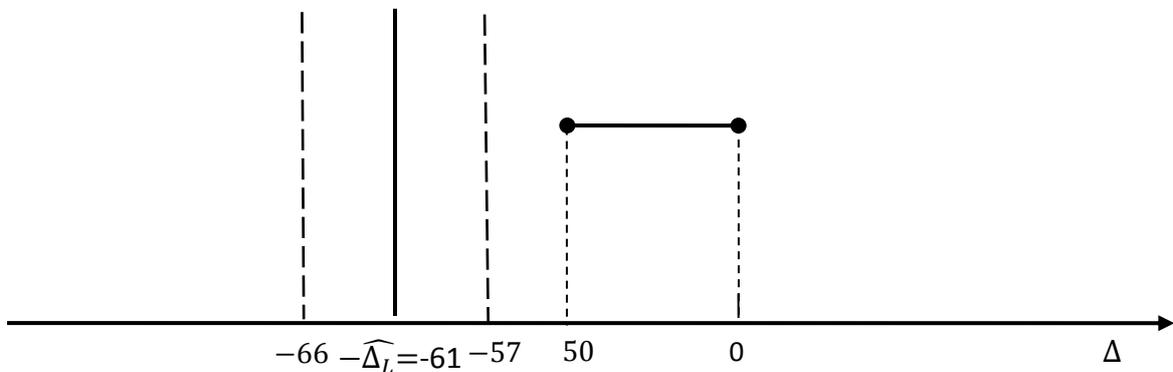


Figure 23. Résultat de l'analyse en intention de traiter du critère principal de l'essai Stratall avec les données disponibles. Représentation de la procédure de décision du test de non-infériorité basé sur les intervalles de confiance à 80 %.

Le résultat des analyses utilisant la méthode LOCF ne permettait pas de conclure que le suivi « clinique exclusif » est non-inférieur au suivi « clinico-biologique ». Au contraire, le résultat de l'analyse limitée aux données disponibles montrait la non-infériorité du suivi clinique exclusif.

4. Discussion

Dans ce chapitre, nous avons proposé deux procédures équivalentes de test de non-infériorité pour des critères quantitatifs continus, basées sur les intervalles de confiance. Nous avons montré que ces deux procédures de test sont valides lorsque la marge de non-infériorité est définie de façon relative (aléatoire et corrélée au paramètre d'intérêt du test). Nous avons établi par des simulations une relation entre le niveau des intervalles de confiance utilisés dans ces procédures de test de non-infériorité et le seuil de ces procédures de test. Nous avons appliqué la première procédure de test proposée (**section 2.2.2.1**) pour l'analyse du critère de jugement principal (gain moyen en lymphocytes T CD4 au 24^{ème} mois de traitement) de l'essai Stratall.

Dans l'essai Stratall, la stratégie de référence choisie n'avait été évaluée dans aucune étude. Il s'agissait en fait de celle pratiquée dans les pays développés et dans des centres de référence de pays du Sud (dont le Cameroun où cet essai a été réalisé).

En ce qui concerne le choix de la marge de non-infériorité, la différence relative maximale acceptée du gain moyen en lymphocytes T CD4 au 24^{ème} mois de traitement entre le groupe bénéficiant d'un suivi clinico-biologique et celui bénéficiant d'un suivi clinique exclusif était de 25 %. Cette différence était considérée cliniquement acceptable compte tenu d'une part des conditions financières et de la limite des infrastructures disponibles pour le suivi biologique et d'autre part du besoin urgent d'étendre la couverture thérapeutique par le biais de la décentralisation de la prise en charge des services spécialisés vers les structures périphériques.

Le choix du critère de jugement principal était motivé par le mécanisme d'action du traitement antirétroviral et par la durée souhaitée de l'étude. Le traitement antirétroviral permet de prévenir les infections opportunistes et de prolonger la vie. Cet objectif passe par la restauration immunitaire elle-même tributaire du contrôle de la réplication virale (**Figure 24**)

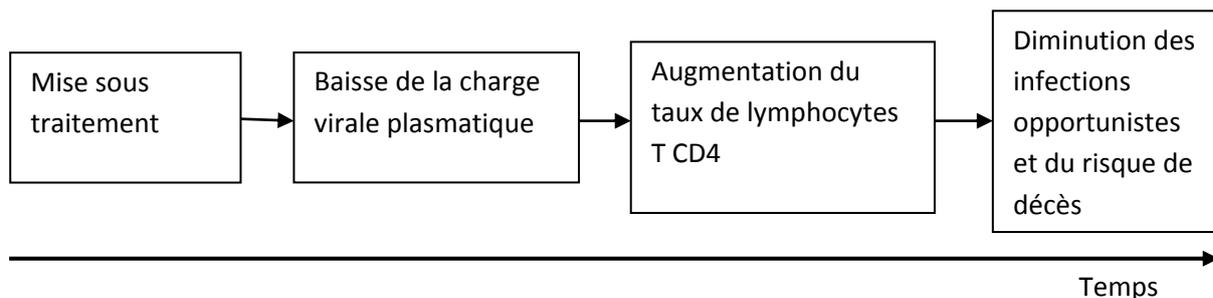


Figure 24. *Impact du traitement antirétroviral*

Compte tenu, d'un côté, de l'urgence à disposer de données permettant d'orienter la prise en charge des patients et, de l'autre côté, de la volonté d'étudier l'émergence des résistances aux antirétroviraux (en objectif secondaire), l'essai Stratall a été programmé pour un suivi des patients de 2 ans. Or, les infections opportunistes et les décès sont peu fréquents après le 6^{ème} mois de traitement quand l'efficacité thérapeutique peut être évaluée. Pour pouvoir utiliser ces événements comme critère de jugement principal, il faut suivre un grand nombre de patients suffisamment longtemps. Quand le suivi est relativement court, on peut utiliser des marqueurs intermédiaires tels que la charge virale plasmatique ou le taux de lymphocytes T CD4. Dans l'essai Stratall, le gain en lymphocytes T CD4 au 24^{ème} mois de traitement a donc été retenu comme critère de jugement principal.

Contrairement aux recommandations pour les essais de non-infériorité, nous n'avons pas pu réaliser d'analyse per protocole pour le critère de jugement principal car il était difficile dans cet essai d'intervention d'identifier les patients à exclure pour cette analyse. En effet, aucun patient n'avait changé de groupe (stratégie de prise en charge) pendant le suivi. Par ailleurs, il n'apparaissait pas opportun (risque de biais) d'exclure de l'analyse les patients du groupe « suivi clinico-biologique » qui n'avaient pas été mis rapidement sous traitement de seconde ligne après l'émergence des résistances. Enfin, l'analyse per protocole ne semblait pas indispensable sachant que la non-infériorité n'avait pas été montrée par l'analyse en intention de traiter. Nous supposons que les mêmes contraintes se sont opposées à l'analyse per protocole d'autres essais d'intervention [39, 146].

Les procédures de décision des tests de non-infériorité basées sur les intervalles de confiance sont les plus utilisées et les plus recommandées pour l'analyse des essais de non-infériorité [39, 128, 131, 138, 140, 145]. Toutefois, toutes les procédures de

décision proposées jusqu'ici dans la littérature faisaient l'hypothèse que la marge de non-infériorité était fixe. La particularité de l'essai Stratall au plan méthodologique tenait à la marge relative de non-infériorité. Nous avons donc proposé deux nouvelles procédures de test pour prendre en compte cette particularité. Nous avons déterminé par simulations la relation entre le seuil de ces procédures de test et le niveau des intervalles de confiances utilisées dans ces tests. Nous avons aussi montré par simulation que ces deux procédures de test étaient équivalentes.

Le critère de jugement principal était partiellement observé au 24^{ème} mois de traitement. Dans une première analyse, nous avons pris en compte cette présence de données manquantes en les remplaçant par les dernières valeurs observées (LOCF). Cette méthode de gestion des données manquantes sous-estime la variabilité du critère étudié car l'incertitude sur les valeurs remplaçantes n'est pas prise en compte. Elle peut donc entraîner une augmentation de l'étendue des intervalles de confiance utilisés dans les procédures de test. Les résultats obtenus en utilisant la méthode LOCF d'un côté et les données disponibles de l'autre côté ont montré qu'il y a un biais important sur les paramètres estimés ; la marge de non-infériorité était de -52 (intervalle de confiance à 80% [-56; -47]) par la méthode LOCF et de -61 (intervalle de confiance à 80 % [-66; -57]) en utilisant les données disponibles. Ce biais peut être attribué soit à la méthode LOCF soit à la sélection de l'échantillon engendrée par l'utilisation des données disponibles, en fonction des raisons de sorties d'étude (informatives ou non informatives). La prise en compte des raisons de sortie d'étude aurait été intéressante pour valider les résultats obtenus, mais les données sur les raisons de sortie d'étude pour les patients perdus de vue n'étaient pas disponibles. La discordance entre les résultats de ces deux méthodes d'analyse permet de poser plus globalement le problème de la gestion des données manquantes. Cette problématique sera abordée dans le chapitre suivant ; nous reviendrons notamment plus en détail sur les deux méthodes d'analyse utilisées ici.

Les procédures de décision des tests de non-infériorité proposées à la **section 2.2.2.1** et **2.2.2.2** ne prennent pas en compte l'ajustement sur d'autres covariables. Lorsqu'on a prévu de prendre en compte des covariables d'ajustement lors du test de non-infériorité, on utilise un modèle correspondant au type du critère de jugement principal pour déterminer une estimation et un intervalle de confiance de Δ ou $\Delta +$

Δ_L ajusté qu'on notera respectivement $\widehat{\Delta}_{ajusté}$ et $\widehat{\Delta + \Delta_L}_{ajusté}$. Si on considère notre exemple d'application, il est plus facile de construire un test ajusté à partir de la deuxième procédure de décision présentée ci-dessus (**section 2.2.2.2**). En effet, on peut estimer la différence de gain moyen en lymphocytes T CD4 au 24^{ème} mois de traitement entre les deux groupes ainsi que son intervalle de confiance à partir d'un modèle de régression linéaire. A l'inverse, on ne peut pas estimer la marge de non-infériorité Δ_L car elle dépend du gain moyen en lymphocytes T CD4 au 24^{ème} mois de traitement du groupe « suivi clinico-biologique » qui n'est pas directement estimé à partir du modèle de régression. Pour pallier à cet inconvénient, on peut directement estimer $\Delta + \Delta_L$. Pour cela, il suffit de remarquer que $\Delta + \Delta_L = \mu_N - \left(1 - \frac{a}{100}\right) * \mu_R$. En pondérant le gain en lymphocytes T CD4 au 24^{ème} mois de traitement des patients du groupe « suivi clinico-biologique » par $\left(1 - \frac{a}{100}\right)$, le coefficient de régression associé à la variable identifiant les stratégies de prise en charge est une estimation de $\Delta + \Delta_L$ ajusté. Une fois $\widehat{\Delta + \Delta_L}_{ajusté}$ et son intervalle de confiance obtenu, on peut donc appliquer la deuxième procédure de décision (**section 2.2.2.2**) pour conclure.

Chapitre 3 : *Données manquantes*

Les données sanitaires recueillies en routine ou, à un moindre degré, au cours d'études épidémiologiques contiennent très souvent des informations partielles. Les raisons de ces informations partielles dans le contexte des pays à ressources limitées sont nombreuses :

- Le manque de systèmes informatiques de suivi de patients : la mise en place de tels systèmes est très récente dans les pays d'Afrique subsaharienne et est le plus souvent encore limitée à quelques centres pilotes. La grande majorité des patients sont suivis à partir des carnets (livrets) individuels de visites ce qui rend la traçabilité des informations sur les patients difficile voire impossible.
- Le manque d'infrastructures : le plateau technique des structures sanitaires sont souvent très limités et, de ce fait, certains éléments ne peuvent être mesurés (données biologiques par exemple).
- Les patients qui ne respectent pas toujours leurs rendez-vous pour des raisons financières notamment.
- les erreurs de report, de saisie des données
- les données aberrantes

Or, la présence de données manquantes constitue un obstacle à l'évaluation des interventions de santé. Il est donc nécessaire de développer des méthodes statistiques permettant de tirer des conclusions valides à partir des données partiellement observées.

Les méthodes inférentielles standards élaborées pour les données complètement observées peuvent être invalides lorsque les données sont incomplètes (manquantes). *Rubin*[147, 148], *Little et Rubin* [149] ont discuté de plusieurs méthodes de gestion des données manquantes basées sur la théorie de la décision et des modèles pour différents mécanismes de données manquantes.

Le travail présenté dans ce chapitre a été motivé par l'étude comparant à long terme l'évolution immuno-virologique, la mortalité, la progression clinique de la maladie et la tolérance biologique entre les personnes infectées par le VIH sous traitement

antirétroviral incluant la didanosine et celles sous traitement incluant la zidovudine (cf **annexe B**). Dans cette étude, plusieurs covariables d'inclusions étaient partiellement observées ; leurs proportions de données manquantes variaient de 1 % à 26 % (**Tableau 7**). Au total, 47 % des 161 patients inclus avaient au moins une donnée manquante.

Tableau 7. Proportion des données manquantes de certaines variables d'inclusion dans l'étude comparant le traitement antirétroviral incluant la didanosine et celui incluant la zidovudine.

Variabiles	Nombre de données manquantes (%)	
Indice de masse corporelle	4	(3 %)
Taux de lymphocytes T CD4	11	(7 %)
Charge virale plasmatique	39	(26 %)
Hémoglobine	2	(1 %)
Neutrophiles	5	(3 %)
Alanine aminotransferase (ALT)	11	(7 %)
Aspartate aminotransferase (AST)	12	(8 %)

On notera dans la suite $Y = (Y_1, \dots, Y_p)$ le tableau de données de dimension $n \times p$, $Y = (Y^{obs}, Y^{miss})$ où Y^{obs} et Y^{miss} sont respectivement les parties observées et manquantes du tableau de données Y , R la matrice indicatrice de données manquantes de même dimension que Y , constituée des variables r_{ij} qui prennent la valeur 1 si y_{ij} est observée et 0 si y_{ij} est manquante et $Y_{-j} = (Y_1, \dots, Y_{j-1}, Y_{j+1}, \dots, Y_p)$. Ainsi, l'ensemble des données observées est $\{Y^{obs}, R\}$.

1. Typologie des données manquantes

La connaissance du mécanisme des données manquantes est un élément central pour le choix de l'analyse statistique appropriée. La nature du mécanisme des données manquantes peut être décrite par la distribution conditionnelle $f(R|Y, \psi)$ de la matrice indicatrice R sachant les données Y , où ψ est un paramètre associé au mécanisme des données manquantes.

1.1. Données manquantes complètement au hasard

Les données sont dites manquantes complètement au hasard (en anglais MCAR pour *missing completely at random*) si

$$f(R|Y, \psi) = f(R|\psi)$$

c'est-à-dire si le mécanisme qui engendre les données manquantes ne dépend pas des données elles-mêmes. La vraisemblance de l'ensemble des données peut donc être simplifiée de la façon suivante :

$$L(\theta, \psi|Y^{obs}, R) = f(Y^{obs}|\theta)f(R|\psi)$$

où θ est le paramètre de la loi de distribution des données. Ce résultat implique que si les paramètres θ et ψ sont distincts, le processus des données manquantes peut être ignoré, et que les estimateurs obtenus à partir de l'analyse des cas complets ou des données disponibles sont sans biais et consistants. Cependant, une proportion importante de données manquantes entraîne une perte de puissance, une diminution de la variabilité et de possibles phénomènes de confusion pouvant invalider l'inférence statistique.

1.2. Données manquantes au hasard

Les données sont dites manquantes au hasard (en anglais MAR pour *missing at random*), si

$$f(R|Y, \psi) = f(R|Y^{obs}, \psi)$$

c'est-à-dire si le mécanisme qui engendre les données manquantes, dépend des données observées. La vraisemblance de l'ensemble des données peut donc être simplifiée de la façon suivante :

$$L(\theta, \psi|Y^{obs}, R) = f(Y^{obs}|\theta)f(R|Y^{obs}, \psi)$$

Ce résultat implique que si les paramètres θ et ψ sont distincts, le mécanisme qui engendre les données manquantes peut être ignoré. L'analyse des cas complets et celle des données disponibles sont valides lorsque l'inférence porte sur la vraisemblance. Cependant, comme pour les données manquantes complètement au hasard, une proportion importante de données manquantes entraîne une perte de puissance, une diminution de la variabilité et de possibles phénomènes de confusion pouvant invalider l'inférence statistique.

1.3. Données manquantes non au hasard

Les données sont dites manquantes non au hasard (en anglais NMAR pour *not missing at random*) lorsque le processus de données manquantes dépend des données non observées :

$$f(R|Y, \psi) = f(R|Y^{obs}, Y^{mis}, \psi)$$

c'est-à-dire le mécanisme qui engendre les données manquantes dépend des données observées et des données non-observées. La vraisemblance ne peut être simplifiée :

$$L(\theta, \psi|Y^{obs}, R) = \int f(Y^{obs}, Y^{miss}|\theta) f(R|Y^{obs}, Y^{miss}, \psi) dY^{miss}$$

Dans ce cas, on dit que les données manquantes sont non ignorables ou informatives. Trois approches paramétriques de modélisation de données longitudinales avec des données manquantes informatives sont présentées dans la littérature :

- L'approche par les modèles de sélection, proposée par *Diggle et Kenward* [150]. Ces modèles correspondent à la factorisation de la distribution de l'ensemble des données observées de la façon suivante :

$$\begin{aligned} f(Y^{obs}, R|\theta, \psi) &= f(Y^{obs}|\theta) f(R|Y^{obs}, \psi) \\ &= \int f(Y^{obs}, Y^{miss}|\theta) dY^{miss} \int f(R|Y^{obs}, Y^{miss}, \psi) dY^{miss} \end{aligned}$$

Le facteur $f(R|Y^{obs}, \psi)$ décrit le mécanisme de sélection des sujets ayant les données manquantes.

- L'approche par les modèles de mélange, proposée par *Little* [151, 152], *Molenberghs et al* [153]. Ces modèles sont basés sur la factorisation de la distribution de l'ensemble des données observées de la façon suivante :

$$f(Y^{obs}, R|\theta, \psi) = f(Y^{obs}|R, \theta) f(R|\psi)$$

Cette approche permet des modèles différents pour chaque disposition des données manquantes. Les données observées sont supposées issues d'un mélange de population pondéré par la probabilité d'avoir des informations partielles.

- L'approche par les modèles à effets aléatoires partagés, proposé par *Wu et al* [154]

$$f(Y^{obs}, R|\theta, \psi) = \int f(Y^{obs}|R, \theta, u) f(R|\psi, u) f(b) du$$

Ici, on inclut explicitement un vecteur d'effets aléatoires u associés aux deux processus. Pour ces modèles, on fait souvent l'hypothèse suivant laquelle les données observées et le mécanisme des données manquantes sont indépendants conditionnellement aux effets aléatoires u .

2. Gestion des données manquantes

Il existe dans la littérature plusieurs approches méthodologiques qui permettent de gérer les problèmes liés aux données manquantes. On peut regrouper ces différentes approches en deux catégories : les approches qui utilisent les tableaux de données non complets (analyse des cas complets, analyse des cases disponibles) et les approches qui utilisent des tableaux de données complétés de façon judicieuse.

Comme nous l'avons présenté dans la **section 1**, l'analyse des cas complets ou des cases disponibles est valide lorsque le mécanisme des données manquantes est MCAR ou MAR et l'inférence est portée sur la vraisemblance. Cependant, nous avons aussi relevé les limites de ces approches d'analyse en termes de diminution de la puissance, et de sous-estimation de la variabilité . En pratique, ces approches d'analyse ne sont utilisées que lorsque la proportion des données manquantes est inférieure à 10 % des données attendues. De plus, lorsque l'inférence ne porte pas sur la vraisemblance, les résultats obtenus peuvent être biaisés (par exemple, l'estimateur des paramètres du modèle GEE est biaisé lorsque le mécanisme des données manquantes est MAR ou NMAR). Une alternative est l'approche basée sur l'imputation des données manquantes. Elles consistent à compléter le tableau de données incomplet en remplaçant les valeurs manquantes par une ou plusieurs valeurs plausibles. Toutefois, les approches basées sur l'imputation simple ou multiple des données manquantes ne sont pas valides lorsque le processus des données manquantes est NMAR.

2.1. Imputation des données manquantes

Deux approches d'imputation de données manquantes existent. La première approche dite « d'imputation simple » consiste à remplacer chaque donnée manquante par une valeur plausible. Plusieurs méthodes d'imputation simple sont proposées dans la littérature : les méthodes hot-deck (chaque donnée manquante est remplacée par la valeur observée pour un individu répondant choisi au hasard), les méthodes cold deck (chaque donnée manquante est remplacée par la valeur observée pour un individu répondant choisit dans une autre enquête), l'imputation par la moyenne et l'imputation par les méthodes de régression, etc. Cependant, cette approche d'imputation ne permet pas de prendre en compte l'incertitude liée à la valeur imputée. Les valeurs imputées sont traitées comme les vraies valeurs non observées ; or, ces valeurs imputées ne sont que des valeurs probables.

Une alternative à cette approche consiste à remplacer chaque donnée manquante par plusieurs valeurs plausibles issues de la distribution prédictive des données qui est construite à partir des modèles d'imputation ; c'est l'approche dite « d'imputation multiple » [147]. Cette approche permet de générer plusieurs tableaux de données complétés. Les résultats obtenus à partir de chaque tableau de données sont combinés pour construire les estimateurs. Des études ont montré que ces estimateurs sont asymptotiquement sans biais et de variance bien estimée si le modèle d'imputation est bien défini [155, 156].

Deux approches paramétriques d'imputation multiple existent dans la littérature : la modélisation jointe [157] et l'approche basée sur les distributions conditionnelles univariées[158].

L'approche par modélisation jointe consiste à spécifier la distribution multivariée des données incomplètes. Les valeurs imputées sont alors issues de la distribution conditionnelle des données manquantes par des techniques de Monte Carlo par chaîne de Markov (MCMC). Cette méthode est intéressante lorsqu'on peut définir la distribution multivariée des données.

L'approche basée sur les distributions conditionnelles univariées est une alternative intéressante lorsqu'on ne peut pas définir explicitement la distribution multivariée des

données. Elle consiste à définir implicitement la distribution multivariée des données en spécifiant un ensemble de densités conditionnelles univariées (une pour chaque variable contenue dans le tableau de données). Ainsi, les valeurs imputées pour chaque variable sont prédites à partir d'un modèle d'imputation issu de la distribution conditionnelle de la variable correspondante.

Le choix du modèle d'imputation est déterminant pour la validité de l'inférence basée sur l'approche d'imputation multiple. La partie structurelle et la distribution de l'erreur du modèle d'imputation doivent être bien spécifiées pour refléter la vraie distribution conditionnelle de la variable à imputer. Ainsi, le choix des covariables à inclure dans la partie structurelle [156, 159-161] et la relation entre la variable à imputer et les covariables sont très importants. Cependant, les modèles d'imputations disponibles dans la majorité des logiciels sont des modèles linéaires, c'est-à-dire qu'ils font l'hypothèse d'une relation linéaire entre la variable à imputer et les covariables du modèle d'imputation. Cette hypothèse de linéarité n'est pas toujours vérifiée sur les données réelles.

Nous proposons dans ce chapitre une méthode d'imputation multiple des variables continues qui permet de prendre en compte des relations non linéaires entre la variable à imputer et les covariables du modèle d'imputation. Cette méthode est basée sur l'approximation de la vraie relation entre la variable à imputer et les covariables du modèle d'imputation par des fonctions B-splines, en préservant l'hypothèse d'additivité du modèle d'imputation.

3. Imputation multiple par les fonctions B-splines

Soit $P(Y, \theta)$ la distribution multivariée du tableau de données Y . Nous supposons que $P(Y, \theta)$ est complètement définie par la connaissance de θ , le vecteur des paramètres inconnus. Pour prendre en compte la complexité posée par les données réelles généralement constituées d'un mélange de variables qualitatives et quantitatives, nous avons utilisé la méthode d'imputation multiple basée sur les densités de probabilités conditionnelles univariées.. Cette méthode permet de définir

implicitement la distribution multivariée $P(Y, \theta)$ en spécifiant séparément les densités de probabilités conditionnelles univariées $P(Y_j|Y_{-j}, \theta_j)$ pour chaque $Y_j, j = 1, \dots, p$; les paramètres $\theta_1, \dots, \theta_p$ sont spécifiques aux densités conditionnelles univariées $P(Y_j|Y_{-j}, \theta_j)$. On peut ainsi définir la distribution prédictive *a posteriori* de Y_j par :

$$P(Y_j|Y_j^{obs}, Y_{-j}) = \int P(Y_j|Y_j^{obs}, Y_{-j}, \theta_j) P(\theta_j|Y_j^{obs}, Y_{-j}) d\theta_j$$

Cette distribution prédictive *a posteriori* est souvent difficile à déterminer analytiquement. Au contraire, la distribution conditionnelle prédictive $P(Y_j|Y_j^{obs}, Y_{-j}, \theta_j)$ est souvent facile à obtenir. Donc, si une réalisation du paramètre θ_j peut être obtenue à partir de sa distribution *a posteriori* $P(\theta_j|Y_j^{obs}, Y_{-j})$, alors l'échantillon obtenu à partir de la distribution prédictive conditionnelle $P(Y_j|Y_j^{obs}, Y_{-j}, \theta_j)$ pour la réalisation de θ_j est un échantillon issu de la distribution prédictive *a posteriori* $P(Y_j|Y_j^{obs}, Y_{-j})$.

Toutefois, la matrice prédictive Y_{-j} utilisée dans le processus d'imputation de Y_j peut aussi être incomplète. Dans ce cas, on généralise la procédure d'échantillonnage présentée ci-dessus par la procédure itérative suivante : on initialise la procédure par un simple tirage issu de la distribution marginale des données observées ; la t^{th} itération conduit à un échantillonnage de Gibbs par tirages successifs de la façon suivante :

$$\begin{aligned} \theta_1^{*(t)} &\sim P(\theta_1|Y_1^{obs}, Y_2^{(t-1)}, \dots, Y_p^{(t-1)}) \\ Y_1^{miss*(t)} &\sim P(Y_1|Y_1^{obs}, Y_2^{(t-1)}, \dots, Y_p^{(t-1)}, \theta_1^{*(t)}) \\ \theta_2^{*(t)} &\sim P(\theta_2|Y_2^{obs}, Y_1^{(t)}, Y_3^{(t-1)}, \dots, Y_p^{(t-1)}) \\ Y_2^{miss*(t)} &\sim P(Y_2|Y_2^{obs}, Y_1^{(t)}, Y_3^{(t-1)}, \dots, Y_p^{(t-1)}, \theta_2^{*(t)}) \\ \theta_p^{*(t)} &\sim P(\theta_p|Y_p^{obs}, Y_1^{(t)}, Y_2^{(t)}, \dots, Y_{p-1}^{(t)}) \\ Y_p^{miss*(t)} &\sim P(Y_p|Y_p^{obs}, Y_1^{(t)}, Y_2^{(t)}, \dots, Y_{p-1}^{(t)}, \theta_p^{*(t)}) \end{aligned}$$

où $Y_j^{(t)} = (Y_j^{obs}, Y_j^{miss*(t)})$ est la variable Y_j complétée à l'itération t . La convergence de cette procédure d'échantillonnage de Gibbs est rapide contrairement à d'autres méthodes MCMC. Le nombre d'itérations nécessaires pour la convergence est souvent compris entre 10 et 20 [162].

Plusieurs modèles d'imputation utilisant cette procédure d'échantillonnage de Gibbs ont été proposés dans la littérature pour estimer les paramètres θ_j^* , pour $j = 1, \dots, p$ et prédire $Y_j^{miss^*}$ sachant Y_j^{obs} et la matrice prédictive Y_{-j} [158, 162, 163]. Aucun des modèles d'imputation proposés à notre connaissance ne permet toutefois de prendre en compte les effets non linéaires des covariables continues. Nous proposons donc de prendre en compte les effets non linéaires des covariables continues du modèle d'imputation pour améliorer la qualité des prédictions.

Pour cela, nous supposons que le modèle d'imputation considéré est de la forme suivante :

$$Y_j = f(Y_{-j}) + \varepsilon_j$$

Nous proposons de maintenir l'hypothèse d'indépendance des covariables par l'utilisation d'un modèle additif :

$$Y_j = \sum_{l=1, l \neq j}^p f_l(Y_l) + \varepsilon_j$$

A l'itération t de la procédure d'échantillonnage de Gibbs présentée ci-dessus, nous proposons d'utiliser le modèle d'imputation suivant pour les variables continues :

$$Y_j^{(t)} = \sum_{l=1}^{j-1} f_l(Y_l^{(t)}) + \sum_{l=j+1}^p f_l(Y_l^{(t-1)}) + \varepsilon_j^{(t)}$$

où $\varepsilon_j^{(t)} = (\varepsilon_{1j}^{(t)}, \dots, \varepsilon_{n_j}^{(t)})$, $\varepsilon_{ij}^{(t)} \sim \mathcal{N}(0, \sigma_j^2)^{(t)}$ pour $i = 1, \dots, n_j$, $f_l(\cdot)$ est une fonction inconnue. Nous proposons de s'affranchir de l'hypothèse de linéarité en utilisant des fonctions B-splines de degré et nœuds fixes pour approximer les fonctions $f_l(\cdot)$.

3.1. Approximation de $f_l(\cdot)$ par des fonctions B-splines

Le mot anglais « spline » désigne une latte flexible utilisée par les dessinateurs pour matérialiser des lignes à courbures variables passant par des points fixés *a priori* ou à proximité de ceux-ci. Ce mot désigne également des familles de fonctions d'interpolation ou de lissage présentant des propriétés « optimales » de régularité.

Lorsque la forme de f_l est inconnue, il est pratique d'utiliser une approximation de la forme d'un polynôme d'ordre $d_l + 1$

$$p_l(x) = \sum_{j=0}^{d_l} \beta_j x^j$$

Un avantage de ce type de modélisation réside dans le fait que le modèle est linéaire en ses coefficients $\beta_0, \dots, \beta_{d_l}$. Ces derniers peuvent donc être estimés par les méthodes habituelles, inférentielles (moindres carrés) ou bayésiennes. De plus, les polynômes permettent d'approcher une grande variété de fonctions.

Si l'on suppose que f_l est $d_l - 1$ fois dérivable sur $[a_l, b_l]$, le théorème de Taylor avec le reste intégral permet d'écrire :

$$f_l(x) = p_l(x) + r_l(x)$$

où $r_l(x) = \int_{a_l}^x \frac{f_l^{(d_l-1)}(\xi)}{d_l!} (x - \xi)^{d_l} d\xi = \int_{a_l}^{b_l} \frac{f_l^{(d_l-1)}(\xi)}{d_l!} (x - \xi)_+^{d_l} d\xi$. On rappelle que u_+ indique la fonction qui vaut u pour $u > 0$ ou zéro par ailleurs.

En supposant que le reste $r_l(x)$ est négligeable sur $[a_l, b_l]$, l'approximation de f_l par $p_l(x)$ est adaptée. Cependant, une telle hypothèse n'est pas toujours satisfaite. En effet, un grand nombre d'exemples montre que l'utilisation des polynômes est inadéquate dans diverses situations. Ils sont globalement déterminés par les données et celles qui sont proches des bornes de l'intervalle influent sur la globalité de l'estimation.

Une alternative, afin d'éviter de tels désagréments, est de ne plus négliger le reste r_l dans l'approximation. Soient $\xi_1^l, \dots, \xi_{K_l}^l$, K_l points tels que $a_l < \xi_1^l < \dots < \xi_{K_l}^l < b_l$ et $\beta_{d_l+1}, \dots, \beta_{d_l+K_l}$, K_l constantes. Le reste $r_l(x)$ peut être approché par $\sum_{j=1}^{K_l} \beta_{d_l+j} (x - \xi_j^l)_+^{d_l}$. Nous obtenons ainsi l'approximation suivante de f_l .

$$s_l(x) = \sum_{j=0}^{d_l} \beta_j x^j + \sum_{j=1}^{K_l} \beta_{d_l+j} (x - \xi_j^l)_+^{d_l}$$

$s_l(x)$ est une spline de degré d_l avec K_l nœuds $\xi^l = (\xi_1^l, \dots, \xi_{K_l}^l)'$.

Pour un vecteur de nœuds ξ^l fixé, on note $S_{\xi^l}^{d_l+1}$ l'ensemble des fonctions splines de degré d_l et de nœuds ξ^l . L'ensemble $S_{\xi^l}^{d_l+1}$ est un espace vectoriel de dimension

$d_l + K_l + 1$ dont les fonctions $1, x, \dots, x^{d_l}, (x - \xi_1^l)_+^{d_l}, \dots, (x - \xi_{K_l}^l)_+^{d_l}$ forment une base appelée base des puissances tronquée [164].

La nature de la fonction spline permet une approximation plus souple de f_l que celle des polynômes. La fonction spline s_l est constituée de polynômes par morceaux reliés par des nœuds, de sorte que la dérivée $d_l - 1$ de s_l soit continue. Avec ces fonctions, chaque donnée n'influe que localement sur l'estimation.

On peut remarquer que l'approximation polynômiale $p_l(x)$ de $f_l(x)$ proposée plus haut n'est rien d'autre qu'un cas particulier de l'approximation par les splines $s_l(x)$ avec $\beta_{d_l+1} = \dots = \beta_{d_l+K} = 0$ ou avec un ensemble de nœuds vides.

A ξ^l fixé, en remplaçant $f_l(x)$ dans le modèle d'imputation par $s_l(x)$, le vecteur β des paramètres de régression peut être estimé par les méthodes habituelles, inférentielles (les moindres carrés) ou bayésiennes. La fonction s_l est appelée représentation de f_l dans la base des puissances tronquées [165]. Comme nous l'avons vu, la base des puissances tronquées est particulièrement simple. Cependant, cette base n'est pas bien adaptée au calcul numérique. Par exemple, deux nœuds confondus entraînent des problèmes d'inversion des matrices.

Les problèmes précédents inhérents à la représentation de f_l dans la base des puissances tronquées peuvent être, en partie, résolus en utilisant les polynômes par morceaux comme base de $S_\xi^{d_l+1}$. On remarque que la représentation s_l de f_l dans la base des puissances tronquées peut encore s'écrire $s_l(x) = \sum_{j=0}^{d_l} \delta_{jr} x^j$, $x \in [\xi_r^l, \xi_{r+1}^l[$, $r = 0, \dots, K_l$, où $\xi_0^l = a_l$ et $\xi_{K_l+1}^l = b_l$ et où les coefficients δ_{jr} satisfont les contraintes de dérivabilité $s_l^{(j)}(\xi_r^l -) = s_l^{(j)}(\xi_r^l +)$, $l = 0, \dots, d_l - 1$, $r = 1, \dots, K_l$ [164]. En remplaçant f_l dans le modèle d'imputation par sa représentation dans la base des polynômes par morceaux, le nombre de paramètres de régression à estimer passe à $\sum_{l=1, l \neq j}^p (K_l + 1)(d_l + 1)$ contre seulement $\sum_{l=1, l \neq j}^p (K_l + d_l + 1)$ lorsqu'on utilise la représentation dans la base des puissances tronquées. Pour cette raison, la base des polynômes par morceaux n'a que peu été utilisée pour la régression spline [166].

Le compromis entre les puissances tronquées et les polynômes par morceaux est l'utilisation des B-splines.

Définissons $2 \times (d_l + 1)$ nœuds supplémentaires $\xi_{-d_l}^l = \dots = \xi_0^l = a_l$ et $b_l = \xi_{K_l+1}^l = \dots = \xi_{K_l+d_l+1}^l$. L'ensemble des fonctions $(B_j^{d_l+1}(x))_{j=1, \dots, K_l+d_l+1}$ définies sur l'intervalle $[a_l, b_l]$ par :

$$B_j^{d_l+1}(x) = (-1)^{d_l+1} (\xi_j^l - \xi_{j-d_l-1}^l) [\xi_{j-d_l-1}^l, \xi_j^l] (x - \xi_j^l)$$

forme une base de $S_\xi^{d_l+1}$ [164]. La fonction $(\xi_j, \dots, \xi_{j+m}) \mapsto [\xi_j, \dots, \xi_{j+m}]f$ est la fonction "différence divisée", définie, pour des ξ_i , deux à deux distincts par

$$[\xi_j, \dots, \xi_{j+m}]f = \sum_{i=j}^{j+m} \frac{f(\xi_i)}{\prod_{l=j, l \neq i}^{j+m} (\xi_i - \xi_l)}$$

Ainsi, la représentation de f_l dans la base des fonctions B-spline peut s'écrire :

$$s_l(x) = \sum_{j=1}^{K_l+d_l+1} \beta_j B_j^{d_l+1}(x), \quad x \in [a_l, b_l]$$

Une des raisons du succès des B-splines est le fait qu'en plus de la qualité d'ajustement, elles sont des éléments d'un espace vectoriel de dimension finie. Tout comme pour les polynômes usuels, une fois la base $\{B_1^{d_l+1}, \dots, B_{K_l+d_l+1}^{d_l+1}\}$ connue, la représentation dans la base spline (fonction spline) s'exprime comme combinaison linéaire des fonctions de base.

Ainsi, pour un vecteur de nœuds ξ^l et de degré d_l fixés, on peut remplacer la fonction f_l par sa représentation dans la base B-spline. Le modèle d'imputation proposé s'écrit alors de la façon suivante :

$$Y_j^{(t)} = \sum_{l=1}^{j-1} \sum_{r=1}^{K_l+d_l+1} \beta_r B_r^{d_l+1}(Y_l^{(t)}) + \sum_{l=j+1}^p \sum_{r=1}^{K_l+d_l+1} \beta_r B_r^{d_l+1}(Y_l^{(t-1)}) + \varepsilon_j^{(t)}$$

Ce modèle étant linéaire par rapport aux paramètres, les paramètres de régression peuvent être estimés par des méthodes usuelles inférentielles ou bayésiennes.

Nous avons rédigé un programme sous le logiciel R permettant de mettre en œuvre cette approche d'imputation multiple proposée, en apportant des modifications nécessaires aux fonctions `mice` et `mice.imput.norm` du package `mice` version 2.2. Nous avons appliqué cette approche d'imputation non-linéaire pour imputer des

covariables continues dans le contexte d'un modèle semi-paramétrique de Cox. Dans cette application, nous avons choisi d'utiliser une approche bayésienne pour estimer les paramètres des modèles d'imputation. Cette application a fait l'objet d'un article dont la version révisée a été soumise et dont le texte intégral est présenté dans l'annexe D.

Conclusion générale et perspectives

Dans ce travail de thèse, nous avons abordé différentes problématiques statistiques rencontrées dans des études sur le traitement antirétroviral des adultes infectés par le VIH en Afrique subsaharienne dont la modélisation de données longitudinales, l'analyse de données issues d'essais de non-infériorité et la gestion des données manquantes. Nous avons notamment proposé deux nouvelles procédures de test de non-infériorité lorsque la marge de non-infériorité est relative et une méthode de prise en compte des effets non linéaires des covariables continues du modèle d'imputation dans la procédure d'imputation multiple par les distributions conditionnelles univariées.

Nous avons ainsi illustré les capacités des modèles linéaires mixtes généralisés, d'équations d'estimation généralisées et de Cox à modéliser l'évolution immuno-virologique et clinique après l'instauration d'un traitement antirétroviral malgré les limites que nous avons discutées. Ces modèles sont de plus en plus utilisés pour évaluer la réponse au traitement antirétroviral mais, en raison de leur complexité, les utilisateurs peu avertis ont souvent du mal à mettre en œuvre ces modèles de façon efficiente. C'est la raison pour laquelle nous nous sommes longuement attardés sur ces modèles. Nous avons également présenté les nouveaux développements de ces modèles pour la prise en compte des sorties d'études informatives et les modèles conjoints. Il serait intéressant d'appliquer ces nouveaux développements à l'étude du traitement antirétroviral en Afrique subsaharienne pour prendre en compte les sorties d'étude informatives dont les causes de perte de vue. Nous avons enfin posé le problème lié à l'inobservation fréquente des dates exactes de survenue des événements lors de leur analyse en termes de biais ; il serait aussi intéressant de réfléchir sur les méthodes d'analyse de survie dans ce contexte.

Nous nous sommes ensuite intéressés à la question des tests de non-infériorité quand la marge de non-infériorité est relative et avons développé deux extensions de la procédure classique des tests de non-infériorité. Nous avons montré par simulations de la relation entre le niveau des intervalles de confiance et le seuil de ces tests que les deux procédures proposées étaient équivalentes. L'ajustement sur différentes covariables dans ces procédures de test pourrait être envisagé en

utilisant, d'une part, un modèle mixte incluant ces covariables et, d'autre part, la deuxième procédure de test proposée.

Enfin, nous nous sommes intéressés à la gestion des données manquantes. Nous avons proposé une extension de la procédure d'imputation multiple des variables continues par les densités conditionnelles univariées. Cette extension consiste à modéliser les effets non linéaires des covariables continues dans les modèles d'imputation à l'aide des fonctions B-splines de degré et nœuds fixes. Nous avons rédigé un programme sous le logiciel R permettant de mettre en œuvre cette approche d'imputation multiple proposée, en apportant des modifications nécessaires aux fonctions `mice` et `mice.imput.norm` du package `mice` version 2.2. Nous avons montré, par simulations et sur des données réelles, l'intérêt de cette procédure dans un modèle semi-paramétrique de Cox. Il serait intéressant dans de futurs travaux d'une part d'optimiser le degré et la position des nœuds des fonctions B-splines et d'autre part d'étendre cette procédure d'imputation multiple non-linéaire pour l'imputer des variables quantitatives discrètes et qualitatives.

Nous espérons que ce travail de thèse contribuera, d'une part, à la compréhension de méthodes statistiques de plus en plus complexes par des utilisateurs peu avertis et, d'autre part, à la disponibilité dans les logiciels de statistique courants des nouvelles procédures de test que nous avons proposées.

Bibliographie

1. World Health Organization. GLOBAL HIV/AIDS RESPONSE : Epidemic update and health sector progress towards Universal Access : Progress Report 2011. In; 2011.
2. Hammer SM, Squires KE, Hugues MD, Grimes JM, Demeter LM, Currier JS, *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**:725-733.
3. Gulick RM, Mellors JW, Havlir D, Eron JJ, Gonzalez C, McMahon D, *et al.* Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. *N Engl J Med* 1997,**337**:734-739.
4. Carrieri P, Spire B, Duran S, Katlama C, Peyramond D, Francois C, *et al.* Health-related quality of life after 1 year of highly active antiretroviral therapy. *J Acquir Immune Defic Syndr* 2003,**32**:38-47.
5. Mocroft A, Ledergerber B, Katlama C, Kirk O, Reiss P, d'Arminio Monforte A, *et al.* Decline in the AIDS and death rates in the EuroSIDA study: an observational study. *Lancet* 2003,**362**:22-29.
6. Palella FJ, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, *et al.* Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *N Engl J Med* 1998,**338**:853-860.
7. Katzenstein D, Laga M, Moatti J-P. The evaluation of the HIV/AIDS drug access initiatives in Côte d'Ivoire, Senegal and Uganda: how access to antiretroviral treatment can become feasible in Africa. *AIDS* 2003,**17**.
8. Laurent C, Diakhaté N, Ngom Gueye NF, Touré MA, Sow PS, Faye MA, *et al.* The Senegalese government's highly active antiretroviral therapy initiative: an 18-month follow-up study. *AIDS* 2002,**16**:1363-1370.
9. Laurent C, Ngom Gueye NF, Ndour CT, Gueye PM, Diouf M, Diakhaté N, *et al.* Long-term benefits of highly active antiretroviral therapy in Senegalese HIV-1-infected adults. *J Acquir Immune Defic Syndr* 2005,**38**:14-17.

10. Weidle PJ, Malamba S, Mwebaze R, Sozi C, Rukundo G, Dowing R, *et al.* Assessment of a pilot antiretroviral drug therapy programme in Uganda: patients' response, survival, and drug resistance. *Lancet* 2002,**360**:34-40.
11. Djomand G, Roels T, Ellerbrock T, Hanson D, Diomande F, Monga B, *et al.* Virologic and immunologic outcomes and programmatic challenges of an antiretroviral treatment pilot project in Abidjan, Côte d'Ivoire. *AIDS* 2003,**17(suppl 3)**:S5-S15.
12. World Health Organization. Towards universal access. Scaling up priority HIV/AIDS interventions in the health sector. Progress report 2010. In; 2010.
13. World Health Organization. Antiretroviral therapy for HIV infection in adults and adolescents. Recommendations for a public health approach. 2010 revision. In; 2010.
14. Institut National de la Statistique (INS) Cameroun. Enquête Démographique et de Santé et à Indicateurs Multiples EDS-MICS 2011: Rapport Préliminaire sur la prévalence de VIH. In. Yaoundé, Cameroun; 2012.
15. UNAIDS. Global report: UNAIDS report on the global AIDS epidemic 2010. In; 2010.
16. Comité Nationale de Lutte contre le Sida (CNLS) Cameroun. Rapport annuel 2010 des activités du programme national de lutte contre le Sida. In. Yaoundé, Cameroun; 2011.
17. Agence Nationale de la Statistique et de la Démographie (ANSD). Sénégal Enquête Démographique et de Santé à Indicateurs Multiples 2010-2011: Rapport Préliminaire. In. Dakar, Senegal: Agence Nationale de la Statistique et de la Démographie (ANSD); 2011.
18. Conseil Nationale de Lutte contre le Sida (CNLS) Sénégal. Rapport de la situation sur la riposte nationale de l'épidémie du VIH/SIDA Sénégal: 2008-2009. In; 2010.
19. Modi AA, Feld JJ. Viral hepatitis and HIV in Africa. *AIDS Rev* 2007,**9**:25-39.
20. Hoffmann CJ, Thio CL. Clinical implications of HIV and hepatitis B co-infection in Asia and Africa. *Lancet Infect Dis* 2007,**7**:402-409.
21. Rockstroh JK, Spengler U. HIV and hepatitis C virus co-infection. *Lancet Infect Dis* 2004,**4**:437-444.
22. Soriano V, Puoti M, Garcia-Gasco P, Rockstroh JK, Benhamou Y, Barreiro P, *et al.* Antiretroviral drugs and liver injury. *AIDS* 2008,**22**:1-13.

23. Stern JO, Robinson PA, Love J, Lanes S, Imperiale MS, Mayers DL. A comprehensive hepatic safety analysis of nevirapine in different populations of HIV infected patients. *J Acquir Immune Defic Syndr* 2003,**34 Suppl 1**:S21-33.
24. Hoffmann CJ, Charalambous S, Martin DJ, Innes C, Churchyard GJ, Chaisson RE, *et al.* Hepatitis B virus infection and response to antiretroviral therapy (ART) in a South African ART program. *Clin Infect Dis* 2008,**47**:1479-1485.
25. Bourgeois A, Laurent C, Mougnotou R, Nkoué N, Lactuock B, Ciaffi L, *et al.* Field assessment of generic antiretroviral drugs: a prospective cohort study in Cameroon. *Antiviral Therapy* 2005,**10**:335-341.
26. Laurent C, Kouanfack C, Koulla-Shiro S, Nkoué N, Bourgeois A, Calmy A, *et al.* Effectiveness and safety of a generic fixed-dose combination of nevirapine, stavudine, and lamivudine in HIV-1-infected adults in Cameroon: open-label multicentre trial. *Lancet* 2004,**364**:29-34.
27. Laurent C, Bourgeois A, Mpoudi-Ngole E, Kouanfack C, Ciaffi L, Nkoué N, *et al.* High rates of active hepatitis B and C co-infections in HIV-1 infected Cameroonian adults initiating antiretroviral therapy. *Hiv Medicine* 2010,**11**:85-89.
28. Colebunders R, Kanya MR, Laurence J, Kambugu A, Byakwaga H, Mwebaze PS, *et al.* First-line antiretroviral therapy in Africa - How evidence-based are our recommendations? *Aids Reviews* 2005,**7**:148-154.
29. Moh R, Danel C, Sorho S, Sauvageot D, Anzian A, Minga A, *et al.* Haematological changes in adults receiving a zidovudine-containing HAART regimen in combination with cotrimoxazole in Cote d'Ivoire. *Antiviral Therapy* 2005,**10**:615-624.
30. Ford N, Calmy A. Improving first-line antiretroviral therapy in resource-limited settings. *Current Opinion in Hiv and Aids* 2010,**5**:38-47.
31. Carr A, Amin J. Efficacy and tolerability of initial antiretroviral therapy: a systematic review. *Aids* 2009,**23**:343-353.
32. Desclaux A, Ciss M, Taverne B, Sow PS, Egrot M, Faye MA, *et al.* Access to antiretroviral drugs and AIDS management in Senegal. *Aids* 2003,**17**:S95-S101.
33. Etard JF, Ndiaye I, Thierry-Mieg M, Gueye NF, Gueye PM, Laniece I, *et al.* Mortality and causes of death in adults receiving highly active antiretroviral therapy in Senegal: a 7-year cohort study. *AIDS* 2006,**20**:1181-1189.

34. Laniece I, Ciss M, Desclaux A, Diop K, Mbodj F, Ndiaye B, *et al.* Adherence to HAART and its principal determinants in a cohort of Senegalese adults. *AIDS* 2003,**17**:S103-S108.
35. WHO. Scaling up antiretroviral therapy in resource-limited settings: treatment guidelines for a public health approach. In: http://www.who.int/hiv/pub/prev_care/en/arvrevision2003en.pdf; 2003.
36. WHO-UNAIDS report. Treating 3 million by 2005: Making it happen. In: <http://www.who.int/3by5/publications/documents/en/3by5StrategyMakingItHappen.pdf>; 2003.
37. Phillips AN, Pillay D, Miners AH, Bennett DE, Gilks CF, Lundgren JD. Outcomes from monitoring of patients on antiretroviral therapy in resource-limited settings with viral load, CD4 cell count, or clinical observation alone: a computer simulation model. *Lancet* 2008,**371**:1443-1451.
38. Mermin J, Ekwaru JP, Were W, Degerman R, Bunnell R, Kaharuza F, *et al.* Utility of routine viral load, CD4 cell count, and clinical monitoring among adults with HIV receiving antiretroviral therapy in Uganda: randomised trial. *BMJ* 2011,**343**:d6792.
39. Mugenyi P, Walker AS, Hakim J, Munderi P, Gibb DM, Kityo C, *et al.* Routine versus clinically driven laboratory monitoring of HIV antiretroviral therapy in Africa (DART): a randomised non-inferiority trial. *Lancet* 2010,**375**:123-131.
40. Chavance M. [Modeling correlated data in epidemiology: mixed or marginal model?]. *Rev Epidemiol Sante Publique* 1999,**47**:535-544.
41. Breslow NE. Analysis of survival data under proportional hazard model. *International Statistical Review* 1975,**43**:45-58.
42. Thiebaut R, Jacqmin-Gadda H, Journot V, Ferchal F, Rancinan C, Molina JM, *et al.* Analysis of undetectable HIV RNA using survival analysis: Results must be interpreted carefully. *Hiv Clinical Trials* 2003,**4**:417-420.
43. Manda SOM. A Nonparametric Frailty Model for Clustered Survival Data. *Communications in Statistics-Theory and Methods* 2011,**40**:863-875.
44. Molenberghs G, Verbeke G. On the Weibull-Gamma frailty model, its infinite moments, and its connection to generalized log-logistic, logistic, Cauchy, and extreme-value distributions. *Journal of Statistical Planning and Inference* 2011,**141**:861-868.
45. McGilchrist CA, Aisbett CW. Regression with frailty in survival analysis *Biometrics* 1991,**47**:461-466.

46. Fan JQ, Li RZ. Variable selection for Cox's proportional hazards model and frailty model. *Annals of Statistics* 2002,**30**:74-99.
47. Collins LM. Analysis of longitudinal data: the integration of theoretical model, temporal design, and statistical model. *Annu Rev Psychol* 2006,**57**:505-528.
48. Zeger SL, Liang KY. Longitudinal Data-Analysis for Discrete and Continuous Outcomes. *Biometrics* 1986,**42**:121-130.
49. Liang KY, Zeger SL. Longitudinal Data-Analysis Using Generalized Linear-Models. *Biometrika* 1986,**73**:13-22.
50. McCulloch CE. Maximum likelihood algorithms for generalized linear mixed models. *Journal of the American Statistical Association* 1997,**92**:162-170.
51. Fukuda J, Johnson KM. Mixed linear-non-linear inversion of crustal deformation data: Bayesian inference of model, weighting and regularization parameters. *Geophysical Journal International* 2010,**181**:1441-1458.
52. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics* 1982,**38**:963-974.
53. Jacqmin-Gadda H, Thiebaut R, Chene G, Commenges D. Analysis of left-censored longitudinal data with application to viral load in HIV infection. *Biostatistics* 2000,**1**:355-368.
54. Thiebaut R, Jacqmin-Gadda H, Leport C, Katlama C, Costagliola D, Le Moing V, *et al.* Bivariate longitudinal model for the analysis of the evolution of HIV RNA and CD4 cell count in HIV infection taking into account left censoring of HIV RNA measures. *J Biopharm Stat* 2003,**13**:271-282.
55. De Beaudrap P, Etard JF, Diouf A, Ndiaye I, Gueye NF, Gueye PM, *et al.* Modeling CD4+ cell count increase over a six-year period in HIV-1-infected patients on highly active antiretroviral therapy in Senegal. *Am J Trop Med Hyg* 2009,**80**:1047-1053.
56. Nothdurft A, Kublin E, Lappi J. A non-linear hierarchical mixed model to describe tree height growth. *European Journal of Forest Research* 2006,**125**:281-289.
57. Labbe L, Verotta D. A non-linear mixed effect dynamic model incorporating prior exposure and adherence to treatment to describe long-term therapy outcome in HIV-patients. *Journal of Pharmacokinetics and Pharmacodynamics* 2006,**33**:519-542.

58. Samson A, Lavielle M, Mentre F. The SAEM algorithm for group comparison tests in longitudinal data analysis based on non-linear mixed-effects model. *Statistics in Medicine* 2007,**26**:4860-4875.
59. Jennrich RI, Schluchter MD. Unbalanced repeated-measures models with structured covariance matrices. *Biometrics* 1986,**42**:805-820.
60. Jiang J. REML Estimation: Asymptotic Behavior and Related Topics. *The Annals of Statistics* 1996,**24**:255-286.
61. Thompson WA, Jr. The Problem of Negative Estimates of Variance Components. *The Annals of Mathematical Statistics* 1962,**33**:273-289.
62. McCullagh P, Nelder JA. *Generalized Linear Models*. London; 1989.
63. Booth JG, Hobert JP. Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo EM algorithm. *Journal of the Royal Statistical Society Series B-Statistical Methodology* 1999,**61**:265-285.
64. Song P XK, Fan YQ, Kalbfleisch JD. Maximization by parts in likelihood inference. *Journal of the American Statistical Association* 2005,**100**:1145-1158.
65. Preisser JS, Lohman KK, Rathouz PJ. Performance of weighted estimating equations for longitudinal binary data with drop-outs missing at random. *Stat Med* 2002,**21**:3035-3054.
66. Beunckens C, Sotto C, Molenberghs G. A simulation study comparing weighted estimating equations with multiple imputation based estimating equations for longitudinal binary data. *Computational Statistics & Data Analysis* 2008,**52**:1533-1548.
67. Lin CY, Barnhart HX, Kosinski AS. The weighted generalized estimating equations approach for the evaluation of medical diagnostic test at subunit level. *Biom J* 2006,**48**:758-771.
68. Bickel P, Diggle PJ, Fienberg S, Gather U, Olkin I, Zeger SL. Parameter Interpretation. In: *Models for Discrete Longitudinal Data*. Edited by Springer. New York: Springer Science+Business Media; 2005. pp. 298.
69. Wong MCM, Lam KF, Lo ECM. Analysis of multilevel grouped survival data with time-varying regression coefficients. *Statistics in Medicine* 2011,**30**:250-259.
70. Chen MH, Gustafson P. Bayesian methods in survival analysis. *Lifetime Data Analysis* 2011,**17**:1-2.

71. Wang QH, Dinse GE. Linear regression analysis of survival data with missing censoring indicators. *Lifetime Data Analysis* 2011,**17**:256-279.
72. Androulakis E, Koukouvinos C, Mylona K, Vonta F. A real survival analysis application via variable selection methods for Cox's proportional hazards model. *Journal of Applied Statistics* 2010,**37**:1399-1406.
73. Andersen PK. Recent developments in survival analysis. *Statistical Methods in Medical Research* 2010,**19**:3-4.
74. Lin DY, Wei LJ, Ying ZL. Accelerated failure time models for counting processes. *Biometrika* 1998,**85**:605-618.
75. Louis TA. Nonparametric analysis of an accelerated failure time model *Biometrika* 1981,**68**:381-390.
76. Robins J, Tsiatis AA. Semiparametric estimation of an accelerated failure time model with time-dependent. *Biometrika* 1992,**79**:311-319.
77. Wei LJ. The accelerated failure time model -A useful alternative to the cox regression-model in survival analysis. *Statistics in Medicine* 1992,**11**:1871-1879.
78. Cox DR. Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)* 1972,**34**:187-220.
79. Lin DY, Wei LJ. The robust inference for the Cox proportional hazards model. *Journal of the American Statistical Association* 1989,**84**:1074-1078.
80. Fleming TR, Harrington DP. *Counting Processes and Survival Analysis*. New-York; 1991.
81. Kalbfleisch JD, Prentice RL. *The statistical Analysis of Failure Time Data*. New-York; 1980.
82. Dafni UG, Tsiatis AA. Evaluating surrogate markers of clinical outcome when measured with error. *Biometrics* 1998,**54**:1445-1462.
83. Hsieh F, Tseng YK, Wang JL. Joint modeling of survival and longitudinal data: Likelihood approach revisited. *Biometrics* 2006,**62**:1037-1043.
84. Hu P, Tsiatis AA, Davidian M. Estimating the parameters in the Cox model when covariate variables are measured with error. *Biometrics* 1998,**54**:1407-1419.
85. Tsiatis AA, Davidian M. Joint modeling of longitudinal and time-to-event data: An overview. *Statistica Sinica* 2004,**14**:809-834.
86. Wulfsohn MS, Tsiatis AA. A joint model for survival and longitudinal data measured with error. *Biometrics* 1997,**53**:330-339.

87. Prentice RL. Covariate Measurement Errors and Parameter-Estimation in a Failure Time Regression-Model. *Biometrika* 1982,**69**:331-342.
88. De Gruttola V, Wulfsohn M, Fischl MA, Tsiatis A. Modeling the relationship between survival and CD4 lymphocytes in patients with AIDS and AIDS-related complex. *J Acquir Immune Defic Syndr* 1993,**6**:359-365.
89. Rizopoulos D, Verbeke G, Lesaffre E. Fully exponential Laplace approximations for the joint modelling of survival and longitudinal data. *Journal of the Royal Statistical Society Series B-Statistical Methodology* 2009,**71**:637-654.
90. Tsiatis A. Nonidentifiability Aspect of Problem of Competing Risks. *Proceedings of the National Academy of Sciences of the United States of America* 1975,**72**:20-22.
91. Lagakos SW, Williams JS. Models for Censored Survival Analysis - Cone Class of Variable-Sum Models. *Biometrika* 1978,**65**:181-189.
92. Gail M. Review and Critique of Some Models Used in Competing Risk Analysis. *Biometrics* 1975,**31**:209-222.
93. Peterson AV. Bounds for a Joint Distribution Function with Fixed Sub-Distribution Functions - Application to Competing Risks. *Proceedings of the National Academy of Sciences of the United States of America* 1976,**73**:11-13.
94. Klein JP, Moeschberger ML. Bounds on Net Survival Probabilities for Dependent Competing Risks. *Biometrics* 1988,**44**:529-538.
95. Williams JS, Lagakos SW. Models for Censored Survival Analysis - Constant-Sum and Variable-Sum Models. *Biometrika* 1977,**64**:215-224.
96. Slud EV, Rubinstein LV. Dependent Competing Risks and Summary Survival Curves. *Biometrika* 1983,**70**:643-649.
97. Link WA. A model for informative censoring. *Journal of the American Statistical Association* 1989,**84**:749-752.
98. Zheng M, Klein JP. Estimates of Marginal Survival for Dependent Competing Risks Based on an Assumed Copula. *Biometrika* 1995,**82**:127-138.
99. Scharfstein D, Robins JM, Eddings W, Rotnitzky A. Inference in randomized studies with informative censoring and discrete time-to-event endpoints. *Biometrics* 2001,**57**:404-413.
100. Scharfstein DO, Robins JM. Estimation of the failure time distribution in the presence of informative censoring. *Biometrika* 2002,**89**:617-634.

101. Hattori S, Kato M. Approximate subject-deletion influence diagnostics for Inverse Probability of Censoring Weighted (IPCW) method. *Statistics & Probability Letters* 2009,**79**:1833-1838.
102. Matsuyama Y, Yamaguchi T. Estimation of the marginal survival time in the presence of dependent competing risks using inverse probability of censoring weighted (IPCW) methods. *Pharmaceutical Statistics* 2008,**7**:202-214.
103. Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics* 2000,**56**:779-788.
104. Siannis F, Copas J, Lu GB. Sensitivity analysis for informative censoring in parametric survival models. *Biostatistics* 2005,**6**:77-91.
105. Molenberghs G, Kenward MG, Goetghebeur E. Sensitivity analysis for incomplete contingency tables: the Slovenian plebiscite case. *Journal of the Royal Statistical Society Series C-Applied Statistics* 2001,**50**:15-29.
106. Akaike H. Likelihood of a model and information criteria. *Journal of Econometrics* 1981,**16**:3-14.
107. Schwarz G. Estimation dimension of a model. *Annals of Statistics* 1978,**6**:461-464.
108. Rissanen J. Modeling by shortest data description. *Automatica* 1978,**14**:465-471.
109. Mallows C. Cp and prediction with many regressors: Response. *Technometrics* 1997,**39**:116-116.
110. Bozdogan H. *Model selection and Akaike Information Criterion (AIC)-the general-theory and its analytical extensions*. *Psychometrika* 1987,**52**:345-370.
111. Burnham KP, Anderson D. *Model selection and multi-model inference*. New York; 2002.
112. Barron A, Birge L, Massart P. Risk bounds for model selection via penalization. *Probability Theory and Related Fields* 1999,**113**:301-413.
113. Birge L, Massart P. Minimal penalties for Gaussian model selection. *Probability Theory and Related Fields* 2007,**138**:33-73.
114. Boufassa F, Saez-Cirion A, Lechenadec J, Zucman D, Avettand-Fenoel V, Venet A, *et al.* CD4 Dynamics over a 15 Year-Period among HIV Controllers Enrolled in the ANRS French Observatory. *Plos One* 2011,**6**:7.

115. Crum-Cianflone NF, Roediger M, Eberly LE, Ganesan A, Weintrob A, Johnson E, *et al.* Impact of Weight on Immune Cell Counts among HIV-Infected Persons. *Clinical and Vaccine Immunology* 2011,**18**:940-946.
116. Keller M, Lu Y, Lalonde RG, Klein MB. Impact of HIV-1 viral subtype on CD4(+) T-cell decline and clinical outcomes in antiretroviral naive patients receiving universal healthcare. *AIDS* 2009,**23**:731-737.
117. Lodi S, Phillips A, Touloumi G, Pantazis N, Bucher HC, Babiker A, *et al.* CD4 decline in seroconverter and seroprevalent individuals in the precombination of antiretroviral therapy era. *AIDS* 2010,**24**:2697-2704.
118. Nash D, Katyal M, Brinkhof MWG, Keiser O, May M, Hughes R, *et al.* Long-term immunologic response to antiretroviral therapy in low-income countries: a collaborative analysis of prospective studies. *AIDS* 2008,**22**:2291-2302.
119. Thiebaut R, Jacqmin-Gadda H, Babiker A, Commenges D. Joint modelling of bivariate longitudinal data with informative dropout and left-censoring, with application to the evolution of CD4+cell count and HIV RNA viral load in response to treatment of HIV infection. *Statistics in Medicine* 2005,**24**:65-82.
120. Toro PL, Katyal M, Carter RJ, Myer L, El-Sadr WM, Nash D, *et al.* Initiation of antiretroviral therapy among pregnant women in resource-limited countries: CD4(+) cell count response and program retention. *AIDS* 2010,**24**:515-524.
121. Binquet C, Chene G, Jacqmin-Gadda H, Journot V, Saves M, Lacoste D, *et al.* Modeling changes in CD4-positive T-lymphocyte counts after the start of highly active antiretroviral therapy and the relation with risk of opportunistic infections - The Aquitaine Cohort, 1996-1997. *American Journal of Epidemiology* 2001,**153**:386-393.
122. Bosch RJ, Wang R, Vaida F, Lederman MM, Albrecht MA. Changes in the slope of the CD4 cell count increase after initiation of potent antiretroviral treatment. *Jaids-Journal of Acquired Immune Deficiency Syndromes* 2006,**43**:433-435.
123. Mezzaroma I, Carlesimo M, Pinter E, Alario C, Sacco G, Muratori DS, *et al.* Long-term evaluation of T-cell subsets and T-cell function after HAART in advanced stage HIV-1 disease. *Aids* 1999,**13**:1187-1193.
124. Pakker NG, Notermans DW, de Boer RJ, Roos MTL, de Wolf F, Hill A, *et al.* Biphasic kinetics of peripheral blood T cells after triple combination therapy in HIV-1 infection: A composite of redistribution and proliferation. *Nature Medicine* 1998,**4**:208-214.

125. Powderly WG, Landay A, Lederman MM. Recovery of the immune system with antiretroviral therapy - The end of opportunism? *Jama-Journal of the American Medical Association* 1998,**280**:72-77.
126. Rizopoulos D, Verbeke G, Molenberghs G. Multiple-Imputation-Based Residuals and Diagnostic Plots for Joint Models of Longitudinal and Survival Outcomes. *Biometrics* 2010,**66**:20-29.
127. Wang Y, Taylor JMG. Jointly modeling longitudinal and event time data with application to acquired immunodeficiency syndrome. *Journal of the American Statistical Association* 2001,**96**:895-905.
128. Elie C, De Rycke Y, Jais JP, Marion-Gallols R, Landais P. Methodological and statistical aspects of equivalence and non inferiority trials. *Revue D Epidemiologie Et De Sante Publique* 2008,**56**:267-277.
129. Li G, Gao S. A group sequential type design for three-arm non-inferiority trials with binary endpoints. *Biometrical Journal* 2010,**52**:504-518.
130. Hida E, Tango T. On the three-arm non-inferiority trial including a placebo with a prespecified margin. *Statistics in Medicine* 2011,**30**:224-231.
131. Hill A, Sabin C. Designing and interpreting HIV noninferiority trials in naive and experienced patients. *AIDS* 2008,**22**:913-921.
132. Nie L, Soon G. A covariate-adjustment regression model approach to noninferiority margin definition. *Statistics in Medicine* 2010,**29**:1107-1113.
133. Vicaut E, Cucherat M. [Noninferiority trials: several simple principles]. *Presse Med* 2007,**36**:531-535.
134. COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE (CMPH). Guideline on the choice of the non-inferiority margin In: European Medicines Agency (EMA); 2005.
135. D'Agostino RB, Sr., Massaro JM, Sullivan LM. Non-inferiority trials: design concepts and issues - the encounters of academic consultants in statistics. *Stat Med* 2003,**22**:169-186.
136. Jones B, Jarvis P, Lewis JA, Ebbutt AF. Trials to assess equivalence: the importance of rigorous methods. *BMJ* 1996,**313**:36-39.
137. Matsuyama Y. A comparison of the results of intent-to-treat, per-protocol, and g-estimation in the presence of non-random treatment changes in a time-to-event non-inferiority trial. *Statistics in Medicine* 2010,**29**:2107-2116.

138. Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA* 2006,**295**:1152-1160.
139. Yoo B. Impact of missing data on type 1 error rates in non-inferiority trials. *Pharmaceutical Statistics* 2010,**9**:87-99.
140. Hilton JF. Noninferiority trial designs for odds ratios and risk differences. *Statistics in Medicine* 2010,**29**:982-993.
141. Neyman J, Pearson ES. On the Problem of the Most Efficient Tests of Statistical Hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 1933,**231**:289-337.
142. Efron B, Tibshirani R. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science* 1986,**1**:54-75.
143. Bickel PJ, Freedman DA. Some Asymptotic Theory for the Bootstrap. *The Annals of Statistics* 1981,**9**:1196-1217.
144. Jason CH, Hwang JTG, Liu H-K, Ruberg SJ. Confidence Intervals Associated with Tests for Bioequivalence. *Biometrika* 1994,**81**:103-114.
145. Committee for Proprietary Medicinal Products (CMPM). Point to Consider on switching between superiority and non-inferiority. In. Edited by (EMEA). EMA; 2000.
146. Sanne I, Orrell C, Fox MP, Conradie F, Ive P, Zeinecker J, *et al.* Nurse versus doctor management of HIV-infected patients receiving antiretroviral therapy (CIPRA-SA): a randomised non-inferiority trial. *Lancet* 2010,**376**:33-40.
147. Rubin DB. Inference and missing data. *Biometrika* 1976,**63**:581-592.
148. Rubin DB. *Multiple imputation for nonreponse in surveys* New York; 1987.
149. Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. New York,; 2002.
150. Diggle PJ, Liang KY, Zeger SL. Analysis of Longitudinal Data. *Oxford Univ. Press*. 1996.
151. Little RJA. Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association* 1995,**90**:1112-1121.
152. Little RJA. Patern-mixture models for multivariate incomplete data *Journal of the American Statistical Association* 1993,**88**:125-134.

153. Molenberghs G, Kenward MG, Lesaffre E. The analysis of longitudinal ordinal data with nonrandom drop-out. *Biometrika* 1997,**84**:33-44.
154. Wu MC, Carroll RJ. Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics* 1988,**44**:175-188.
155. Harel O, Zhou XH. Multiple imputation: review of theory, implementation and software. *Stat Med* 2007,**26**:3057-3077.
156. Zhang P. Multiple Imputation: Theory and Method. *International Statistical Review* 2003,**71**:581-592.
157. Schafer JL. *Analysis of Incomplete Multivariate Data*. London; 1997.
158. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research* 2007,**16**:219-242.
159. Mackinnon A. The use and reporting of multiple imputation in medical research - a review. *J Intern Med* 2010.
160. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, *et al*. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009,**338**:b2393.
161. White IR, Royston P. Imputing missing covariate values for the Cox model. *Stat Med* 2009,**28**:1982-1998.
162. van Buuren S, Oudshoorn KG. MICE: Multivariate Imputation by Chained Equation in R. *journal of statistical software, forthcoming* 2010.
163. Royston P. Multiple imputation of missing values. *Stata J* 2004,**4**:227-231.
164. de Boor C. *A practical guide to spline*. New-York; 1978.
165. Fuller WA. Grafted polynomials as approximating functions. *Australian J. of Agr. Econ.* 1969,**13**:35-46.
166. Buse A, Lim L. Cubic Splines as a Special Case of Restricted Least Squares. *JASA* 1977,**72**:64-68.

Annexe A: article « Hepatotoxicity and effectiveness of a Nevirapine-based antiretroviral therapy in HIV-infected patients with or without viral hepatitis B or C infection in Cameroon »

RESEARCH ARTICLE

Open Access

Hepatotoxicity and effectiveness of a Nevirapine-based antiretroviral therapy in HIV-infected patients with or without viral hepatitis B or C infection in Cameroon

Jules B Tchatchueng Mbougua^{1,2}, Christian Laurent^{1*}, Charles Kouanfack³, Anke Bourgeois^{1,4}, Laura Ciaffi⁵, Alexandra Calmy⁵, Henri Gwet², Sinata Koulla-Shiro³, Jacques Ducos⁶, Eitel Mpoudi-Ngolé⁷, Nicolas Molinari^{8,9}, Eric Delaporte^{1,4}

Abstract

Background: Coinfection with hepatitis B virus (HBV) or hepatitis C virus (HCV) in HIV-infected patients receiving a commonly used nevirapine-based antiretroviral therapy is a major concern for African clinicians owing to its high prevalence, the infrequent testing and treatment of viral hepatitis, and the impact of liver disease on the tolerability and effectiveness of anti-HIV treatment. We compared the hepatotoxicity and the immunological, virological and clinical effectiveness of a nevirapine-based antiretroviral therapy between patients infected with HIV only and patients coinfecting with hepatitis B or C virus in Cameroon.

Methods: A retrospective cohort study was conducted among HIV-1-infected patients. Plasma HBV DNA and HCV RNA were tested in positive or indeterminate samples for HBsAg or HCV antibodies, respectively. All patients received nevirapine and lamivudine plus stavudine or zidovudine.

Results: Of 169 HIV-1-infected patients with a median baseline CD4 count of 135 cells/mm³ (interquartile range [IQR] 67-218), 21% were coinfecting with HBV or HCV. In coinfecting patients, the median viral load was 2.47×10^7 IU/mL for HBV (IQR $3680-1.59 \times 10^8$) and 928 000 IU/mL for HCV (IQR 178 400- 2.06×10^6). Multivariate analyses showed that the risk of hepatotoxicity was 2-fold higher in coinfecting patients ($p < 0.01$). The response to antiretroviral therapy was however comparable between mono-infected and coinfecting patients in terms of CD4 cell count increase ($p = 0.8$), HIV-1 viral load below 400 copies/mL ($p = 0.9$), death ($p = 0.3$) and death or new AIDS-defining event ($p = 0.1$). Nevirapine was replaced by a protease inhibitor in 4 patients owing to hepatotoxicity.

Conclusion: This study suggests that the nevirapine-based antiretroviral therapy could be used safely as first-line treatment in patients with low CD4 cell count in Africa despite frequent coinfections with HBV or HCV and infrequent testing of these infections. Although testing for HBV and HCV should be systematically performed before initiating antiretroviral therapy, transaminases elevations at baseline or during treatment should be a decisive argument for testing when hepatitis status is unknown.

* Correspondence: Christian.Laurent@mpl.ird.fr

¹Institut de Recherche pour le Développement, University Montpellier 1, UMR 145, Montpellier, France

Background

Coinfection with hepatitis B virus (HBV) or hepatitis C virus (HCV) in HIV-infected patients receiving a nevirapine-based antiretroviral therapy is a major concern for African clinicians owing to its high prevalence, the infrequent testing and treatment of viral hepatitis, and the impact of liver disease on the tolerability and effectiveness of anti-HIV treatment [1]. While HIV-related morbidity and mortality are decreasing thanks to the scaling-up of antiretroviral therapy, the impact of liver disease is likely to increase in Africa. In addition, initiation of antiretroviral therapy has been suggested in all patients coinfecting with HIV and hepatitis B or C virus irrespective of the CD4 cell count.

Nevirapine, the non nucleosidic reverse transcriptase inhibitor (NNRTI) most often used in first-line treatment in combination with two nucleosidic reverse transcriptase inhibitors, is associated with early hypersensitivity reactions (generally in the first 12 weeks after treatment initiation) which can cause fulminant hepatitis leading to hepatic failure and death, and with later onset of direct drug-related hepatotoxicity leading to liver enzymes elevations [2]. The risk of hepatotoxicity is increased in patients coinfecting with HBV or HCV [3-5]. The World Health Organization (WHO) therefore recommends to use nevirapine with caution and regular monitoring in patients who have baseline grade 1, 2 or 3 elevations of liver enzymes and positive or unknown HBV or HCV testing [6]. In addition, nevirapine is not recommended in patients with a grade 4 elevation of liver enzymes.

The impact of HBV or HCV coinfection on the effectiveness of a nevirapine-based antiretroviral therapy remains unclear owing to the lack of specific studies especially in the African context. A recent South African study in HBV-coinfecting patients receiving the less hepatotoxic efavirenz (the second NNRTI used in first-line treatment) found similar response to antiretroviral therapy between monoinfected and coinfecting patients despite higher hepatotoxicity in the latter [7]. Regardless of antiretroviral regimens (often not reported), studies on HBV or HCV coinfection in Western and Asian countries provided conflicting results with respect to CD4 cell increase, HIV suppression, AIDS progression and mortality [8-21].

We therefore compared the hepatotoxicity and the immunological, virological and clinical effectiveness of a nevirapine-based antiretroviral therapy between patients infected with HIV only and patients coinfecting with HIV and hepatitis B or C virus in Cameroon [22].

Methods

Study design

A retrospective cohort study was conducted in two major hospitals (the Military Hospital and the Central Hospital) in Yaoundé, the capital of Cameroon, among HIV-1-infected patients enrolled from 2001 to 2003 in two clinical research projects designed to assess antiretroviral treatments [23-25]. The National Ethics Committee of Cameroon approved the study protocols, and the patients gave their written informed consent. The eligibility criteria, follow-up methods, medical and social staff, and coordinators were similar in the two projects. Briefly, patients over 18 years were eligible if they had confirmed HIV-1 infection, AIDS or a CD4 count below 350 cells/mm³, a Karnofsky score over 50%, and no contraindications to antiretroviral treatment, including serum liver enzyme levels less than five times the upper limit of normal (ULN) in the first project or less than three times the ULN in the second project. All patients received nevirapine and lamivudine plus stavudine or zidovudine. Tenofovir was not available. Efavirenz was temporarily substituted for nevirapine in case of concomitant tuberculosis therapy. Care (visits and laboratory exams) and drugs (antiretrovirals, and preventive and curative treatments for opportunistic infections) were provided free of charge.

Clinical and laboratory procedures

Hepatitis B and C markers were assessed on baseline blood samples frozen at -80°C. Enzyme immunoassays (EIA) were used to detect hepatitis B surface antigens (HBsAg; Monolisa Ag HBs Plus, Bio-rad, Marnes la coquette, France) and antibodies to hepatitis C virus (anti-HCV; Ortho HCV EIA 3.0, Ortho-clinical Diagnostics, Riratan, NJ, USA). Plasma HBV DNA and HCV RNA were tested in positive or indeterminate samples for HBsAg or anti-HCV, respectively, using the Cobas Ampliprep/Cobas TaqMan quantitative assay (Roche Diagnostics GmbH, Mannheim, Germany; quantification range of 12 to 2.2 × 10⁸ IU/mL for HBV and 15 to 6.9 × 10⁷ IU/mL for HCV).

The CD4 cell counts were measured with a FACS-Count device (Becton Dickinson, Mountain View, California, USA) at baseline and then every 6 months. The HIV-1 RNA load was determined using the Roche Amplicor HIV-1 Monitor assay (Roche Molecular Systems, Branchburg, New Jersey, USA) or the Bayer bDNA HIV-1 Quantiplex assay (Bayer Diagnostics, Emeryville, California, USA) at baseline, months 3 and 6, and then every 6 months. Serum liver enzymes (alanine aminotransferase [ALT] and aspartate aminotransferase [AST]) levels were assayed at baseline, weeks 2, 4

and 6, and then months 2, 3, 6, 9, 12, 18 and 24. Hepatotoxicity was graded with the French National Agency for Research on AIDS and Viral Hepatitis (ANRS) toxicity scale in which an ALT or AST level of 1.25-2.5 times the ULN defines grade 1, >2.5-5 times the ULN defines grade 2, >5-10 times the ULN defines grade 3 and >10 times the ULN defines grade 4 <http://www.anrs.fr/content/download/2242/12805/file/ANRS-GradeEI-V1-En-2008.pdf>. At each time point, hepatotoxicity grade was defined on the basis of the higher value of either ALT or AST. The HIV disease stage was determined according to the 1993 Centers for Disease Control and Prevention (CDC) classification.

Statistical analysis

Patients were classified as coinfecting if they had positive HBV DNA or HCV RNA. The analyses were based on an intention-to-treat approach. Data were censored at the time of each patient's month 24 visit or, if follow-up was shorter, at the time of the last visit or death. The χ^2 test and, when sample sizes were too small, Fisher's exact test were used to compare the distribution of categorical variables between the infection groups (HBV or HCV-coinfecting patients *versus* HIV-monoinfecting patients). For continuous variables, comparisons were based on the non parametric Mann-Whitney two-sample test.

Hepatotoxicity was first assessed by using two end points: 1) occurrence of at least one episode of grade 2 or higher hepatotoxicity, and 2) occurrence of at least one episode of grade 3 or 4 hepatotoxicity. Kaplan-Meier curves were plotted and differences between monoinfecting and coinfecting patients were checked for significance by the log-rank test. Cox proportional hazard models were used to compare hepatotoxicity between groups adjusted on baseline covariates. The proportional hazards assumption was evaluated by both graphic and statistical (based on Schoenfeld residuals) methods. A multivariate Poisson regression was then used to compare between groups the highest grade of hepatotoxicity reached by each patient during follow-up. The evolution of the CD4 cell count during follow-up was estimated using a mixed-effect linear regression model. Taking into account the biphasic evolution of the CD4 cell count during antiretroviral therapy [26], we allowed the slope to change after 6 months and included a quadratic effect in the second phase. In order to compare the CD4 cell counts evolution in monoinfecting and coinfecting patients, terms of interaction between infection status and each phase's slope or quadratic effect were tested. The CD4 cell counts were square root-transformed to approximate a normal distribution. The time to reach a viral load below 400 copies/mL, death, and death or new AIDS-defining event from

treatment initiation were compared between groups by using multivariate Cox models.

Multivariate analyses were adjusted on the following baseline covariates: gender (women *versus* men), age (≥ 42 *versus* <42 years), time since diagnosis of HIV infection (<24 *versus* ≥ 24 months), body mass index (≥ 21 *versus* <21 kg/m²), Karnofsky score ($<100\%$ *versus* 100%), CDC clinical stage, CD4 cell count (≥ 100 *versus* <100 cells/mm³), HIV-1 viral load (<5.0 *versus* ≥ 5.0 log₁₀ copies/mL), hemoglobin level (≥ 10 *versus* <10 g/dL), total lymphocytes count (≥ 1700 *versus* <1700 cells/mm³), liver enzymes levels (≥ 1.25 *versus* $<1.25 \times$ ULN), antiretroviral therapy and cotrimoxazole prophylaxis. Independent covariates associated with outcomes with a conservative *p* value of < 0.25 in univariate analysis were subsequently tested in multivariate analysis. A backward elimination procedure was used to determine the final model containing only the infection group, together with significant covariates and potential confounders. The incidence rates of hepatotoxicity, deaths, and deaths or new AIDS-defining events were expressed as the number of patients with at least one episode of the given event per 100 person-years of follow-up. For the incidence rates as for the survival analyses, data were censored at the time of the first episode of the given event. All analyses were performed using STATA 10.01 software (STATA Corporation, college Station, TX, USA).

Results

Patients

Of 169 HIV-1-infected patients enrolled between January 2001 and April 2003, 35 (21%) were coinfecting with HBV ($n = 14$) or HCV ($n = 21$). The baseline patient's characteristics are shown in table 1. Two thirds of patients were women. Most patients were at an advanced HIV clinical stage (42% were at the CDC stage B and 44% were at stage C). The median CD4 count was 135 cells/mm³ (interquartile range [IQR] 67-218). Fifteen women (13%; 12 monoinfecting with HIV and 3 coinfecting with HBV or HCV) had a CD4 count above 250 cells/mm³. No man had a CD4 count above 400 cells/mm³. Most characteristics were similar between HIV-monoinfecting patients and those coinfecting with an hepatitis virus. However, coinfecting patients were older ($p < 0.001$), and had higher liver enzyme levels ($p < 0.001$). The median viral load was 2.47×10^7 IU/mL for HBV (IQR 3680- 1.59×10^8 ; range 270 - $>2.2 \times 10^8$) and 928 000 IU/mL for HCV (IQR 178 400- 2.06×10^6 ; range 640- 5.5×10^6) in patients coinfecting with the respective virus. The median time of follow-up was 23.9 months (IQR 17.2-24.0) in monoinfecting patients and 24.0 months (IQR 20.0-24.0) in coinfecting patients. Half the patients received zidovudine, lamivudine and nevirapine at baseline, and the other half received

Table 1 Baseline characteristics of patients by infection group

	HIV monoinfected patients (n = 134)	HBV or HCV coinfecting patients (n = 35)	p
Women (no.)	91 (68%)	22 (63%)	0.6
Age (years)			
Median (IQR)	34.5 (28.4-39.7)	41.6 (33.7-48.8)	<0.001
<42 (no.)	113 (84%)	19 (54%)	<0.001
Time since diagnosis of HIV seropositivity (months)			
Median (IQR)	29.4 (13.5-53.4)	22.1 (8.3-43.6)	0.2
<24 (no.)	56 (42%)	18 (51%)	0.3
Body weight (Kg) [median (IQR)]	65 (55-70)	63 (54-70)	0.7
Body mass index (Kg/m ²)*			
median (IQR)	23.2 (21.6-24.7)	22.2 (20.6-25.6)	0.2
≥ 21 (no.)	108 (82%)	25 (74%)	0.3
Karnofsky score >90% (no.)	72 (54%)	15 (43%)	0.3
CDC clinical stage (no.)			0.7
A	19 (14%)	4 (11%)	
B	58 (43%)	13 (37%)	
C	57 (43%)	18 (51%)	
CD4 cell count (/mm ³)			
Median (IQR)	138 (67-222)	135 (68-216)	0.7
<100 (no.)	80 (60%)	22 (63%)	0.7
HIV-1 viral load (log ₁₀ copies/mL) [†]			
Median (IQR)	5.3 (4.7-5.5)	5.3 (4.8-5.7)	0.5
<5.0 (no.)	50 (37%)	11 (31%)	0.5
Hemoglobin (g/dL)			
Median (IQR)	11.0 (9.8-12.1)	11.1 (9.8-12.9)	0.6
<10 (no.)	36 (27%)	10 (29%)	0.8
Total lymphocyte count (/mm ³)			
Median (IQR)	1578 (1080-2100)	1408 (1000-1900)	0.4
<1700 (no.)	77 (57%)	23 (66%)	0.4
ALT level (xULN)			
Median (IQR)	0.6 (0.4-0.7)	0.8 (0.6-1.2)	<0.001
≥ 1.25 (no.)	8 (6%)	8 (30%)	0.002
ASAT (xULN) [‡]			
Median (IQR)	0.8 (0.6-1.1)	1.3 (0.9-1.7)	<0.001
≥ 1.25 (no.)	21 (17%)	17 (52%)	<0.001
History of antiretroviral treatment (no.)			0.6
None	131 (98%)	34 (97%)	
HAART	1 (1%)	1 (3%)	
Nevirapine (PMTCT)	2 (2%)	0 -	
Baseline antiretroviral treatment (no.)			0.3
Zidovudine+lamivudine+nevirapine	70 (52%)	15 (43%)	
Stavudine+lamivudine+nevirapine	64 (48%)	20 (57%)	
Cotrimoxazole prophylaxis (no.)	124 (93%)	35 (100%)	0.1

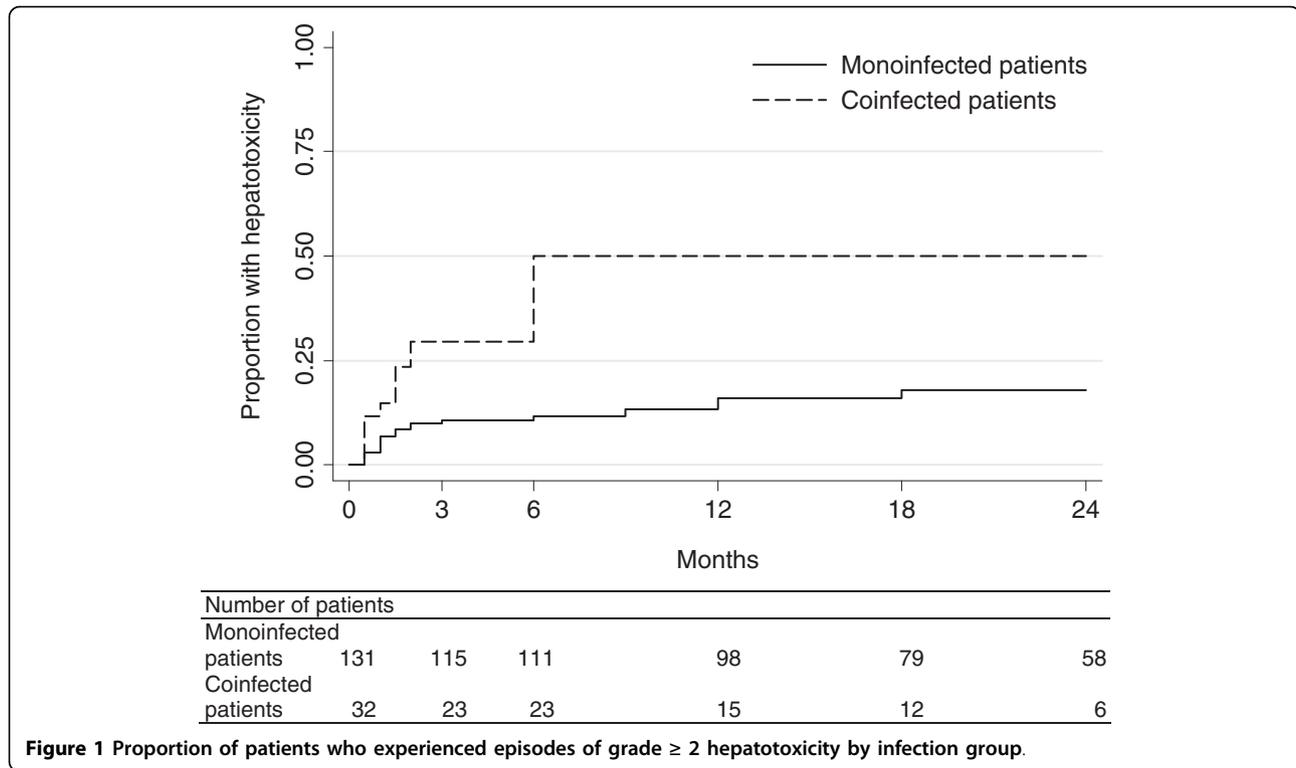
ALT, alanine aminotransferase; AST, aspartate aminotransferase; ULN, upper limit of normal; PMTCT, prevention of mother-to-child transmission; HAART, highly active antiretroviral therapy

* 3 missing values; [†] 1 missing value; [‡] 9 missing values

stavudine, lamivudine and nevirapine. Nevirapine was interrupted during follow-up in 16 monoinfected patients (12%) and 4 coinfecting patients (11%) because of tuberculosis (n = 9 and 2, respectively), adverse effects (n = 4 and 2, respectively) and antiretroviral drug resistance (n = 3 and 0, respectively).

Hepatotoxicity

A total of 1588 measures of liver enzymes were available. The number of measures per patient was similar in both infection groups (median 10, IQR 9-11, p = 0.2). Two patients only (both were coinfecting) had a grade 2 hepatotoxicity at baseline (90 IU/L and 97 IU/L,



respectively); at week 2, hepatotoxicity dropped to grade 1 in one patient and remained at grade 2 in the other patient. No patient had a baseline grade 3 or 4 hepatotoxicity.

During follow-up, 39 patients experienced at least one episode of grade 2 or higher hepatotoxicity (22 monoinfected patients [16%] and 17 coinfecting patients [49%]). The corresponding incidence rate was 12.4 per 100 person-years (95% confidence interval [CI] 8.2-18.9) in monoinfected patients and 54.0 per 100 person-years (CI 33.6-89.5) in coinfecting patients. The Kaplan-Meier analysis showed the higher risk in coinfecting patients (p

< 0.001 ; figure 1). Of note, all episodes of hepatotoxicity in the coinfecting patients occurred in the first six months. In monoinfected patients, such episodes occurred throughout the follow-up (although more frequently in the first months). In multivariate Cox analysis, the risk of grade ≥ 2 hepatotoxicity remained higher in coinfecting patients (hazard ratio [HR] 2.94, CI 1.49-5.81, $p = 0.002$; table 2) after adjustment on age (HR 0.90, CI 0.43-1.91, $p = 0.8$), Karnofsky score (HR 0.41, CI 0.20-0.84, $p = 0.01$) and liver enzymes levels (HR 1.45, CI 0.74-2.82, $p = 0.3$). Of the 17 coinfecting patients who experienced a grade ≥ 2 hepatotoxicity, 4

Table 2 Adjusted risks of outcomes associated with the hepatitis coinfection*

Outcome	Ratio or coefficient	95% confidence interval	p
Hepatotoxicity			
Grade $\geq 2^\dagger$	2.94	1.49-5.81	0.002
Grade $\geq 3^\dagger$	2.18	0.61-7.75	0.2
Highest grade [‡]	1.83	1.28-2.60	0.001
CD4 cell count increase			
Primary slope	0.02	-0.20 to 0.24	0.9
Secondary slope [¶]	-0.04	-0.38 to 0.31	0.8
Quadratic effect [¶]	0.00	-0.00 to 0.01	0.8
Viral load < 400 copies/mL [†]	1.00	0.68-1.47	0.9
Death [†]	0.52	0.14-1.93	0.3
Death or new AIDS-defining event [†]	0.37	0.11-1.29	0.1

* Coinfecting patients versus monoinfected patients. [†] Cox regression. [‡] Poisson regression. ^{||} Mixed-effect linear regression of the square root-transformed CD4 cell count. [¶] From 6 months to 24 months, combined with the primary slope.

were positive for HBV (median viral load 1.80×10^8 IU/mL, IQR 1.04×10^8 - 2.10×10^8) and 13 were positive for HCV (median viral load 970 000 IU/mL, IQR 173 800- 2.06×10^6).

Only 12 patients (7%) experienced a grade 3 or 4 hepatotoxicity (7 monoinfected patients [5%] and 5 coinfecting patients [14%]). The incidence rates were 3.6 (CI 1.7-7.6) and 9.9 per 100 person-years (CI 4.1-23.7) in monoinfected and coinfecting groups, respectively. The risk tended to be higher in coinfecting patients but the difference was not significant (HR 2.18, CI 0.61-7.75, $p = 0.2$; table 2) after adjustment on age (HR 1.28, CI 0.36-4.53, $p = 0.8$) and liver enzymes levels (HR 1.61, CI 0.47-5.46, $p = 0.8$). All 5 coinfecting patients who experienced a grade 3 or 4 hepatotoxicity were HCV positive; their median viral load was 970 000 IU/mL (IQR 320 000- 1.22×10^6).

The higher risk for hepatotoxicity in coinfecting patients was confirmed by the multivariate Poisson regression analysis (incidence risk ratio [IRR] 1.83, CI 1.28-2.60, $p = 0.001$; table 2) after adjustment on age (IRR 0.83, CI 0.56-1.24, $p = 0.4$) and Karnofsky score (IRR 0.54, CI 0.39-0.76, $p < 0.001$).

Three patients developed a rash in the first four weeks (two women with baseline CD4 counts of 31 and 307 cells/mm³, respectively, and one man with a baseline CD4 count of 18 cells/mm³). Only the male patient was coinfecting (HBV viral load 4.88×10^7 IU/mL). Nevirapine was replaced by a protease inhibitor in all three patients. The woman with a high CD4 cell count experienced simultaneously a grade 3 hepatotoxicity. Nevirapine was also replaced by a protease inhibitor in one other (monoinfected) patient with an isolated grade 3 hepatotoxicity. No patient had clinical hepatitis.

Immunological response

The number of CD4 cell counts per patient was comparable between the monoinfected group (median 5, IQR 3-5) and the coinfecting group (median 4, IQR 4-5; $p = 0.5$). As shown in figure 2A, the CD4 cell count increased after treatment initiation from 143 to 325 cells/mm³ in monoinfected patients and from 136 to 297 cells/mm³ in coinfecting patients. The square root CD4 cell count increase was not significantly different between monoinfected and coinfecting patients (either in the first 6 months [$p = 0.9$] and thereafter [$p = 0.8$]; table 2) after adjustment on gender (coefficient 1.59, CI 0.41 to 2.77, $p = 0.008$), age (coefficient -1.07, CI -2.43 to 0.29, $p = 0.1$), HIV-1 viral load (coefficient 1.55, CI 0.40 to 2.70, $p = 0.008$), hemoglobin level (coefficient 2.43, CI 1.18 to 3.67, $p < 0.001$) and total lymphocyte count (coefficient 2.93, CI 1.83 to 4.03, $p < 0.001$).

Virological response

The number of HIV-1 RNA measures per patient was similar in both infection groups (median 5, IQR 4-6, $p = 0.8$). The proportion of patients with viral load below 400 copies/mL was 91% and 85% after 3 months and 91% and 78% after 24 months in monoinfected and coinfecting patients, respectively (figure 2B). The time for reaching a viral load below 400 copies/mL did not differ between the groups (HR 1.00, CI 0.68-1.47, $p = 0.9$; table 2). No other variable remained in the model.

Deaths or new AIDS-defining events

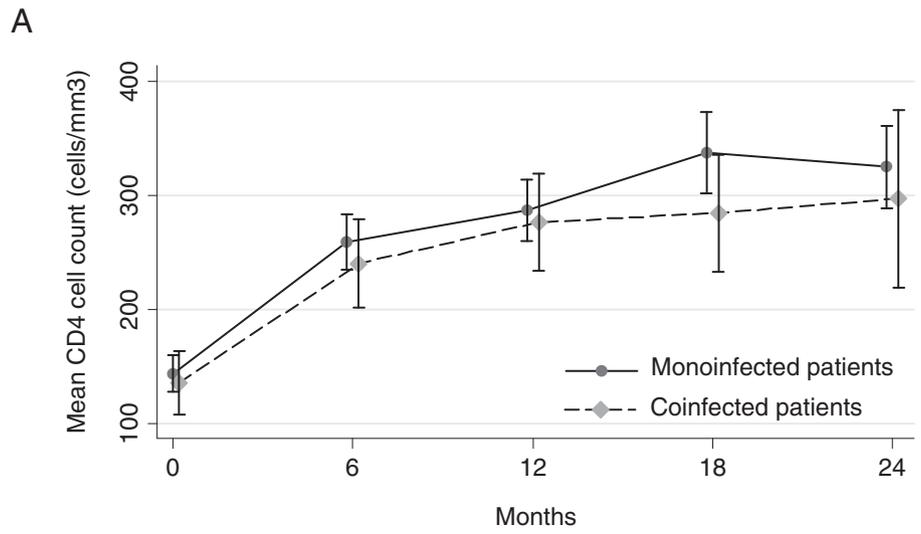
There were 16 deaths (12%) in monoinfected patients and 3 (9%) in those coinfecting. The mortality rates were 7.2 (CI 4.4-11.7) and 4.8 (CI 1.4-15.0) per 100 person-years, respectively. The Kaplan-Meier curves did not differ significantly between the groups ($p = 0.6$, figure 3A). In multivariate analysis, death was not associated with coinfection (HR 0.52, CI 0.14-1.93, $p = 0.3$; table 2) after adjustment on gender (HR 0.29, CI 0.11-0.77, $p = 0.01$), age (HR 1.90, CI 0.68-5.29, $p = 0.2$), CD4 cell count (HR 0.15, CI 0.04-0.55, $p = 0.004$) and hemoglobin level (HR 0.28, CI 0.11-0.74, $p = 0.01$). In the monoinfected group, the deaths were related to advanced HIV disease ($n = 5$), poor general health ($n = 5$), multifocal tuberculosis, pulmonary infection, wasting, malaria, pancreatitis and hepatic carcinoma ($n = 1$ each). In the coinfecting group, death was related to advanced HIV disease, poor general health and persistent fever of unknown origin ($n = 1$ each); the first two patients had taken traditional medicines.

Deaths or new AIDS-defining events occurred in 25 monoinfected patients (19%) and 5 coinfecting patients (14%). The respective incidence rates were 10.6 (CI 7.0-16.1) and 6.7 per 100 person-years (CI 2.5-17.8). The Kaplan-Meier curves showed no significant difference between the groups ($p = 0.5$, figure 3B). In multivariate analysis, the occurrence of deaths or new AIDS-defining events was not associated with coinfection (HR 0.37, CI 0.11-1.29, $p = 0.1$; table 2) after adjustment on age (HR 1.33, CI 0.53-3.32, $p = 0.5$) and body mass index (HR 4.40, CI 1.98-9.78, $p < 0.001$).

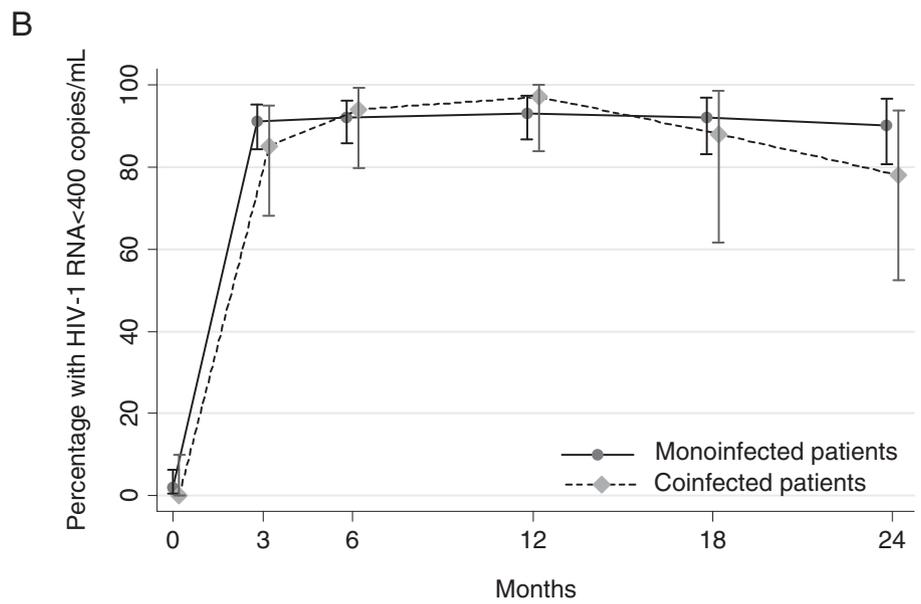
Discussion

This study in Cameroon showed a higher hepatotoxicity in HBV or HCV-coinfecting patients receiving a nevirapine-based antiretroviral therapy than in their HIV-monoinfected homologues. However, this adverse event did not impact negatively on the effectiveness of treatment.

Altogether, our analyses showed that the risk of hepatotoxicity was 2-fold higher in hepatitis B or C coinfecting patients. This finding is consistent with previous reports [11,17,27,28], especially with a study in patients

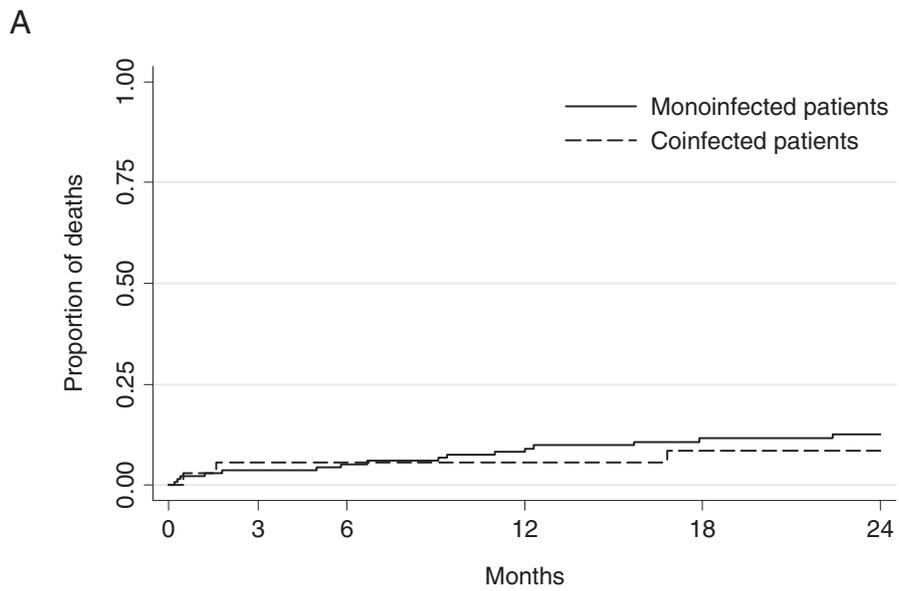


Number of patients						
Monoinfected patients	134	123	113	93	72	
Coinfected patients	35	32	32	25	21	

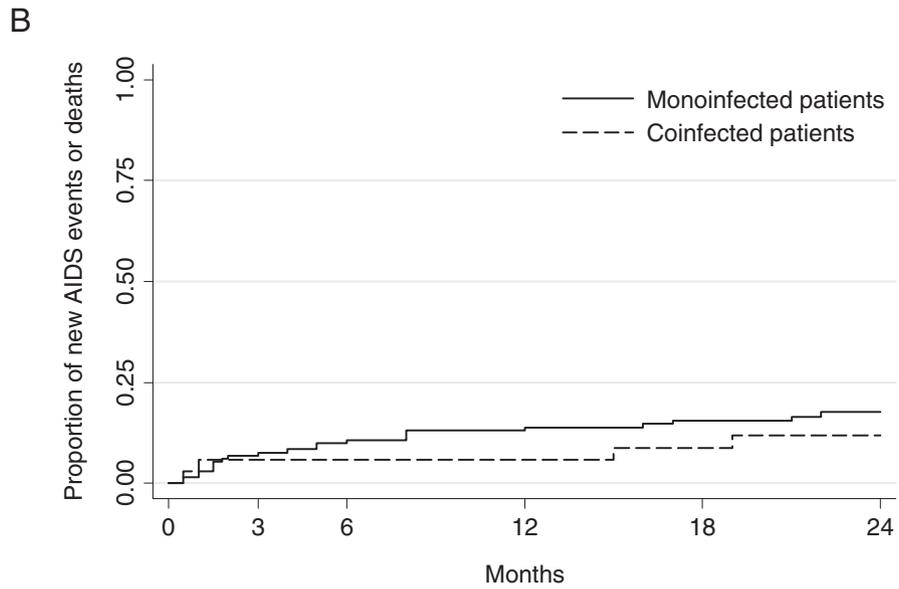


Number of patients						
Monoinfected patients	133	128	125	104	73	64
Coinfected patients	35	33	33	32	16	18

Figure 2 Immunological and virological responses to antiretroviral therapy by infection group: mean CD4 cell count (A), and percentage of patients with plasma HIV-1 viral load below 400 copies/mL (B). Bars indicate 95% confidence intervals.



Number of patients						
Monoinfected patients	134	129	126	115	96	63
Coinfecting patients	35	33	33	33	30	18



Number of patients						
Monoinfected patients	131	122	117	106	90	59
Coinfecting patients	34	32	32	32	29	18

Figure 3 Clinical progression by infection group: progression to death (A), and progression to death or new AIDS-defining event (B).

receiving a nevirapine-based treatment in United States [29]. Of note, HCV-coinfected patients accounted for 76% of overall coinfecting patients with liver enzymes elevations >2.5 times the ULN (ALT and/or AST levels) during treatment and 100% of those with liver enzymes elevations >5 times the ULN (as compared with 60% of the study's coinfecting group). It should however be noted that HCV infection was not treated while HBV infection was treated with lamivudine monotherapy. Although anti-HBV lamivudine monotherapy has been shown to lead to frequent emergence of drug resistance [30] and, consequently, to possible acute hepatitis, fulminant hepatic failure, and death [3], lamivudine was likely to be effective against HBV in the first months of treatment when most episodes of hepatotoxicity occurred. Importantly, up to half the coinfecting patients experienced liver enzymes elevations >2.5 times the ULN.

Overall, the hepatic tolerability of the nevirapine-based antiretroviral therapy was as expected. Seven percents of patients experienced liver enzymes elevations >5 times the ULN. In a comprehensive analysis of 17 randomized clinical trials of nevirapine, the rate was 10% after one year of treatment [31]. A higher rate of 17% has been found in South African patients who had a mean baseline CD4 count of 398 cells/mm³ [32]. Our lower rate should be explained by the profound immunodeficiency of our patients at the time of treatment initiation (median CD4 count of 135 cells/mm³). Indeed, transaminases elevations >5 times the ULN in patients receiving nevirapine has been associated with sex-dependant CD4 cell count [5]. Thus, nevirapine should be avoided in women with a CD4 count above 250 cells/mm³ and in men with a CD4 count above 400 cells/mm³. In our study, only 13% of women and no man had a CD4 cell count above these respective cut-offs. Most transaminases elevations in our patients were asymptomatic. Only 2% of patients had a rash following treatment initiation. In their analysis, Dieterich *et al* found a rash in 2.2% of patients [31]. On the other hand, the incidence rates of liver enzymes elevations >5 times the ULN in our mono-infected and coinfecting patients were comparable to those reported in South African patients receiving an efavirenz-based antiretroviral therapy [7].

The immunological response to the nevirapine-based antiretroviral therapy was comparable between coinfecting and mono-infected patients, in accordance with several studies irrespective of the treatment prescribed [7,11,13,14,18,20]. In contrast, other studies found delayed CD4 cell count recovery in coinfecting patients especially in those coinfecting with HCV [8-10,21]. However, this lower immunological response was not sustained beyond the first few months in the Thai and Swiss studies [10,33].

The virological response also was comparable between patients coinfecting with an hepatitis virus and patients infected with HIV only. Increased risk for virological failure has been reported in HBV-coinfected patients in Taiwan [11]. However, most other studies like ours did not find any difference [7,9,10,17,21].

The clinical progression did not differ significantly between coinfecting and mono-infected patients but treatment was initiated late (median CD4 count of 135 cells/mm³). Surprisingly, deaths or new AIDS-defining events were less common in coinfecting patients. A possible explanation could be a close medical supervision of coinfecting patients owing to frequent liver enzymes elevations. Higher rates of deaths (from any cause, liver-related causes and/or AIDS-related causes) and/or new AIDS-defining events in patients coinfecting with HBV or HCV has been reported by some [8,11-16,18,20] but not all studies [7,9,10,17]. In our study, only one death was from a liver-related cause and this was recorded in a mono-infected patient. Intake of traditional medicines which have a potential for hepatic toxicity was however reported in two of the three coinfecting patients who died.

The main limitation to our study was the relatively small sample size. However, a significant higher risk of hepatotoxicity in coinfecting patients was demonstrated. Also, the absence of viral hepatitis impact on immunological and virological responses to antiretroviral therapy was strongly suggested by the regression coefficients and hazard ratio, respectively. Finally, the risk of clinical progression tended to decrease in coinfecting patients which is likely to be related to an unmeasured confounding factor. The exclusion of patients presenting with serum liver enzyme levels higher than three or five times the ULN values (depending on the initial study) could have led to a selection bias (exclusion of patients with an advanced hepatitis clinical stage) but this is concordant with the recommendations for nevirapine use. Classification biases for hepatitis status could also not be ruled out because seronegative blood samples were not tested for HBV DNA or HCV RNA. Higher rates of negative HBsAg or anti-HCV results in viraemic samples have been observed in immunocompromised HIV-infected patients [3,4]. In addition, unrepeated testing of HBV and HCV during follow-up did not allow to identify new or reactivated hepatitis infections. New infections, if any, were however likely to be rare owing to the study's duration and, for HBV, the usual childhood infections in Africa through close contacts within households and, to a lesser extent, vertical transmission (unclear modes of transmission for HCV) [1]. On the other hand, no symptomatic reactivated hepatitis infections were recorded. In contrast, classification of patients on the basis of HBV DNA and HCV RNA is clearly a strength

of our study while most other studies only used serological assays.

Conclusion

This study suggests that the nevirapine-based antiretroviral therapy could be used safely as first-line treatment in patients with low CD4 cell count in Africa despite frequent coinfections with HBV or HCV and infrequent testing of these infections. Although testing for HBV and HCV should be systematically performed before initiating antiretroviral therapy, transaminases elevations at baseline or during treatment should be a decisive argument for testing when hepatitis status is unknown.

Acknowledgements

We thank all the patients and staff of the Military Hospital and Central Hospital who participated in the study, and the National AIDS Programme, Yaoundé, Cameroon. The study was supported by the French National Agency for Research on AIDS and viral hepatitis (ANRS 1274), Institut de Recherche pour le Développement (France) and Médecins Sans Frontières (Switzerland). Jules Brice Tchatchueng Mbougua was the recipient of a doctoral fellowship from the Institut de Recherche pour le Développement.

Author details

¹Institut de Recherche pour le Développement, University Montpellier 1, UMR 145, Montpellier, France. ²National advanced school of engineering, University Yaoundé 1, Yaoundé, Cameroon. ³Central Hospital, Yaoundé, Cameroon. ⁴University Hospital, Department of Infectious and Tropical Diseases, Montpellier, France. ⁵Médecins Sans Frontières, Geneva, Switzerland. ⁶University Hospital, Laboratory of viral hepatitis, Montpellier, France. ⁷Military hospital, Yaoundé, Cameroon. ⁸Department of Biostatistics, University Montpellier 1, Montpellier, France. ⁹University Hospital, Department of Biostatistics, Nîmes, France.

Authors' contributions

JBTM analysed the data and, with CL, wrote the first draft of the report. CL also designed the study and contributed to data collection, analysis and interpretation of data. HG and NM contributed to data analysis. CK, AB, LC, AC, SKS, JD, EMN and ED contributed to data collection, interpretation of data and drafting of manuscript. All authors read and approved the final draft of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 1 December 2009

Accepted: 1 March 2010 Published: 1 March 2010

References

1. Modi AA, Feld JJ: **Viral hepatitis and HIV in Africa.** *AIDS Rev* 2007, **9**(1):25-39.
2. Soriano V, Puoti M, Garcia-Gasco P, Rockstroh JK, Benhamou Y, Barreiro P, McGovern B: **Antiretroviral drugs and liver injury.** *Aids* 2008, **22**(1):1-13.
3. Hoffmann CJ, Thio CL: **Clinical implications of HIV and hepatitis B co-infection in Asia and Africa.** *Lancet Infect Dis* 2007, **7**(6):402-9.
4. Rockstroh JK, Spengler U: **HIV and hepatitis C virus co-infection.** *Lancet Infect Dis* 2004, **4**(7):437-44.
5. Stern JO, Robinson PA, Love J, Lanes S, Imperiale MS, Mayers DL: **A comprehensive hepatic safety analysis of nevirapine in different populations of HIV infected patients.** *J Acquir Immune Defic Syndr* 2003, **34**(Suppl 1):S21-33.
6. World Health Organization: **Antiretroviral therapy for HIV infection in adults and adolescents in resource-limited settings: towards universal access. Recommendations for a public health approach.** Geneva: World Health Organization 2006, 140.
7. Hoffmann CJ, Charalambous S, Martin DJ, Innes C, Churchyard GJ, Chaisson RE, Grant AD, Fielding KL, Thio CL: **Hepatitis B virus infection and response to antiretroviral therapy (ART) in a South African ART program.** *Clin Infect Dis* 2008, **47**(11):1479-85.
8. Greub G, Ledergerber B, Battegay M, Grob P, Perrin L, Furrer H, Burgisser P, Erb P, Boggian K, Piffaretti JC, Hirschel B, Janin P, Francioli P, Flepp M, Telenti A: **Clinical progression, survival, and immune recovery during antiretroviral therapy in patients with HIV-1 and hepatitis C virus coinfection: the Swiss HIV Cohort Study.** *Lancet* 2000, **356**(9244):1800-5.
9. Lincoln D, Petoumenos K, Dore GJ: **HIV/HBV and HIV/HCV coinfection, and outcomes following highly active antiretroviral therapy.** *HIV Med* 2003, **4**(3):241-9.
10. Law WP, Duncombe CJ, Mahanontharit A, Boyd MA, Ruxrungtham K, Lange JM, Phanuphak P, Cooper DA, Dore GJ: **Impact of viral hepatitis co-infection on response to antiretroviral therapy and HIV disease progression in the HIV-NAT cohort.** *AIDS* 2004, **18**(8):1169-77.
11. Sheng WH, Chen MY, Hsieh SM, Hsiao CF, Wang JT, Hung CC, Chang SC: **Impact of chronic hepatitis B virus (HBV) infection on outcomes of patients infected with HIV in an area where HBV infection is hyperendemic.** *Clin Infect Dis* 2004, **38**(10):1471-7.
12. Bonacini M, Louie S, Bzowej N, Wohl AR: **Survival in patients with HIV infection and viral hepatitis B or C: a cohort study.** *AIDS* 2004, **18**(15):2039-45.
13. Konopnicki D, Mocroft A, de Wit S, Antunes F, Ledergerber B, Katlama C, Zilmer K, Vella S, Kirk O, Lundgren JD: **Hepatitis B and HIV: prevalence, AIDS progression, response to highly active antiretroviral therapy and increased mortality in the EuroSIDA cohort.** *AIDS* 2005, **19**(6):593-601.
14. Rockstroh JK, Mocroft A, Soriano V, Tural C, Losso MH, Horban A, Kirk O, Phillips A, Ledergerber B, Lundgren J: **Influence of hepatitis C virus infection on HIV-1 disease progression and response to highly active antiretroviral therapy.** *J Infect Dis* 2005, **192**(6):992-1002.
15. Backus LI, Phillips BR, Boothroyd DB, Mole LA, Burgess J, Riggsby MO, Chang SW: **Effects of hepatitis C virus coinfection on survival in veterans with HIV treated with highly active antiretroviral therapy.** *J Acquir Immune Defic Syndr* 2005, **39**(5):613-9.
16. Weis N, Lindhardt BO, Kronborg G, Hansen AB, Laursen AL, Christensen PB, Nielsen H, Moller A, Sorensen HT, Obel N: **Impact of hepatitis C virus coinfection on response to highly active antiretroviral therapy and outcome in HIV-infected individuals: a nationwide cohort study.** *Clin Infect Dis* 2006, **42**(10):1481-7.
17. Zhou J, Dore GJ, Zhang F, Lim PL, Chen YM: **Hepatitis B and C virus coinfection in The TREAT Asia HIV Observational Database.** *J Gastroenterol Hepatol* 2007, **22**(9):1510-8.
18. Omland LH, Weis N, Skinhoj P, Laursen A, Christensen PB, Nielsen HI, Moller A, Engsig F, Sorensen HT, Obel N: **Impact of hepatitis B virus coinfection on response to highly active antiretroviral treatment and outcome in HIV-infected individuals: a nationwide cohort study.** *HIV Med* 2008, **9**(5):300-6.
19. Carmo RA, Guimaraes MD, Moura AS, Neiva AM, Versiani JB, Lima LV, Freitas LP, Rocha MO: **The influence of HCV coinfection on clinical, immunological and virological responses to HAART in HIV-patients.** *Braz J Infect Dis* 2008, **12**(3):173-9.
20. Hoffmann CJ, Seaberg EC, Young S, Witt MD, D'Acunato K, Phair J, Thio CL: **Hepatitis B and long-term HIV outcomes in coinfecting HAART recipients.** *AIDS* 2009, **23**(14):1881-9.
21. De Luca A, Bugarini R, Lepri AC, Puoti M, Girardi E, Antinori A, Poggio A, Pagano G, Tositti G, Cadeo G, Macor A, Toti M, D'Arminio Monforte A: **Coinfection with hepatitis viruses and outcome of initial antiretroviral regimens in previously naive HIV-infected subjects.** *Arch Intern Med* 2002, **162**(18):2125-32.
22. Laurent C, Bourgeois A, Mpoudi-Ngole E, Kouanfack C, Ciuffi L, Nkoue N, Mougnutou R, Calmy A, Koulla-Shiro S, Ducos J, Delaporte E: **High rates of active hepatitis B and C co-infections in HIV-1 infected Cameroonian adults initiating antiretroviral therapy.** *HIV Med* 2010, **11**(1):85-9.
23. Laurent C, Kouanfack C, Koulla-Shiro S, Nkoué N, Bourgeois A, Calmy A, Lactuock B, Nzeusseu V, Mougnutou R, Peytavin G, Liegeois F, Nerrienet E, Tardy M, Peeters M, Andrieux-Meyer I, Zekeng L, Kazatchkine M, Mpoudi-Ngole E, Delaporte E: **Effectiveness and safety of a generic fixed-dose combination of nevirapine, stavudine, and lamivudine in HIV-1-infected adults in Cameroon: open-label multicentre trial.** *Lancet* 2004, **364**(9428):29-34.

24. Bourgeois A, Laurent C, Mognutou R, Nkoué N, Lactuock B, Ciaffi L, Liégeois F, Andrieux-Meyer I, Zekeng L, Calmy A, Mpoudi-Ngolé E, Delaporte E: **Field assessment of generic antiretroviral drugs: a prospective cohort study in Cameroon.** *Antivir Ther* 2005, **10**(2):335-41.
25. Laurent C, Kouanfack C, Koulla-Shiro S, Njoume M, Mawanba Nkene Y, Ciaffi L, Brulet C, Peytavin G, Vergne L, Calmy A, Mpoudi-Ngolé E, Delaporte E: **Long-term safety, effectiveness and quality of a generic fixed-dose combination of nevirapine, stavudine and lamivudine.** *AIDS* 2007, **21**(6):768-71.
26. De Beaudrap P, Etard JF, Diouf A, Ndiaye I, Gueye NF, Gueye PM, Sow PS, Mboup S, Ndoye I, Ecochard R, Delaporte E: **Modeling CD4+ cell count increase over a six-year period in HIV-1-infected patients on highly active antiretroviral therapy in Senegal.** *Am J Trop Med Hyg* 2009, **80**(6):1047-53.
27. Hoffmann CJ, Charalambous S, Thio CL, Martin DJ, Pemba L, Fielding KL, Churchyard GJ, Chaisson RE, Grant AD: **Hepatotoxicity in an African antiretroviral therapy cohort: the effect of tuberculosis and hepatitis B.** *Aids* 2007, **21**(10):1301-8.
28. den Brinker M, Wit FW, Wertheim-van Dillen PM, Jurriaans S, Weel J, van Leeuwen R, Pakker NG, Reiss P, Danner SA, Weverling GJ, Lange JM: **Hepatitis B and C virus co-infection and the risk for hepatotoxicity of highly active antiretroviral therapy in HIV-1 infection.** *AIDS* 2000, **14**(18):2895-902.
29. Sulkowski MS, Thomas DL, Mehta SH, Chaisson RE, Moore RD: **Hepatotoxicity associated with nevirapine or efavirenz-containing antiretroviral therapy: role of hepatitis C and B infections.** *Hepatology* 2002, **35**(1):182-9.
30. Matthews GV, Bartholomeusz A, Locarnini S, Ayres A, Sasaduesz J, Seaberg E, Cooper DA, Lewin S, Dore GJ, Thio CL: **Characteristics of drug resistant HBV in an international collaborative study of HIV-HBV-infected individuals on extended lamivudine therapy.** *Aids* 2006, **20**(6):863-70.
31. Dieterich DT, Robinson PA, Love J, Stern JO: **Drug-induced liver injury associated with the use of nonnucleoside reverse-transcriptase inhibitors.** *Clin Infect Dis* 2004, **38**(Suppl 2):S80-9.
32. Sanne I, Mommeja-Marin H, Hinkle J, Bartlett JA, Lederman MM, Maartens G, Wakeford C, Shaw A, Quinn J, Gish RG, Rousseau F: **Severe hepatotoxicity associated with nevirapine use in HIV-infected subjects.** *J Infect Dis* 2005, **191**(6):825-9.
33. Kaufmann GR, Perrin L, Pantaleo G, Opravil M, Furrer H, Telenti A, Hirschel B, Ledergerber B, Vernazza P, Bernasconi E, Rickenbach M, Egger M, Battegay M: **CD4 T-lymphocyte recovery in individuals with advanced HIV-1 infection receiving potent antiretroviral therapy for 4 years: the Swiss HIV Cohort Study.** *Arch Intern Med* 2003, **163**(18):2187-95.

Pre-publication history

The pre-publication history for this paper can be accessed here:<http://www.biomedcentral.com/1471-2458/10/105/prepub>

doi:10.1186/1471-2458-10-105

Cite this article as: Mbougua et al.: Hepatotoxicity and effectiveness of a Nevirapine-based antiretroviral therapy in HIV-infected patients with or without viral hepatitis B or C infection in Cameroon. *BMC Public Health* 2010 **10**:105.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



AnnexeB :article« *Long-term effectiveness and safety of didanosine combined with lamivudine and efavirenz or nevirapine in antiretroviral-naive patients: a 9-year cohort study in Senegal* »

Long-term effectiveness and safety of didanosine combined with lamivudine and efavirenz or nevirapine in antiretroviral-naïve patients: a 9-year cohort study in Senegal

Christian Laurent¹, Jules Brice Tchatchueng Mbougua^{1,2}, Ndèye Fatou Ngom Guèye³, Jean-François Etard¹, Assane Diouf⁴, Roland Landman^{5,6}, Nicolas Molinari^{7,8}, Pierre-Marie Girard^{5,6}, Papa Salif Sow^{4,9}, Ibra Ndoeye¹⁰ and Eric Delaporte^{1,11} for the ANRS 1215/1290 Study Group*

1 *Institut de Recherche pour le Développement, University Montpellier 1, UMR 145, Montpellier, France*

2 *National Advanced School of Engineering, University Yaoundé 1, Yaoundé, Cameroon*

3 *Ambulatory Care Unit, Fann University Teaching Hospital, Dakar, Senegal*

4 *Regional Research and Training Centre for HIV/AIDS, Fann University Teaching Hospital, Dakar, Senegal*

5 *Institut de Médecine et d'Epidémiologie Appliquée, Paris, France*

6 *Department of Infectious Diseases, Bichat-Claude Bernard Hospital, Paris, France*

7 *Department of Biostatistics, University Montpellier 1, Montpellier, France*

8 *Department of Biostatistics, University Hospital, Nîmes, France*

9 *Department of Infectious Diseases, Fann University Teaching Hospital, Dakar, Senegal*

10 *National AIDS Program, Ministry of Health, Dakar, Senegal*

11 *Department of Infectious and Tropical Diseases, University Hospital, Montpellier, France*

Summary

OBJECTIVE The use of didanosine (ddI) in first-line antiretroviral therapy has been recently promoted for resource-limited settings. We therefore compared the long-term effectiveness and safety of the regimen combining ddI, lamivudine, and efavirenz or nevirapine with that of the WHO-recommended regimen of zidovudine (ZDV), lamivudine, and efavirenz or nevirapine in antiretroviral-naïve patients in Senegal. **METHODS** Observational cohort study of patients enrolled between January 2000 and April 2002 in the Senegalese antiretroviral drug access initiative. Multivariate analyses were performed to compare, between the ddI and ZDV groups, the proportion of patients with a viral load <500 copies/ml during follow-up; the increase in the CD4 cell count; survival; treatment changes and severe adverse events. **RESULTS** Of 151 patients, 71 received the ddI-based treatment and 80 received the ZDV-based treatment. Throughout follow-up, 80–95% of patients had a viral load below 500 copies/ml in both the ddI and ZDV groups ($P = 0.5$). The CD4 cell count increased after treatment initiation from 176 to 497 cells/mm³ in the ddI group and from 176 to 567 cells/mm³ in the ZDV group ($P > 0.3$). The rate of death tended to be higher in the ddI group ($P = 0.06$). ddI was less commonly discontinued than ZDV ($P = 0.03$).

CONCLUSION The combination of ddI, lamivudine, and efavirenz or nevirapine resulted in sustained viral suppression and immunological recovery.

keywords Africa, HIV, antiretroviral therapy, didanosine, zidovudine

Introduction

First-line combination antiretroviral therapy (ART) recommended by WHO in resource-limited countries consists of one non-nucleoside reverse transcriptase inhibitor (NNRTI; efavirenz or nevirapine) and two nucleoside reverse transcriptase inhibitors (NRTI). The preferred NRTI

backbones are now zidovudine (ZDV) plus lamivudine or tenofovir plus either lamivudine or emtricitabine (World Health Organization, 2010). But the use of ZDV, although common, is hindered by frequent haematological disorders (especially severe anaemia) of various origins, including the HIV infection itself, other infectious diseases and malnutrition (Colebunders *et al.* 2005; Moh *et al.* 2005). Tenofovir has only been adopted by few African programmes because of its high cost until recently and concerns about

*See Appendix.

C. Laurent *et al.* Long-term use of didanosine in Senegal

renal toxicity (Ford & Calmy 2010). Recently, the use of didanosine (ddI) in resource-limited settings has been promoted owing to higher efficacy of regimens including ddI plus either lamivudine or emtricitabine than regimens of ZDV plus lamivudine, and ddI's low cost and wide availability (Carr & Amin 2009). However, the authors acknowledged as limitations of their study the ecological design, the short duration of follow-up (14 months) and the fact that only 12% of patients were from resources-limited settings. Here, we compared the long-term effectiveness and safety of the regimen combining ddI, lamivudine, and efavirenz or nevirapine with that of the WHO-recommended regimen of ZDV, lamivudine, and efavirenz or nevirapine in ART-naïve patients treated in Senegal.

Methods

From August 1998 to April 2002, 404 HIV-1-infected patients (age >15 years) were enrolled in the Senegalese antiretroviral drug access initiative (ISAARV) in Dakar and included in an observational cohort study (Desclaux *et al.* 2003). The methods of the cohort study have been described extensively elsewhere (Etard *et al.* 2006). Of the 404 patients, 80 received ZDV, lamivudine, and efavirenz or

nevirapine and 71 received enteric-coated ddI, lamivudine, and efavirenz or nevirapine. Of these latter, 40 patients participated for the first 18 months in a prospective, open-label, one-arm trial (Landman *et al.* 2003). Antiretroviral treatment was provided every month or 2 months free of charge for the 40 patients initially included in the clinical trial or commensurate with their incomes for the other patients, but became free for all patients as of December 2003. Patients attended clinical examinations at least every 2 months and biological examinations at least every 6 months including plasma HIV-1 viral load (Amplicor HIV-1 1.5 or 2.0 assay, Roche Molecular Systems or Bayer bDNA HIV-1 2.0 or 3.0 assay; Bayer Diagnostic) and CD4 cell count (FACSCount; Becton Dickinson). Phone calls and, if necessary, home visits were made to all patients who did not attend scheduled visits to discuss reasons for non-adherence or to identify deaths (Etard *et al.* 2006). The Senegalese national ethics committee approved the study, and the patients gave their informed consent.

Between-group comparisons (ddI group *vs.* ZDV group) were performed using the chi-square test, Fisher's exact test and the non-parametric Mann-Whitney two-sample test as appropriate for baseline patients' characteristics; Cox models for treatment changes and deaths; generalised

Table 1 Baseline characteristics of the 151 patients by treatment group

	Didanosine (<i>n</i> = 71)	Zidovudine (<i>n</i> = 80)	<i>P</i>
Women, no. (%)	41 (58)	40 (50)	0.3
Age (years), median (IQR)	37.1 (31.7–45.5)	38.3 (31.5–44.1)	0.7
Body weight (kg), median (IQR)	58 (49–66)	57 (50–68)	0.8
Body mass index (kg/m ²), median (IQR)*	19.6 (17.7–22.3)	19.8 (17.9–22.3)	0.9
Karnofsky score 100%, no. (%)	36 (51)	25 (31)	0.02
CDC clinical stage, no. (%)			
A	6 (8)	6 (8)	0.9
B	26 (37)	32 (40)	
C	39 (55)	42 (52)	
CD4 cell count (/mm ³), median (IQR)†	176 (98–233)	182 (88–261)	0.9
HIV-1 viral load (log ₁₀ copies/ml)‡			
Median (IQR)	5.4 (5.0–5.7)	4.9 (4.5–5.3)	<0.001
<4.0, no. (%)	2 (4)	5 (9)	<0.001
4.0–4.9, no. (%)	10 (18)	27 (48)	
≥5.0, no. (%)	44 (78)	24 (43)	
Haemoglobin (g/dl), median (IQR)§	10.3 (9.5–11.5)	11.1 (9.8–12.5)	0.04
Neutrophil count (/mm ³), median (IQR)¶	2385 (1716–3053)	1845 (1417–2725)	0.01
Alanine aminotransferase (U/l), median (IQR)†	20 (11–29)	21 (13–31)	0.3
Aspartate aminotransferase (U/l), median (IQR)**	29 (19–42)	23 (17–32)	0.06
HBsAg, no. †† (%)	7 (11)	13 (17)	0.3
Anti-HCV, no. ‡‡ (%)	3 (5)	5 (7)	0.5
Non-nucleoside reverse transcriptase inhibitor, no. (%)			
Efavirenz	63 (89)	52 (65)	0.001
Nevirapine	8 (11)	28 (35)	
Cotrimoxazole prophylaxis, no. (%)	43 (61)	68 (85)	0.001

Numbers of missing data were: *4, †11, ‡39, 2§, ¶5, **12, ††14 and ‡‡15.

C. Laurent *et al.* Long-term use of didanosine in Senegal

estimating equations models for the proportion of patients with plasma HIV-1 viral load below 500 copies/ml; mixed-effect linear regression models for the evolution of the CD4 cell count and negative binomial models for grade 3 or 4 adverse events (classification of the French National Agency for Research on AIDS and Viral Hepatitis; <http://www.anrs.fr/content/download/2242/12805/file/ANRS-GradeEI-V1-En-2008.pdf>). We allowed the slope of the CD4 cell count to change after 6 months and included quadratic and cubic effects in the second phase. The multivariate analyses were adjusted on baseline variables associated with either the outcome ($P < 0.25$ in univariate analysis) or the treatment group. Multiple imputations based on both baseline and outcomes data were used to fill out incomplete baseline variables assuming that data were missing at random (Rubin 1987). Data were censored at the time of ddI or ZDV change for the analyses of treatment responses.

Results

Of the patients enrolled between January 2000 and April 2002, 71 received ddI and 80 received ZDV (in addition to lamivudine and a NNRTI). Most patient characteristics were comparable between the ddI and ZDV groups (Table 1). However, compared to patients in the ZDV group, those in the ddI group had a higher Karnofsky score (51% *vs.* 31% had a score of 100%), viral load (median

5.4 *vs.* 4.9 \log_{10} copies/ml) and neutrophil count (median 2385 *vs.* 1845 cells/mm³). In contrast, they had a lower haemoglobin level (median 10.3 *vs.* 11.1 g/dl) and were less likely to receive a cotrimoxazole prophylaxis (61% *vs.* 85%). Study follow-up was 470 person-years for the ddI group [median follow-up time of 7.7 years, interquartile range (IQR) 5.1–9.0] and 473 person-years for the ZDV group (median follow-up time of 7.0 years, IQR 5.9–7.4).

During follow-up, ddI was discontinued in 10 patients (14%) and ZDV in 15 patients (19%). Treatment changes occurred after a median of 7.5 years (IQR 7.0–8.8) and 5.7 years (IQR 5.3–6.4), respectively ($P = 0.01$). The incidence rates of treatment changes were 2.2 per 100 person-years [95% confidence interval (CI) 1.2–4.1] in the ddI group and 3.4 per 100 person-years (CI 2.0–5.6) in the ZDV group. The rate of change was significantly lower in the ddI group [hazard ratio (HR) 0.35, CI 0.13–0.91, $P = 0.03$]. Treatment changes were owing to adverse effects ($n = 5$; digestive intolerance in one patient and unspecified intolerance in the others), treatment failure ($n = 4$) and 5-year loss to follow-up ($n = 1$) for ddI, and to treatment failure ($n = 10$), adverse effects ($n = 4$; anaemia in 3 patients and unspecified intolerance in the other) and loss to follow-up ($n = 1$) for ZDV. ddI was replaced by ZDV ($n = 5$) or tenofovir ($n = 5$), and ZDV was replaced by tenofovir ($n = 11$), ddI ($n = 2$) or stavudine ($n = 2$). Lamivudine and the NNRTI in the antiretroviral treatment were also switched in case of therapeutic failure.

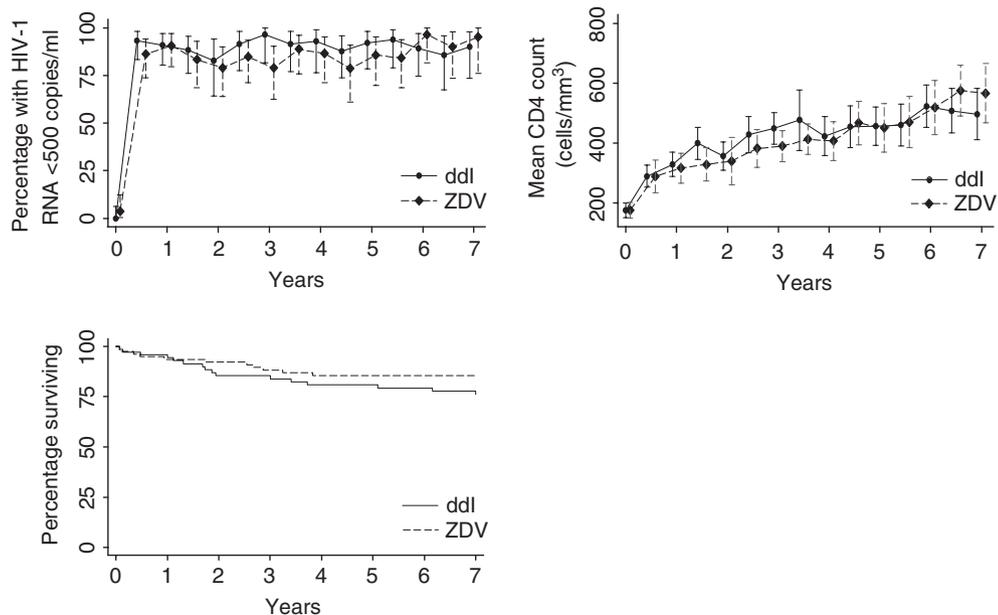


Figure 1 Outcomes in the 151 patients by treatment group: percentage of patients with plasma HIV-1 viral load below 500 copies/ml, mean CD4 cell count and survival. Bars indicate 95% confidence intervals.

C. Laurent *et al.* Long-term use of didanosine in Senegal

Throughout follow-up, 80–95% of patients had a viral load below 500 copies/ml in both the ddI and ZDV groups (Figure 1). There was no significant difference between the treatment groups (ddI *vs.* ZDV group; adjusted odds ratio 1.27, CI 0.64–2.51, $P = 0.5$). As shown in Figure 1, the CD4 cell count increased after treatment initiation from 176 to 497 cells/mm³ in the ddI group and from 176 to 567 cells/mm³ in the ZDV group. The increase was not significantly different between the two groups ($P > 0.3$).

There were 16 deaths (23%) in the ddI group and 11 (14%) in the ZDV group. The incidence rates of death were 3.5 (CI 2.2–5.7) and 2.5 (CI 1.4–4.4) per 100 person-years, respectively. The rate of death tended to be higher in the ddI group (Figure 1; adjusted HR 2.09, 0.96–4.54, $P = 0.06$). In the ddI group, the deaths were related to meningitis ($n = 3$), tuberculosis ($n = 3$), HIV encephalopathy ($n = 1$), isosporiasis ($n = 1$), septicaemia ($n = 1$), unspecified infectious disease ($n = 1$), non-infectious liver disease ($n = 1$), non-infectious respiratory insufficiency ($n = 1$) or unknown cause ($n = 4$; three patients were lost to follow-up). In the ZDV group, the deaths were related to tuberculosis ($n = 2$), unspecified infectious disease ($n = 2$), HIV encephalopathy ($n = 1$), central nervous system toxoplasmosis ($n = 1$), pneumopathy ($n = 1$), cardiac failure ($n = 1$), anaemia ($n = 1$) or unknown cause ($n = 2$; one patient was lost to follow-up).

In the ddI group, 22 patients (31%) experienced at least one episode of grade 3 or 4 biological adverse event: once ($n = 15$), twice ($n = 6$) or eight times ($n = 1$). In the ZDV group, 19 patients (24%) did once ($n = 15$) or twice ($n = 4$). The rate tended to be higher in the ddI group but the difference was not significant (adjusted incidence rate ratio 1.79, CI 0.91–3.52, $P = 0.09$). The predominant events were anaemia ($n = 19$), neutropenia ($n = 5$) and thrombopenia ($n = 5$) in the ddI group, and leucopenia ($n = 5$), anaemia ($n = 4$) and neutropenia ($n = 3$) in the ZDV group.

Discussion

This cohort study in Senegal showed comparable long-term virological and immunological effectiveness of a regime consisting of ddI, lamivudine, and efavirenz or nevirapine as a regime consisting of ZDV, lamivudine, and the same NNRTIs in ART-naive patients. ddI was less commonly discontinued than ZDV.

Didanosine has been used in first-line ART in Africa combined with other NRTIs, namely stavudine and ZDV (Laurent *et al.* 2002; Seyler *et al.* 2003). However, these combinations were abandoned owing to poor efficacy and/or tolerability (Canestri *et al.* 2007; Busmann *et al.* 2009; Ratsela *et al.*, 2009). To our knowledge, data on ddI–lamivudine–NNRTI in ART-naive patients in Africa

are mainly limited to our Senegalese cohort. The clinical trial in 40 of our patients showed promising results for the first 15 months but there was no comparison regimen (Landman *et al.* 2003). Our present study therefore extends the knowledge in the long term and with comparison to one of the WHO-recommended regimens.

Overall effectiveness and tolerability with both ddI- and ZDV-based regimens were similar or even better to those seen initially in other African settings (Coetzee *et al.* 2004; Laurent *et al.* 2008) or western countries (Berenguer *et al.* 2008; Crespo *et al.* 2009). It is noteworthy that the good virological effectiveness was maintained in the long term and that the CD4 cell count continued to increase with both regimens. It was therefore surprising that the deaths tended to be more frequent in the ddI group. Unreported ddI toxicity was unlikely to explain the excess deaths because most deaths were related to infectious diseases or, to a lesser extent, loss to follow-up. Similarly to our study, one of the western studies reported a twofold higher risk of treatment change in patients who received a ZDV–lamivudine backbone as in those treated with ddI–lamivudine (Crespo *et al.* 2009).

The main limitation of our study is the lack of treatment randomisation. However, this observational cohort study in which all analyses were adjusted on baseline patients' characteristics and which was performed over a 9-year period in Africa adds useful information to the current debate on the use of ddI in first-line ART in routine resource-limited settings (Benson & Hartz 2000; Concato *et al.* 2000; Concato & Horwitz 2004). The number of patients was also relatively small, potentially leading to fail demonstrating significant differences between the treatment groups (e.g. for death).

In conclusion, the combination of ddI, lamivudine, and efavirenz or nevirapine resulted in sustained viral suppression and immunological recovery. However, additional data on disease progression in these settings are needed.

Acknowledgements

We thank all the patients and staff who participated in the study. The study was funded by the French National Agency for Research on AIDS and Viral Hepatitis, IMEA, the European Union and the Institut de Recherche pour le Développement. Jules Brice Tchatchueng Mbougua was the recipient of a doctoral fellowship from IRD.

References

- Benson K & Hartz AJ (2000) A comparison of observational studies and randomized, controlled trials. *New England Journal of Medicine* **342**, 1878–1886.

C. Laurent *et al.* Long-term use of didanosine in Senegal

- Berenguer J, Gonzalez J, Ribera E *et al.* (2008) Didanosine, lamivudine, and efavirenz versus zidovudine, lamivudine, and efavirenz for the initial treatment of HIV type 1 infection: final analysis (48 weeks) of a prospective, randomized, noninferiority clinical trial, GESIDA 3903. *Clinical Infectious Diseases* 47, 1083–1092.
- Bussmann H, Wester CW, Thomas A *et al.* (2009) Response to zidovudine/didanosine-containing combination antiretroviral therapy among HIV-1 subtype C-infected adults in Botswana: two-year outcomes from a randomized clinical trial. *Journal of Acquired Immune Deficiency Syndromes* 51, 37–46.
- Canestri A, Sow PS, Vray M *et al.* (2007) Poor efficacy and tolerability of stavudine, didanosine, and efavirenz-based regimen in treatment-naïve patients in Senegal. *Medscape General Medicine* 9, 7.
- Carr A & Amin J (2009) Efficacy and tolerability of initial antiretroviral therapy: a systematic review. *AIDS* 23, 343–353; discussion 55–6.
- Coetzee D, Hildebrand K, Boule A *et al.* (2004) Outcomes after two years of providing antiretroviral treatment in Khayelitsha, South Africa. *AID* 18, 887–895.
- Colebunders R, Kanya MR, Laurence J *et al.* (2005) First-line antiretroviral therapy in Africa – how evidence-base are our recommendations? *AIDS Reviews* 7, 148–154.
- Concato J & Horwitz RI (2004) Beyond randomised versus observational studies. *Lancet* 363, 1660–1661.
- Concato J, Shah N & Horwitz RI (2000) Randomized, controlled trials, observational studies, and the hierarchy of research designs. *New England Journal of Medicine* 342, 1887–1892.
- Crespo M, Ribera E, Suarez-Lozano I *et al.* (2009) Effectiveness and safety of didanosine, lamivudine and efavirenz versus zidovudine, lamivudine and efavirenz for the initial treatment of HIV-infected patients from the Spanish VACH cohort. *Journal of Antimicrobial Chemotherapy* 63, 189–196.
- Desclaux A, Ciss M, Taverne B *et al.* (2003) Access to antiretroviral drugs and AIDS management in Senegal. *AIDS* 17(Suppl. 3), S95–S101.
- Etard JF, Ndiaye I, Thierry-Mieg M *et al.* (2006) Mortality and causes of death in adults receiving highly active antiretroviral therapy in Senegal: a 7-year cohort study. *AIDS* 20, 1181–1189.
- Ford N & Calmy A (2010) Improving first-line antiretroviral therapy in resource-limited settings. *Current Opinion in HIV and AIDS* 5, 38–47.
- Landman R, Schiemann R, Thiam S *et al.* (2003) Once-a-day highly active antiretroviral therapy in treatment-naïve HIV-1-infected adults in Senegal. *AIDS* 17, 1017–1022.
- Laurent C, Diakhaté N, Ngom Gueye NF *et al.* (2002) The Senegalese government's highly active antiretroviral therapy initiative: an 18-month follow-up study. *AIDS* 16, 1363–1370.
- Laurent C, Bourgeois A, Mpoudi-Ngole E *et al.* (2008) Tolerability and effectiveness of first-line regimens combining nevirapine and lamivudine plus zidovudine or stavudine in Cameroon. *AIDS Research and Human Retroviruses* 24, 393–399.
- Moh R, Danel C, Sorho S *et al.* (2005) Haematological changes in adults receiving a zidovudine-containing HAART regimen in combination with cotrimoxazole in Cote d'Ivoire. *Antiviral Therapy* 10, 615–624.
- Ratsela A, Polis M & the Phidisa II Study Group (2009) Phidisa II: a randomized 2x2 factorial trial comparing initial therapy of efavirenz with lopinavir/ritonavir and zidovudine + didanosine with stavudine + lamivudine in treatment-naïve HIV-infected persons with < 200 CD4+ cells/mm³ or a prior AIDS diagnosis. 16th Conference on Retroviruses and Opportunistic Infections. Montréal, Canada.
- Rubin DB (1987) *Multiple Imputation for Nonresponse in Surveys*. Wiley, New York.
- Seyler C, Anglaret X, Dakoury-Dogbo N *et al.* (2003) Medium-term survival, morbidity and immunovirological evolution in HIV-infected adults receiving antiretroviral therapy, Abidjan, Côte d'Ivoire. *Antiviral Therapy* 8, 385–393.
- World Health Organization. Antiretroviral therapy for HIV infection in adults and adolescents. Recommendations for a public health approach. 2010 revision. [cited October 12, 2010]; Available at: http://whqlibdoc.who.int/publications/2010/9789241599764_eng.pdf.

Corresponding Author Christian Laurent, Institut de Recherche pour le Développement (UMR 145), 911 avenue Agropolis, BP 64501, 34394 Montpellier Cedex 5, France. E-mail: Christian.Laurent@ird.fr

Appendix

ANRS 1215/1290 Study Group

I. Ndiaye (Multisectorial AIDS Program, Dakar, Senegal), P. de Beaudrap, E. Delaporte, J. F. Etard, C. Laurent, B. Taverne (UMR 145, Institut de Recherche pour le Développement (IRD)/University Montpellier 1, Montpellier, France), M. M. Bitèye, A. B. Dieng, A. Diouf, A. Sarr, L. Zié (Regional Research and Training Center for HIV/AIDS, Fann University Teaching Hospital, Dakar, Senegal), V. Cilote, I. Lanièce (French Ministry of Foreign Affairs, Dakar, Senegal), I. Ndiaye, A. Ndir, C. T. Ndour, C. S. Senghor, P. S. Sow (Department of Infectious Diseases, Fann University Teaching Hospital,

C. Laurent *et al.* **Long-term use of didanosine in Senegal**

Dakar, Senegal), M. Basty Fall, N. M. Dia Badiane, N. Diakhaté, M. Diallo, L. M. Diouf, N. F. Ngom Guèye (Ambulatory Care Unit-Red-Cross, Fann University Teaching Hospital, Dakar, Senegal), K. Ba Fall, P. M. Guèye, C. Périno (Military Hospital of Dakar, Senegal), P. A. Diaw, H. Diop Ndiaye, S. Mboup, N. C. Touré Kane (Le Dantec University Teaching Hospital, Virology and Bacteriology Laboratory, Dakar, Senegal), K. Sow (AIDS Control Division, Ministry of Health, Dakar, Senegal), K. Diop, B. Ndiaye (Central Pharmacy, Fann University Teaching Hospital, Dakar, Senegal), M. Ciss (Ministry of Health, National Drug Laboratory, Dakar, Senegal), and A. Desclaux (Paul Cezanne University, CReCSS, Aix-en-Provence, France).

Annexe C :article« *Monitoring of HIV viral loads, CD4 cell counts and clinical assessments versus clinical monitoring alone for antiretroviral therapy in rural district hospitals in Cameroon (Stratall ANRS 12110/ESTHER): a randomised non-inferiority trial. »*

Monitoring of HIV viral loads, CD4 cell counts, and clinical assessments versus clinical monitoring alone for antiretroviral therapy in rural district hospitals in Cameroon (Stratall ANRS 12110/ESTHER): a randomised non-inferiority trial



Christian Laurent, Charles Kouanfack, Gabrièle Laborde-Balen, Avelin Fobang Aghokeng, Jules Brice Tchatchueng Mbougua, Sylvie Boyer, Maria Patrizia Carrieri, Jean-Marc Mben, Marlise Dontsop, Serge Kazé, Nicolas Molinari, Anke Bourgeois, Eitel Mpoudi-Ngolé, Bruno Spire, Sinata Koulla-Shiro, Eric Delaporte, for the Stratall ANRS 12110/ESTHER study group*

Summary

Background Scaling up of antiretroviral therapy in low-resource countries is done on the basis of decentralised, integrated HIV care in rural facilities; however, laboratory monitoring is generally unavailable. We aimed to assess the effectiveness and safety of clinical monitoring alone (CLIN) in terms of non-inferiority to laboratory and clinical monitoring (LAB).

Methods We did a randomised, open-label, non-inferiority trial in nine rural district hospitals in Cameroon. Eligible participants were adults (≥ 18 years) infected with HIV-1 group M (WHO disease stage 3–4) who had not previously received antiretroviral therapy, and were followed-up for 2 years by health-care workers in routine activities. We randomly assigned participants (1:1) to CLIN or LAB (counts of HIV viral load and CD4 cell every 6 months) groups with a computer-generated list. The primary outcome was non-inferiority of CLIN to LAB in terms of increase in CD4 cell count with a non-inferiority margin of 25%. We did all analyses in participants who attended at least one follow-up visit. This trial is registered with ClinicalTrials.gov, number NCT00301561.

Findings 238 (93%) of 256 participants assigned to CLIN and 221 (93%) of 237 assigned to LAB were eligible for analysis. CLIN was not non-inferior to LAB; the mean increase in CD4 cell count was 175 cells per μL (SD 190, 95% CI 151–200) with CLIN and 206 (190, 181–231) with LAB (difference -31 [-63 to 2] and non-inferiority margin -52 [-58 to -45]). Furthermore, in the predefined secondary outcome of treatment changes, 13 participants (6%) in the LAB group switched to second-line regimens whereas no participants in the CLIN group did so ($p < 0.0001$). By contrast, other predefined secondary outcomes were much the same in both groups—viral suppression (< 40 copies per mL; 465 [49%] of 952 measurements in CLIN vs 456 [52%] of 884 in LAB), HIV resistance (23 [10%] of 238 participants vs 22 [10%] of 219 participants), mortality (44 [18%] of 238 vs 32 [14%] of 221), disease progression (85 [36%] of 238 vs 64 [29%] of 221), adherence (672 [63%] of 1067 measurements vs 621 [61%] of 1011), loss to follow-up (21 [9%] of 238 vs 17 [8%] of 221), and toxic effects (46 [19%] of 238 vs 56 [25%] of 221).

Interpretation Our findings support WHO's recommendation for laboratory monitoring of antiretroviral therapy. However, the small differences that we noted between the strategies suggest that clinical monitoring alone could be used, at least temporarily, to expand antiretroviral therapy in low-resource settings.

Funding French National Agency for Research on AIDS (ANRS) and Ensemble pour une Solidarité Thérapeutique Hospitalière En Réseau (ESTHER).

Introduction

Scaling up of antiretroviral therapy in low-resource countries has mostly been achieved through the public-health approach promoted by WHO.^{1,2} Taking into account the need for such treatment and the financial and infrastructural restrictions of health facilities, this approach recommends treatment of all patients with an advanced stage of HIV disease even if viral loads and CD4 cell counts are not available. A mathematical modelling study suggested that use of antiretroviral therapy without monitoring of viral loads or CD4 cell counts would not have distinct detrimental effects on survival of patients or development of resistance.³ Nevertheless, in the home-based AIDS care (HBAC) trial⁴

in Uganda disease progression to death or new AIDS-defining event was faster in patients who were provided antiretroviral therapy without monitoring of viral load and CD4 cell count than in those who had such tests. Furthermore, the Development of AntiRetroviral Therapy in Africa (DART) trial⁵ in Uganda and Zimbabwe showed that routine monitoring of the CD4 cell counts has a small but significant benefit in terms of disease progression and mortality.

The revised 2010 WHO guidelines⁶ recommend use of viral-load measurements (if available) to detect or confirm treatment failure and to inform the decision to switch to second-line regimens (in addition to immunological and clinical monitoring). However, evidence to support this

Lancet Infect Dis 2011; 11: 825–33

Published Online

August 9, 2011

DOI:10.1016/S1473-3099(11)70168-2

See [Comment](#) page 803

*Members listed at end of paper

Institut de Recherche pour le Développement (IRD), University Montpellier 1, UMI 233, Montpellier, France (C Laurent PhD, A F Aghokeng PhD, J B T Mbougua MSc, A Bourgeois MD, Prof E Delaporte MD); Central Hospital (C Kouanfack MD, Prof S Koulla-Shiro MD), Virology laboratory IRD/IMP/CREMER (A F Aghokeng, E Mpoudi-Ngolé MD), IRD (J-M Mben MD, M Dontsop MD, S Kazé MD), and University Yaoundé 1 (S Koulla-Shiro), UMI 233, Yaoundé, Cameroon; French Ministry of Foreign Affairs, Yaoundé, Cameroon (G Laborde-Balen MSc); National Advanced School of Engineering, University Yaoundé 1, Yaoundé, Cameroon (J B T Mbougua); INSERM, IRD, University Aix Marseille, UMR 912, Marseille, France (S Boyer PhD, M P Carrieri PhD, B Spire MD); UMR 729 MISTEA, Montpellier, France (N Molinari PhD); Department of Biostatistics, University Hospitals, Nîmes and Montpellier, France (N Molinari); and Department of Infectious and Tropical Diseases, University Hospital, Montpellier, France (A Bourgeois, E Delaporte)

Correspondence to:

Dr Christian Laurent, Institut de Recherche pour le Développement (UMI 233), 911 Avenue Agropolis, BP 64501, 34394 Montpellier cedex 5, France
christian.laurent@ird.fr

recommendation is scarce in low-resource settings. Moreover, the WHO scale-up strategy is based on decentralised, integrated delivery of HIV care in rural districts where most people live, but local health facilities generally have low-grade equipment.¹⁷ In this study, we aimed to assess non-inferiority of an exclusively clinical monitoring strategy compared with a clinical monitoring strategy plus laboratory monitoring (including of viral load) in terms of effectiveness and safety in patients who had not had antiretroviral therapy and were followed up in rural district hospitals in Cameroon.

Methods

Study design and participants

We did a randomised, controlled, open-label, non-inferiority trial in nine rural district hospitals in Cameroon. Participants were recruited between May 23, 2006, and Jan 31, 2008, and followed up for 24 months by hospital health-care workers involved in routine activities. Patients were eligible if they were 18 years or older and had confirmed HIV-1 group M infection, WHO clinical stage 3 or 4, or WHO clinical stage 2 with a total lymphocyte count of fewer than 1200 cells per μL . Patients were ineligible if they were unlikely to attend follow-up in district hospitals regularly, for example because they lived far away from the hospital or had difficulty in accessing it. Other exclusion criteria were HIV-1 serotypes O or N or HIV-2 infection, active tuberculosis and total lymphocyte count of more than 1200 cells per μL , active malignant disease (apart from mucocutaneous Kaposi's sarcoma), active psychiatric disorders, hepatocellular insufficiency, previous antiretroviral therapy, present intake of corticosteroids, immunomodulators, or other experimental drugs, or pregnancy. The protocol was approved by the National Ethics Committee of Cameroon and the Institutional Ethics Committee of the French Institut de Recherche pour le Développement. All patients provided written informed consent.

Randomisation and masking

Participants were randomly allocated (1:1) at enrolment to receive either laboratory and clinical monitoring (LAB; measurement of viral load and CD4 cell count every 6 months, plus clinical monitoring every 3 months) or clinical monitoring alone (CLIN; every 3 months); both groups had haematological and biochemical tests. The computer-generated, sequentially numbered randomisation list with block sizes of 100 was prepared centrally and only visible to the project supervisor (GL-B). For every eligible participant, the project supervisor allocated a monitoring strategy to the local doctor who initiated antiretroviral therapy. Participants, health-care workers, and data analysers were not masked to group allocation.

Procedures

Participants attended clinical visits at weeks 0 and 2, months 1 and 3, and every 3 months thereafter. At week 2

and months 3, 9, 15, and 21, nurses assessed participants in the CLIN group and, in the case of adverse events, referred them to doctors for care. Systematic appointments with doctors were scheduled for all other visits. Participants in the LAB group were supposed to be seen by doctors for all study visits. Participants could also attend clinics at any time that they felt unwell. Clinical staging of HIV disease was based on the 2006 revised WHO classification.⁸ Treatment failure was defined in the LAB group as a persistent viral load of more than 5000 copies per mL on two consecutive samples, 3 months apart, and by persistent (new or recurrent) WHO stage 3 or 4 adverse event in the CLIN group after the first 6 months. Clinical and biological adverse events were graded according to the French National Agency for Research on AIDS and Viral Hepatitis (ANRS) toxic effects scale.⁹

CD4 cell counts (FACSCount device, Becton Dickinson, Mountain View, CA, USA) and plasma viral loads (RealTime HIV-1 assay, Abbott Molecular, Des Plaines, IL, USA) were recorded at baseline and every 6 months thereafter. We assessed genotypic mutations associated with antiretroviral drug resistance with the Abbott Viroseq assay (Celera Diagnostics, Alameda, CA, USA) when the viral load was higher than 5000 copies per mL on two consecutive samples or when the patient's last viral load was higher than this threshold. We tested samples that had a viral load of more than 5000 copies per mL and, if positive, corresponding baseline samples. Mutations were classified as minor or major by use of the ANRS consensus statements on antiretroviral drug resistance version from July, 2010.¹⁰ We assessed CD4 cell counts, viral load, and resistance in a reference HIV laboratory in Yaoundé (Cameroon). All results for participants in the LAB group were returned immediately to doctors. Results for participants in the CLIN group were retained at the reference laboratory until month 24.

The first-line antiretroviral regimen was stavudine (or zidovudine), lamivudine, and nevirapine (or efavirenz), at the treating doctor's discretion. Treatment change was allowed in case of drugs-related adverse events or failures. The second-line regimen was lopinavir, ritonavir, and two nucleoside reverse transcriptase inhibitors (NRTIs). Tenofovir, emtricitabine, didanosine, and abacavir became available during the study. Treatment was dispensed once per month at district hospitals' pharmacies. We assessed adherence to treatment at months 1, 3, 6, and every 6 months thereafter through community health workers on the basis of patients' responses about dose-taking for the previous 4 days, adjusted for responses about the preceding 4 weeks.^{11,12} Participants who did not attend scheduled appointments were phoned or visited at home.

The primary outcome was mean increase in CD4 cell count from treatment initiation to month 24. Secondary outcomes were viral-load suppression, mortality, new or

recurrent WHO stage 3 or 4 adverse events, HIV resistance mutations, loss to follow-up, adherence, treatment changes, clinical and biological side-effects related to antiretroviral drugs, effect on patients' daily life, acceptability by patients and health-care workers of the particular strategy, and cost-effectiveness. Effects on patients' daily life, acceptability by patients and health-care workers, and cost-effectiveness will be reported separately.

Statistical analysis

We initially calculated a total sample size of 340 participants to obtain 80% power to show a non-inferior increase in CD4 cell counts in the CLIN group compared with the LAB group, at a 5% significance level.¹³ We made this calculation on the basis of results from a previous study in Yaoundé in which the mean increase in CD4 cell count with laboratory monitoring was 140 cells per μL (SD 130).¹⁴ The Shapiro ($p=0.78$) and skewness ($p=0.56$) tests confirmed that the distribution of the CD4 increase was normally distributed. CD4 cell count increases in the CLIN group were prespecified as non-inferior to those in the laboratory-monitoring group if the difference was 25% or less. This difference was regarded as clinically acceptable because of expected financial and infrastructural difficulties in implementation of laboratory monitoring in rural facilities for many years to come, and because of the urgent need to scale-up antiretroviral therapy. We initially assumed a maximum of 20% of participants would die or be lost to follow-up and the total sample size was increased to 430. After the trial began, we noted a higher-than-anticipated rate of deaths in both groups and reassessed the rate of participant discontinuation as 25%; 454 participants were then needed to maintain statistical power.

The independent data safety and monitoring board (based in Paris, France) met every 12 months to review interim data for recruitment of patients and follow-up, safety, and effectiveness. An interim analysis of the increase in CD4 cell count and mortality was done on June 9, 2009, when 50% of randomly allocated participants had reached month 24. The data safety and monitoring board subsequently advised the investigators to continue the trial without any change.

We did all analyses for participants who were randomly allocated to monitoring groups and had at least one follow-up visit. No patients switched groups. We compared the difference between the two groups in terms of mean increase in CD4 cell count from baseline to 24 months by comparison of the lower limit of the 95% CI for the reported difference to the upper limit of the 95% CI for the non-inferiority margin (ie, -25% of the mean increase in CD4 cell counts in the LAB group). Because such an analysis with two bilateral 95% CIs resulted in an α coefficient of less than 5% (predefined value for the sample size calculation), we repeated it with 90% CIs to obtain an α close to 5%.

The primary analysis was done with all participants who attended at least one follow-up visit and a last-

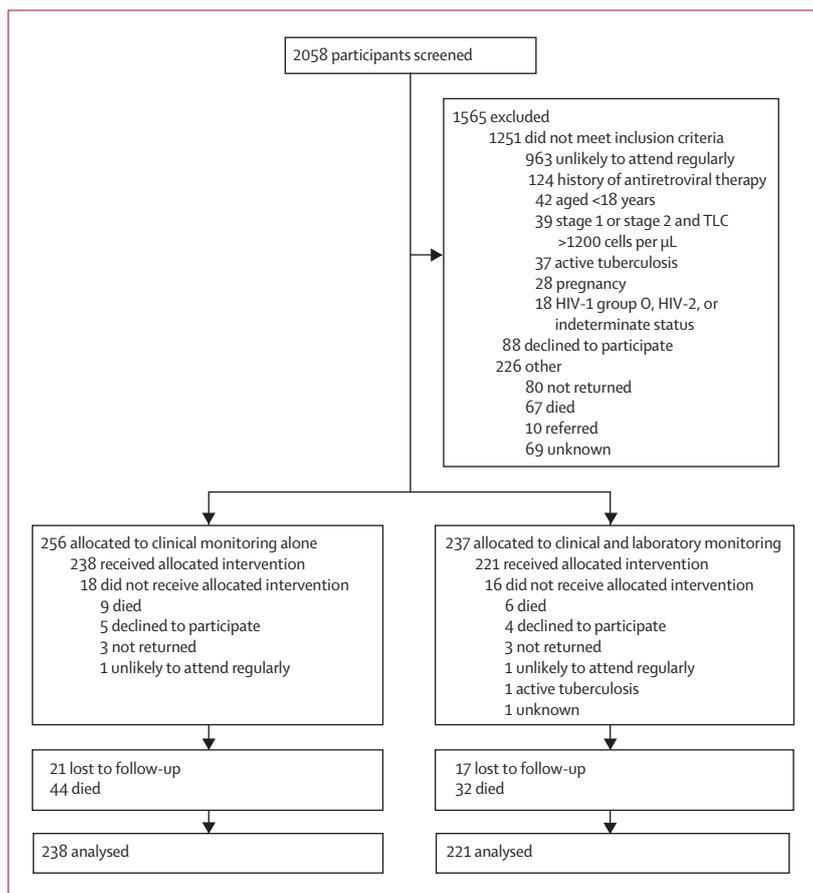


Figure 1: Trial profile
TLC=total lymphocyte count.

observation carried forward imputation was applied for participants whose CD4 data were missing at 24 months. The secondary analysis only involved participants with available CD4 data at 24 months. The proportion of participants with viral load of fewer than 40 copies per mL was compared between both groups with mixed-effect logistic regression models. The primary analysis regarded missing data as failure whereas the secondary analysis only included available data. We compared proportions at 24 months with standard logistic regression. We did survival analyses with the Kaplan-Meier method and Cox proportional hazard models. We double-entered and checked data with Microsoft Access and analysed data with Stata version 10.1.

This trial is registered with ClinicalTrials.gov, number NCT00301561.

Role of the funding source

The sponsors of the study had no role in study design, data collection, data analysis, data interpretation, or writing of 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.

Results

We randomly allocated 493 patients to intervention groups (figure 1); of whom 238 (93%) of 256 in the CLIN group and 221 (93%) of 237 in the LAB group attended at least one follow-up visit. The most frequent reason for ineligibility was presumed difficulty in regular attendance of the district hospital, especially for patients living far away (60%). Baseline characteristics were balanced between groups (table 1).

Median follow-up was 24 months (IQR 19–24) for patients in the CLIN group and 24 months (23–24) in the LAB group ($p=0.47$). Overall follow-up in the CLIN group was 385.9 person-years and 369.0 person-years in the LAB group. Completeness of visits was lower in the CLIN group (2106 [80%] of 2618 visits) than it was in the LAB group (2035 [84%] of 2431; $p=0.003$). 21 (9%) of

238 participants in the CLIN group and 17 (8%) of 221 participants in the LAB group were lost to follow-up (incidence of 5.2 per 100 person-years for the CLIN group and 4.3 per 100 person-years for the LAB group; incidence rate ratio [IRR] 1.20, 95% CI 0.59–2.47; $p=0.60$). In the CLIN group, 1087 (93%) of 1172 visits at months 0, 1, 6, 12, 18, and 24 were done by doctors (ie, 7% by nurses alone), and 736 (79%) of 934 visits at 2 weeks and 3, 9, 15, and 21 months were done by nurses alone (21% by doctors alone or with nurses). In the LAB group, 1846 (91%) of 2035 visits were done by doctors (9% by nurses alone).

37 (16%) of 238 participants in the CLIN group and 37 (17%) of 221 participants in the LAB group substituted at least one antiretroviral drug because of treatment-related adverse events, yielding an incidence of 10.8 per 100 person-years for CLIN and 11.1 per 100 person-years for LAB (IRR 0.97, 95% CI 0.60–1.58; $p=0.90$). 13 (6%) of 221 participants in the LAB group switched to second-line regimens after a median follow-up of 18 months (IQR 17–20; incidence 3.6 per 100 person-years) because of treatment failure, whereas no participants in the CLIN group did so ($p<0.0001$). Switching occurred after a median time of 8 months (IQR 6–10) from the first viral load of more than 5000 copies per mL. Overall, 33 (15%) of 221 participants in the LAB group had confirmed viral loads of more than 5000 copies per mL, of whom 22 (67%) had resistance (including the 13 participants who switched). Virological failure was due to adherence issues only for the 11 participants without resistance. Self-reported complete adherence, for the whole study period, was 63% (672 of 1067 measures) in the CLIN group and 61% (621 of 1011) in the LAB group ($p=0.47$) and, at 24 months, was 57% (91 of 161) in the CLIN group and 56% (87 of 155) in the LAB group ($p=0.94$).

In the primary analysis, the mean increase in CD4 cell count from baseline to month 24 was 175 cells per μL (SD 190, 95% CI 151–200) in participants in the CLIN group and 206 cells per μL (190, 181–231) in participants in the LAB group (figure 2); the mean difference was therefore -31 cells per μL (95% CI -63 to 2). The lower limit of the 95% CI for the reported difference was lower than was the upper limit of the 95% CI for the non-inferiority margin (-52 cells per μL , -58 to -45). The analysis on the basis of 90% CIs also did not show non-inferiority (-58 to -3 for the reported difference and -57 to -46 for the non-inferiority margin).

At month 24, 165 (69%) of 238 participants in the CLIN group and 169 (76%) of 221 participants in the LAB group had available CD4 cell counts and were therefore included in the secondary analysis. The mean increase in CD4 cell count was 220 cells per μL (SD 187, 95% CI 192–249) in the CLIN group and 245 cells per μL (187, 217–274) in the LAB group (figure 2). The lower limit of the 95% CI for the reported difference (mean -25 cells per μL , 95% CI -64 to 14) was also lower than the upper

	Clinical monitoring group (n=238)	Laboratory plus clinical monitoring group (n=221)
Sex		
Female	166 (70%)	158 (71%)
Male	72 (30%)	63 (29%)
Age (years)	36 (30–44)	37 (31–45)
Level of education*		
No formal education	5 (2%)	5 (2%)
Primary school	106 (46%)	95 (45%)
Secondary school or higher	117 (51%)	112 (52%)
Marital status†		
Married or cohabiting	78 (34%)	62 (29%)
Divorced or separated	5 (2%)	8 (4%)
Widowed	28 (12%)	31 (15%)
Single	118 (52%)	111 (53%)
Time taken to reach district hospital‡		
<30 min	110 (49%)	83 (40%)
30–60 min	60 (27%)	74 (36%)
>60–120 min	37 (16%)	31 (15%)
>120 min	18 (8%)	17 (8%)
Head of the household§		
No	138 (61%)	122 (58%)
Yes	88 (39%)	90 (42%)
Economically active¶		
No	99 (49%)	91 (47%)
Yes	105 (51%)	104 (53%)
Water supply*		
Private tap or mineral water	41 (18%)	27 (13%)
Public tap	14 (6%)	20 (9%)
Fitted well or bored well with pump	121 (53%)	121 (57%)
Purchase or unfitted well or spring	43 (19%)	41 (19%)
Other	8 (4%)	4 (2%)
WHO clinical stage		
2	1 (1%)	0
3	174 (73%)	163 (74%)
4	63 (26%)	58 (26%)
Bodyweight (kg)	55 (49–60)	55 (49–61)

(Continues on next page)

limit of the 95% CI for the non-inferiority margin (−61 cells per μL , 95% CI −68 to −54). The analysis done on the basis of 90% CIs again did not show non-inferiority (−58 to 7 for the reported difference and −67 to −56 for the non-inferiority margin).

During follow-up, viral load was fewer than 40 copies per mL in 465 (49%) of 952 measurements in the CLIN group and in 456 (52%) of 884 measurements in the LAB group, if missing data were regarded as failure (odds ratio [OR] 0.78, 95% CI 0.44–1.39, $p=0.40$; figure 2). At month 24, 111 (47%) of 238 participants in the CLIN group and 115 (52%) of 221 participants in the LAB group had an undetectable viral load (OR 0.81, 95% CI 0.56–1.16; $p=0.25$). In the secondary analysis done on the basis of available data only, the proportion of follow-up viral loads of fewer than 40 copies per mL was 65% in both groups (465 of 717 measurements in CLIN and 456 of 703 measurements in LAB; OR 0.96, 95% CI 0.61–1.52, $p=0.88$; figure 2). At month 24, the proportion of undetectable viral loads was 66% in participants in the CLIN group and 68% in participants in the LAB group (OR 0.93, 95% CI 0.59–1.47; $p=0.76$). Of the 13 participants in the LAB group who switched to second-line regimens, six had a viral load of more than 5000 copies per mL at month 24, one had a viral load of 2385 copies per mL, one had a viral load of 41 copies per mL, and three had a viral load of fewer than 40 copies per mL (one patient was lost to follow-up and testing was not done for the other patient).

Genotypic resistance was assessed in 34 (14%) of 238 participants in the CLIN group and 33 (15%) of 221 participants in the LAB group with viral loads of more than 5000 copies per mL during follow-up. Major mutations associated with antiretroviral drug resistance were detected from the baseline samples of two male participants in the LAB group and were excluded from further analyses. Analyses of the patients' last samples with high viral load showed the presence of major mutations in 23 (10%) of 238 participants in the CLIN group and 22 (10%) of 219 participants in the LAB group ($p=0.89$, table 2). All 45 participants had resistance to non-nucleoside reverse transcriptase inhibitors (NNRTIs), specifically nevirapine and efavirenz. 17 (7%) of 238 participants in the CLIN group and 18 (8%) of 219 participants in the LAB group also had resistance to NRTIs, specifically to lamivudine and emtricitabine. Groups did not differ in terms of the number of drugs to which the participants had resistance.

44 (18%) of 238 participants in the CLIN group died (11.5 deaths per 100 person-years), as did 32 (14%) of 221 in the LAB group (8.8 deaths per 100 person-years)—hazard ratio of 1.31 (95% CI 0.83–2.06, $p=0.25$; figure 3). Survival at month 24 was 81% (95% CI 75–85) in the CLIN group and 85% (80–89) in the LAB group.

85 (36%) of 238 participants in the CLIN group and 64 (29%) of 221 participants in the LAB group had

	Clinical monitoring group (n=238)	Laboratory plus clinical monitoring group (n=221)
(Continued from previous page)		
Body mass index (kg/m^2)	19.9 (18.4–21.8)	20.1 (18.2–22.1)
CD4 cell count (cells per μL)	179 (68–323)	182 (96–345)
HIV-1 viral load (\log_{10} copies per mL)	5.6 (5.3–6.1)	5.6 (5.0–6.0)
Haemoglobin (g/L)	98 (85–110)	96 (85–112)
Neutrophils (cells per μL)	1665 (1050–2890)	1750 (960–2700)
Platelets (cells per μL)	204 000 (157 000–268 000)	211 000 (169 000–278 000)
Aspartate aminotransferase (IU/L)	28 (20–42)	30 (20–44)
Alanine aminotransferase (IU/L)	19 (12–31)	20 (13–32)
Amylase (IU/L)	81 (60–108)	79 (60–105)
Creatinine ($\mu\text{mol}/\text{L}$)	80 (62–97)	80 (62–106)
Blood glucose (mmol/L)	5.0 (4.3–5.8)	5.1 (4.3–5.9)
Total cholesterol (mmol/L)	3.1 (1.9–4.2)	3.1 (1.8–4.3)
Triglycerides (mmol/L)	1.1 (0.6–1.6)	1.0 (0.6–1.7)
Initial antiretroviral regimen		
Stavudine, lamivudine, and nevirapine	153 (64%)	151 (68%)
Stavudine, lamivudine, and efavirenz	44 (18%)	38 (17%)
Zidovudine, lamivudine, and nevirapine	22 (9%)	12 (5%)
Zidovudine, lamivudine, and efavirenz	19 (8%)	20 (9%)
Co-trimoxazole prophylaxis	230 (97%)	211 (95%)

Data are n (%) or median (IQR). *Data missing for 19 participants. †Data missing for 18 participants. ‡Data missing for 29 participants. §Data missing for 21 participants. ¶Data missing for 60 participants. ||Data missing for one participant.

Table 1: Baseline characteristics of participants

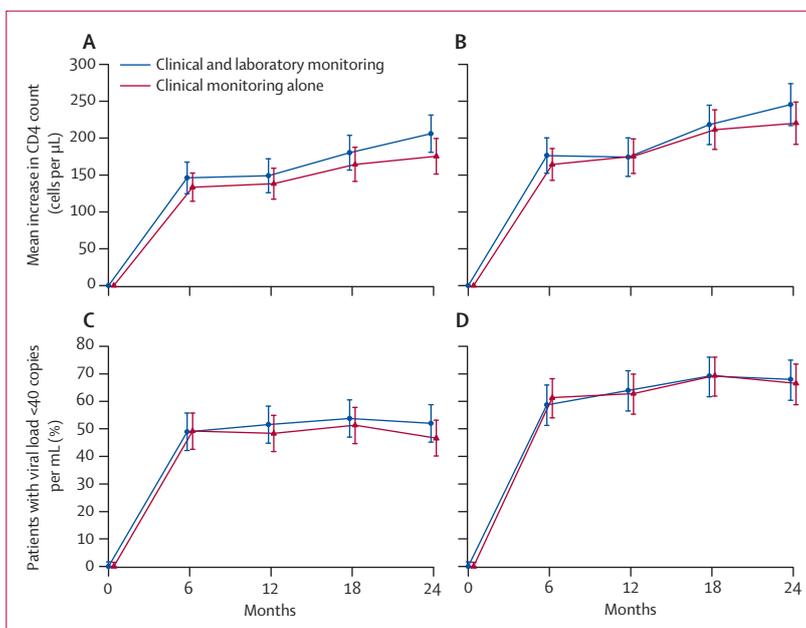


Figure 2: Immunovirological responses to antiretroviral therapy

Mean increase in CD4 cell count from the last-observation carried forward imputation method (A) and from reported data only (B). Proportion of participants with plasma HIV-1 viral load below 40 copies per mL, regarding missing data as failure (C) and from reported data only (D). Vertical bars are 95% CIs.

disease progression to new or recurrent WHO stage 3 or 4 adverse events or death; incidence was 26.6 per 100 person-years for CLIN and 19.9 per 100 person-years

	Clinical monitoring group (n=238)	Laboratory plus clinical monitoring group (n=219)
Overall	23 (10%)	22 (10%)
Non-nucleoside reverse transcriptase inhibitors	23 (10%)	22 (10%)
Nevirapine	3	0
Nevirapine and efavirenz	15	17
Nevirapine, efavirenz, and etravirine	5	4
Nevirapine, efavirenz, and etravirine*	0	1
Nucleoside reverse transcriptase inhibitors	17 (7%)	18 (8%)
Lamivudine/emtricitabine	11	15
Lamivudine/emtricitabine and abacavir	1	0
Lamivudine/emtricitabine, stavudine, and zidovudine	2	2
Lamivudine/emtricitabine, stavudine, zidovudine, abacavir, and tenofovir*	1	0
Lamivudine/emtricitabine, stavudine, zidovudine, abacavir, and didanosine	0	1
Lamivudine/emtricitabine, stavudine, abacavir, tenofovir, and didanosine*	2	0

Data are n (%) or n. *Possible resistance.

Table 2: Genotypic resistance to antiretroviral therapy

for LAB (HR 1.30, 95% CI 0.94–1.80, $p=0.11$; figure 3). Survival without new or recurrent WHO stage 3 or 4 adverse events at month 24 was 63% (95% CI 56–69) in the CLIN group and 69% (62–75) in the LAB group. 51 (21%) of 238 participants in the CLIN group and 40 (18%) of 221 participants in the LAB group had WHO stage 3 adverse events, whereas 18 (8%) of 238 participants and 12 (5%) of 221 participants, respectively, had WHO stage 4 adverse events. 28 (12%) of 238 participants in the CLIN group had WHO stage 3 or 4 adverse events after the first 6 months but none of these events persisted as defined for clinical failure.

Grade 3 or 4 toxic effects related to antiretroviral therapy were reported in 46 (19%) of 238 participants in the CLIN group (60 events; incidence 13.9 per 100 person-years) and 56 (25%) of 221 participants in the LAB group (69 events; 18.2 per 100 person-years); hazard ratio 0.77 (95% CI 0.52–1.14, $p=0.19$; figure 3). Laboratory side-effects predominated in both groups (52 [87%] of 60 events in the CLIN group and 58 [84%] of 69 events in the LAB group).

Discussion

Clinical monitoring alone is not non-inferior to clinical monitoring plus laboratory monitoring in terms of mean increase in CD4 cell count to 2 years. Moreover, more patients followed up with the LAB strategy switched to second-line treatment than did those in the CLIN group; antiretroviral resistance was detected in all cases. By contrast, the two monitoring strategies did not differ in terms of viral suppression, emergence of HIV resistance, mortality, disease progression, adherence, loss to follow-up, or incidence of toxic effects. Survival at 24 months was very similar between groups.

Thus, although the immunological recovery was good with both monitoring strategies compared with other studies in Africa,^{14–17} we cannot confirm that CLIN would not result in substantially reduced effectiveness when compared with that of LAB follow-up.¹⁸ Notably, the difference in methodology (non-inferiority test for immunological outcome vs superiority test for viral suppression) explains the apparent contradiction between the failure to show the non-inferiority of the CD4 cell count increase with CLIN and the equivalent virological effectiveness between the two strategies.

Detection of treatment failure with CLIN, leading to switching of antiretroviral therapy is a key issue. Although 12% of participants in the CLIN group had WHO stage 3 or 4 adverse events after the first 6 months and 10% of participants had HIV resistance, none was switched to a second-line regimen on the basis of our clinical definition—several possible reasons for this finding exist. First, the adverse events that occurred during the study were reversed because of adherence interventions combined with adequate care. Second, the patients with adverse events subsequently died or were lost to follow-up. Third, our study only lasted 2 years, but clinical diagnosis of treatment failure frequently occurs after 2 years. Finally, adverse events might have been underdiagnosed because of the difficulty of use of the WHO clinical staging of HIV disease for health-care staff who are inexperienced (most of them in our trial) or had little experience in HIV-related care¹⁹ and a heavy workload or little laboratory support. The poor sensitivity of clinical criteria for detection of treatment failure has already been reported, for example in South Africa.²⁰

In participants monitored with the LAB strategy, switching to second-line regimens occurred late after the first detectable viral load (median time 8 months) because of the difficulty for the health-care workers and participants to adapt the scheduled follow-up and because of laboratory constraints (ie, remoteness of our rural district hospitals from the laboratory in Yaoundé and logistical constraints). Furthermore, eight of 13 participants who switched to second-line regimens still had virological failure at month 24 (two were not assessed), probably because of adherence issues. Nonetheless, although the LAB strategy was not perfect, failure to show non-inferiority of CLIN suggests a potential benefit of LAB.

In the absence of LAB, especially monitoring of viral load, late clinical diagnosis of treatment failure leads to the accumulation of genotypic drug-resistant mutations that might compromise the efficacy of second-line regimens and favour the transmission of antiretroviral-resistant HIV strains.^{21,22} In our trial, the rates of resistance were similar between the two monitoring strategies but notably these results were only reported in the first 2 years. As expected in a trial (although the patients were managed by non-specialist health-care

workers in rural district hospitals), the rates of overall resistance (10%) were lower than were those in a routine HIV/AIDS outpatient clinic in Yaoundé (17% at month 24 in one study²³ and 16% after a median follow-up of 10 months in another²⁴).

Participants in both groups were reviewed by nurses instead of doctors in accordance with the WHO recommendation of task shifting.¹ This practice was unlikely to bias our comparison because participants in both groups were mostly reviewed by doctors, almost all doctors (and nurses) were initially inexperienced in HIV care, many doctors were transferred to other health facilities during the study and were replaced by others (who in turn were generally inexperienced), nurses were trained before and during the study with the doctors, and task shifting is effective.^{25–28} The joint management of patients by doctors and nurses in our trial suggests applicability of our findings to routine care.

In addition to our trial, the HBAC⁴ and DART⁵ trials have compared CLIN with different strategies that combine laboratory monitoring for ART in low-resource countries (panel). Nevertheless, our study is distinct for several reasons. First, viral load and CD4 cell count were measured every 6 months, as recommended by WHO.⁶ Second, participants were managed by local non-specialist health-care workers in rural district hospitals in which scale-up of antiretroviral therapy is underway. Third, antiretroviral therapy was started on the basis of WHO clinical criteria alone (ie, CD4 cell count was not used), in accordance with routine clinical practice in most settings. Finally, the emergence of HIV resistance was assessed.

Weaknesses of our study were the fairly short follow-up and small sample size. Inferiority of immunological recovery with clinical monitoring or significant differences between the two monitoring strategies for secondary outcomes (eg, disease progression) might have been reported with a longer follow-up or more participants. The exclusion of patients who were unlikely to attend regularly might reduce external validity because these patients are supposed to be less adherent and at higher risk of treatment interruption and HIV resistance; laboratory monitoring would therefore have allowed early detection of treatment failure and promote the switch to second-line regimens in a higher proportion of patients. However, increased adherence to these more complex regimens and subsequent follow-up is doubtful as shown by the fact that only three of 13 patients who switched to second-line regimens had undetectable viral loads at month 24. The large size of randomisation blocks could explain the difference in the number of participants between the clinical and laboratory groups. However, the groups were similar regarding all baseline characteristics. The study nurses and doctors were not blinded to study group; this is unlikely to have introduced any bias into assessment of outcomes (including the more subjective

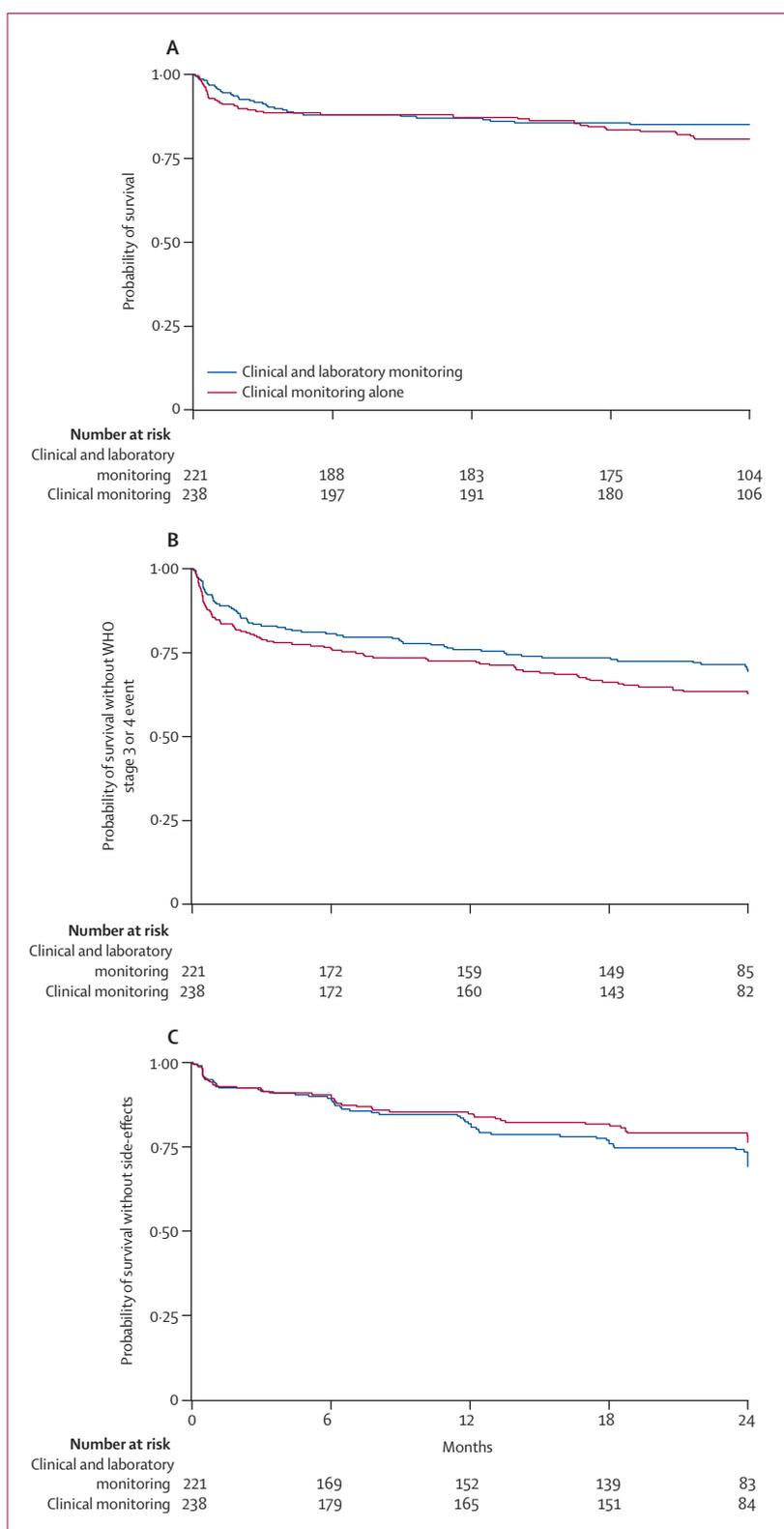


Figure 3: Kaplan-Meier curves of survival probability

Overall survival probability (A). Survival without new WHO stage 3 or 4 adverse events (B). Survival without grade 3 or 4 antiretroviral drugs-related side-effects (C).

Panel: Research in context**Systematic review**

We searched the Medline database from Aug 1, 1998, to Nov 23, 2010, without language restrictions, for randomised controlled trials that compared clinical monitoring of antiretroviral therapy with laboratory monitoring with the search terms “antiretroviral”, “laboratory” OR “biologic” OR “biological” OR “viral load” OR “CD4”, “monitoring”, “Africa” OR “resource-limited” OR “resource-poor” OR “resource-constrained” OR “developing”, and “trial”. We identified two randomised trials, the DART trial⁵ in Uganda and Zimbabwe and a cluster trial²⁹ in Zambia. We discussed the findings of the DART trial here, but the report of the second trial only included the study design, implementation, and baseline cohort characteristics. We also searched abstracts from major conferences and identified the home-based AIDS care (HBAC) trial⁶ in Uganda.

Interpretation

Our results support the WHO recommendation for laboratory monitoring of antiretroviral therapy when possible to improve quality of HIV care. This finding underlines the need for the development of simple and affordable point-of-care tests or alternative methods for viral-load measurement. In the meantime, the few overall differences that we noted between clinical and laboratory monitoring suggest that clinical monitoring alone could be used, at least temporarily, to expand antiretroviral therapy (which is the top priority) in low-resource settings that have financial and infrastructural constraints.

outcomes such as WHO stage 3 or 4 events) owing to local organisation.

Thus, the inability to show non-inferiority of immunological recovery and detect a need to switch to second-line treatment with CLIN supports the WHO recommendation for laboratory monitoring of antiretroviral therapy when possible to improve the quality of HIV care. Nevertheless, we suggest that the overall differences between the two strategies are small, and that CLIN could be used in patients who are antiretroviral therapy naive to expand treatment (which is the top priority) and to take into account financial and infrastructural constraints in low-resource settings. Some amount of laboratory monitoring at the programmatic scale should almost always be undertaken. Operational research (including laboratory measurements) to establish the overall success of the programmes is at least as crucial as provision of individual scale clinical monitoring.

Contributors

CL, CK, and ED designed and coordinated the trial. CL did the statistical analysis and wrote the first draft of the report. CK and GL-B coordinated the implementation of the trial. GL-B, SB, MPC, J-MM, MD, and SK contributed to data collection. AFA did the main laboratory analyses. JBTM and NM contributed to the data analysis of the primary outcome. All authors contributed to the interpretation of data and reviewed the report.

Stratall ANRS 12110/ESTHER study group

C Kouanfack, S Koulla-Shiro (Central Hospital, Yaoundé, Cameroon); A Bourgeois, E Delaporte, C Laurent, M Peeters (Institut de Recherche pour le Développement [IRD], University Montpellier 1, UMI 233, Montpellier, France); G Laborde-Balen (French Ministry of Foreign Affairs, Yaoundé, Cameroon); M Dontsop, S Kazé, J-M Mben (IRD, Yaoundé, Cameroon); A Aghokeng, M G Edoul, E Mpoudi-Ngolé, M Tongo (Virology Laboratory, IRD/IMP/CREMER, UMI 233, Yaoundé, Cameroon); J Blanche, S Boyer, M P Carrieri, F Marcellin, J-P Moatti, B Spire (INSERM, IRD, University Marseille, UMR 912, Marseille, France); C Abé, S-C Abega, C-R Bonono, H Mimcheu, S Ngo Yebga, C Paul Bile (IRSA, Catholic University of Central Africa, Yaoundé, Cameroon); S Abada, T Abanda, J Baga, P Bilobi Fouda, P Etong Mve, G Fetsé Tama, H Kemo, A Ongodo, V Tadewa, H D Voundi (District Hospital, Ayos, Cameroon); A Ambani, M Atangana, J-C Biaback, M Kennedy, H Kibedou, F Kounga, M Maguip Abanda, E Mamang, A Mikone, S Tang, E Tchuangue, S Tchuenco, D Yakan (District Hospital, Bafia, Cameroon); J Assandje, S Ebana, D Ebo'o, D Etoundi, G Ngama, P Mbarga Ango, J Mbezele, G Mbong, C Moug, N Ekotto, G Nguemba Balla, G Ottou, M Tigougmo (District Hospital, Mbalmayo, Cameroon); R Beyala, B Ebene, C Effemba, F Eyebe, M-M Hadjaratou, T Mbarga, M Metou, M Ndam, B Ngoa, E B Ngock, N Obam (District Hospital, Mfou, Cameroon); A M Abomo, G Angoula, E Ekassi, Essama, J J Lentchou, I Mvilongo, J Ngapou, F Ntokombo, V Ondoua, R Palawo, S Sebe, E Sinou, D Wankam, I Zobo (District Hospital, Monatélé, Cameroon); B Akono, A L Ambani, L Bilock, R Bilo'o, J Boombhi, F X Fouda, M Guitonga, R Mad'aa, D R Metou'ou, S Mgbih, A Noah, M Tadena, Ntcham (District Hospital, Nanga Eboko, Cameroon); G Ambassa Elime, A A Bonongnaba, E Foaleng, R M Heles, R Messina, O Nana Ndankou, S A Ngono, D Ngono Menounga, S S Sil, L Tchouamou, B Zambou (District Hospital, Ndikimeme, Cameroon); R Abomo, J Ambomo, C Beyomo, P Eloundou, C Ewole, J Fokom, M Mvoto, M Ngadena, R Nyolo, C Onana, A Oyie (District Hospital, Obala, Cameroon); and P Antyimi, S Bella Mbatonga, M Bikomo, Y Molo Bodo, S Ndi Ntang, P Ndoudoumou, L Ndzomo, S O Ngolo, M Nkengue, Nkoa, Y Tchinda (District Hospital, Sa'a, Cameroon).

Conflicts of interest

We declare that we have no conflicts of interest.

Acknowledgments

The results of this study have been presented orally (late breaker) at the 18th Conference on Retroviruses and Opportunistic Infections; Boston, MA, USA; Feb 27–March 2, 2011. We thank all patients and staff of the district hospitals who participated in the study, Vincent Le Moing (Department of Infectious and Tropical Diseases, University Hospital, Montpellier, France) for his helpful comments during the preparation of the paper, and Jude Sweeney (Rome, Italy) for the English revision and editing of the manuscript. The study was supported by a grant from the French National Agency for Research on AIDS (ANRS 12110) and Ensemble pour une Solidarité Thérapeutique Hospitalière En Réseau (ESTHER). Jules Brice Tchatchueng Mbougua was the recipient of a doctoral fellowship from the Institut de Recherche pour le Développement (IRD, France) and Marlise Dontsop was the recipient of a master's fellowship from ANRS.

References

- 1 Gilks CF, Crowley S, Ekpini R, et al. The WHO public-health approach to antiretroviral treatment against HIV in resource-limited settings. *Lancet* 2006; **368**: 505–10.
- 2 WHO. Towards universal access. Scaling up priority HIV/AIDS interventions in the health sector. Progress report 2010. http://whqlibdoc.who.int/publications/2010/9789241500395_eng.pdf (accessed Oct 13, 2010).
- 3 Phillips AN, Pillay D, Miners AH, Bennett DE, Gilks CF, Lundgren JD. Outcomes from monitoring of patients on antiretroviral therapy in resource-limited settings with viral load, CD4 cell count, or clinical observation alone: a computer simulation model. *Lancet* 2008; **371**: 1443–51.
- 4 Coutinho A, Mermin J, Ekwari J, et al. Utility of routine viral load, CD4 cell count, and clinical monitoring among HIV-infected adults in Uganda: a randomized trial. 15th Conference on Retroviruses and Opportunistic Infections; Boston, MA, USA; Feb 3–6, 2008.

- 5 Mugenyi P, Walker AS, Hakim J, et al. Routine versus clinically driven laboratory monitoring of HIV antiretroviral therapy in Africa (DART): a randomised non-inferiority trial. *Lancet* 2010; **375**: 123–31.
- 6 WHO. Antiretroviral therapy for HIV infection in adults and adolescents. Recommendations for a public health approach. 2010 revision. http://whqlibdoc.who.int/publications/2010/9789241599764_eng.pdf (accessed Oct 13, 2010).
- 7 Ferradini L, Jeannin A, Pinoges L, et al. Scaling up of highly active antiretroviral therapy in a rural district of Malawi: an effectiveness assessment. *Lancet* 2006; **367**: 1335–42.
- 8 WHO. Antiretroviral therapy for HIV infection in adults and adolescents: Recommendations for a public health approach. 2006 revision. <http://www.who.int/hiv/pub/guidelines/artadultguidelines.pdf> (accessed Aug 25, 2010).
- 9 French National Agency for Research on AIDS and Viral Hepatitis. ANRS scale to grade the severity of adverse events in adults. 2008. <http://www.anrs.fr/content/download/2242/12805/file/ANRS-GradeEI-VI-En-2008.pdf> (accessed Jan 27, 2011).
- 10 French National Agency for Research on AIDS and Viral Hepatitis. HIV-1 genotypic drug resistance interpretation's algorithms 2010. <http://www.hivfrenchresistance.org/tab2010.html> (accessed Jan 27, 2011).
- 11 Spire B, Duran S, Souville M, Lepout C, Raffi F, Moatti JP. Adherence to highly active antiretroviral therapies (HAART) in HIV-infected patients: from a predictive to a dynamic approach. *Soc Sci Med* 2002; **54**: 1481–96.
- 12 Moatti JP, Carrieri MP, Spire B, Gastaut JA, Cassuto JP, Moreau J. Adherence to HAART in French HIV-infected injecting drug users: the contribution of buprenorphine drug maintenance treatment. The Manif 2000 study group. *AIDS* 2000; **14**: 151–55.
- 13 Machin D, Campbell M, Fayers P, Pinol A. Sample size tables for clinical studies. Oxford, UK: E Blackwell Science, 1987.
- 14 Bourgeois A, Laurent C, Mougnotou R, et al. Field assessment of generic antiretroviral drugs: a prospective cohort study in Cameroon. *Antivir Ther* 2005; **10**: 335–41.
- 15 Laurent C, Meilo H, Guiard-Schmid JB, et al. Antiretroviral therapy in public and private routine health care clinics in Cameroon: lessons from the Douala antiretroviral (DARVIR) initiative. *Clin Infect Dis* 2005; **41**: 108–11.
- 16 Laurent C, Ngom Gueye NF, Ndour CT, et al. Long-term benefits of highly active antiretroviral therapy in Senegalese HIV-1-infected adults. *J Acquir Immune Defic Syndr* 2005; **38**: 14–17.
- 17 Ratsela A, Polis M, Dhlomo S, et al. A randomized factorial trial comparing 4 treatment regimens in treatment-naïve HIV-infected persons with AIDS and/or a CD4 cell count <200 cells/ μ L in South Africa. *J Infect Dis* 2010; **202**: 1529–37.
- 18 Hill A, Sabin C. Designing and interpreting HIV noninferiority trials in naïve and experienced patients. *AIDS* 2008; **22**: 913–21.
- 19 Kouanfack C, Laborde-Balen G, Aghokeng AF, et al. WHO clinical criteria-based initiation of antiretroviral therapy: lessons from rural district hospitals in Cameroon with regard to 2009 revised WHO recommendations. *Trop Med Int Health* 2010; **15**: 580–83.
- 20 Mee P, Fielding KL, Charalambous S, Churchyard GJ, Grant AD. Evaluation of the WHO criteria for antiretroviral treatment failure among adults in South Africa. *AIDS* 2008; **22**: 1971–77.
- 21 Gupta RK, Hill A, Sawyer AW, et al. Virological monitoring and resistance to first-line highly active antiretroviral therapy in adults infected with HIV-1 treated under WHO guidelines: a systematic review and meta-analysis. *Lancet Infect Dis* 2009; **9**: 409–17.
- 22 Hosseinipour MC, van Oosterhout JJ, Weigel R, et al. The public health approach to identify antiretroviral therapy failure: high-level nucleoside reverse transcriptase inhibitor resistance among Malawians failing first-line antiretroviral therapy. *AIDS* 2009; **23**: 1127–34.
- 23 Kouanfack C, Montavon C, Laurent C, et al. Low levels of antiretroviral-resistant HIV infection in a routine clinic in Cameroon that uses the World Health Organization (WHO) public health approach to monitor antiretroviral treatment and adequacy with the WHO recommendation for second-line treatment. *Clin Infect Dis* 2009; **48**: 1318–22.
- 24 Laurent C, Kouanfack C, Vergne L, et al. Antiretroviral drug resistance and routine therapy, Cameroon. *Emerg Infect Dis* 2006; **12**: 1001–04.
- 25 Sanne I, Orrell C, Fox MP, et al. Nurse versus doctor management of HIV-infected patients receiving antiretroviral therapy (CIPRA-SA): a randomised non-inferiority trial. *Lancet* 2010; **376**: 33–40.
- 26 Shumbusho F, van Griensven J, Lowrance D, et al. Task shifting for scale-up of HIV care: evaluation of nurse-centered antiretroviral treatment at rural health centers in Rwanda. *PLoS Med* 2009; **6**: e1000163.
- 27 Bedelu M, Ford N, Hilderbrand K, Reuter H. Implementing antiretroviral therapy in rural communities: the Lusikisiki model of decentralized HIV/AIDS care. *J Infect Dis* 2007; **196** (suppl 3): S464–68.
- 28 Jaffar S, Amuron B, Foster S, et al. Rates of virological failure in patients treated in a home-based versus a facility-based HIV-care model in Jinja, southeast Uganda: a cluster-randomised equivalence trial. *Lancet* 2009; **374**: 2080–89.
- 29 Koethe JR, Westfall AO, Luhanga DK, et al. A cluster randomized trial of routine HIV-1 viral load monitoring in Zambia: study design, implementation, and baseline cohort characteristics. *PLoS One* 2010; **5**: e9680.

AnnexeD :article« Nonlinear multiple imputation for continuous covarite within semi-parametric Cox model : application to HIV data in Senegal. »

Nonlinear multiple imputation for continuous covariate within semi-parametric Cox model: application to HIV data in Senegal

Jules Brice Tchatchueng Mbougua^{1,2}, Christian Laurent¹, Ibra Ndoye³, Eric Delaporte^{1,4}

Henri Gwet², Nicolas Molinari⁵

¹Institut de Recherche pour le Développement (IRD), Université Montpellier 1, UMI 233, Montpellier, France.

²Ecole Nationale Supérieure Polytechnique (ENSP), Université Yaoundé 1, Yaoundé, Cameroun.

³National AIDS Program, Ministry of Health, Dakar, Senegal

⁴Department of Infectious Diseases, University Hospital, Montpellier, France

⁵Université Montpellier 1, CHU Montpellier-Nîmes, UMR 729 MISTEA, Montpellier, France.

SUMMARY

Multiple imputation is commonly used to impute missing covariate in Cox semi-parametric regression setting. It is to fill each missing data with more plausible values, via a Gibbs sampling procedure, specifying an imputation model for each missing variable. This imputation method is implemented in several softwares which offer imputation models steered by the shape of the variable to be imputed; but all these imputation models make an assumption of linearity on covariates effect. However this assumption is not often verified in practice as the covariates can have a nonlinear effect in imputation model. Such a linear assumption can lead to a misleading conclusion because imputation model should be constructed to reflect the true distributional relationship between the missing values and the observed values. In order to estimate nonlinear effects of continuous covariates in imputation model, we propose a method based on B-splines function. To assess the performance of this method, we conducted a simulation study, where we compared the multiple imputation method using Bayesian splines imputation model to multiple imputation using Bayesian linear imputation model in survival analysis setting. We evaluated the proposed method on the motivated data set collected in HIV-infected patients enrolled in an observational cohort study in Senegal, which contains several incomplete variables. We found that our method performs well estimate hazard ratio compared to the linear imputation methods, when data is missing completely at random (MCAR), or missing at random (MAR).

1. INTRODUCTION

Partially observed information frequently occurs in epidemiological studies of chronic diseases, like AIDS, particularly in countries with limited resources. Several reasons can lead to missing data, namely exams not performed, data entry errors, outliers, or even discontinuation. If the missing information is not taken into account in statistical analysis then, misleading conclusion can be taken because validity and efficiency of complete-data based methods cannot be guaranteed [1]. It may be necessary in this context to implement a method for imputing missing data.

The impact of missing data and ways to handle incomplete data depend much upon the patterns of incompleteness. Three observation processes can be distinguished: data missing completely at random (MCAR), when the conditional probability of incomplete observation given data is independent of any characteristics of the subjects (it's constant); data missing at random (MAR), when the conditional probability of incomplete observation given data can be specified using observed variables; and data not missing at random (NMAR), when the conditional probability of incomplete observation given data depends on the unobserved data. It is worth noting that all proposed imputation methods do not provide correct statistical inference under the NMAR hypothesis [2, 3].

There are several methods to deal with missing data, including complete cases analysis [4, 5], single imputation methods (mean substitution, nearest neighbor, etc) and multiple imputation. However, the single imputation method does not take into account appropriately the uncertainty associated with missing data. Multiple imputation is a simple and efficient method for analyzing incomplete observations, based on the generation of several plausible complete data sets using an imputation model, while taking into account all levels of uncertainty regarding missing values.

Multiple imputation method developed by Little and Rubin [1] provide an approximation of the maximum likelihood estimator. A prediction model (imputation model) must be built, taking into consideration the relationships between the incomplete variable and the other variables, as observed in the complete part of the data. The missing data is not replaced by their expectation, but by a random sample of plausible values drawn from the distribution posited by the imputation model to generate several completed data. Each completed data is analyzed by using standard procedure. The results from each complete data are combined to construct the multiple imputation estimator. When the imputation model is correctly defined, the multiple imputation estimator is asymptotically unbiased and its variance is correctly estimated. There are many bibliographic references and review papers [6-8]. Multiple imputation for missing data in epidemiological and clinical research is presented in [5, 9-13].

For multivariate multiple imputation, two general parametric approaches have emerged: joint modeling [2] and Fully Conditional Specification (FCS) [14]. Joint modeling involves specifying a multivariate distribution for the missing data, and drawing imputation from their conditional distributions by Markov Chain Monte Carlo (MCMC) techniques. This

methodology is attractive if the multivariate distribution is a reasonable description of the data. By contrast FCS specifies the multivariate imputation model on a variable-by-variable basis, by a set of conditional densities, one for each incomplete variable. FCS is attractive as an alternative to joint modeling in common cases where no suitable multivariate distribution can be found.

Thus, the choice of the imputation model determines the validity of the inference based on multiple imputation method. Both the structural part of the imputation model and the error distribution must be well specified. To reflect the true distributional relationship between the missing values and the observed values, the choice of covariates in structural part of the imputation model being crucial [7, 11, 12, 15] the true relationships between imputed variable and covariates in structural part of imputation model is generally unknown in practice. So, the choice of the relationships between imputed variable and covariates in structural part of imputation model is also important. Several authors have implemented multivariate multiple imputation method by FCS in several software and offer several imputation models, but all these models make an assumption of linear relation between the covariates and imputed variable [16, 17]. This assumption is not often verified in practice, the covariates can have an arbitrary relation.

In survival analysis setting, many approaches such as likelihood based approach[18] and imputation approaches [19] have been proposed for handling missing covariates when fitting a Cox proportional hazards model. But likelihood based approach generally requires problem-specific programs and therefore is not generally readily available. In this paper, we considered analysis of a censored survival outcome using Cox proportional hazard model with one or more incomplete covariates and we used multivariate multiple imputation by FCS approach to handle incomplete covariates. Each imputation model involves regression of an incomplete variable on the other covariates or survival outcome.

The objective of this study is to establish multiple imputation in the Cox model without any linearity assumption. We propose a method based on B-splines function, to estimate arbitrary continuous covariates effects in imputation model. Such use is very appropriate [20], and can be seen as an extension of the linear assumption.

Our work will be presented as follows: the next section presents the HIV data set that motivated this work. Section 3 presents multiple imputation method with extension of imputation model taking into account nonlinear continuous covariates effects. In section 4 we present the results of the simulation study, in order to compare our method with the one widely available in standard software, based on linear imputation models. Application on motivated data set is presented in section 5, and in section 6 we end with a discussion on this method in the setting of survival analyses.

2. THE DATA

This work was motivated by a study of the long-term effectiveness and safety of a first-line antiretroviral treatment including didanosine, lamivudine, and efavirenz or nevirapine in HIV-infected patients monitored in Senegal [21]. This regimen was compared to a WHO-recommended regimen of zidovudine, lamivudine, and efavirenz or nevirapine. A total of 151 patients were included, of whom 71 received the didanosine-based treatment and 80 received the zidovudine-based treatment. The total observation time was 943 person-years, and median follow-up time was 7.1 years (interquartile range [IQR] 5.1-7.9). Twenty-seven patients died during follow-up (16 in the didanosine group and 11 in the zidovudine group).

In this study, most baseline covariates showed missing data, proportion of missing values ranging from one to 26 per cent (Table 1). Overall, 47 per cent of the sample population had at least one missing data (the sample size would be reduced from 151 to 86). An initial complete case analysis of survival showed that baseline viral load was a confounder. Several long-term survival studies in HIV-infected patients showed that baseline viral load is a risk factor for mortality [22, 23]. It was, therefore, important to impute the baseline covariates with missing data, so as to better understand the effect of these covariates on the long-term survival, and also to give more power to our study. These data were initially imputed and then analyzed [21], using a linear multiple imputation method. In this work, we will use this data set to illustrate our proposed method.

3. MULTIPLE IMPUTATION BY SPLINES

Let Y be an $n \times p$ data matrix, where $Y = (Y_1, \dots, Y_p)$. The observed and missing part of vector Y_j with $j = 1, \dots, p$ are denoted by Y_j^{obs} and Y_j^{miss} , respectively, so $Y^{obs} = (Y_1^{obs}, \dots, Y_p^{obs})$ and $Y^{miss} = (Y_1^{miss}, \dots, Y_p^{miss})$ are the observed and missing part of data in Y , $Y_{-j} = (Y_1, \dots, Y_{j-1}, Y_{j+1}, \dots, Y_p)$. Let T and D be respectively vector of times to follow-up, and vector of event indicator, T and D or a function of T can be components of the predictor matrix of imputation model. Let m be number of values filled in for each missing value.

Let $P(Y, \theta)$ the p -multivariate distribution of Y . We assume that $P(Y, \theta)$ is completely specified by θ , a vector of unknown parameters. In order to address the issues posed by the real-life complexities of the data, we used FCS method, commonly named chained equations method. This method defines implicitly $P(Y, \theta)$, by specifying a separate conditional density $P(Y_j | Y_{-j}, \theta_j)$ for each Y_j , the parameters $\theta_1, \dots, \theta_p$ are specific to the respective conditional densities. The conditional probability distribution of Y_j , given Y_j^{obs} and Y_{-j} can be derived by integrating over the parameter space of θ_j .

$$P(Y_j | Y_j^{obs}, Y_{-j}) = \int P(Y_j | Y_j^{obs}, Y_{-j}, \theta_j) P(\theta_j | Y_j^{obs}, Y_{-j}) d\theta_j$$

The posterior predictive distribution $P(Y_j | Y_j^{obs}, Y_{-j})$ can rarely be expressed in a closed form, due to the integral form in the above equation. So, it will be difficult to draw samples directly from this distribution. However, the conditional predictive distribution $P(Y_j | Y_j^{obs}, Y_{-j}, \theta_j)$ is often easy to obtain once the observed data and the values of the parameters θ_j are given. Therefore, if the values of the parameters θ_j can be drawn from their posterior distribution $P(\theta_j | Y_j^{obs}, Y_{-j})$, then the corresponding draws from the conditional predictive distribution $P(Y_j | Y_j^{obs}, Y_{-j}, \theta_j)$ are the draws from the posterior predictive distribution $P(Y_j | Y_j^{obs}, Y_{-j})$. The predictors Y_{-j} used in the imputation model may themselves be incomplete, which leads to an iterative algorithm with two steps: an estimation step and an imputation step. In the estimation step, an imputation model is separately specified for each covariate, involving the other covariates as predictors. In the imputation step, an imputation is generated for the missing variable, and this imputed value is used for the imputation of the next covariate. Starting from a simple draw from observed marginal distributions, the t^{th} iteration of this iterative algorithm is a Gibbs sampler that successively draws.

$$\begin{aligned}
\theta_1^{*(t)} &\sim P(\theta_1 | Y_1^{obs}, Y_2^{(t-1)}, \dots, Y_p^{(t-1)}) \\
Y_1^{miss*(t)} &\sim P(Y_1 | Y_1^{obs}, Y_2^{(t-1)}, \dots, Y_p^{(t-1)}, \theta_1^{*(t)}) \\
\theta_2^{*(t)} &\sim P(\theta_2 | Y_2^{obs}, Y_1^{(t)}, Y_3^{(t-1)}, \dots, Y_p^{(t-1)}) \\
Y_2^{miss*(t)} &\sim P(Y_2 | Y_2^{obs}, Y_1^{(t)}, Y_3^{(t-1)}, \dots, Y_p^{(t-1)}, \theta_2^{*(t)}) \\
\theta_p^{*(t)} &\sim P(\theta_p | Y_p^{obs}, Y_1^{(t)}, Y_2^{(t)}, \dots, Y_{p-1}^{(t)}) \\
Y_p^{miss*(t)} &\sim P(Y_p | Y_p^{obs}, Y_1^{(t)}, Y_2^{(t)}, \dots, Y_{p-1}^{(t)}, \theta_p^{*(t)})
\end{aligned}$$

Where $Y_j^{(t)} = (Y_j^{obs}, Y_j^{miss*(t)})$ is the j^{th} imputed variable at iteration t , and $Y_k = Y_k^{(t)}$ if Y_k is a complete variable for k in $\{1, \dots, p\}$. Convergence of this Gibbs sampler is quite fast, the number of iteration can be often between 10-20 [16].

In this paper, we are particularly interested by multiple imputation procedure in survival analysis setting. We use a semi-parametric Cox model. Let be hazard function defined by:

$$h(t, Y_1^*, \dots, Y_p^*) = h_0(t) \exp \left\{ \sum_{i=1}^p \alpha_i Y_i^* \right\}$$

where $Y_j^* = (Y_j^{obs}, Y_j^{miss*})$ is a variable to be competed for j in $\{1, \dots, p\}$ and $h_0(t)$ baseline hazard function.

Several imputation models have been proposed in the literature to predict Y_j^{miss*} given Y_j^{obs} and predictor matrix Y_{-j} using the Gibbs sampler scheme beyond, but none of these models hitherto does take into account the nonlinear effects of continuous predictors. We propose to relax linear assumption and retain independence by using an additive imputation model. At the

t^{th} iteration of the Gibbs sampler scheme beyond, we used the following imputation model for j^{th} imputed variable.:

$$Y_j^{(t)} = \sum_{l=1}^{j-1} f_l(Y_l^{(t)}) + \sum_{l=j+1}^p f_l(Y_l^{(t-1)}) + f_{p+1}(T) + \beta_j^{(t)} * D + \varepsilon_j^{(t)}$$

Where $\varepsilon_j^{(t)} = (\varepsilon_{1j}^{(t)}, \dots, \varepsilon_{nj}^{(t)})$, $\varepsilon_{ij}^{(t)} \sim \mathcal{N}(0, \sigma_j^{2(t)})$ for $i = 1, \dots, n$, $f_l(\cdot)$ is an unspecified function of $Y_l^{(t)}$, $Y_l^{(t-1)}$ or T , $l \neq j$, reflecting the true relationship between continuous covariates and imputed variable $Y_j^{(t)}$ at iteration t . To simplify notations, in what follows we will denote T by Y_{p+1} , $T = Y_{p+1} = Y_{p+1}^{(t-1)}$ since T isn't an imputed variable. Thus, imputation model above can be written as:

$$Y_j^{(t)} = \sum_{l=1}^{j-1} f_l(Y_l^{(t)}) + \sum_{l=j+1}^{p+1} f_l(Y_l^{(t-1)}) + \beta_j^{(t)} * D + \varepsilon_j^{(t)}$$

We used splines functions with fixed knots and degree, to approximate the unknown function $f_l(\cdot)$. Splines functions refer to the families of interpolation or smoothing functions with optimal regularity properties. Using of splines in imputation model allows the investigation of non-linear effects of continuous covariates. In fact a splines function belongs to a finite dimensional linear space. To choose a basis for this linear space, B-splines basis functions are very appropriate due to the fact that they are numerically well conditioned and also because they achieve a local sensitivity to data.

To determine an approximation of $f_l(\cdot)$ by splines using B-splines function basis, we assumed that $Y_l^{(t)}$ takes their values in interval $[a^{l(t)}; b^{l(t)}]$. Let, let $\xi^{l(t)} = (\xi_1^{l(t)}, \dots, \xi_{k^l}^{l(t)})$ be a vector named *knots vector* of k^l distinct points, such as $(\xi_0^{l(t)} =) a^{l(t)} < \xi_1^{l(t)} < \xi_2^{l(t)} < \dots < \xi_{k^l}^{l(t)} < b^{l(t)} (= \xi_{k^l+1}^{l(t)})$ with $k^l \in \mathbb{N}$. Splines approximation function of $f_l(Y_l^{(t)})$ denoted $f_l(Y_l^{(t)}, \xi^{l(t)}, d^l)$, is a polynomial function of degree d^l on any interval $[\xi_i^{l(t)}; \xi_{i+1}^{l(t)}]$ with $i = 0, \dots, k^l$, and has $d^l - 1$ continuous derivatives on the open interval $]a^{l(t)}; b^{l(t)}[$, $d^l \in \mathbb{N}^*$. For a fixed sequence of knots $\xi^{l(t)} = (\xi_1^{l(t)}, \dots, \xi_{k^l}^{l(t)})$, the set of such splines is a linear space of functions with $k^l + 1 + d^l$ free parameters [20]. A useful basis $\{B_r(\cdot, \xi^{l(t)})\}_{r=1, \dots, k^l+1+d^l}$ for this linear space is given by Schoenberg's B-splines, or Basic-splines [24]. So, we can write splines approximation function of $f_l(Y_l^{(t)})$ in B-spline basis as:

$$f_l(Y_l^{(t)}, \xi^{l(t)}, d^l) = \sum_{r=1}^{k^l+1+d^l} \beta_r^{(t)} B_r(Y_l^{(t)}, \xi^{l(t)}, d^l)$$

And the proposed imputation model as:

$$Y_j^{(t)} = \sum_{l=1}^{j-1} \sum_{r=1}^{k^l+1+d^l} \beta_{j_r}^{(t)} B_r \left(Y_l^{(t)}, \xi^{l(t)}, d^l \right) + \sum_{l=j+1}^{p+1} \sum_{r=1}^{k^l+1+d^l} \beta_{j_r}^{(t)} B_r \left(Y_l^{(t-1)}, \xi^{l(t)}, d^l \right) + \beta_{j_D}^{(t)} * D + \varepsilon_j^{(t)}$$

Where $\beta_{j_r}^{(t)}$ and $\beta_{j_D}^{(t)}$ are regression parameters at t^{th} iteration.

We used a Bayesian approach to estimate the vector of regression parameters $\beta_j^{(t)} = \left((\beta_{j_r}^{(t)})_{r=1, \dots, k^l+1+d^l}^{l=1, \dots, p+1}; \beta_{j_D}^{(t)} \right)$ and variance parameter $\sigma_j^{2(t)}$, $\theta_j^{(t)} = \left(\beta_j^{(t)}, \sigma_j^{2(t)} \right)$. We assumed a non-informative prior distribution for $\beta_j^{(t)}$ and $\sigma_j^{2(t)}$, $P(\beta_j^{(t)}) \propto 1$ and $P(\sigma_j^{2(t)}) \propto 1$. We calculated their posterior distribution and obtained:

$$\begin{aligned} \sigma_j^{2(t)} | Y_j^{(t-1)}, B^{(t)}(Y_{-j}) &\sim \mathcal{IG} \left(\frac{n - R_j}{2}, \frac{n - R_j}{2} S_{n-R_j} \right) \\ \beta_j^{(t)} | \sigma_j^{2(t)}, Y_j^{(t-1)}, B^{(t)}(Y_{-j}) &\sim \mathcal{N} \left(\begin{aligned} & \left[B^{(t)}(Y_{-j})^T B^{(t)}(Y_{-j}) \right]^{-1} B^{(t)}(Y_{-j})^T Y_j^{(t-1)} \\ & \sigma_j^{2(t)} \left[B^{(t)}(Y_{-j})^T B^{(t)}(Y_{-j}) \right]^{-1} \end{aligned} \right) \end{aligned}$$

where $R_j = 1 + \sum_{l=1; l \neq j}^{p+1} (k^l + 1 + d^l)$,

$$\begin{aligned} B^{(t)}(Y_{-j}) &= \left(\left[B_r \left(Y_l^{(t)}, \xi^{l(t)}, d^l \right) \right]_{r=1, \dots, k^l+1+d^l}^{l=1, \dots, j-1}, \left[B_r \left(Y_l^{(t-1)}, \xi^{l(t+1)}, d^l \right) \right]_{r=1, \dots, k^l+1+d^l}^{l=j+1, \dots, p}, D \right) \\ S_{n-R_j} &= \frac{1}{n - R_j} \left(Y_j^{(t-1)} - B^{(t)}(Y_{-j}) \left[B^{(t)}(Y_{-j})^T B^{(t)}(Y_{-j}) \right]^{-1} B^{(t)}(Y_{-j})^T Y_j^{(t-1)} \right)^T \\ &\quad * \left(Y_j^{(t-1)} - B^{(t)}(Y_{-j}) \left[B^{(t)}(Y_{-j})^T B^{(t)}(Y_{-j}) \right]^{-1} B^{(t)}(Y_{-j})^T Y_j^{(t-1)} \right) \end{aligned}$$

So we can draw m samples of $Y_j^{miss^{(t)}}$, at t^{th} iteration from the following sampler scheme:

- Draw $\sigma_j^{*2(t)}$ from its posterior distribution
- Draw $\beta_j^{*(t)}$ from its posterior distribution
- Draw $Y_j^{miss^{(t)}} \sim \mathcal{N} \left(B^{(t)}(Y_{-j}) \beta_j^{*(t)}, \sigma_j^{*2(t)} \right)$

Thereby generating m complete data set. Analysis is performed on each data set separately and then combined by Rubin's rules [1].

4. SIMULATIONS STUDY

To assess the performance of the method proposed, we conducted a simulation study comparing the multiple imputation method by splines presented above to multiple imputation method using Bayesian linear imputation model, in survival analysis setting. We

consider the simple case where there is a single continuous variable to impute Y_3 , which is explained by two other complete and independent covariates Y_1 and Y_2 .

4.1. Study design

We generated three variables Y_1 , Y_2 and Y_4 from normal distribution with $m_{Y_1}, m_{Y_2}, m_{Y_4}, \sigma_{Y_1}, \sigma_{Y_2}, \sigma_{Y_4}$ respectively mean and standard deviation of Y_1 and, Y_2 and Y_4 . The variable $Y_3 = f(Y_1, Y_2)$ where $f(\cdot)$ is the true relationship between Y_3 and Y_1, Y_2 generally unknown in practice. Survival times T^* were drawn from exponential distribution $h_{T^*}(t) = \lambda_{T^*} \exp(\beta_{Y_1} Y_1 + \beta_{Y_2} Y_2 + \beta_{Y_3} Y_3)$, and random censored times C were also drawn from exponential distribution $h_c(t) = \lambda_c [25]$, thus the vector of time to follow-up was obtained by $T = \min(T^*, C)$ and event indicator vector D by $D_i = 1$ if $T^* \leq C$ for i^{th} observation and zero else. The next step consisted to generate missing values on Y_3 .

To explore how the missing data mechanism affects the results, two mechanisms were considered to generate missing values: MCAR and MAR. Assuming MCAR, the probability of incomplete observation is constant, we took this probability equal to c . And under the MAR assumption, we took probability that a value of the variable of interest is missing equal to $\frac{\exp(\alpha_4 Y_4)}{1 + \exp(\alpha_4 Y_4)}$.

Sample size of the simulated data was $n = 500$, and number of simulations was 1000 to ignore the sampling variation. For each missingness mechanism considered above we used three imputation models $MCAR_i$ and MAR_i for $i = 1, 2, 3$. Respectively, the first imputation models considered ($MCAR_1$ and MAR_1) did not include response variables (T and D), the second imputation model ($MCAR_2$ and MAR_2) included the covariates T and D , the third model ($MCAR_3$ and MAR_3) included $\log(T)$ and D as covariates. To avoid the identifiability problem in Bayesian splines imputation model, we centered the covariates on their mean, and remove intercept in splines function. For each simulated data, we used multiple imputation methods as presented above for each missingness mechanisms and linear imputation models to construct $m = 10$ imputed data set. The analysis model was a Cox regression model, and results from the imputed data sets were combined using Rubin's rules.

To evaluate the performance of our analysis method, we compared the results of an analysis based on multiple imputation using Bayesian linear imputation model and an analysis based on multiple imputation using methods described above (we limited the study to the splines of degree two and with two knots). Comparisons were based on the following criteria:

- *Bias of the estimates*: the average difference between estimate and true specified parameter value.
- *The accuracy*: we used means square error to assess the accuracy.
- *The coverage of the confidence interval*: the proportion of time that the obtained confidence interval contains the true specified parameter value.

- *The empirical power of test:* the proportion of simulation samples in which the null hypothesis of no effect is rejected[26].

We will consider successively the cases where the true relationship between imputed variables Y_3 and explanatory variables Y_1, Y_2 is polynomial and nonlinear effects.

4.2. Polynomial case

In this case, $Y_3 = \alpha_1 Y_1^2 + \alpha_2 Y_1 + \alpha_3 Y_2^3 + e_1$, where e_1 represent the measurement error, $e_1 \sim N(0; 0.01)$. The parameter values used were: $m_{Y_1} = 0.2, m_{Y_2} = m_{Y_4} = 1, \sigma_{Y_1} = \sigma_{Y_4} = \alpha_1 = 1, \sigma_{Y_2} = 2, \alpha_4 = 0.5, \alpha_2 = -2, \alpha_3 = 2, \beta_{Y_1} = \beta_{Y_2} = 1, \beta_{Y_3} = -0.5, c = 0.5, \lambda_T = \lambda_C = 2$. In sensitivity analysis overall proportion of missing values was fixed: 30, 40 and 50 per cent (result not shown for 50 per cent of missing values)

For 30 and 40 per cent of missing data, the results of the analysis based on multiple imputations by Bayesian linear imputation model given in Table 2 lead to biased estimates, when the missing mechanism is MCAR or MAR. The introduction of the response variables (T and D) has improved the bias and accuracy. Under coverage shows that simulations have led to estimates significantly different from true values.

However, with the analysis based on multiple imputation by Bayesian splines imputation model (Table 2), the coverage was greater, but remain unacceptable. The introduction of response variable does not improve the evaluation criteria. The analysis based on multiple imputation by Bayesian splines imputation model are generally the most satisfactory ones, in term of bias. The obtained results were consistent with the increase of the proportion of missing data up to 50 per cent.

4.3. Nonlinear case

In this case, $Y_3 = \sqrt{e^{Y_1}} + \frac{1}{\sqrt{e^{(Y_2^2-1)}}} + e_1$, where e_1 represent the measurement error, $e_1 \sim N(0; 0.01)$. The parameter values used were : $m_{Y_1} = 0.2, m_{Y_2} = m_{Y_4} = 1, \sigma_{Y_1} = \sigma_{Y_4} = \beta_{Y_1} = \beta_{Y_2} = 1, \sigma_{Y_2} = 2, \alpha_4 = 0.5, \beta_{Y_3} = -1, c = 0.5, \lambda_T = \lambda_C = 2$. In sensitivity analysis overall proportion of missing values was fixed: 30, 40 and 50 per cent (result not shown for 50 per cent of missing values) and overall location knots and degree of splines have been regarded in Bayesian splines imputation model.

Concerning the estimated parameters of survival analysis model, we have unsurprisingly noticed that the obtained results were similar to the previous case (Table 3). In multiple imputation based on Bayesian linear imputation model, the bias and accuracy of the estimators obtained towards the null when we introduced the response variables in the imputation model. Empirical power and coverage were small for estimates associated with

variables Y_2 and Y_3 . This under coverage shows that more simulations have led to estimates significantly different from true values.

However, in multiple imputation method based on Bayesian splines imputation model, the bias and accuracy tend towards zero. The coverage was approximately equal to the nominal coverage (95 per cent). The introduction of the response variables does not improve the evaluation criteria.

We also used a log transformation of the imputed variable in multiple imputation procedure using Bayesian linear imputation model, but the results remained similar to the results obtained with Bayesian linear multiple imputation without transformation (Table 3).

In this analysis the results remain consistent with the increase of the proportion of missing data

We conducted a sensitivity analysis on degree and location knots of splines function used for Y_1 and Y_2 in Bayesian splines imputation model. In this analysis we generated missing values on Y_3 (30% of missing data) following MAR mechanism and in multiple imputation procedure we included response variables (T , D) as predictors in imputation model. The obtained results (Table 4), remain consistent regardless the degree and location knots of the splines function used for Y_1 and Y_2 .

5. APPLICATION TO HIV DATA

In order to evaluate the proposed method on real data, an application on the HIV data set presented in section 2 was carried out. We attempted to build a prognostic model for survival before antiretroviral treatment changes, using baseline covariates: treatment group, gender, age, body mass index, Karnofsky index, clinical stage, cotrimoxazole prophylaxis, CD4 cell count, HIV viral load, hemoglobin level, neutrophil cell count, alanine aminotransferase (ALT) level, and aspartate aminotransferase (AST) level. Several continuous covariates were categorized, namely: age (≥ 40 versus < 40 years); CD4 cell count (< 200 versus ≥ 200 cells/mm³); HIV viral load (≥ 5 versus < 5 log₁₀ copies/mL); hemoglobin level (≥ 12 versus < 12 g/dL); and ALT and AST level (≥ 37 versus < 37 IU/L). We also recoded clinical stage (stage C versus stage A or B) and Karnofsky index (> 90 versus $\leq 90\%$). We chose this recoding for clinical and statistical reasons (verification of the proportional hazard assumption, and maximizing the Akaike information criterion). Some among these variables were incomplete and did not support Gaussian assumption (Table 1). We used a passive imputation procedure for these variables.

We transformed the continuous variables to be imputed (body mass index, CD4 cell count, HIV viral load, hemoglobin level, neutrophil cell count, ALT and AST levels), using Box-cox transformation, to approximate a normal distribution. Then, imputations were performed on the transformed variables, using Bayesian linear regression [16] and Bayesian splines regression in R 2.9 software. As in the simulation study, we used three imputation models for

each methods: the first did not included response variable (T, D), second included T and D as covariates and the last include $\log(T)$ and D as covariates. All other predictors were selected following the strategy proposed by *S. van Buuren et al*[16]. In imputation model for transformed CD4 cell count, we also took into account the longitudinal measures of CD4 cell count during the first three years. After imputation, all these transformed variables were converted into their original scale before being used in the analysis model.

The Bayesian splines regression imputation model was implemented by an additional program, and the mice function in MICE package in R software [16] has been changed to take into account the splines parameters. The results obtained with complete cases analysis, and after multiple imputation using Bayesian linear regression and Bayesian splines regression are presented in Table 5.

The results given in Table 5 have shown significant differences between estimated parameters in complete cases analysis and multiple imputation analysis. The hazard ratio associated with the hemoglobin was not available in complete case analysis: indeed all patients who died and had a baseline hemoglobin level ≥ 12 g/dL, had missing data therefore, were excluded from complete case analysis. The use of multiple imputation reduced the standard error. Most hazard ratio obtained by linear multiple imputation and multiple imputation by splines for different imputation models were similar. However hazard ratio associated with HIV viral load was only similar between linear multiple imputation using $(\log(T), D)$ and multiple imputation by splines. Overall the results obtained by linear imputation using $(\log(T), D)$ were roughly similar to those obtained with multiple imputation by splines.

6. DISCUSSION

In this article we propose a new method for multiple imputations by FCS, using B-splines functions with fixed degree and knots in Bayesian imputation model; this method allows to take into account the non-linear effects of continuous covariates in imputation models in survival regression setting, when the missing mechanism is MCAR or MAR. Simulations studies have shown that multiple imputation by splines is generally more effective than commonly used multiple imputation method, and does not require the introduction of response variable (T and D). This finding was confirmed on a real data set.

Several methods handling missing covariates in survival analysis have been proposed in the past two decades [15, 27-31], but only recent studies have used multiple imputation by FCS method [15, 27, 31]. These studies have shown that, for handling missing covariates in survival analysis, performing a multiple imputation may be preferred for missing data approach to the complete cases analysis or the single imputation method. But no one has investigated the nonlinear continuous covariate effect in imputation model. We chose to use B-splines functions to take into account this nonlinear effect, because they are numerically well conditioned and they achieve a local sensitivity to data [20]. Several authors have also used B-splines function to approximate nonlinear function in others contexts [32, 33].

Despite the arbitrary choice of degree and splines location knots, our example study showed similar results when performing either our proposed multiple imputation method or linear multiple imputation using $(\log(T), D)$. This result is very interesting because recent studies have shown that the use of outcome as predictor in imputation model improved results in Cox proportional hazard model [11, 27]. However, this arbitrary choice is a limitation of our study. Several algorithms have been proposed in the literature [32, 34, 35] to choose the optimal degree and splines location knots, and we believe that we can improve our proposed multiple imputation method by considering degree and splines location knots as free parameters, estimated using a selection model method. However this hypothesis has not yet been tested and will be assessed in a new study. Another limitation of our method is that it is effective only for large samples or when the imputation models do not contain many explicative variables. Indeed the use of splines leads to over-parameterization of the imputation models; thus, if the sample size is small, it can lead to a problem of identifiability of imputation models. However, the solution would be to find a compromise between the number of covariates to include in imputation model and the splines parameters, in relation to the sample size.

In conclusion, there is a gain to use multiple imputation by splines to fill missing values in survival analysis setting. Nevertheless a compromise is needed between the number of covariates to include in imputation model and splines parameters when the sample size is small.

Acknowledgments

We thank all the investigators of the HIV study in Senegal for allowing us to use the data in the present work. Jules Brice TchatchuengMbougua was the recipient of a doctoral fellowship from the Institut de Recherche pour le Développement.

1. Rubin DB. Inference and missing data. *Biometrika* 1976,**63**:581-592.
2. Schafer JL. *Analysis of Incomplete Multivariate Data*. London; 1997.
3. Rubin DB. *Multiple imputation for nonresponse in surveys* New York; 1987.
4. Vach W, Blettner M. Biased estimation of the odds ratio in case-control studies due to the use of ad hoc methods of correcting for missing values for confounding variables. *Am J Epidemiol* 1991,**134**:895-907.
5. Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. *Am J Epidemiol* 1995,**142**:1255-1264.
6. Harel O, Zhou XH. Multiple imputation: review of theory, implementation and software. *Stat Med* 2007,**26**:3057-3077.
7. Zhang P. Multiple Imputation: Theory and Method. *International Statistical Review* 2003,**71**:581-592.
8. Andridge RR, Little RJA. A Review of Hot Deck Imputation for Survey Non-response. *International Statistical Review* 2010,**78**:40-64.
9. White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. *Statistics in Medicine* 2010,**29**:2920-2931.
10. Taljaard M, Donner A, Klar N. Imputation strategies for missing continuous outcomes in cluster randomized trials. *Biometrical Journal* 2008,**50**:329-345.
11. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, *et al*. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009,**338**:b2393.
12. Mackinnon A. The use and reporting of multiple imputation in medical research - a review. *J Intern Med* 2010.
13. Klebanoff MA, Cole SR. Use of multiple imputation in the epidemiologic literature. *American Journal of Epidemiology* 2008,**168**:355-357.
14. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research* 2007,**16**:219-242.
15. White IR, Royston P. Imputing missing covariate values for the Cox model. *Stat Med* 2009,**28**:1982-1998.
16. van Buuren S, Oudshoorn KG. MICE: Multivariate Imputation by Chained Equation in R. *Journal of Statistical Software, forthcoming* 2010.
17. Royston P. Multiple imputation of missing values. *Stata J* 2004,**4**:227-231.
18. Herring AH, Ibrahim JG. Likelihood-based methods for missing covariates in the Cox proportional hazards model. *Journal of the American Statistical Association* 2001,**96**:292-302.
19. Marshall A, Altman DG, Royston P, Holder RL. Comparison of techniques for handling missing covariate data within prognostic modelling studies: a simulation study. *Bmc Medical Research Methodology* 2010,**10**:16.
20. de Boor C. *A practical guide to spline*. New-York; 1978.
21. Laurent C, Tchatchueng Mbougua JB, Ngom Gueye NF, Etard JF, Diouf A, Landman R, *et al*. Long-term effectiveness and safety of didanosine combined with lamivudine and efavirenz or nevirapine in antiretroviral-naïve patients: a 9-year cohort study in Senegal. *Trop Med Int Health* 2011,**16**:217-222.

22. Hansmann A, Schim van der Loeff MF, Kaye S, Awasana AA, Sarge-Njie R, O'Donovan D, *et al.* Baseline plasma viral load and CD4 cell percentage predict survival in HIV-1- and HIV-2-infected women in a community-based cohort in The Gambia. *J Acquir Immune Defic Syndr* 2005,**38**:335-341.
23. The Antiretroviral Therapy Cohort Collaboration. Causes of death in HIV-1-infected patients treated with antiretroviral therapy, 1996-2006: collaborative analysis of 13 HIV cohort studies. *Clin Infect Dis* 2010,**50**:1387-1396.
24. Curry RJ, Schoenberg IJ. On Polya frequency function. IV: The fundamental splines and their limits. *Journal of Analyse Mathematics* 1966,**17**:71-107.
25. Bender R, Augustin T, Blettner M. Generating survival times to simulate Cox proportional hazards models. *Stat Med* 2005,**24**:1713-1723.
26. Burton A, Altman DG, Royston P, Holder RL. The design of simulation studies in medical statistics. *Stat Med* 2006,**25**:4279-4292.
27. Marshall A, Altman DG, Holder RL. Comparison of imputation methods for handling missing covariate data when fitting a Cox proportional hazards model: a resampling study. *Bmc Medical Research Methodology* 2010,**10**:10.
28. Lin DY, Ying Z. Cox Regression with Incomplete Covariate Measurements. *Journal of the American Statistical Association* 1993,**88**:1341-1349.
29. Pons O. Estimation in the Cox model with missing covariate data. *Journal of Nonparametric Statistics* 2002,**14**:223-247.
30. Paik MC, Tsai WY. On using the Cox proportional hazards model with missing covariates. *Biometrika* 1997,**84**:579-593.
31. Giorgi R, Belot A, Gaudart J, Launoy G. The performance of multiple imputation for missing covariate data within the context of regression relative survival analysis. *Stat Med* 2008,**27**:6310-6331.
32. Holmes CC, Mallick BK. Generalized nonlinear modeling with multivariate free-knot regression splines. *Journal of the American Statistical Association* 2003,**98**:352-368.
33. Molinari N, Daures JP, Durand JF. Regression splines for threshold selection in survival data analysis. *Statistics in Medicine* 2001,**20**:237-247.
34. Molinari N, Durand JF, Sabatier R. Bounded optimal knots for regression splines. *Computational Statistics & Data Analysis* 2004,**45**:159-178.
35. Denis M, Molinari N. Free Knot Splines with RJMCMC in Survival Data Analysis. *Communications in Statistics-Theory and Methods* 2010,**39**:2617-2629.

Table 1 : Incomplete variables at baseline in the HIV data set.

Variables	Median (IQR)		Missing values numbers (%)	
Body mass index (kg/m ²)	19.8	(17.8-22.3)	4	(3%)
CD4 cell count (/mm ³)	177	(94-246)	11	(7%)
HIV viral load (log ₁₀ copies/ml)	5.2	(4.7-5.6)	39	(26%)
Hemoglobin (g/dl)	10.6	(9.6-11.9)	2	(1%)
Neutrophil count (/mm ³)	2192	(1500-2900)	5	(3%)
Alanine aminotransferase (ALT) (UI/l)	21	(13-31)	11	(7%)
Aspartate aminotransferase (AST) (UI/l)	25	(17-41)	12	(8%)

Table 2: Simulation results for logarithm hazard ratio when true relation between Y_3 and others covariates is polynomial

Percentage of missing values		30%						40%					
variables		Y_1		Y_2		Y_3		Y_1		Y_2		Y_3	
Model and mechanism	missing	Lin	Spli	Lin	Spli	Lin	Spli	Lin	Spli	Lin	Spli	Lin	Spli
MCAR1													
<i>biais</i>		0.19	1.27e-03	-0.97	-0.009	0.10	4.57e-04	0.21	1.48e-03	-0.11	-1.06e-02	0.11	4.92e-04
<i>accuracy</i>		5.55e-03	2.01e-05	1.31e-02	3.79e-04	1.41e-03	2.52e-06	6.75e-03	2.85e-05	1.65e-02	5.58e-04	1.74e-03	3.75e-06
<i>coverage</i>		14.3	80	2.2	22.7	0.00	1.2	11.2	73.9	1.1	21.3	0.0	1.4
<i>power</i>		100	100	100	100	99.8	100	100	100	100	100	99.8	99.9
MCAR2													
<i>biais</i>		0.021	1.85e-03	-0.011	-0.014	0.011	1.06e-03	0.02	1.92e-03	-0.012	-1.64e-02	0.012	1.22e-03
<i>accuracy</i>		6.23e-04	2.17e-05	1.68e-03	5.08e-04	1.64e-04	3.87e-06	7.28e-04	2.98e-05	2.13e-03	7.81e-04	1.93e-04	5.92e-06
<i>coverage</i>		9.2	78	0.9	17.3	0.1	1.1	9.2	73.8	2.2	15.7	0.0	0.8
<i>power</i>		100	100	100	100	99.8	100	100	100	100	100	99.7	99.9
MCAR3													
<i>biais</i>		0.020	1.77e-03	-0.011	-0.014	0.011	1.08e-03	0.022	2.36e-03	-0.013	-1.81e-02	0.011	1.36e-03
<i>accuracy</i>		6.17e-04	1.99e-05	1.65e-03	5.39e-04	1.21e-04	4.07e-06	7.50e-04	3.44e-05	2.09e-03	8.58e-04	1.87e-04	6.30e-06
<i>coverage</i>		10.7	78	2	18.7	00	0.8	9.8	71.1	1.1	13.3	0.0	08
<i>power</i>		100	100	100	100	100	100	100	100	100	100	99.7	100
MAR1													
<i>biais</i>		0.18	1.34e-03	-0.98	-9.63e-03	0.1	5.34e-04	0.19	1.54e-03	-0.10	-9.71e-03	0.10	4.37e-04

<i>accuracy</i>	5.53e-03	2.00e-05	1.34e-02	4.36e-04	1.43e-03	3.02e-06	5.72e-03	2.80e-05	1.48e-02	5.17e-04	1.51e-03	3.43e-06
<i>coverage</i>	13.9	77.8	2.1	21.8	0.2	0.8	14.2	74.6	2.2	18.5	0.0	0.8
<i>power</i>	100	100	100	100	99.9	100	100	100	100	100	99.7	100
MAR2												
<i>biais</i>	0.02	1.68e-03	-0.011	-1.34e-02	0.011	9.86e-04	0.02	2.15e-03	-0.013	-1.69e-02	0.012	1.26e-03
<i>accuracy</i>	2.04e-05	2.10e-05	5.10e-02	1.69e-05	3.76e-04	4.23e-06	7.03e-04	3.01e-05	2.07e-03	7.90e-04	1.91e-04	5.93e-06
<i>coverage</i>	12	80.1	1.6	20.5	0.1	1.1	8.9	71.9	1.0	14.1	0.0	0.6
<i>power</i>	100	100	100	100	99.8	99.9	100	100	100	100	99.8	100
MAR3												
<i>biais</i>	0.02	1.72e-03	-0.011	-1.39e-02	0.011	1.05e-03	0.02	4.98e-04	-0.01	-6.26e-03	8.33e-03	3.42e-06
<i>accuracy</i>	5.96e-04	2.06e-05	1.63e-03	5.37e-04	1.62e-04	4.06e-06	7.97e-04	3.06e-05	1.35e-03	4.39e-04	1.13e-04	3.31e-06
<i>coverage</i>	13.7	79.9	1.3	19.3	0.0	1.0	11	70.2	2.8	22	0.0	1.2
<i>power</i>	100	100	100	100	99.7	100	100	100	100	100	99.3	99.9

Lin: analysis based on multiple imputation using Bayesian linear imputation model; Spli: analysis based on multiple imputation using Bayesian splines imputation model; MCAR_i, $i = 1, 2, 3$ analyses under MCAR assumption, using respectively the following imputation models: $Y_3 = f(Y_1, Y_2)$, $Y_3 = f(Y_1, Y_2, T, D)$, $Y_3 = f(Y_1, Y_2, \log(T), D)$; MAR_i, $i = 1, 2, 3$ analyses under MAR assumption, using respectively the following imputation models: $Y_3 = f(Y_1, Y_2, Y_4)$, $Y_3 = f(Y_1, Y_2, Y_4, T, D)$, $Y_3 = f(Y_1, Y_2, Y_4, \log(T), D)$; where f is a linear function in analyses using linear imputation models, and splines function of degree two, with two knots $\xi = (\text{quartile to } 33\%, \text{quartile to } 66\%)$, T is the vector of follow-up times, and D is event indicator variable.

Table 3: Simulation results for logarithm hazard ratio when true relation between Y_3 and others covariates is non-linear

Percentage of missing values variables		30%									40%								
		Y_1			Y_2			Y_3			Y_1			Y_2			Y_3		
Model	and missing mechanism	LN	Lin	Spli	LN	Lin	Spli	LN	Lin	Spli	LN	Lin	Spli	LN	Lin	Spli	LN	Lin	Spli
MCAR1																			
<i>biais</i>		-	-7.75e-01	-4.23e-04	-	-1.73e-02	-2.68e-05	3.32e-01	1.36e-01	6.13e-04	-	-1.36e-01	-1.29e-03	-	-1.65e-02	-1.08e-06	3.71e-01	2.41e-01	2.08e-03
<i>accuracy</i>		3.86e-02	2.59e-02	1.08e-04	2.39e-04	1.99e-04	7.58e-08	1.19e-01	8.11e-02	3.24e-04	-	3.81e-02	1.48e-04	2.71e-04	2.08e-04	1.09e-07	1.57e-01	1.19e-01	4.48e-04
<i>coverage power</i>		26.981	33.981.1	92.788.3	66.899.5	6999.9	99.299.7	37.166.1	48.465.2	97.680.4	26.778.2	30.577.7	88.786.6	63.299.6	6799.3	99.199	36.361.6	4360.4	96.279.1
MCAR2																			
<i>biais</i>		-	-3.79e-03	-2.94e-04	-	-1.88e-05	-3.54e-06	9.64e-03	6.55e-03	5.01e-04	-	-8.0e-03	-1.15e-03	-	-2.38e-04	-1.01e-05	1.81e-02	1.37e-02	1.91e-03
<i>accuracy</i>		4.18e-03	3.95e-03	1.38e-04	2.88e-06	2.48e-06	8.54e-08	1.24e-02	1.22e-02	4.21e-04	6.00e-03	6.03e-03	2.21e-04	4.32e-06	3.85e-06	1.39e-07	1.75e-02	1.86e-02	6.73e-04
<i>coverage power</i>		28.184.7	32.282.6	89.887.3	62.799.7	65.399.4	99.499.5	39.573.9	45.673.2	95.879.2	28.783	32.883.4	79.489.4	54.899.7	58.599.8	9999.6	38.869.5	45.370.8	90.280
MCAR3																			
<i>biais</i>		-	-3.87e-03	-3.70e-04	-	-1.23e-05	-3.04e-06	1.26e-02	6.85e-03	6.41e-04	-	-8.01e-03	-1.58e-03	-	-2.38e-04	-1.01e-05	2.46e-02	1.37e-02	1.91e-03
<i>accuracy</i>		4.28e-03	3.58e-03	1.26e-04	2.78e-06	2.26e-06	7.96e-08	1.2_e-02	1.04e-02	3.80e-04	5.63e-03	6.03e-03	2.21e-04	4.09e-06	3.85e-06	1.39e-07	1.66e-02	1.86e-02	6.73e-04
<i>coverage power</i>		27.783.2	31.783.2	90.487.3	63.899.7	67.699.6	99.699.5	3870.9	44.971.5	96.678.8	27.481.1	32.883.4	79.489.4	56.699.8	58.599.8	9999.6	38.670.3	45.370.8	90.280
MAR1																			
<i>biais</i>		-	-1.36e-01	-1.29e-03	-	-1.65e-04	-1.08e-04	2.87e-01	2.41e-01	4.48e-03	-	-1.62e-01	-3.97e-04	-	-9.14e-03	-9.95e-06	4.29e-01	2.86e-01	6.02e-04

	-01			-03						-01			-02					
<i>accuracy</i>	3.54e	3.81e-	1.48e-	1.85e	2.08e-	1.09e-	1.10e	1.19e-	4.48e-	5.35e	4.13e-	1.24e-	2.88e	2.32e-	9.10e-	1.66e	1.27e-	3.78e-
	-02	03	04	-04	04	07	-01	01	04	-02	02	04	-04	04	08	-01	01	04
<i>coverage</i>	30.4	30.5	88.7	71.4	67	99.1	397	43	96.2	25.1	28.2	90	57.4	66.7	99.6	34.2	41	96.9
<i>power</i>	81.2	77.7	86.6	99.5	99.3	99	62.2	60.4	79.1	81.1	80.9	86	99.2	99.2	99.6	60.5	62.1	78.1
MAR2																		
<i>biais</i>	-	-9.21e-	-1.03e-	-	-1.93e-	-1.36e-	1.46e	1.60e-	9.86e-	-	-3.67e-	-2.29e-	-	-7.99e-	-1.70e-	2.11e	6.81e-	4.82e-
	7.21e	03	03	1.2'e-	04	05	-02	02	04	1.09e	03	04	2.46e	07	05	-02	03	04
	-03			04						-02			-04					
<i>accuracy</i>	4.48e	5.32e-	2.10e-	3.26e	3.56e-	1.69e-	1.37e	1.64e-	6.32e-	6.17e	5.68e-	2.21e-	5.50e	4.20e-	1.44e-	1.81e	1.75e-	6.81e-
	-03	03	08	-06	06	07	-02	02	04	-03	03	04	-06	06	07	-02	02	04
<i>coverage</i>	29.1	29.3	79.4	60.5	60.6	98.7	40.7	42.4	91.3	25.6	28	80.6	51	54.3	98.3	37.9	44	90.2
<i>power</i>	83.8	82.0	89.3	99.4	99.7	99.6	73.6	69.9	79.3	81.6	82.8	80.6	99.4	99.9	99.7	69.4	73.2	81.1
MAR3																		
<i>biais</i>	-	-1.25e-	-7.23e-	-	-1.51e-	-2.66e-	1.93e	2.21e-	1.09e-	-	-5.31e-	-1.78e-	-	-2.25e-	-1.06e-	1.84e	9.74e-	3.87e-
	9.86e	02	04	1.02e	04	06	-02	02	03	9.06e	03	04	8.64e	05	05	-02	03	04
	-03			-04						-03			-05					
<i>accuracy</i>	4.19e	2.75e-	9.65e-	2.65e	1.89e-	7.21e-	1.26e	8.52e-	2.91e-	5.10e	5.06e-	1.97e-	2.96e	3.77e-	1.36e-	1.51e	1.56e-	6.08e-
	-03	03	05	-06	06	08	-02	05	04	-03	03	04	-06	06	07	-02	02	04
<i>coverage</i>	28.5	33.2	93.5	62.5	69.7	99.4	39.0	46.8	98.9	28.2	28.8	81.4	57.6	56.6	98.5	38.9	42.5	92
<i>power</i>	83.5	82.2	89.2	99.1	99.7	99.7	71.8	65	78.3	82.3	82.9	88.2	99.3	99.9	99.7	69.1	72	80.5

LN: analysis based on multiple imputation using Bayesian linear imputation model with log transformation of imputed variable Y_3 ; Lin: analysis based on multiple imputation using Bayesian linear imputation model without transformation; Spli: analysis based on multiple imputation using Bayesian splines imputation model; MCAR $_i$, $i = 1, 2, 3$ analyses under MCAR assumption, using respectively the following imputation models: $Y_3 = f(Y_1, Y_2)$, $Y_3 = f(Y_1, Y_2, T, D)$, $Y_3 = f(Y_1, Y_2, \log(T), D)$; analyses under MCAR assumption, using respectively the following imputation models: $Y_3 = f(Y_1, Y_2, Y_4)$, $Y_3 = f(Y_1, Y_2, Y_4, T, D)$, $Y_3 = f(Y_1, Y_2, Y_4, \log(T), D)$; where f is a linear function in analyses using linear imputation models, and splines function of degree two, with two knots $\xi = (\text{quartile to } 33\%, \text{quartile to } 66\%)$, T is the vector of follow-up times, and D is event indicator variable.

Table 4: Sensibility analysis results on degree and location knots of Y_1 and Y_2

Location knots variables	$Y_1=(1,1)$ $Y_2=(1,1)$			$Y_1=(1,1)$ $Y_2=(1,2)$			$Y_1=(2,1)$ $Y_2=(1,2)$			$Y_1=(2,1)$ $Y_2=(2,1)$			$Y_1=(2,1)$ $Y_2=(2,2)$			$Y_1=(3,3)$ $Y_2=(2,1)$		
	Y_1	Y_2	Y_3															
Model and missing mechanism	Spli	Spli	Spli															
MAR3																		
<i>biais</i>	-	-	1.49e	-	-	2.89e	-	-	5.34e	-	-	5.25e	-	-	1.80e	-2.88e-	-	5.41e
	7.68e	4.66e	-03	1.34e	1.30e	-03	2.86e	1.27e	-03	2.81e	1.22e	-03	8.98e	2.66e	-03	03	1.77e	-03
	-04	-05		-03	-04		-03	-04		-03	-04		-04	-06			-04	
<i>accuracy</i>	1.29e	1.47e	4.74e	1.38e	1.52e	5.06e	1.53e	1.44e	5.81e	1.53e	1.45e	5.84e	1.49e	1.49e	5.69e	2.73e_0	2.96e	1.05e
	-03	-07	-03	-03	-06	-03	-03	-06	-03	-03	-06	-03	-03	-06	-03	3	-06	-02
<i>coverage</i>	39.3	78.8	57.7	41.4	79.4	56.9	36.2	78.3	53.8	36.9	78.6	53.1	3.85	78.3	53.3	31.8	63.9	46.3
<i>power</i>	93.2	100	83.1	93.5	100	80.4	92	100	79	92.5	100	79	92.9	100	79.2	90.2	100	75.4

$Y_i = (d_i, k_i)$ means (d_i, k_i) are parameters splines for Y_i , where d_i is the degree of freedom, and k_i is the number of knots of splines function. The vector of location knots for Y_i is $(\text{quartile to } \frac{100}{k_i+1} \%, \text{quartile to } \frac{2*100}{k_i+1} \%, \dots, \text{quartile to } \frac{k_i*100}{k_i+1} \%)$.

Table 4: Application to HIV data: multivariate COX analyses results - hazard ratio (standard deviation)

	Complete case N=86	Linear N=151			Splines N=151		
		Normal	(<i>T,D</i>)	(<i>log(T),D</i>)	Normal	(<i>T,D</i>)	(<i>log(T),D</i>)
Treatment group (didanosinevszidovudine)	6.50 (1.27)	2.02 (0.47)	2.01 (0.47)	1.97 (0.46)	2.00 (0.47)	1.97 (0.47)	1.97 (0.47)
Women vs Men	0.29 (1.07)	0.72 (0.51)	0.73 (0.53)	0.74 (0.51)	0.71 (0.51)	0.76 (0.51)	0.75 (0.51)
Age (≥ 40 vs <40 years)	1.33 (0.84)	2.47 (0.50)	2.38 (0.51)	2.48 (0.50)	2.42 (0.49)	2.49 (0.49)	2.60 (0.50)
Body mass index*	1.06 (0.15)	1.02 (0.07)	1.02 (0.07)	1.02 (0.06)	1.02 (0.07)	1.01 (0.07)	1.01 (0.07)
Karnofsky index (100 vs <100)	0.10 (1.04)	0.12 (0.68)	0.13 (0.68)	0.12 (0.68)	0.12 (0.68)	0.12 (0.68)	0.11 (0.69)
Clinical stage (C vs A&B)	0.36 (0.76)	0.55 (0.43)	0.57 (0.43)	0.55 (0.43)	0.57 (0.43)	0.55 (0.43)	0.54 (0.44)
CD4 cell count (<200 vs ≥ 200 cells/mm ³)*	0.50 (0.73)	1.24 (0.52)	1.24 (0.54)	1.20 (0.52)	1.18 (0.52)	1.36 (0.55)	1.20 (0.55)
HIV viral load (≥ 5 vs <5 log ₁₀ copies/mL)*	0.77 (1.13)	1.19 (0.58)	1.04 (0.52)	1.22 (0.55)	1.23 (0.53)	1.34 (0.62)	1.59 (0.67)
Hemoglobin (≥ 12 vs <12 g/dL)*	NA (0.00)	0.39 (0.82)	0.37 (0.84)	0.36 (0.83)	0.41 (0.81)	0.43 (0.81)	0.43 (0.81)
Neutrophil count*	1.00 (4.2e-04)	1.00 (2.23e-04)	1.00 (2.20e-04)	1.00 (2.25e-04)	1.00 (2.24e-04)	1.00 (2.18e-04)	1.00 (2.22-04)
ALT (≥ 37 vs <37 IU/L)*	0.81 (1.42)	0.85 (0.73)	0.78 (0.74)	0.73 (0.69)	0.83 (0.71)	(0.86) (0.72)	0.78 (0.69)
AST (≥ 37 vs <37 IU/L)*	3.45 (1.04)	1.07 (0.50)	1.18 (0.52)	1.09 (0.51)	1.09 (0.49)	1.08 (0.53)	1.06 (0.51)
Cotrimoxazole prophylaxis (no vs yes)	2.33 (0.93)	1.66 (0.48)	1.67 (0.48)	1.65 (0.48)	1.65 (0.48)	1.69 (0.48)	1.70 (0.49)

*Incomplete variable

Résumé

Partant de problématiques statistiques rencontrées dans l'étude du traitement antirétroviral des adultes infectés par le virus de l'immunodéficience humaine (VIH) en Afrique subsaharienne, cette thèse cherche, d'une part, à favoriser la vulgarisation d'outils méthodologiques relativement récents auprès d'un public d'utilisateurs moins avertis et, d'autre part, à participer au développement de nouveaux outils. Le premier chapitre présente différentes méthodes de modélisation des données longitudinales dont des méthodes d'analyse de l'évolution d'un critère au cours du temps (les modèles linéaires mixtes généralisés et les modèles d'équations d'estimation généralisées) ou de la survenue d'un événement au cours du temps (le modèle semi-paramétrique de Cox et ses extensions à la prise en compte des covariables dépendantes du temps et de la censure informative). Le deuxième chapitre s'intéresse aux tests de non-infériorité et propose deux développements de la procédure classique de ces tests pour les cas où la marge de non-infériorité est relative. Enfin, le troisième chapitre aborde la question des données manquantes et propose une extension de la méthode d'imputation multiple par les distributions conditionnelles univariées qui consiste à prendre en compte des effets non-linéaires des covariables dans les modèles d'imputation par des fonctions B-splines. Ces méthodes sont illustrées par des études sur le VIH au Cameroun et au Sénégal.

Title : *Statistical problems encountered in the study of antiretroviral treatment of adults infected with HIV in sub-Saharan Africa*

Abstract

On the basis of statistical challenges encountered in study of antiretroviral treatment of adults infected with human immunodeficiency virus (HIV) in sub-Saharan Africa, this thesis aims to promote the dissemination of relatively recent methodological tools of less aware audience of users on one hand and to participate to development of new tools on the other hand. The first chapter presents various methods for modeling longitudinal data of which analysis methods for changing of a criterion over time (the generalized linear mixed models and models of generalized estimating equations) or the occurrence of an event over time (the semi-parametric Cox model and its extensions to take into account time-dependent covariates and informative censoring). The second chapter focuses on non-inferiority test and provides two developments of the classical procedure of these tests in cases where the non-inferiority margin is relative. The third chapter addresses the question of missing data and proposes an extension of the multiple imputation method based on fully conditional specification, to take into account nonlinear effects of covariates in the imputation models using B-splines functions. These methods are illustrated by studies on HIV in Cameroon and Senegal.

Discipline : *Mathématiques appliquées-Biologie santé (Biostatistiques)*

Mots clés : Méthodes statistiques, données longitudinales, analyse de survie, données manquantes, non-infériorité, VIH/SIDA.

Laboratoires :

Laboratoire Informatique, Modélisation, Simulations et Systèmes (LIMSS), Ecole Nationale Supérieure Polytechnique, Université de Yaoundé 1, BP 8390 Yaoundé Cameroun

&

IRD UMI 233, Transitions épidémiologiques, recherches translationnelles appliquées au VIH et aux Maladies Infectieuses (TransVIHMI), 911 avenue agropolis BP 64501 34394 Montpellier cedex 5