A NEW METHOD OF INTEGRATING EPIDEMIOLOGICAL AND HEALTH SERVICES MODELLING TECHNIQUES FOR STUDYING INFECTIOUS DISEASES: AN EXAMPLE USING HIV/AIDS

by

KARISSA MICHELLE JOHNSTON

BSc. University of Victoria, 2003
MSc. University of British Columbia, 2005

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Health Care and Epidemiology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2009

© Karissa Michelle Johnston, 2009

Abstract

Since the introduction in 1996 of highly active antiretroviral therapy (HAART) for treating individuals infected with HIV, morbidity and mortality among individuals who access care in a timely fashion have dramatically decreased. As a result, patterns of health services utilization have shifted from hospitalizations and acute care services to outpatient services and medications. An additional implication of increased HAART coverage has been a reduction in HIV-1 viral load amongst treated individuals. Individuals with lower viral load are less likely to transmit HIV infection to uninfected individuals. A resulting hypothesis is that increasing treatment with HAART—either by treating individuals earlier in their infection or by expanding treatment coverage rate to a greater proportion of infected individuals—could potentially lead to reduced rates of HIV transmission. To date, the results of ecological studies and mathematical modelling studies have been consistent with this hypothesis. The objectives in this dissertation were: to build a computational tool that could be used to answer complex questions regarding the economics of HIV during the HAART era in British Columbia (BC), Canada; and to address specific economic questions relevant to the current treatment of HIV in BC. The particular questions of interest included: the monthly costs associated with various categories of health services utilization and the correlation between categories (Chapter 3); the expected long-term incidence and costs associated with hospitalizations after initiation of HAART (Chapter 4); the lifetime direct medical costs associated with an individual infected with HIV during the HAART era (Chapter 5); and the cost-effectiveness of a strategy to increase HAART coverage from 50% to 75% of clinically eligible individuals, accounting for individual-level and population-level effects of HAART (Chapter 6). These questions were addressed using a variety of statistical and

mathematical techniques, based on a BC-specific population-based data source. A key finding with important public health relevance was that use of HAART was associated with decreases in other categories of direct medical costs. These cost offsets were due to both an improvement in morbidity resulting in a reduced need for acute health care services and, more importantly, a reduction in HIV incidence.

Table of contents

Abstract....ii

Table of contents....iv

List of tables....ix

List of figures....x

Acknowledgements....xii

Dedication....xiii

Co-authorship statement....xiv

1. Introduction....1

1.1. Context....1

1.2. Burden of HIV in British Columbia....2

1.3. Treatment for Individuals Infected with HIV....2

1.4. Objective....4

1.5. Literature Review....5

1.5.1. Direct medical costs associated with HIV....5

1.5.2. Lifetime direct medical costs of HIV....6

1.5.3. Impact of HAART on HIV transmission....7

1.6. Dissertation Overview....10

1.7. Ethical Considerations....13

1.8. Original Contributions....13

1.8.1. Methodological contributions....13

1.8.2. Health policy contributions....14

1.9. References....16

2. Characterizing trajectories of plasma HIV-1 RNA viral load and CD4 count in individuals treated with highly active antiretroviral therapy 23
2.1. Introduction 23
2.2. Methods 25
2.2.1. Study population 25
2.2.2. Variable measurement 26
2.2.3. Statistical methodology 27
2.3. Results 29
2.4. Discussion 32
2.4.1. Interpretation 32
2.4.2. Strengths and limitations 33
2.4.3. Potential applications 35
2.5. Conclusion 36
2.6. References 45
3. The impact of highly active antiretroviral therapy on hospitalization and other medical costs for community-dwelling individuals infected with HIV 49
3.1. Introduction 49
3.2. Methods 51
3.2.1. Data sources 51
3.2.2. Economic variables 53
3.2.3. Demographic and clinical variables 54
3.2.4. Statistical methods 55
3.3. Results 58

3.4. Discussion 61
3.5. Conclusion 66
3.6. References 79
4. Accounting for time-dependent CD4-cell count trajectories in estimating cumulative utilization and cost of hospitalizations in individuals infected with HIV 83
4.1. Introduction 83
4.2. Methods 85
4.2.1. Data 85
4.2.2. Statistical models 86
4.2.3. Cumulative incidence and costs 90
4.3. Results 91
4.4. Discussion 93
4.5. References 102
5. Estimating the lifetime direct medical costs associated with HIV/AIDS in British Columbia, Canada 105
5.1. Introduction 105
5.2. Methods 107
5.2.1. Data 107
5.2.2. Costs 108
5.2.3. Statistical models 108
5.2.4. Microsimulation structure 111
5.2.5. Justification for microsimulation approach 113
5.2.6. Assumptions 114

5.2.7. Uncertainty and sensitivity analysis 114
5.2.8. Internal validation 116
5.2.9. Computational details 116
5.3. Results 117
5.4. Discussion 120
5.5. References 134
6. Cost-Effectiveness of increasing access to highly active antiretroviral therapy for treating HIV/AIDS 137
6.1. Introduction 137
6.2. Methods 139
6.2.1. Model structure 139
6.2.2. Cost and utility estimates 143
6.2.3. Scenario analyses 144
6.2.4. Model outcomes 144
6.3. Results 145
6.4. Discussion 147
6.5. Conclusion 151
6.6. References 159
7. Discussion 164
7.1. Study Results and Implications 164
7.1.1. Summary of findings 164
7.1.2. Cost effectiveness of expanded access to HAART 165
7.1.3. International variability in HIV direct medical costs 166

7.1.4. Importance of individual-level variability 167
7.2. Study Strengths 169
7.3. Study Limitations 171
7.3.1. Data limitations 171
7.4. Future Research 174
7.5. Contributions to Original Knowledge 176
7.5.1. Methodological contributions and implications 176
7.5.2. Policy contributions and implications 178
7.6. References 179
Appendix A: derivation of conditional DISTRIBUTION of haart costs 181
Appendix B: Derivation of the cumulative cost function 183
APPENDIX C: TIME UNTIL GAINING AND LOSING RESISTANCE MUTATIONS FOR INDIVIDUALS RECEIVING HAART 186
Summary 186
Methods 186
Data 186
Estimating time at risk 187
Explanatory variables 187
Statistical methods 187
Results 188
References 194
APPENDIX D: Ethics approval certificate 195

List of tables

Table 2.1: Description of the HOMER cohort (1996-2005). 43

Table 2.2: Estimated linear coefficients for multivariable generalized additive models. 44

Table 3.1: Descriptive characteristics of study sample. 73

Table 3.2: Estimated regression coefficients and standard errors. 75

Table 3.3: Estimated regression coefficients and standard errors. 76

Table 3.4: Measures of variability for statistical models. 77

Table 3.5: Predicted monthly HAART costs. 77

Table 3.6: Correlations between random intercepts. 78

Table 4.1: Descriptive characteristics of study sample. 99

Table 4.2: Regression coefficients and 95% confidence intervals 100

Table 4.3: Covariance, correlation and conditional expectation of random effects. 101

Table 5.1: Observed demographic and treatment characteristics in the HOMER cohort. 131

Table 5.2: Empirical covariance matrix of random effects 132

Table 5.3: Estimated lifetime costs, life-years, and quality-adjusted life years 133

Table 6.1: Statistical methods used to describe processes included in microsimulation. 157

Table 6.2: Parameters associated with transmission model. 158

Table C.1: Descriptive characteristics of study sample 191

Table C.2: Regression coefficients estimated using Weibull survival model 192

Table C.3: Estimated hazard ratios estimated using Weibull survival model. 193

List of figures

Figure 2.1: Untreated history of CD4 cell count and plasma viral load. 37

Figure 2.2: Non-parametric results of generalized additive models of plasma viral load. 38

Figure 2.3: Non-parametric results of generalized additive models of CD4 cell count..... 39

Figure 2.4: Hypothetical "typical" trajectories for (a) log viral load, and (b) CD4 cell count. ... 40

Figure 2.5: Observed and predicted trajectories of viral load..... 41

Figure 2.6: Observed and predicted trajectories of CD4 cell count..... 42

Figure 3.1: Number of subjects included from the HOMER and CHASE cohorts, Vancouver BC. 68

Figure 3.2: Available data and follow-up information. 69

Figure 3.3: Estimated mean monthly costs. 70

Figure 3.4: Estimated total direct medical costs. 71

Figure 3.5: Estimated association between highly active antiretroviral therapy (HAART) cost random effects and expected hospitalization costs. 72

Figure 4.1: Estimated (a) cost and (b) incidence functions. 97

Figure 4.2: Cumulative (a) cost and (b) incidence functions for selected quantiles of the empirical frailty distribution. 98

Figure 5.1: Microsimulation structure 125

Figure 5.2: Estimated monthly costs (2005 CDN$)..... 126

Figure 5.3: Example CD4 trajectories from microsimulation..... 127

Figure 5.4: Common CD4 trajectory from cohort approach..... 128

Figure 5.5: Distribution of lifetime costs. 129

Figure 5.6: Observed vs. randomly generated log-costs used for internal validation purposes. 130

Figure 6.1: Schematic of model structure. 152

Figure 6.2: Incremental net benefit. 153

Figure 6.3: Cost-effectiveness acceptability curve. 154

Figure 6.4: Overall population and patient-centered incremental net benefit 155

Figure 6.5: Estimated cumulative number of new infections over time. 156

Figure C.1: Kaplan-Meier curves describing time until gaining resistance mutation, stratified by (a) medication adherence, and (b) medication category. 190

Acknowledgements

I wish to extend my gratitude to Adrian Levy, my supervisor and mentor, for inspiring me to undertake a PhD in Epidemiology and for providing invaluable guidance and support throughout the entire process. I would also like to thank my entire supervisory committee, all of whom were generous with their time and knowledge: Andrew Briggs for patient advice in health economics, Paul Gustafson for extensive statistical advice, Robert Hogg for guidance on the demographics of HIV in BC, Julio Montaner for unparalleled clinical expertise, and Mark Tyndall for insight into the HIV epidemic in the downtown east side of Vancouver. I also greatly benefitted from input and advice from Viviane Dias Lima, P Richard Harrigan, Jeannie Shoveller, Anita Palepu, and Douglas James. I am extremely grateful to the staff of the Centre for Excellence in HIV/AIDS Research Drug Treatment Program for allowing me to have access to the HOMER cohort data source, and to the investigators of the CHASE study for allowing me access to the CHASE cohort data source. I would like to thank Benita Yip and Dragan Lesovski for assisting with the data extraction process, and Rita Sobolev for providing advice on database-related issues. I would like to thank fellow students who have provided critical evaluation and support over the years; in particular, Kathy Lepik, Jeffrey Quon, Sarah Goring, Shelagh Szabo, Elisa Lloyd-Smith, Lindsay Hedden, and Beth Rachlis. It would not have been possible for me to complete my PhD without generous funding from the Michael Smith Foundation for Health Research and the Canadian Institutes for Health Research, for which I am very grateful. Last but not least I would like to acknowledge the bravery displayed by all those living with HIV, in BC and globally.

Dedication

This thesis is dedicated to my amazing and supportive family: my husband Tim Fry, my parents Mike and Ulrike Johnston, my sister Sabrina Sexsmith, and the world's strongest baby, Maria Hope Johnston-Fry.

Co-authorship statement

The studies described in Chapters 1-5 were conceptualized and designed by Karissa M Johnston (KMJ) with input and guidance from Adrian R Levy (ARL). The study described in Chapter 6 was designed by KMJ based on a concept suggested by Julio S Montaner (JSM).

KMJ was responsible for all data analysis and the interpretation of results. KMJ wrote the original drafts of all chapters, then edited based on feedback from the entire supervisory committee: ARL, JSM, Andrew Briggs (AB), Paul Gustafson (PG), Robert S Hogg (RSH), and Mark W Tyndall (MWT). In addition, P Richard Harrigan (PRH) provided input on Chapter 2 and Appendix C, while Viviane Dias Lima (VDL) provided input into Chapters 3-6. The mathematical model described in Chapter 6 was based on an earlier published mathematical model created by VDL and KMJ, updated by KMJ to meet the objectives of this study.

ARL provided guidance regarding the epidemiology and design of all studies. JSM, RSH, and MT provided guidance regarding the clinical, demographic, and epidemiological details of HIV and AIDS in British Columbia. PRH provided guidance regarding HIV virology. AB provided guidance regarding health economics theory and practice. PG and VDL provided guidance regarding statistical analysis. All co-authors made suggestions regarding writing content and style.

1. INTRODUCTION

1.1. Context

Widely acknowledged as "Canada's poorest postal code", Vancouver's downtown eastside (DTES) has been described as being "torn between the forces of gentrification and ghettoisation".(1) In 1997, the neighbourhood was home to the highest reported rates of human immunodeficiency virus (HIV) infection in the developed world,(2) which have regrettably remained similarly high since then.(3) The HIV prevalence in the DTES is currently estimated to be approximately 17% amongst active drug users in the community.(4) While British Columbia (BC) is not home to a generalized epidemic, in which high rates of HIV are endemic throughout the general population, particular geographic regions have experienced concentrated epidemics of similar magnitude to the generalized epidemics of Sub-Saharan Africa and other developing regions.(5, 6) Within BC, the particular subpopulations most affected by HIV include some of the most marginalized segments of society, including injection drug users, Aboriginals, and sex trade workers.(7-9) Following HIV infection, it is these marginalized subgroups who are also least likely to seek and receive optimal treatment for their disease.(10) This dissertation addresses the economic implications of expanding access to HIV treatment to those in medical need in BC. Due to the socioeconomic distribution amongst individuals currently in need of HIV treatment in BC, expanding HIV treatment to those who have a clinical indication but are not currently receiving it could potentially help to address the spread of disease within the province.

1.2. Burden of HIV in British Columbia

HIV and Acquired Immune Deficiency Syndrome (AIDS) present a major burden to healthcare systems throughout the world, including BC and Canada. At the end of 2005, it was estimated that 58,000 Canadians were living with HIV, a 16% increase since 2002.(11) Over 95% of HIV and AIDS diagnoses in Canada occur in the provinces of BC, Ontario, Quebec, and Alberta.(11) In 2003, the number of HIV cases in BC was estimated to be 13,000, or 22.4% of 58,000.(12) Only 13% of Canada's population resides in BC, meaning that HIV is disproportionately represented in this province. A rough estimate of the magnitude of direct costs associated with treating HIV in Canada in 1997, updated to adjust for increases in prevalence between 1997 and 2005, is over $800 million per year.(13)

1.3. Treatment for Individuals Infected with HIV

Beginning in 1987, standard antiretroviral treatment for individuals infected with HIV comprised zidovudine ("AZT") monotherapy;(14) as other antiretroviral agents were developed, the benefits of dual-therapy were documented.(15) Prior to 1995, all FDA-approved antiretroviral medications belonged to the non-nucleoside reverse transcriptase inhibitor (NRTI) class.(14) In 1995, the first protease inhibitor (PI), saquinavir, was approved by the FDA, followed by nevirapine, a non-nucleoside reverse transcriptase inhibitor (NNRTI), which was approved in 1996.(14)

Antiretroviral therapy was revolutionized in 1996 when highly active antiretroviral therapy (HAART) was first described at the Vancouver International AIDS Society conference with the

unveiling of the first International AIDS Society antiretroviral therapy guidelines.(16) This was due to the recognition of the key role of plasma HIV-1 RNA viral load as a prognostic marker(17) and the results of two randomized clinical trials: Merck 035(18) and INCAS(19). Mellors et al.(17) showed that low plasma HIV-1 RNA viral load was associated with reduced disease progression for at least three years. Gulick et al.(18) and Montaner et al.(19) showed that a consistently low plasma HIV-1 RNA viral load could be achieved using triple-drug therapy, also referred to as HAART, in treatment-naïve individuals using two NRTIs and either a PI(18) or a NNRTI.(19)

Widespread uptake of HAART in developed countries has materially altered the epidemiology of HIV, with a dramatic decrease in morbidity and mortality from HIV,(20) a prolonged period of undetectable viral load in treated individuals,(21) and a shift in health resource utilization from inpatient services to outpatient services.(22, 23) In addition to the individual-level health benefits provided by HAART, the potential for HAART to curb the spread of HIV infection has been suggested.(24) The mechanism for this benefit relates to the ability of HAART to reduce viral load amongst treated individuals,(25) which leads to a reduced likelihood of transmitting HIV to uninfected individuals.(26, 27)

For a number of reasons, the shifting epidemiology of HIV during the HAART era necessitates an in-depth analysis of the direct medical costs associated with HIV and the economic implications of expanding treatment coverage rates. Firstly, with stable incidence of HIV and increasing life expectancy amongst infected individuals, the prevalence of HIV is expected to increase. With a growing group of individuals requiring life-long treatment and medical care, it

is becoming increasingly meaningful for health planners and policy makers to use accurate estimates of the costs required to provide such care. Secondly, it has been suggested that access to HAART be expanded in BC, through providing treatment earlier in the disease history as well as by providing outreach to individuals currently in need of treatment who for social, cultural, or co-morbid reasons fail to access therapy.(24, 28) The upfront costs of this strategy mean that downstream economic consequences need to be well characterized in advance, including the acquisition costs of therapy, the long-term costs of providing medical care, and the offsets over time brought about by improved health and a potentially reduced number of infections.

1.4. Objective

The objective of this dissertation is to develop a comprehensive framework for evaluating economic scenarios describing the treatment and transmission of HIV during the HAART era. The framework is designed to address questions specific to BC, and is also potentially generalizable to other settings with similar treatment patterns.

In addition to developing the framework, four particular economic questions are addressed:

1. What are the monthly costs associated with various categories of health services expenditure, stratified by current health status and demographics?
2. What is the expected long-term utilization and expenditure associated with hospitalizations during the HAART era?
3. What is the lifetime cost associated with medical care for an individual infected with HIV?

4. What are the comprehensive economic implications of expanding access to HAART, accounting for individual-level health benefits and population-level benefits associated with reduced transmission?

1.5. Literature Review

1.5.1. Direct medical costs associated with HIV

The direct costs of treating an individual infected with HIV have been estimated by a number of investigators in several different countries. A systematic review of studies reporting monthly treatment costs during the HAART era(23) identified studies conducted in Canada,(29) England,(30) the United States,(31-33) Italy,(34, 35) and France.(22) Studies published since then include two from the United States(36, 37) and one from Canada.(38) One consistent trend observed in all studies was an inverse relationship between monthly direct medical costs and CD4 cell count, with individuals in the lowest CD4 cell count strata (typically defined as 0-50 cells/mm^3) incurring the greatest expenditure. However, beyond the comparison of broad trends, between-study comparisons are challenging for a number of reasons, including: international variation in treatment costs and funding structures; different costing methodologies used by different investigators; differences in study populations with respect to disease stage distribution; and different levels of aggregation in reporting results.(23) For these reasons, comparisons of absolute monthly cost figures across studies are not meaningful. In particular, the observed international variation highlights the need for costing studies to be repeated across different jurisdictions, so that any inference made based on assumed monthly costs is relevant to the jurisdiction of interest.

Investigators in Alberta recently published estimates of direct medical costs in this province between 1997/1998 and 2005/2006.(38) As expected, they found that costs were inversely related to CD4 cell counts. It was also observed that, within particular CD4 cell count strata, HIV-related costs per person per month tended to either increase or remain stable between 1997/1998 and 2005/2006. However, the distribution of CD4 cell counts shifted over time, with an increasing proportion of individuals in the highest CD4 strata. Thus, the observed net result was a 1.8% decrease in HIV-related monthly costs over time, from $953 in 1997/1998 to $936 in 2005/2006.

Although comprehensive health services utilization databases are maintained for the province of BC,(39) BC-specific HIV monthly cost estimates are not currently available.

1.5.2. Lifetime direct medical costs of HIV

Compared to studies describing monthly costs, relatively few studies have been published describing the lifetime costs associated with treating an individual infected with HIV during the HAART era. One reason for this is that the life expectancies conferred by HAART exceed its period of widespread availability,(40) meaning that any lifetime-based inference regarding HIV treated with HAART requires the use of extrapolation, and cannot be done directly using empirical data. Any systematic variability in monthly cost estimates will be compounded when extrapolating those costs over the course of a lifetime, highlighting the importance of taking the specific setting of interest into account when estimating lifetime costs.

To date, two studies have been published describing the lifetime costs associated with treating an individual infected with HIV during the HAART era: one in France(22) and one in the United States.(41) Both studies combined a simulation model of HIV disease with country-specific costs, stratified by CD4 cell count and other clinical indicators (e.g. final month of life, history of opportunistic infections, currently experiencing an opportunistic infection). Substantial between-study variability was observed: converting all costs to Canadian $2005, the estimated lifetime cost, discounted at 3%, was $311,831 in France and $486,321 in the United States. Again, the between-country differences suggest the importance of producing country-specific estimates; when estimating lifetime treatment costs in Canada, results from other countries may not be directly relevant. The small number of lifetime cost studies and the degree of inter-study variability implies a need for similar studies to be repeated in different settings.

In order to fill the research gap related to monthly and lifetime direct medical costs for individuals infected with HIV in BC, producing BC-specific estimates of these quantities represents a key objective of this dissertation.

1.5.3. Impact of HAART on HIV transmission

Ideally, the impact of expanding treatment with HAART on HIV transmission would be evaluated using a randomized controlled trial (RCT), as this study design is considered the "gold standard" when evaluating a medical intervention.(42) However, an experimental study design is not plausible for evaluating the impact of a population-level policy decision. In the absence of an RCT, observational ecological data are one source of information that can be used to estimate the effects of increasing treatment with HAART. Ecological comparisons can be made across

time periods or geographic locations. For example, it was observed that after implementing a policy of free access to HAART in Taiwan, HIV infection rates fell by 53%, despite a stable incidence of syphilis which suggested that there was no decrease in sexual risk behaviour that could have explained the reduced HIV transmission.(43) This is consistent with observed patterns in new HIV infections observed in BC after the widespread use of HAART in 1996.(24) Further, an international gradient has been observed between access to treatment and likelihood of transmission from prevalent individuals, with regions that have the greatest access to treatment showing the lowest rates of transmission relative to the size of the infected population.(24)

While ecological evidence motivates the potential for increased HAART use to lead to reduced HIV transmission, results are highly context specific, and it is possible that important confounders are unknown or unmeasured. Further, the practice of assuming that aggregate-level factors observed in a population apply to individual population members can lead to flawed inference, a phenomenon known as the ecological fallacy.(44) An alternative method that has been widely applied in this setting is mathematical modelling. The advantages of a mathematical model include: it can be designed to incorporate all factors that are known or suspected to affect results; and parameters and assumptions can be varied in order to assess the impact of uncertainty and to generalize results.

Previous studies based on mathematical models have reported the impact of increased HAART use on HIV transmission.(28, 45-48) While all studies have found some relationship between the intervention and outcome, the specific results have varied widely. A model based on

observed population data in BC(28) and another model of Sub-Saharan Africa developed by the AIDS program of the World Health Organization(45) both predicted dramatic reductions in new transmission when treatment with HAART was increased. A model based on the San Francisco population further predicted that full eradication of the epidemic may be possible.(46) Conversely, two models based on hypothetical populations consistent with observed resource-poor settings both predicted relatively modest effects of increased HAART use on transmission.(47, 48) Two of these studies reported simplified economic results,(28, 45) with all studies focussed on new infections as the outcome of interest, and none reported a full economic evaluation of expanded HAART use.

There are several broad reasons why the above-described discrepancies in mathematical model results may occur. The first is the influence of population-specific epidemic properties, such as the level of risk behaviour or the relative size of the infected population, on the potential impact of prevention efforts.(49) As such, a single model might predict a dramatic reduction in transmission in one population, but a more modest effect in an alternative population with distinct transmission dynamics. A second reason is that "expansion of HAART" does not have a uniform definition, and different models are likely to be evaluating the impact of slightly different expansion strategies. A third reason is that different models may be based on different assumptions leading to alternative conclusions. In light of these potential sources of variability, it is advisable that results obtained for one population not be directly applied to another population unless the specific epidemics are thought to be highly similar. It is further desirable for investigators to be explicit and comprehensive regarding the scenarios compared and assumptions made within a mathematical model, so that results can be critically evaluated.

1.6. Dissertation Overview

In quantitative health research, the term "modelling" is used in several distinct contexts.(50) Three such contexts will be addressed here: (1) *Statistical* modelling, in which individual-level data are analyzed under the assumption that they follow specified probability distributions, and parameters of interest are extracted; (2) *Mathematical* modelling, in which differential or difference equations are used to describe the transmission of an infectious disease, and mathematical theory is used to derive long-term summary parameters; and (3) *Decision* modelling, in which various forms of information and data are synthesized into a single framework, such as a Markov process(51) or a discrete-event simulation,(52) in order to describe the course of a disease in terms of costs and outcomes.

The topics addressed within this dissertation require the use of all three forms of quantitative modelling. An overarching framework for addressing economic issues in HIV treatment must account for the impact of HAART at both the individual level and the population transmission level, and, as such, both decision modelling and mathematical modelling techniques are applied. Statistical models are used to translate individual-level data sources into the parameters needed to populate the integrated mathematical and decision models.

This dissertation is divided into seven chapters, which focus on the different statistical techniques, their integration into a single comprehensive framework, and the resulting health policy implications.

Chapter 2 discusses the development of non-linear statistical models for describing longitudinal trajectories of CD4 cell count and HIV-1 plasma viral load (pVL), two important prognostic markers in HIV infection, during the HAART era. CD4 cell count is a predictor of HIV-related morbidity and mortality,(17, 53, 54) while pVL is a predictor of infectivity.(26, 27) Based on these relationships, the results of the corresponding statistical models are important inputs into both the decision model describing lifetime medical costs and the mathematical model describing disease transmission.

Chapter 3 provides a detailed description of the overlapping cohorts that form the basis of the data analyses underlying the decision and mathematical models, and presents the results of longitudinal statistical models which describe monthly direct medical costs for individuals infected with HIV during the HAART era. These statistical models are adjusted for relevant demographics and time-varying CD4 cell count in order to quantify the impact of these variables on direct medical costs. Specific methodological approaches are employed in order to account for common characteristics of cost data that violate standard regression assumptions. Within Chapter 3, the empirical correlations between different categories of health services utilization (e.g. inpatient, outpatient, pharmaceutical) are also derived. These quantities can be used within decision and mathematical models to account for the fact that an individual's use of a particular health services category may be informative of their use of other health services categories, based on an underlying level of health-seeking behaviour.

Chapter 4 describes a novel statistical method for predicting the long-term incidence and costs of hospitalizations during the HAART era. The statistical model integrates a linear regression

model for estimating the cost of a single hospitalization, a recurrent events model for estimating the incidence of hospitalizations for a given CD4 cell count, and a non-linear model for estimating CD4 cell count trajectories over time. These three models are formally combined to yield a global estimator of long-term expected hospitalization costs. In the decision and mathematical models described in other chapters, the methods of Chapter 3 are used to incorporate the costs of hospitalizations. The methods described in Chapter 4 represent an alternative method for incorporating the cost of hospitalizations, which could be applied in order to incorporate the specific timing of individual hospitalizations.

Chapter 5 describes the integration of the statistical models described in Chapters 2 and 3 into a microsimulation decision model which can be used to describe the lifetime direct medical costs accrued by individuals infected with HIV in BC during the HAART era. Briefly, the non-linear statistical models are used to generate individualized trajectories for CD4 cell count and pVL, accounting for individual-level variability. Each month, based on an individual's demographics and their current CD4 cell count, direct medical costs are randomly generated and added to a cumulative total. A Cox proportional hazards survival model is used to generate mortality probabilities. This procedure is repeated for a large number of individuals under several sets of assumptions, in order to estimate the average lifetime medical costs as well as the level of variability observed between individuals and the potential impact of uncertainty in key model assumptions. Within this chapter, the need for employing a computationally-intensive microsimulation—which simulates heterogeneous individuals—as opposed to a simpler approach is also explored.

Chapter 6 extends the microsimulation model of Chapter 5 to incorporate a mathematical model of disease transmission. Individual disease histories are generated as described above, with a set of difference equations used to estimate new infections each month, accounting for the relationship between pVL and infectivity. This integrated framework between the decision and mathematical models is then used to estimate the economic impact of expanding access to HAART.

Chapter 7 provides a discussion of all the studies undertaken within the dissertation, and describes the individual and collective implications of the findings described in Chapters 2 to 6.

1.7. Ethical Considerations

This study has been recognized as ethically sound by the University of British Columbia Providence Health Care Research Institute Office of Research Services. A copy of the ethics approval certificate is included as Appendix D.

1.8. Original Contributions

This dissertation provides several contributions to original knowledge, which can be classified as either methodological or policy-related contributions.

1.8.1. Methodological contributions

The research described in Chapters 2 and 3 is based on existing statistical methodologies, and they are applied in novel contexts here. The specific non-linear statistical modelling method

used in Chapter 2, generalized additive modelling, has not been used previously to describe prognostic markers in HIV, and, as described, is found to be an effective method for accurately describing short- and long-term dynamics in the absence of explicit theoretical knowledge about functional forms. In Chapter 3, random effects modelling of longitudinal data is used in conjunction with two-stage modelling techniques for health care cost data, in order to make full use of a longitudinal data source with a time dependent covariate. To our knowledge, these forms of statistical modelling have not been combined previously in the published literature.

The methodology described in Chapter 4 is statistically novel, and provides a method for estimating the cumulative value of a marked point process that is not independently and identically distributed in the presence of a time-dependent covariate.

The framework described in Chapter 6 is a novel integration of decision modelling and mathematical modelling techniques. This method is directly applicable to other infectious disease areas.

1.8.2. Health policy contributions

The lifetime direct medical costs estimated within Chapter 5 provide the first such estimates in a Canadian setting during the HAART era. This information can be used to make economic projections regarding the health resources required to provide care for individuals infected with HIV over the course of their lifetime. The model can also be adapted to similar jurisdictions by updating the relevant input parameters.

The results of the integrated decision and mathematical models described in Chapter 6 provide evidence of the economic benefits associated with expanding treatment with HAART. Previously, the empirical associations between HAART and viral load(25) and viral load and infectivity,(26) respectively, have been derived, and mathematical models have been used to estimate the expected impact of increasing treatment with HAART on the rate of new infections.(28, 45-48) The study described here is the first to further incorporate economic quantities and provide an economic evaluation of HAART expansion that accounts for expenditure and benefit at both the individual and population level.

1.9. References

1. Smith HA. Planning, policy and polarisation in Vancouver's Downtown Eastside. *Tijdschrift Voor Economische En Sociale Geografie* 2003;94(4):496-509.
2. Strathdee SA, Patrick DM, Currie SL, et al. Needle exchange is not enough: Lessons from the Vancouver injecting drug use study. *AIDS* 1997;11(8):F59-F65.
3. British Columbia Centre for Disease Control. STI/HIV Prevention and Control Annual Report. 2008.
4. Tyndall M, Wood E, Zhang R, et al. HIV seroprevalence among participants at a Supervised Injection Facility in Vancouver, Canada: implications for prevention, care and treatment. *Harm Reduction Journal* 2006;18:3-36.
5. World Health Organization. 2007 Report on the global AIDS epidemic. 2007.
6. Tyndall MW, Craib KJP, Currie S, et al. Impact of HIV infection on mortality in a cohort of injection drug users. *Journal of Acquired Immune Deficiency Syndromes* 2001;28(4):351-7.
7. Craib KJP, Spittal PM, Wood E, et al. Risk factors for elevated HIV incidence among Aboriginal injection drug users in Vancouver. *Canadian Medical Association Journal* 2003;168(1):19-24.
8. Spittal PM, Craib KJP, Wood E, et al. Risk factors for elevated HIV incidence rates among female injection drug users in Vancouver. *Canadian Medical Association Journal* 2002;166(7):894-9.

9. Spittal PM, Hogg RS, Li K, et al. Drastic elevations in mortality among female injection drug users in a Canadian setting. *AIDS Care-Psychological and Socio-Medical Aspects of AIDS/HIV* 2006;18(2):101-8.

10. Wood E, Montaner JSG, Bangsberg DR, et al. Expanding access to HIV antiretroviral therapy among marginalized populations in the developed world. *AIDS* 2003;17(17):2419-27.

11. Public Health Agency of Canada. HIV/AIDS Epi Updates 2007. Public Health Agency of Canada, 2007.

12. BC Ministry of Health Planning. Priorities for action in managing the HIV epidemics: 2003-2007. 2003.

13. Albert T, Williams G. The economic burden of HIV/AIDS in Canada. Canadian Policy Research Networks, 1997.

14. Chopra KF, Tyring SK. Current antiretroviral therapy in the treatment of HIV infection. *Seminars in Cutaneous Medicine and Surgery* 1997;16(3):224-34.

15. Meng TC, Fischl MA, Boota AM, et al. Combination therapy with zidovudine and dideoxycytidine in patients with advanced human immunodeficiency virus infection: A phase I-II study. *Annals of Internal Medicine* 1992;116(1):13-20.

16. Carpenter CCJ, Fischl MA, Hammer SM, et al. Antiretroviral therapy for HIV infection in 1996 - Recommendations of an international panel. *JAMA-Journal of the American Medical Association* 1996;276(2):146-54.

17. Mellors JW, Munoz A, Giorgi JV, et al. Plasma viral load and CD4(+) lymphocytes as prognostic markers of HIV-1 infection. *Annals of Internal Medicine* 1997;126(12):946-54.

18. Gulick RM, Mellors JW, Havlir D, et al. Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. *New England Journal of Medicine* 1997;337(11):734-9.

19. Montaner JSG, Reiss P, Cooper D, et al. A randomized, double-blind trial comparing combinations of nevirapine, didanosine, and zidovudine for HIV-infected patients - The INCAS trial. *JAMA-Journal of the American Medical Association* 1998;279(12):930-7.

20. Palella FJ, Delaney KM, Moorman AC, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *New England Journal of Medicine* 1998;338(13):853-60.

21. Chu H, Gange SJ, Li X, et al. HIV viral load trajectory among treatment naive men and women initiating long-term HAART: An application using a segmental Bernoulli/lognormal random effects model with left censoring. *American Journal of Epidemiology* 2006;163(11):S228-S.

22. Yazdanpanah Y, Goldie SJ, Losina E, et al. Lifetime cost of HIV care in France during the era of highly active antiretroviral therapy. *Antiviral Therapy* 2002;7(4):257-66.

23. Levy AR, James D, Johnston KM, et al. The direct costs of HIV/AIDS care. *Lancet Infectious Diseases* 2006;6(3):171-7.

24. Montaner JSG, Hogg R, Wood E, et al. The case for expanding access to highly active antiretroviral therapy to curb the growth of the HIV epidemic. *Lancet* 2006;368(9534):531-6.

25. Hogg RS, Rhone SA, Yip B, et al. Antiviral effect of double and triple drug combinations amongst HIV-infected adults: lessons from the implementation of viral load-driven antiretroviral therapy. *AIDS* 1998;12(3):279-84.

26. Quinn TC, Wawer MJ, Sewankambo N, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1. *New England Journal of Medicine* 2000;342(13):921-9.

27. Cohen MS, Gay C, Kashuba ADM, et al. Narrative review: Antiretroviral therapy to prevent the sexual transmission of HIV-1. *Annals of Internal Medicine* 2007;146(8):591-601.

28. Lima VD, Johnston K, Hogg RS, et al. Expanded access to highly active antiretroviral therapy: A potentially powerful strategy to curb the growth of the HIV epidemic. *Journal of Infectious Diseases* 2008;198(1):59-67.

29. Johnston K, Levy A, Briggs A, et al. The impact of heterogeneity in estimating the lifetime direct medical costs associated with HIV/AIDS in British Columbia, Canada. 2009.

30. Beck EJ, Tolley K. Financing HIV service provision in England: estimated impact of the cost of antiretroviral combination therapy. *International Journal of STD & AIDS* 1998;9(9):512-7.

31. Crane L, Crowe R, Fine S, et al. Hospital and outpatient health services utilization among HIV-infected patients in care in 1999. *Journal of Acquired Immune Deficiency Syndromes* 2002;30(1):21-6.

32. Bozzette SA, Joyce G, McCaffrey DF, et al. Expenditures for the care of HIV-infected patients in the era of highly active antiretroviral therapy. *New England Journal of Medicine* 2001;344(11):817-23.

33. Gebo KA, Chaisson RE, Folkemer JG, et al. Costs of HIV medical care in the era of highly active antiretroviral therapy. *AIDS* 1999;13(8):963-9.

34. Garattini L, Tediosi F, Di Cintio E, et al. Resource utilization and hospital cost of HIV/AIDS care in Italy in the era of highly active antiretroviral therapy. *AIDS Care-Psychological and Socio-Medical Aspects of AIDS/HIV* 2001;13(6):733-41.

35. Tramarin A, Campostrini S, Postma MJ, et al. A multicentre study of patient survival, disability, quality of life and cost of care - Among patients with AIDS in northern Italy. *Pharmacoeconomics* 2004;22(1):43-53.

36. Roberts RR, Kampe LM, Hammerman M, et al. The cost of care for patients with HIV from the provider economic perspective. *AIDS Patient Care and STDs* 2006;20(12):876-86.

37. Chen RY, Accortt NA, Westfall AO, et al. Distribution of health care expenditures for HIV-infected patients. *Clinical Infectious Diseases* 2006;42(7):1003-10.

38. Krentz HB, Gill MJ. Cost of medical care for HIV-infected patients within a regional population from 1997 to 2006. *HIV Medicine* 2008;9(9):721-30.

39. Chamberlayne R, Green B, Barer ML, et al. Creating a population-based linked health database: A new resource for health services research. *Canadian Journal of Public Health-Revue Canadienne De Sante Publique* 1998;89(4):270-3.

40. King JT, Justice AC, Roberts MS, et al. Long-term HIV/AIDS survival estimation in the highly active antiretroviral therapy era. *Medical Decision Making* 2003;23(1):9-20.

41. Schackman BR, Gebo KA, Walensky RP, et al. The lifetime cost of current human immunodeficiency virus care in the United States. *Medical Care* 2006;44(11):990-7.

42. Barton S. Which clinical studies provide the best evidence? The best RCT still trumps the best observational study. *British Medical Journal* 2000;321(7256):255-6.

43. Fang CT, Hsu HM, Twu SJ, et al. Decreased HIV transmission after a policy of providing free access to highly active antiretroviral therapy in Taiwan. *Journal of Infectious Diseases* 2004;190(5):879-85.

44. Piantadosi S, Byar DP, Green SB. THE ECOLOGICAL FALLACY. *American Journal of Epidemiology* 1988;127(5):893-904.

45. Granich R, Gilks C, Dye C, et al. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. *Lancet* 2009;373:48-57.

46. Velasco-Hernandez JX, Gershengorn HB, Blower SM. Could widespread use of combination antiretroviral therapy eradicate HIV epidemics? *Lancet Infectious Diseases* 2002;2(8):487-93.

47. Baggaley RF, Garnett GP, Ferguson NM. Modelling the impact of antiretroviral use in resource-poor settings. *PLoS Medicine* 2006;3:493-504.

48. Salomon JA, Hogan DR. Evaluating the impact of antiretroviral therapy on HIV transmission. *AIDS* 2008;22:S149-S59.

49. Wasserheit JN, Aral SO. The dynamic topology of sexually transmitted disease epidemics: Implications for prevention strategies. *Journal of Infectious Diseases* 1996;174:S201-S13.

50. Stahl JE. Modelling methods for pharmacoeconomics and health technology assessment - An overview and guide. *Pharmacoeconomics* 2008;26(2):131-48.

51. Sonnenberg FA, Beck JR. Markov models in medical decision making: A practical guide. *Medical Decision Making* 1993;13(4):322-38.

52. Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. *Health Economics* 2003;12(10):837-48.

53. Mellors JW, Margolick JB, Phair JP, et al. Prognostic value of HIV-1 RNA, CD4 cell count, and CD4 cell count slope for progression to AIDS and death in untreated HIV-1 infection. *JAMA-Journal of the American Medical Association* 2007;297(21):2349-50.

54. Beral V, Coutinho R, Darbyshire J, et al. Short-term risk of AIDS according to current CD4 cell count and viral load in antiretroviral drug-naive individuals and those treated in the monotherapy era. *AIDS* 2004;18(1):51-8.

2. CHARACTERIZING TRAJECTORIES OF PLASMA HIV-1 RNA VIRAL LOAD AND CD4 COUNT IN INDIVIDUALS TREATED WITH HIGHLY ACTIVE ANTIRETROVIRAL THERAPY*

2.1. Introduction

Progression of the human immunodeficiency virus (HIV) can be characterized using both plasma HIV-1 RNA viral load (pVL) and CD4 cell count. These prognostic variables are strong predictors of morbidity and mortality,(1, 2, 3, 4) and are used to initiate and monitor treatment with antiretroviral therapy.(5)

The trajectories of these variables have relevance for individuals and for populations. For individuals infected with HIV, CD4 cell counts can be used to estimate the time elapsed between seroconversion and diagnosis and to estimate when antiretroviral therapy should be initiated based on threshold guidelines,(6) while pVL measurements are used to estimate the efficacy of antiretroviral therapy.(3) Both CD4 and pVL are used to predict short-term risk of progression to acquired immunodeficiency syndrome (AIDS) and death.(1, 2, 7, 8) Greater insight into the trajectories of these variables over time and their relationships with individual-level characteristics could allow clinicians to extrapolate longer-term predictions.

---

* A version of this chapter will be submitted for publication. Johnston KM, Levy AR, Tyndall M, Hogg RS, Gustafson P, Montaner JS. Characterizing trajectories of plasma HIV-1 RNA viral load and CD4 count in individuals treated with highly active antiretroviral therapy.

At the population level, better understanding of the average trajectories of pVL and CD4 cell count under current treatment practices could allow for improved ability to predict long-term population distributions of these variables, which has implications for health services utilization.(9) Simulation models used to make economic projections require estimates of pVL and CD4 over time, e.g.(10, 11) To date, the incorporation of pVL and CD4 into economic models has been through stratification of the variables into broad categories. The ability to estimate individual-level trajectories of CD4 and pVL on a continuous scale has the potential to lead to more flexible simulation methods.

The typical trajectory of pVL and CD4 for an untreated individual has been characterized previously, and is shown in Figure 2.1.(12) Both trajectories are non-linear in time, making them difficult to describe via standard parametric statistical models. The introduction of effective antiretroviral medications—specifically, highly active antiretroviral therapy (HAART)—has dramatically altered these trajectories, with individuals experiencing extended periods of suppressed pVL and increased CD4 cell count following treatment initiation.(13) The form of trajectories has become more difficult to describe via statistical models following the introduction of HAART. Firstly, treatment failure may result in decreases in CD4 and increases in pVL, followed by subsequent improvement as the individual is placed on a successful new treatment regimen; this process may be repeated several times over the course of treatment,(14) potentially resulting in several localized maxima and minima to be incorporated into a model. Secondly, during periods of suppressed pVL, individuals may experience unpredictable transient periods ("blips") of detectable pVL levels.(15) These blips are more difficult than rounds of

treatment failure to incorporate into the functional form of a statistical model, since the causes and timing are not well understood theoretically.

Several statistical models have been proposed to describe pVL and CD4 trajectories for individuals treated with HAART. These models have either included an assumption of linearity(16) or allowed for non-linearity, but pre-specifed a functional form *a priori*,(17, 18) which may not capture all dynamics of actual trajectories.

The objective of this study was to characterize the trajectories over time of pVL and CD4 for individuals infected with HIV and treated with HAART. In particular, the goal was to describe average trajectories over time at the population level, as well as to develop the ability to generate individual-level projections that would capture both short-term dynamics and long-term trends of pVL and CD4 over the course of an individual infection. The latter individualized trajectories can then be used within a simulation model of the clinical and economic course of HIV in order to quantify health services implications during the modern HAART era.

2.2. Methods

2.2.1. Study population

The study population comprised a prospective population-based cohort of individuals infected with HIV in British Columbia (BC), Canada. In BC, all antiretroviral medications are centrally distributed through the BC Centre for Excellence in HIV/AIDS Research Drug Treatment Program. The HAART Observational Medical Evaluation and Research cohort consists of all antiretroviral-naïve individuals who initiated treatment with HAART between August, 1996 and

September, 2003 (n=2,217).(19) For these individuals, all pVL and CD4 measurements collected through the Drug Treatment Program were available, including those taken prior to initiation of HAART. This resulted in 44,332 pVL measurements collected between April 1996 and March 2005 and 49,466 CD4 count measurements collected between May 1986 and March 2005.

2.2.2. Variable measurement

pVL and CD4 cell count were monitored according to a frequency set out in clinical guidelines, approximately every three months.(20) pVL was measured using the Roche Amplicator Monitor Assay (Roche Molecular Systems, Mississauga, Canada).The sensitivity of pVL testing varied over time, with lower detection limits ranging from 50-500 copies/mL and upper detection limits ranging from 100,000-1,000,000 copies/mL. Observations falling outside the limits of detection were imputed assuming a normal distribution.(21) A sensitivity analysis was performed in which censored viral load observations were truncated to the limits of detection rather than imputed, and results (not shown) were similar. Adherence was defined by the percentage of months in which antiretroviral medications were dispensed during the first year of HAART.(22, 23)

Genotypic medication resistance testing was performed approximately every three months for a subset of 1,617 individuals, with resistance separated into four categories: lamivudine (3TC), other nucleoside reverse-transcriptase inhibitors (NRTI), non-nucleoside reverse-transcriptase inhibitors (NNRTI), and protease inhibitors (PI). Resistance was defined longitudinally as a

factor with levels determined by cumulative number of resistance categories detected, plus a separate level for individuals never tested.

2.2.3. Statistical methodology

Statistical models for CD4 cell count and log-transformed pVL were fit using the generalized additive modelling framework to account for non-linear trajectories over time.(24, 25) A generalized additive model can be thought of as analogous to a linear regression model, but with some or all linear coefficients replaced by non-parametric, non-linear functions relating an outcome to a set of covariates. For a normally-distributed outcome variable y, a matrix $\boldsymbol{X}$ of covariates for which a linear relationship is assumed, and covariates $x^{(1)}, \ldots, x^{(p)}$ for which non-parametric relationships are assumed, a generalized additive model is written:

$$y_i = X\theta + \sum_{j=1}^{p} s_j(x_i^{(j)}) + \varepsilon_i \quad (1)$$

for parameters $\boldsymbol{\theta}$ and functions $s_j(x_i^{(j)})$. Under the generalized additive modelling framework, no functional restrictions are placed on these relationships, although the functions are fit using penalized splines to avoid over-fitting the data and are restricted to be centered around zero so that the model intercept is well-defined. Penalized splines are a methodology for fitting a function to data that is a middle-ground between the two extremes of either letting the data determine the function exactly and assuming a linear relationship. The amount of smoothing was optimized using the generalized cross-validation score, which is a goodness-of-fit criteria that accounts for the additional error associated with generalizing a model to a new population.(25) A

constant multiplier was used to inflate the model effective degrees of freedom to check for overfitting.(25, 26, 27)

After fitting several models with various covariate combinations including interaction terms, final models for pVL and CD4 cell count were chosen based on a combination of statistical significance and clinical interest. Both models contained linear terms for age at cohort entry, sex, use of injection drugs (binary), year of first antiretroviral use, first HAART combination (NNRTI, PI, or ritonavir-boosted PI), and cumulative number of resistance mutations detected (time-varying). These variables were included via linear terms, as opposed to smooth functions, either because they were categorical in nature, or because exploratory analyses suggested a linear relationship with the outcome variables.

The pVL model included smooth terms for adherence, time relative to first antiretroviral treatment, and most recent previous pVL. The CD4 model included smooth terms for adherence, most recent pVL measurement, time relative to first antiretroviral treatment, and most recent previous CD4 cell count. The motivation for including recent pVL in the CD4 model and not including recent CD4 in the pVL model was the fact that there are several biological mechanisms by which pVL directly affects CD4 cells,(28) while the converse is less important following the initial stages of primary infection. In addition to the multivariable models, unadjusted models were fit for each smoothed term individually to estimate the crude relationships with pVL and CD4, respectively.

Because both multivariable models included terms for most recent measurement, the correlation between multiple observations from the same individual was implicitly taken into account. However, when fitting crude models for variables other than most recent measurement, auto-correlated error terms were used to account for within-subject correlation of measurements. All analyses were performed using version 2.3.1 of the R package,(27) including the mgcv 1.3-17 library.(25)

2.3. Results

The cohort of 2,217 individuals provided a median of 19 pVL measurements and 21 CD4 measurements per patient over a median follow-up of 64 months (interquartile range (IQR): 38-90). Median pVL at start of antiretroviral therapy was 5.09 log copies/mL (IQR: 4.65-5.51), while median CD4 count was 210 cells/mm^3 (IQR: 90-360). The median interval of time between consecutive pVL measurements was 2.7 months (IQR: 1.6-3.7 months); for CD4 it was 2.3 months (IQR: 1.1-3.6 months). Table 2.1 reports further descriptive characteristics for the study sample.

Results from the generalized additive models are summarized in Table 2.2 and Figures 2.2 and 2.3. As described by equation (1), each covariate in the model is related to the outcome via a linear parameter θ or a smooth function $s()$. Coefficients associated with covariates for which a linear relationship was assumed are given in Table 2.2. Interpretation of these coefficients is analogous to the interpretation of coefficients from a standard linear model. For the remaining variables, smooth functions were used to describe their relationships with pVL and CD4, respectively. In this case, the relationship between a variable and outcome cannot be described

by a single parameter, and is best described graphically. Note that the functions shown in Figures 2.2 and 2.3 do *not* represent predicted trajectories, but, rather, show the estimated non-linear relationship between explanatory and outcome variables. Predicted trajectories of pVL and CD4 can be produced by summing all linear covariates and non-linear functional values for a particular combination of covariates.

The smooth functions estimated within the crude and multivariable pVL models are shown in Figures 2.2(a)-(f). The left-hand column displays crude, unadjusted curves resulting from a model with only the single smooth function, while the right-hand column displays smooth functions resulting from the full multivariable model. A monotonic decreasing relationship was observed between adherence and pVL (Figure 2.2(a)-(b)), although this relationship was dampened in the adjusted model (Figure 2.2(b)). Average pVL trajectory over time was characterized by a steady increase until treatment initiation, followed by a large drop immediately after treatment initiation. In the first five to eight years of treatment, the trajectory displayed several brief rebounds (Figure 2.2(d)), potentially representing treatment failures and subsequent switching to a successful regimen. Overall, pVL appeared to remain at relatively constant levels for more than eight years of antiretroviral treatment, the longest follow-up available for study subjects (Figure 2.2(c)-(d)). For pVL between about 2-6 log copies/mL, most recent pVL measurement was predictive of current pVL in a positive, approximately linear fashion. For a most recent pVL greater than 6 log copies/mL, there was a slightly negative relationship, suggesting that extremely high levels of pVL tended to be transient phenomena. For pVL less than 2 log copies/mL, most recent pVL was observed to have a constant effect on current pVL (Figure 2.2(e)-(f)).

The smooth functions estimated within the crude and multivariable CD4 models are shown in Figure 2.3(a)-(h). The crude relationship between CD4 count and adherence to antiretroviral medication was monotonic increasing (Figure 2.3(a)), and, similarly to pVL, when adjusted for other variables, the relationship became roughly constant(Figure 2.3(b)). Recent pVL levels less than about 4 log copies/mL had a constant, positive, effect on CD4, with a steep decreasing trend for pVL between 4 and 6 log copies/mL. Above 6 copies/mL, the relationship was positive, and characterized by greater uncertainty (Figure 2.3(c)-(d)). CD4 cell count decreased over time until the initiation of antiretroviral therapy; following initiation, the crude trajectory of CD4 over time displayed an increase (Figure 2.3(e)), while the adjusted relationship was roughly constant over time (Figure 2.3(f)). The most recent CD4 count was related to current CD4 count in a positive, approximately linear manner for all but extremely high values of most recent CD4 (Figure 2.3(g)-(h)).

Figure 2.4 displays hypothetical trajectories of (a) pVL and (b) CD4 for an individual with typical covariate values. The estimated trajectories are based on a male, non-injection drug user, initiating a PI-based HAART regimen. Age and adherence were assumed to be the mean values observed within the cohort (Table 2.1). Different curves are shown for different numbers of resistance mutations over the course of infection. For trajectories based on one or two resistance mutations, the time of mutation development was assumed to be the average time observed within the HOMERcohort (1.5 years for first mutation, 2.1 years for second mutation).The initiation of HAART was associated with dramatic decreases in pVL and increases in CD4, respectively. Over the duration of follow-up, CD4 cell count remained relatively stable, while

pVL displayed localized maxima at approximately three years and five years following HAART initiation, as well as a small increasing trend over time. During the first seven years following HAART initiation, CD4 cell count was estimated to reach a minimum of approximately 500 cells/mm^3, while pVL was estimated to reach a maximum of approximately 3.0 log copies/mL.

The pVL and CD4 multivariable models were assessed based on the percentage of null deviance described by the full models. In the context of a generalized additive model, this statistic is analogous to the R^2 statistic resulting from a linear model. The deviance explained by the pVL model was 63%; for the CD4 model it was 77%, suggesting that the models were able to describe the majority of variability displayed by both prognostic variables. Examples of predictive ability of the models are given in Figures 2.5 and 2.6, which display actual and predicted pVL and CD4 trajectories for 15 randomly-selected individuals. Both models were successful at predicting short- and long-term trends for individuals who provided more measurements; there were greater discrepancies between the curves for individuals providing only two or three measurements or with long intervals between measurements.

2.4. Discussion

2.4.1. Interpretation

In this study, we used generalized additive models to describe population-average and individual-level trajectories of pVL and CD4 trajectories, and to elucidate dynamic relationships between these prognostic laboratory variables and other relevant covariates. In the multivariable pVL model, the smooth functions associated with medication adherence and time on HAART were both dampened after adjusting for most recent pVL measurement; smooth

functions associated with these variables were also dampened in the multivariable CD4 model after adjusting for most recent pVL and CD4 measurements. This does not imply that medication adherence and time on HAART do not influence pVL and/or CD4; at any given time point, recent pVL and CD4 measurements are themselves likely to be influenced by multiple factors such as medication adherence or time on HAART, so that these variables are already implicitly accounted for. The multivariable models suggest that, given recent pVL and CD4 measurements, explicitly adjusting for adherence and time on HAART offers limited additional information in predicting future pVL and CD4. For this reason, the primary utility of the multivariable models is with respect to predictive ability. The crude models, as displayed in Figure 2.2(a)/(c)/(e) and Figure 2.3(a)/(c)/(e)/(g), are better tools for exploring the relationship between a particular variable and CD4 and/or pVL.

2.4.2. Strengths and limitations

With respect to predictive ability, the multivariable generalized additive models appear to adequately describe pVL and CD4 trajectories at the individual level, both long-term trends and short-term dynamics. The latter are not always well understood theoretically, particularly temporary and dramatic changes in measurements, suggesting advantages to using non-parametric statistical models in this context. Generalized additive models have been suggested as a method for reducing bias in estimates when the exact functional form of a relationship is not known.(29) However, a potential drawback to data-driven methods is the risk of over-fitting a model to the data at hand such that model results are not generalizable. To address this, we re-fit the models inflating the contribution of each effective degree of freedom by a factor of 1.4, as suggested to correct for potential over-fitting,(25, 26) and found no difference in the

interpretation. In addition, we fit the same models using only randomly-selected subsets of individuals, and compared results to those obtained using the entire dataset. Results (not shown here) were not found to be appreciably different, and fitted vs. actual trajectory plots (analogous to those shown in Figures 2.5 and 2.6) showed similar accuracy for individuals who were included in the model fit compared to those who were not. Thus, given that data were population-based for BC, it seems likely that these results are generalizable to other populations following similar treatment guidelines.(20)

There were several limitations to this study. Longitudinal data may produce overly optimistic estimates of pVL and CD4, because individuals who do not respond to treatment would be expected to experience higher pVL and lower CD4 than those whose disease is successfully suppressed.(31) Such individuals would also be expected to live shorter lives and thus contribute fewer measurements and have less of an impact on model fit. This is empirically supported here because the models were less predictive of actual trajectories for individuals providing fewer measurements. However, the success rate of HAART was high as shown by the 91% of individuals who provided at least one year of follow-up after initiating HAART, and 79% providing at least two years. We therefore expect that the models are reliable for the majority of individuals accessing HAART. Future refinements could be to adapt methods of accounting for this bias that have been developed for alternative modelling techniques e.g.(16, 32, 33, 34) to a generalized additive modelling framework.

The measure of antiretroviral adherence used here was a cross-sectional measure taken one year following initiation of antiretroviral medication. While this measure has been validated and used

previously,(22, 23) greater insight might be gained by including a dynamic measure of adherence.

2.4.3. Potential applications

Trajectories generated from the multivariable models can be applied in health services applications. Average trajectories for specific covariate combinations can be used to make population-level projections regarding disease-stage distribution or expected morbidity and mortality over time, based on the relationship between these outcomes and pVL and CD4. Individual trajectories can be used within microsimulation models designed to estimate economic parameters such as cost-effectiveness of a particular therapy or economic burden of illness. In addition to health services applications, the ability to accurately predict future pVL and CD4 trajectories at the individual level has potential implications for clinical practice. For example, when deciding when to initiate treatment with HAART in an asymptomatic individual, a clinician must estimate when the individual's CD4 count is expected to drop to levels such that they are eligible for treatment based on current guidelines, without allowing it to reach a point where they are at an increased risk for clinical events.(35) Predictive models of CD4 and pVL that account for individual-level characteristics would allow for more precise estimation of this time point. A potential future extension of this work is to estimate positive and negative predictive values of the model-based estimated time until particular critical CD4 thresholds are reached and to develop risk equations analogous to those created for other disease areas.(36)

2.5. Conclusion

We used population-based observational data to describe longitudinal trajectories of pVL and CD4 cell count in individuals infected with HIV. We found that the method was successful for describing both long- and short-term trends, at both the population-average and individual levels. Improved ability to predict short-term pVL and CD4 measures has the potential to lead to better informed treatment decisions, while knowledge of individual trajectories and population trends of prognostic disease markers can help in predicting future demand for HIV-related health services.

Figure 2.1: Untreated history of CD4 cell count and plasma viral load.
(Copyright © 2003 International Medical Press. All rights reserved. Reprinted with permission.)

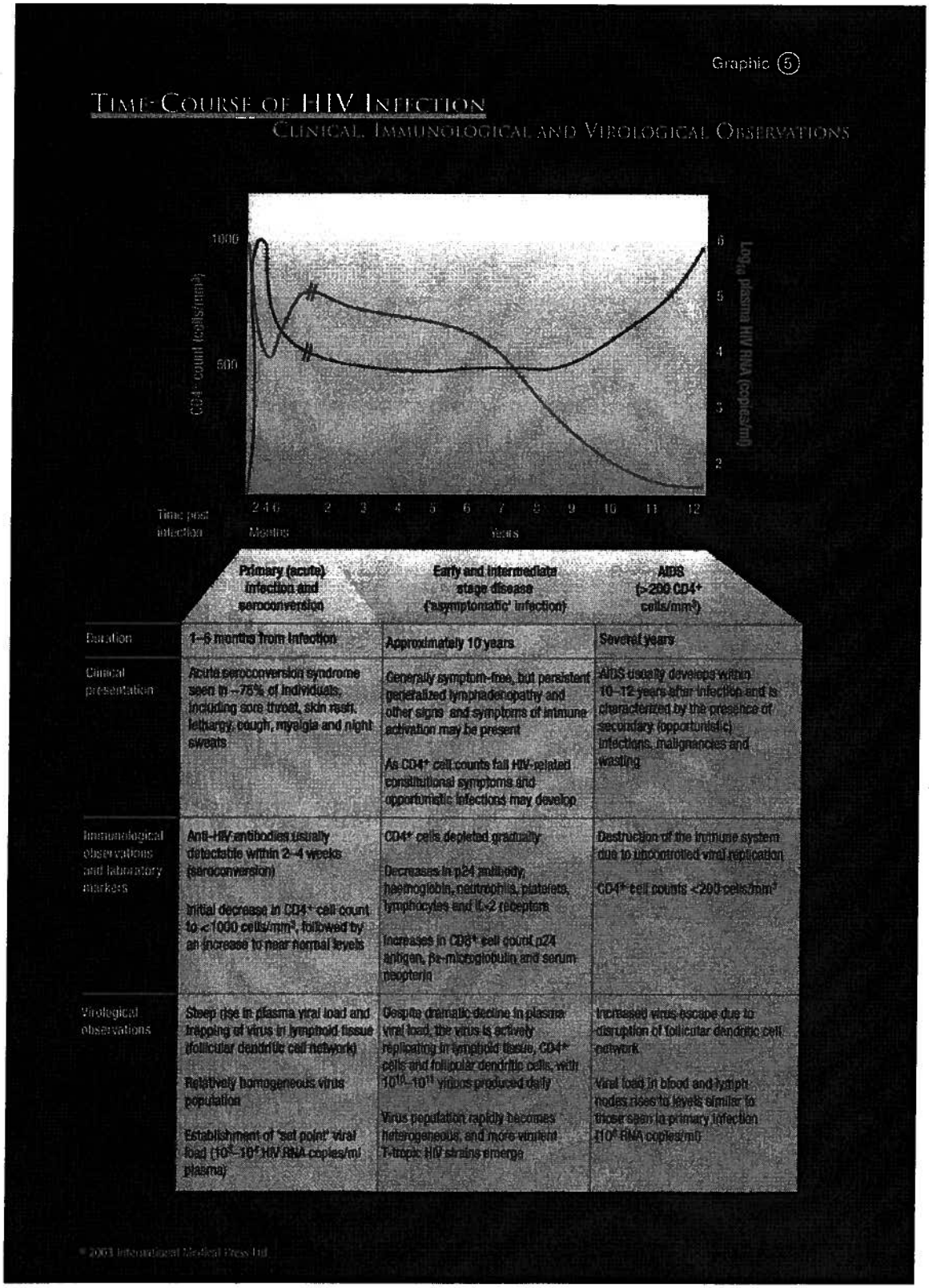

Figure 2.2: Non-parametric results of generalized additive models of plasma viral load.
Smooth terms (with 95% confidence intervals) included in viral load models, both crude and multivariable: (a/b) adherence to antiretroviral medications, (c/d) time relative to first antiretroviral treatment, (e/f) most recent viral load. Smooth terms represent non-linear contributions of covariates to viral load. Note that the scale is meaningful, but the zero is arbitrary, so absolute values should not be interpreted. Multivariable models also adjusted for linear relationships with age, year of HAART initiation, initial HAART regimen, injection drug use, sex, and age. HAART = highly active antiretroviral therapy.

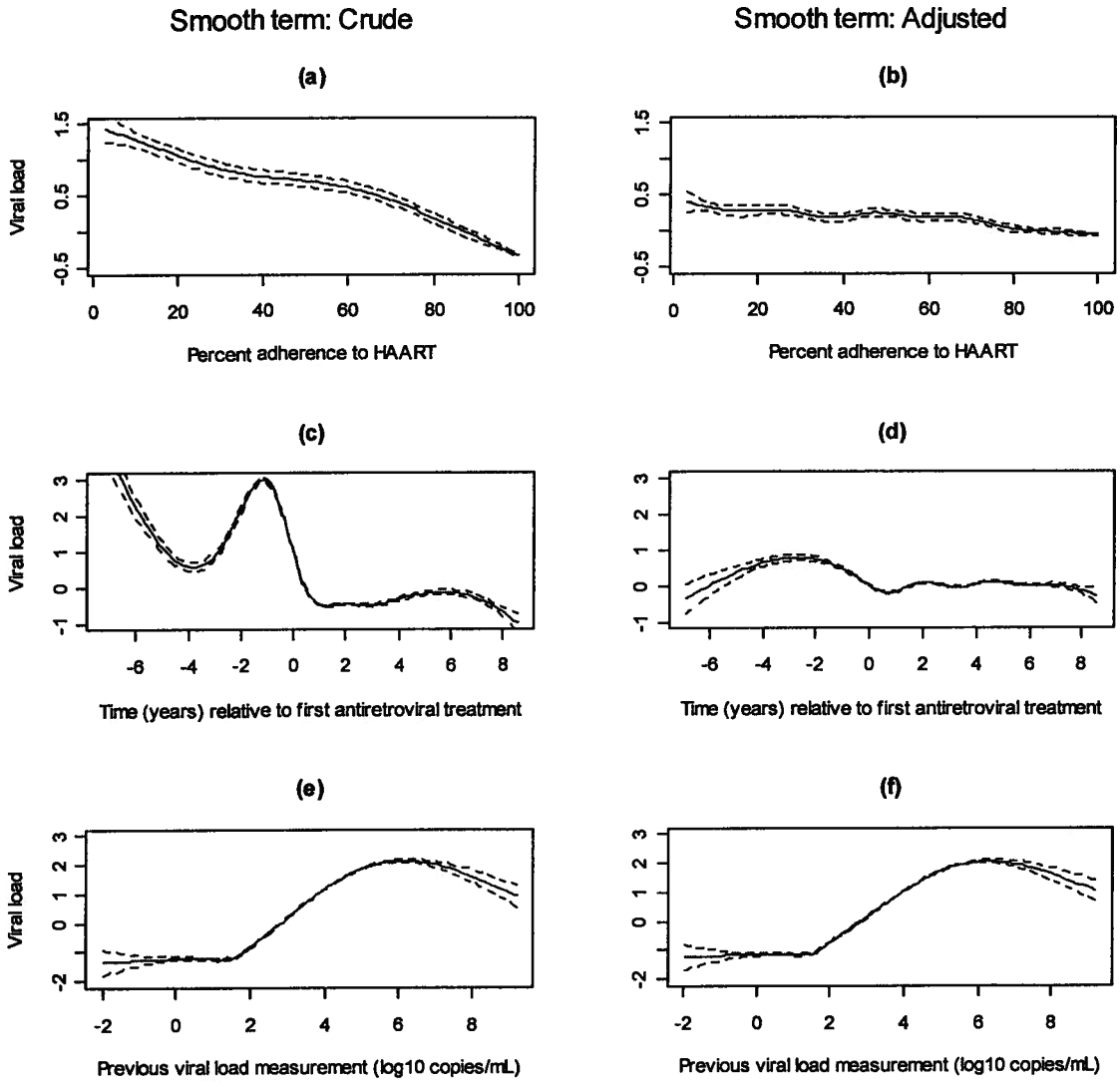

Figure 2.3: Non-parametric results of generalized additive models of CD4 cell count.
Smooth terms (with 95% confidence intervals) included in CD4 cell count models, both crude and multivariable: (a/b) adherence to antiretroviral medications, (c/d) most recent viral load measurement, (e/f) time relative to first antiretroviral treatment, (g/h) most recent CD4 cell count. Smooth terms represent non-linear contributions of covariates to CD4 cell count. Note that the scale is meaningful, but the zero is arbitrary, so absolute values should not be interpreted. Multivariable models also adjusted for linear relationships with age, year of HAART initiation, initial HAART regimen, injection drug use, sex, and age. HAART = highly active antiretroviral therapy.

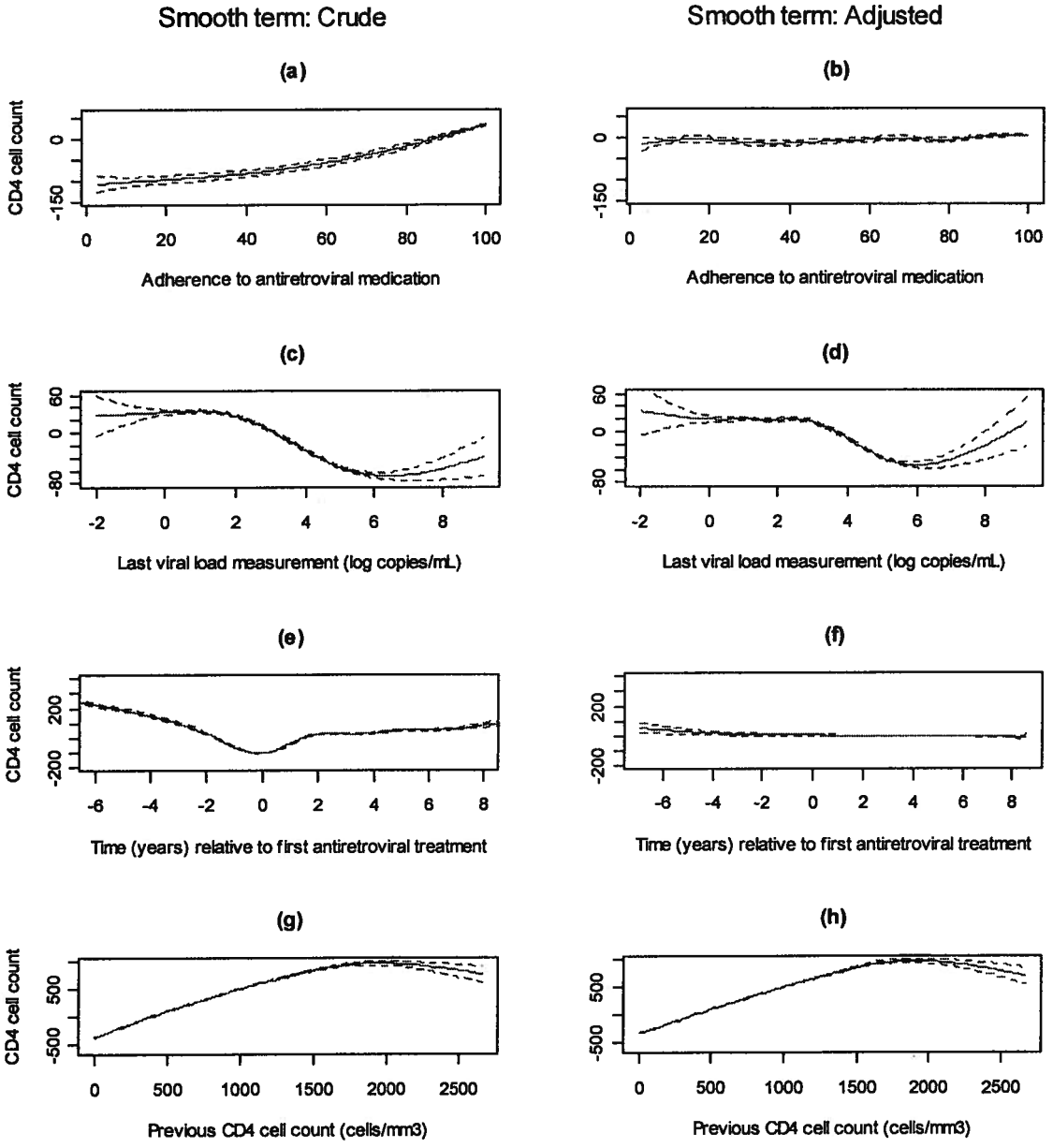

Figure 2.4: Hypothetical "typical" trajectories for (a) log viral load, and (b) CD4 cell count. Separate trajectories for individuals with no resistance mutations, individuals who develop one resistance mutation, and individuals who develop two resistance mutations. Resistance mutations assumed to occur at average times observed in HOMER cohort.

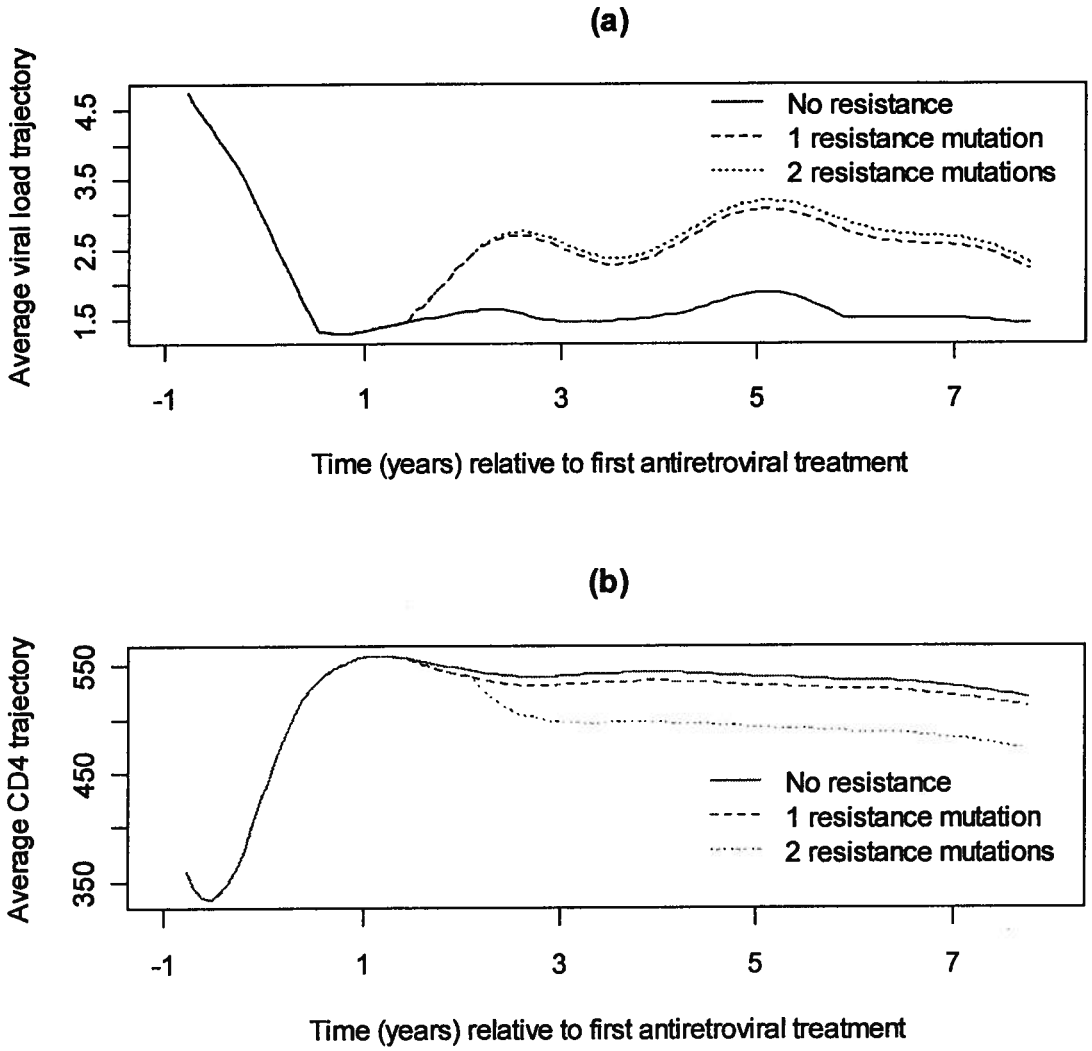

Figure 2.5: Observed and predicted trajectories of viral load.
15 randomly selected HIV-positive individuals being treated with antiretroviral medication in British Columbia, Canada.

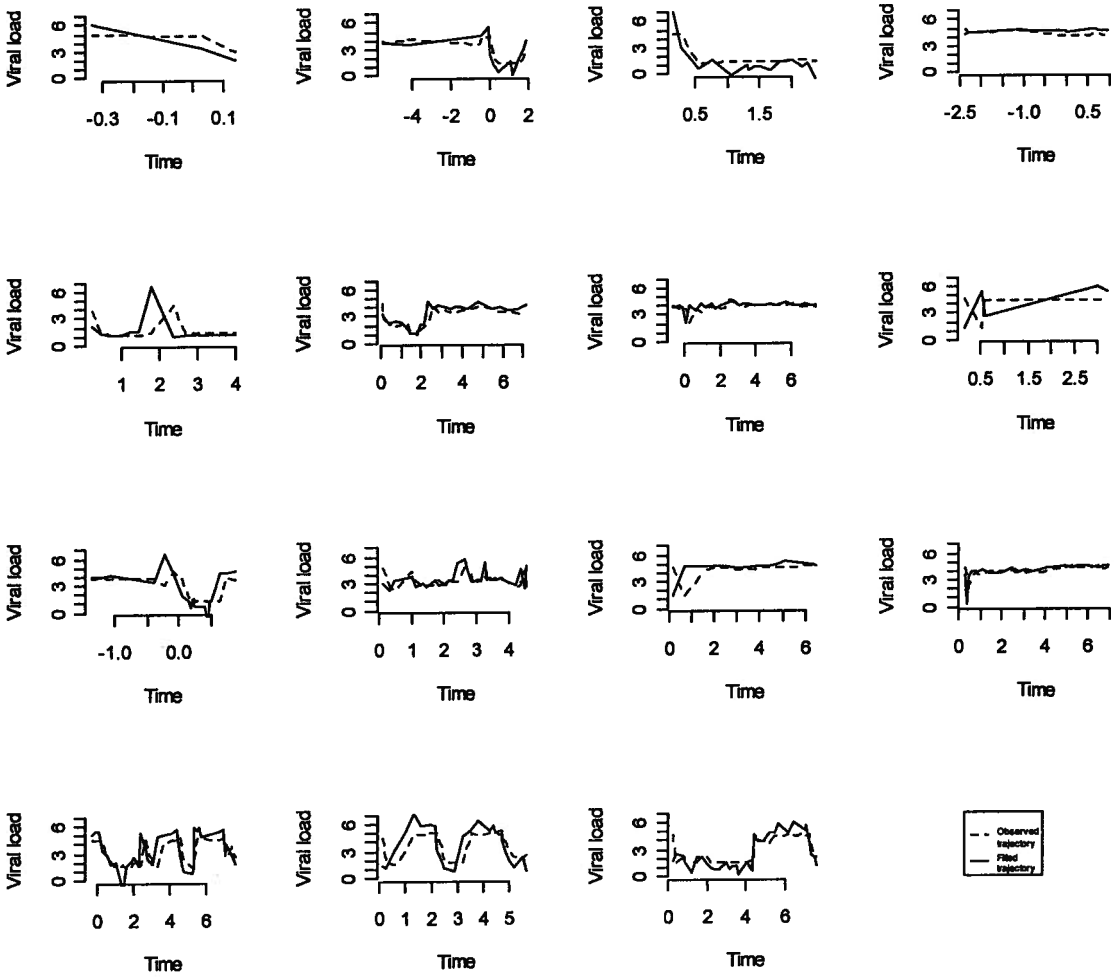

Figure 2.6: Observed and predicted trajectories of CD4 cell count.
15 randomly selected HIV-positive individuals being treated with antiretroviral medication in British Columbia, Canada.

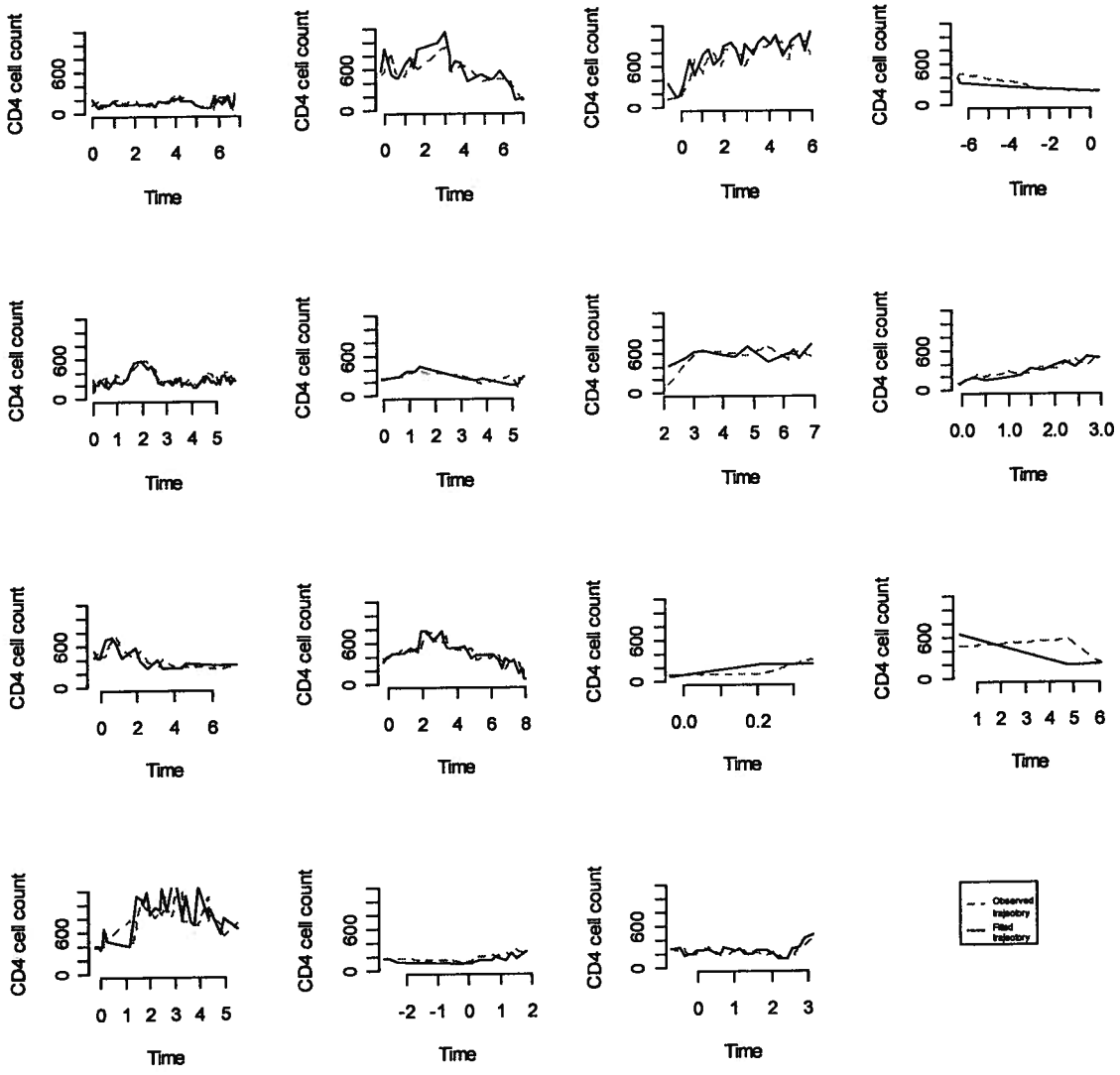

Table 2.1: Description of the HOMER cohort (1996-2005).
Antiretroviral-naïve individuals initiating treatment with highly active antiretroviral therapy. (n=2,217)

Mean age (y) (standard deviation [s.d.])	37.7 (9.8)
Sex (% male)	82.0
Mean % adherence to antiretroviral therapy (s.d.)	78.5 (29.8)
Injection drug user (%)	24.7
Year of first antiretroviral therapy (%) 1996-1998 1999-2001 2002-2003	 41.6 39.6 18.8
First antiretroviral regimen* (%) NNRTI PI Ritonavir-boosted PI	 38.2 47.2 14.6
Maximum number of resistance categories detected (%) 0 1 2 3 4 Not Tested	 47.1 10.5 8.7 4.5 1.0 28.1

*NNRTI= non-nucleoside reverse transcriptase inhibitor, PI=protease inhibitor

Table 2.2: Estimated linear coefficients for multivariable generalized additive models.

Variable	Viral load model (log copies/mL) (Standard error)	CD4 cell count model (cells/mm^3) (Standard error)
Year of first antiretroviral use (relative to 1996)	0.019* (0.004)	-4.16* (0.48)
Sex		
Female	0.000 (---)	0.000 (---)
Male	0.009* (0.015)	1.75 (1.86)
Injection drug user		
No	0.000 (---)	0.000 (---)
Yes	0.034* (0.013)	-2.97* (1.47)
First antiretroviral regimen		
NNRTI	0.000 (---)	0.000 (---)
PI	0.022* (0.014)	0.17 (1.67)
Ritonavir-boosted PI	-0.002 (0.018)	-8.77* (2.11)
Age at cohort entry (per 10 y)	-0.033* (0.006)	-1.03 (0.67)
Cumulative number of resistance categories**		
0	0.000 (---)	0.000 (---)
1	0.155* (0.018)	-1.63 (2.13)
2	0.169* (0.021)	-8.92* (2.49)
3	0.190* (0.030)	-18.08* (3.63)
4	0.146* (0.053)	-20.57* (6.58)

* Statistically significant at $\alpha = 0.05$ level

**Resistance categories comprised of all NRTIs, all NNRTIs, all PIs, and lamivudine alone

2.6. References

1. Mellors JW, Rinaldo CR, Gupta P, et al. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. Science 1996;272:1167-70.
2. Mellors JW, Munoz A, Giorgi JV, et al. Plasma viral load and CD4(+) lymphocytes as prognostic markers of HIV-1 infection. Annals of Internal Medicine 1997;126:946-54.
3. Obrien WA, Hartigan PM, Martin D, et al. Changes in plasma HIV-1 RNA and CD4+ lymphocyte counts and the risk of progression to AIDS. New England Journal of Medicine 1996;334:426-31.
4. Fahey JL, Taylor JMG, Detels R, et al. The prognostic value of cellular and serologic markers in infection with human immunodeficiency virus type-1. New England Journal of Medicine 1990;322:166-72.
5. De Milito A, Titanji K, Zazzi M. Surrogate markers as a guide to evaluate response to antiretroviral therapy. Current Medicinal Chemistry 2003;10:349-65.
6. AIDSInfo. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. 2007.
7. Mellors JW, Margolick JB, Phair JP, et al. Prognostic value of HIV-1 RNA, CD4 cell count, and CD4 cell count slope for progression to AIDS and death in untreated HIV-1 infection. Journal of the American Medical Association 2007;297:2349-50.
8. Beral V, Coutinho R, Darbyshire J, et al. Short-term risk of AIDS according to current CD4 cell count and viral load in antiretroviral drug-naive individuals and those treated in the monotherapy era. AIDS 2004;18:51-8.

9. Levy AR, James D, Johnston KM, et al. The direct costs of HIV/AIDS care. Lancet Infectious Diseases 2006;6:171-7.

10. Freedberg KA, Losina E, Weinstein MC, et al. The cost effectiveness of combination antiretroviral therapy for HIV disease. New England Journal of Medicine 2001;344:824-31.

11. Yazdanpanah Y, Goldie SJ, Losina E, et al. Lifetime cost of HIV care in France during the era of highly active antiretroviral therapy. Antiviral Therapy 2002;7:257-66.

12. Richman D. Human Immunodeficiency Virus. International Medical Press, 2003.

13. Hunt PW, Deeks SG, Rodriguez B, et al. Continued CD4 cell count increases in HIV-infected adults experiencing 4 years of viral suppression on antiretroviral therapy. AIDS 2003;17:1907-15.

14. Mocroft A, Ledergerber B, Viard JP, et al. Time to virological failure of 3 classes of antiretrovirals after initiation of highly active antiretroviral therapy: Results from the EuroSIDA Study Group. Journal of Infectious Diseases 2004;190:1947-56.

15. Pierson T, McArthur T, Siliciano RF. Reservoirs for HIV-1: Mechanisms for viral persistence in the presence of antiviral immune responses and antiretroviral therapy. Annual Review of Immunology 2000;18:665-708.

16. Thiebaut R, Jacqmin-Gadda H, Babiker A, et al. 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.

17. Chu H, Gange SJ, Li X, et al. HIV viral load trajectory among treatment naive men and women initiating long-term HAART: An application using a segmental

Bernoulli/lognormal random effects model with left censoring. American Journal of Epidemiology 2006;163:S228-S.

18. Raab GM, Parpia T. Random effects models for HIV marker data: practical approaches with currently available software. Statistical Methods in Medical Research 2001;10:101-16.

19. Hogg RS, Heath KV, Yip B, et al. Improved survival among HIV-infected individuals following initiation of antiretroviral therapy. Journal of the American Medical Association 1998;279:450-4.

20. British Columbia Centre for Excellence in HIV/AIDS. Therapeutic guidelines: Antiretroviral therapy for HIV-1 infected adults. 2006.

21. Lynn HS. Maximum likelihood inference for left-censored HIV RNA data. Statistics in Medicine 2001;20:33-45.

22. Hogg RS, Heath K, Bangsberg D, et al. Intermittent use of triple-combination therapy is predictive of mortality at baseline and after 1 year of follow-up. AIDS 2002;16:1051-8.

23. Low-Beer S, Yip B, O'Shaughnessy MV, et al. Adherence to triple therapy and viral load response. Journal of Acquired Immune Deficiency Syndromes 2000;23:360-1.

24. Hastie T, Tibshirani R. Generalized Additive Models. Chapman & Hall/CRC, 1990.

25. Wood S. Generalized Additive Models: An Introduction with R. Chapman & Hall/CRC, 2006.

26. Kim YJ, Gu C. Smoothing spline Gaussian regression: more scalable computation via efficient approximation. Journal of the Royal Statistical Society Series B-Statistical Methodology 2004;66:337-56.

27. R: A Language and Environment for Statistical Computing. 2006.

28. McCune JM. The dynamics of CD4(+) T-cell depletion in HIV disease. Nature 2001;410:974-9.

29. Benedetti A, Abrahamowicz M. Using generalized additive models to reduce residual confounding. Statistics in Medicine 2004;23:3781-801.

30. Harrigan PR, Hogg RS, Dong WWY, et al. Predictors of HIV drug-resistance mutations in a large antiretroviral-naive cohort initiating triple antiretroviral therapy. Journal of Infectious Diseases 2005;191:339-47.

31. Touloumi G, Pocock SJ, Babiker AG, et al. Impact of missing data due to selective dropouts in cohort studies and clinical trials. Epidemiology 2002;13:347-55.

32. Brown ER, Ibrahim JG, DeGruttola V. A flexible B-spline model for multiple longitudinal biomarkers and survival. Biometrics 2005;61:64-73.

33. Lipsitz SR, Fitzmaurice GM, Molenberghs G, et al. Quantile regression methods for longitudinal data with drop-outs: Application to CD4 cell counts of patients infected with the human immunodeficiency virus. Applied Statistics-Journal of the Royal Statistical Society Series C 1997;46:463-76.

34. Degruttola V, Tu XM. Modeling progression of CD4-lymphocyte count and its relationship to survival time. Biometrics 1994;50:1003-14.

35. Henry WK, Tebas P. Presenting plasma HIV RNA level and rate of CD4 T-cell decline - Reply. Journal of the American Medical Association 2007;297:807-.

36. Wilson PWF, D'Agostino RB, Levy D, et al. Prediction of coronary heart disease using risk factor categories. Circulation 1998;97:1837-47.

3. THE IMPACT OF HIGHLY ACTIVE ANTIRETROVIRAL THERAPY ON HOSPITALIZATION AND OTHER MEDICAL COSTS FOR COMMUNITY-DWELLING INDIVIDUALS INFECTED WITH HIV*

3.1. Introduction

The introduction of highly active antiretroviral therapy (HAART) for treating individuals infected with HIV has brought about dramatic decreases in morbidity and mortality.(1, 2) Treatment with HAART is lifelong and, given the increased survival and stable rates of new infections,(3) the net effect is an increase in prevalence of HIV in western counties. In Canada, approximately 58,000 individuals were estimated to be infected with HIV at the end of 2005, a 16% increase from 2002.(3) As such, the costs of treating persons infected with HIV present a growing burden to the Canadian health care system.

Major sources of direct medical costs include: medication costs, hospitalization costs, and outpatient costs (including physician visits, emergency room visits, and laboratory and diagnostic tests).(4) Hospitalizations accounted for approximately 90% of HIV-related medical costs during the early years of the epidemic,(5) and continue to comprise a substantial proportion of overall medical costs.(6) However, the use of HAART has resulted in a decrease in the frequency and cost of hospitalizations,(7) and an overall decrease in non-antiretroviral sources of medical costs.(8) It has been suggested that use of HAART(9) and routine outpatient care(10)

---

* A version of this chapter will be submitted for publication. Johnston KM, Levy AR, Briggs AH, Tyndall M, Lima VD, Gustafson P, Hogg RS, Montaner JSM. The impact of highly active antiretroviral therapy on hospitalization and other medical costs for community-dwelling individuals infected with HIV

are both inversely related to risk of an HIV-related hospitalization. It is important to understand and quantify the impact of HAART on other categories of expenditures because of the implications when allocating relative budgets for HIV-related services.

Both before and after the introduction n of HAART, the direct medical costs associated with HIV were characterized as a "moving target", based on observed variability over time.(5, 11) The impact of HAART on medical costs has previously been estimated by comparing mean costs in the pre- and post-HAART eras.(7, 12) These comparisons may conflate changes brought about directly by the use of HAART with other, unrelated, temporal changes in medical costs. In Vancouver, British Columbia (BC), despite the free-of-charge availability of HAART, 40% of individuals who died of HIV-related causes between 1997 and 2005 were never dispensed HAART.(13) This suggests the existence of a group of individuals experiencing "modern day pre-HAART" conditions, and comparing the direct medical costs incurred by such individuals with those incurred by individuals receiving HAART could potentially account for the impact of HAART availability on other direct medical costs without the need for historical controls.

The aim of this study was to characterize the interplay between different categories of direct medical expenditure in individuals infected with HIV in BC. This expenditure included all sources of medical care, and was not restricted to HIV-related care only. The specific objective was to quantify pairwise associations between categories of utilization and expenditure—particularly the relationships between use of HAART and hospitalization frequency and cost—after adjusting for clinical and sociodemographic covariates.

3.2. Methods

3.2.1. Data sources

Data were obtained from two observational cohorts based in Vancouver, BC: the HAART Observational Medical Evaluation and Research (HOMER) cohort and the Community Health and Safety Evaluation (CHASE) cohort. Each of these cohorts has been linked to the British Columbia Linked Health Database (BCLHD), which contains administrative records of hospitalizations and outpatient physician visits.(14)

In BC, HAART is provided free of charge through a centralized program based out of the Drug Treatment Program (DTP) of the BC Centre for Excellence in HIV/AIDS. The HOMER cohort is comprised of all antiretroviral-naïve individuals in BC who initiated treatment with HAART between 1996 and 2005 (n=2,583). For this cohort, data are available describing baseline sociodemographic characteristics, longitudinal CD4 cell count and plasma viral load, longitudinal antiretroviral prescription records, and longitudinal laboratory test records. A subset of 1,979 individuals in the HOMER cohort were receiving HAART prior to 2002 and were linked to the BCLHD such that administrative health records were available between April 1995 and March 2001.

The CHASE cohort consists of a representative sample of low-income individuals residing in Vancouver's downtown east side (n=3,484). Individuals participating in CHASE completed a baseline questionnaire describing sociodemographic and health services utilization information and provided informed consent to linkages with various health services database, including BCLHD and the St. Paul's Emergency Room database, which captures approximately 75% of

emergency room visits amongst individuals infected with HIV.(15) The CHASE cohort has been linked to the BC Centre for Disease Control's HIV testing database, allowing for the identification of individuals with confirmed HIV infection (n=304), and to the BCLHD between January 1998 and December 2005.

For this study, the HOMER and CHASE cohorts were linked in order to identify individuals included in both. Study inclusion criteria and cohort overlap are shown in Figure 1 and the periods of follow-up and CD4 monitoring based on cohort inclusion are shown in Figure 3.2.

All individuals in the HOMER cohort (n=2,583) were included in the analysis of antiretroviral costs. For the analyses of other health services categories, individuals must have been included in the linkage to BCLHD and had at least one laboratory measurement on record during the period of the linkage. This resulted in 1,831 individuals from HOMER being included in all analyses.

Individuals in the CHASE cohort were considered for inclusion if they had at least one positive HIV test on record with the BC Centre for Disease Control. Individuals were excluded if they had ever accessed antiretroviral treatment but did not meet the inclusion criteria for HOMER, because the majority of such individuals were treatment-experienced with mono- or dual-therapy prior to initiating HAART. In total, 183 individuals from the CHASE cohort were included: 119 were in both CHASE and HOMER, while 64 had never presented for treatment and were thus considered to be treatment-naïve in all analyses. Individuals in CHASE characterized as "HIV- or unknown" either had a recorded negative HIV test or had never been tested by the BC

Centre for Disease Control. Thus, it is likely that some undiagnosed individuals infected with HIV were excluded from this study because their HIV status could not be verified.

This resulted in 2,583 individuals included in the antiretroviral cost analysis, and 1,895 individuals included in all other health services analysis (1,712 in HOMER only, 119 in both HOMER and CHASE, and 64 in CHASE only).

3.2.2. Economic variables

The perspective taken in this analysis was that of the BC Ministry of Health, the third party payer for a large majority of health care services and antiretroviral therapy in BC. Costs were estimated using a combination of charges paid and unit costs. Hospitalization costs were estimated by multiplying the length of each hospital stay in days by a daily unit cost of $610.33, which was estimated using a fully-allocated costing model developed for the province of BC.(16) For physician visits, charges paid by the Ministry of Health were available for each visit recorded in the BCLHD, and these charges were applied directly. The unit costs associated with laboratory tests were obtained directly from a centralized testing facility, and were to be $60 for a plasma viral load assay, $65 for a CD4 cell count test, and $230 for a genotype based HIV drug resistance assay. The cost of an emergency room visit was estimated to be $430, based on the charges applied to non-residents of Canada.(Reference: Personal Communication) For HAART medications, records describing medication name and number of days dispensed were available. Actual doses were not recorded, so days dispensed were combined with unit costs(17) and standard dosing guidelines.(18) All costs were converted to 2005 Canadian dollars.(19)

3.2.3. *Demographic and clinical variables*

Treatment status was included as a binary variable that did not change once an individual initiated treatment with HAART. For individuals with low adherence or treatment interruptions this may include extended periods during which treatment was not accessed. In the observational setting considered here, the treatment status indicator is comparable to the treatment status indicator in an intent-to-treat analysis of a randomized controlled trial, as it reflects the actual experiences of individuals who have initiated HAART.

CD4 cell count was included as a time-dependent categorical variable, and, for each monthly interval, the CD4 cell count measured closest to the interval start date was attributed to the interval. Categories of CD4 cell count (cells/mm^3) were: 0-49, 50-99, 100-199, 200-349, 350-499, and $\geq$500. A separate level was included to indicate a time interval during which there was no CD4 cell count available within one year of the interval start date. This was done because failure to present for routine laboratory visits may be informative with respect to other categories of health services utilization and may affect other categories of direct medical costs as a result. For the 64 individuals not included in the HOMER cohort, CD4 counts were not available. Similarly, for individuals who contributed person-time through the CHASE cohort prior to enrolling in HOMER, CD4 cell counts were not available prior to enrollment in HOMER (Figure 2). However, based on study inclusion criteria, these individuals were known to be antiretroviral-naïve, and the date of first positive HIV test was used in conjunction with the average untreated history of HIV(20) to extrapolate expected CD4 cell counts for each time period.

Other variables considered in all models were sex, self-reported history of injection drug use, and adherence to antiretroviral therapy as measured by number of months during the first year of HAART in which HAART medications were dispensed. In the model describing HAART costs, other variables considered were: year of HAART initiation, inclusion of a protease inhibitor (PI) in current HAART regimen, and number of distinct medications included in current HAART regimen (excluding Ritonavir for boosted-PI regimens).

3.2.4. Statistical methods

Each individual's follow-up time was divided into monthly intervals, and the costs associated with each category of expenditure were recorded for each interval.

Statistical modelling of health services cost data is often complicated by a high proportion of person-time during which no costs are incurred, and a small number of person-time units associated with very high costs.(21) The former results in "zero-inflation" of data, while the latter can result in positively skewed data; both phenomena violate the assumptions of standard linear regression models. The issue of zero-inflation can be addressed by modelling costs in two stages: in the first stage, a logistic regression model is used to estimate the probability of non-zero use, while in the second stage, only non-zero costs are modelled.(21) The issue of positive skew can often be addressed through a log transformation.

For all categories of costs besides HAART medications, a two-stage modelling approach was used.(21) For the outpatient, hospitalization, laboratory, and emergency categories, costs were log-transformed in the second-stage model due to right-skewed data. For HAART, costs were

left untransformed, as the data were consistent with a normal distribution. In each stage, random intercepts were included to account for the correlation between multiple observations from a single individual.

Results were expressed as monthly costs. For categories in which the two-stage approach was used, expected monthly costs were derived using the formula:

$$E[C^{(j)}|X,U,V]=P(C^{(j)}>0|X,U)E[C^{(j)}|C^{(j)}>0;X,V] \quad \text{(Eq. 3.1)}$$

where $C^{(j)}$ refers to monthly costs associated with cost category j, X is a vector of covariates, U is a random intercept in the utilization model, and V is a random intercept in the conditional cost model.

An interaction between CD4 cell count category and treatment status was included to differentiate between individuals not receiving HAART because their CD4 cell count was sufficiently high that HAART initiation was not clinically indicated(22) and individuals not receiving HAART despite having CD4 cell counts low enough that HAART was likely to be clinically indicated.

For log-transformed data, $E[C^{(j)}]$ was estimated from $E[\log(C^{(j)})]$ using the Duan smearing factor.(23) The Duan smearing factor is the average of the exponentiated residual error terms, and is used to correct for bias induced by retransforming results from the log scale back to the original scale. Predictive ability of the combined two-stage models was assessed using the root

mean-square error and mean absolute error, which were estimated by fitting models to 50% of the data, selected at random, and comparing model-predicted values to actual observations for the 50% of observations not used in fitting the models.(21) These diagnostic statistics confirmed predictive advantages to the above described modelling technique, relative to alternative methods such as single-stage approximation or use of a generalized linear model with a log-link rather than log-transforming the data.

For HAART medications, a two-stage approach was not used, because treatment with HAART continues for life, suggesting that HAART costs should be incurred during all months following treatment initiation. Because, by definition, HAART costs are only incurred by people receiving treatment, they were only modelled following initiation of treatment with HAART, and the treatment status variable was not included in the model. CD4 cell count was not included as a covariate in this model. Although the initial decision to begin HAART treatment is guided by CD4 cell count, it is not recommended that subsequent treatment decisions be based on CD4 cell count,(24, 25) so it was not felt that CD4 cell count would be a meaningful predictor in this context.

Within each model, the random intercept can be interpreted as an individual's tendency to utilize a particular category of health services, after adjusting for their CD4 cell count, sex, and injection drug use status. We used pairwise correlations between the random intercepts from the eleven utilization and cost models to assess relationships between categories of health services utilization and costs, beyond what would be expected based on covariate values. The relationships between random effects associated with HAART cost and those associated with

inpatient use and cost, respectively were explored in greater detail using formulae given in Appendix A.

3.3. Results

Descriptive characteristics of the study population are given in Tables 3.1(a) and 3.1(b). The quantities reported in Table 3.1(a) are stratified by HAART treatment status, and refer to the inpatient, outpatient, laboratory, and emergency room models, which all used a common study sample. Sex and injection drug use status are expressed as percent person-time because individuals were eligible to contribute person-time to both the treatment-experienced and treatment-naïve arms at different points in their disease and treatment history. The majority of person-time was contributed at higher CD4 cell count strata, although the distribution differed by treatment status, with the treatment-experienced group contributing a greater proportion of person-time to the highest strata than the treatment-naive group. Women and injection drug users contributed relatively more time to the treatment-naïve group.

Quantities reported in Table 3.1(b) refer to the study sample utilized in the HAART cost model. There was a small decreasing trend in number of individuals initiating HAART over time. The majority of individuals initiated HAART with a non-nucleoside reverse transcriptase inhibitor (NNRTI)-based regimen and maintained a three-drug regimen. The number of drugs included in a HAART regimen was estimated using dispensing records, so regimens including less than three drugs likely reflect refill patterns resulting from sub-optimal adherence, rather than prescription of mono- or dual-therapy.

Coefficients and associated standard errors for all models are reported in Tables 3.2-3.4. All first-stage model coefficients are reported on the logit scale. For second-stage models, costs associated with HAART were left untransformed, so coefficients are reported on the original cost scale. Costs associated with other categories of health services (i.e. inpatient, outpatient, laboratory, and emergency room visits) were log-transformed in the second stage of analysis, so second-stage coefficients associated with these categories are reported on the log scale. Several model diagnostic statistics are reported for each model in Table 3.4. Root mean-square error and mean absolute error were calculated on the scale of predicted costs, and can be interpreted on this scale for all categories of health services. For all use models, the variance of random intercepts is reported on the logit scale. For cost models, the scale of the residual standard error and the variance of random intercepts, respectively, are dependent on the nature of the dependent variable. For HAART, costs were left untransformed, and these diagnostics are reported on the original cost scale, while for all other categories they are reported on the log cost scale.

The model results were converted to the more interpretable quantity of predicted monthly costs via equation (3.1). Predicted costs for HAART, stratified by regimen type and demographic group, are given in Table 3.5. Estimated monthly HAART costs ranged from $922-$1,130. The majority of variability was observed between HAART regimen types (PI- vs. NNRTI-based); for each regimen type, costs were similar between demographic groups. Predicted costs associated with all cost categories besides HAART, stratified by demographic group, treatment status, and CD4 cell count, are displayed graphically in Figure 3.3. The costs displayed in Figure 3.3 are based on a medication adherence rate of 79%, which was the mean level of adherence observed in HOMER, estimated using refill compliance.(26) In all demographic groups, a dramatic

increase in monthly costs—comprised largely of hospitalization costs—was observed in individuals having a CD4 cell count below 50 cells/mm^3. Within this CD4 strata, estimated monthly costs ranged from $945 in HAART-experienced male non-users of injection drugs to $3,659 in HAART-naïve women who used injection drugs.

In Figure 3.4, the inverse relationship between direct medical costs (excluding HAART) and adherence is shown for a male non-injection drug user, stratified by levels of CD4 cell count. Each line covers the interquartile range of adherence observed within the corresponding CD4 cell count stratum. Similar inverse relationships between adherence and costs were observed for all other demographic groups (i.e. women, injection drug users).

The majority of predicted monthly costs were lower in individuals who had received treatment with HAART, and this difference was most pronounced in the lower CD4 cell count stratum. However, overall predicted monthly costs were still higher in the treatment-experienced group, after including the costs associated with HAART. Predicted costs were consistently higher for women and injection drug users.

Table 3.6 reports pairwise correlation coefficients between the random effects associated with the eleven utilization and cost models considered here. Correlations greater than 0.20 in absolute value are bolded. There was a consistent pattern of positive correlations for all categories of utilization and cost with outpatient utilization. The majority of correlations—including all correlations that were relatively large in absolute value—were positive, suggesting that individuals who tended to utilize one type of service were also likely to utilize other services.

The majority of negative correlations were associated with laboratory utilization, laboratory cost, and HAART cost, respectively. This provides evidence that regular laboratory monitoring and treatment with HAART were associated with reduced need for other health care services.

Figure 3.5 displays the relationship between HAART random effects and expected inpatient cost for hypothetical individuals (male, non-user of injection drugs, non-aboriginal, with CD4 cell counts of 49, 199, and 499 cells/mm^3, respectively). Both inpatient use and inpatient cost were inversely related to HAART cost, as characterized by the correlation between random effects given in Table 3.6. As expected, the scale of expected inpatient costs differed dramatically across CD4 cell counts, but within each CD4 stratum, there was an inverse relationship with HAART random effects, indicating that individuals who incurred greater HAART costs tended to have lower inpatient costs. Because HAART costs were left untransformed, a random effect of -200 indicates that an individual tended to incur $200 less in monthly HAART costs relative to the population average, while a random effect of 200 indicates that an individual tended to incur $200 more than average in monthly HAART costs. Within each of the three CD4 strata shown in Figure 5, expected inpatient costs were more than two-fold greater for an individual with a HAART random effect of -200 relative to a HAART random effect of 200.

3.4. Discussion

In this study we used random-effects models to estimate the direct medical costs associated with various categories of health services utilization in individuals infected with HIV, accounting for HAART treatment status and CD4 cell count. We observed that regular laboratory monitoring and treatment with HAART were associated with reduced need for other health care services.

For all cost categories other than HAART, we observed an inverse relationship between costs and CD4 cell count. We also observed a general trend of higher costs in untreated individuals, although overall costs were estimated to be higher in the treated group after if aggregated to include costs associated with HAART. Compared to studies based on historical comparisons in the United States, the cost reductions associated with HAART estimated here were less than those reported by Bozzette et al.(12), and greater than those reported by Crane et al.(7).

Common practice in estimating current costs associated with HIV has been to define a study population based on access to specialized HIV treatment clinics, e.g.(7, 12, 27, 28) which *a priori* excludes individuals who have not accessed treatment. Through linkages of two cohorts, we were able to access health services records for individuals who were infected with HIV but had never accessed treatment. The wider population included here allowed for further examination of the economic impact of accessing HAART. The results are consistent with the hypothesis that increased use of HAART is associated with a decrease in other direct medical costs: after adjusting for CD4 cell count, individuals who had accessed HAART incurred lower overall medical costs across all expenditure categories and demographic goups (Figure 3.3); increased adherence to HAART—as defined by first-year refill percentage—was inversely related to overall direct medical costs (Figure 3.4); and increased monthly expenditure on HAART—as defined by individual random intercepts in the HAART cost model—was inversely related to both hospitalization frequency and hospitalization cost (Figure 3.5).

The predicted costs shown in Figures 3.3(a) – 3.3(d) suggest that the costs of all non-HAART health services are higher in women and injection drug users, after adjusting for treatment status

and CD4 cell count. This may be connected to the observation that both women and injection drug users are relatively over-represented in the group not treated with HAART (Table 3.1). The descriptive results reported in Table 3.1 do not account for treatment eligibility, so it is possible that in this population, women and injection drug users were less likely to be eligible for treatment with HAART based on current clinical status. However, as can be seen in Figure 3.3, cost differences between treated and untreated individuals were more pronounced in women and injection drug users, and particularly the intersection of women who were injection drug users (Figure 3.3(d)). This may be due to non-HIV-related care, and is also consistent with the hypothesis that, for these demographic groups, there is a greater proportion of individuals who remain untreated despite being clinically eligible for treatment. These results are consistent with evidence that women and injection drug users may be less likely to receive optimal treatment for HIV.(29) This may lead to to greater use of other health services, in order to deal with acute care needs that could have potentially been prevented through optimal prescription of HAART.

All models adjusted for CD4 cell count, suggesting that the higher costs incurred by women and injection drug users were not explained entirely by disease status. For injection drug users, this is plausible due to increased risk of comorbidities and health complications directly related to use of injection drugs, as well as a decreased use of preventive services and early interventions, leading to more intensive use of hospitalizations and emergency room services.(16) Women have also been consistently observed to utilize health care services at a higher rate than men, independent of HIV status.(30, 31, 32, 33) It is further possible that the factors that place women at increased risk of HIV infection, such as poverty, marginalization, gender power

inequities, sex work, and vulnerability to violence,(34, 35) may also result in a greater need for acute health care services following infection.

The majority of correlation coefficients between random effects estimated within the utilization and cost models were positive, suggesting that individuals who tended to utilize a particular category of health services were likely to utilize other categories as well. This is not inconsistent with the hypothesis that use of regular outpatient care may reduce the need for other acute care health services such as inpatient stays or emergency room visits.(10) An individual's random effect associated with a particular category of health services utilization or cost is not just reflective of their health status. In addition to health status, increased use of health care services may reflect an individual's underlying level of health-seeking behavior which defines the threshold of ill health at which they present for that particular category of medical care. Thus, the effects associated with different categories would be expected to exhibit positive relationships with one another. However, while, for example, there was a positive correlation observed between inpatient utilization and outpatient utilization, it is plausible that for a given individual, their inpatient utilization would have been even higher had they not received routine outpatient care.

A limitation to this study was that not all data sources were available for all study participants. The emergency room linkage was only available through the CHASE cohort, so no emergency room information was available for subjects who were included in the HOMER cohort but not the CHASE cohort. The impact of this was likely small and would not be likely to change the interpretations. In addition, the costs of non-HAART medications were not included because

utilization data were only available for individuals eligible for public assistance and this status was not known. We performed a sensitivity analysis in which these costs were included and the effect on overall costs was negligible. The costs of hospitalizations were based on a common per-diem cost. This cost was based on IDUs infected with HIV in Vancouver, and likely reflects an overall average relevant to the population of interest. Ideally, however, more detailed costing information could be used to stratify hospitalization costs by demographic characteristics and length of hospital stay.

The categories of health services utilization reported in the BCLHD were available for all study participants, but the available follow-up period varied. For individuals in the HOMER cohort, the linkage to the BCLHD was only available to March 2001, meaning that monthly cost estimates were based on utilization patterns observed prior to this time. All unit costs were updated to 2005 Canadian dollars, however, so that final estimates reflect an updated cost of these utilization patterns.

An additional limitation was the fact that CD4 cell counts were not available for some individuals prior to initiation of HAART. Based on study inclusion criteria, all individuals who were missing CD4 counts were participants of the CHASE cohort, meaning that date of first positive HIV test was known (Figure 3.2). This date was used as a proxy for seroconversion date, and CD4 cell counts over time were extrapolated based on a typical natural history.(20) Due to individual-level variability in CD4 trajectories, some person-time intervals may have been associated with misclassified CD4 strata. Despite this limitation, we felt that any potential misclassification introduced by including these person-time intervals was outweighed by the

information that was added by including follow-up across the entire spectrum of actual monitoring and treatment practices occurring in the community. The CHASE cohort utilized a community-based sampling scheme, allowing for the inclusion of some individuals infected with HIV who had never accessed treatment. This is in contrast to the practice of selecting study participants from the population accessing an HIV clinic, which by definition excludes all individuals who are infected with HIV but not seeking treatment. The CHASE cohort therefore provides a unique opportunity to observe health services utilization amongst individuals infected with HIV and not receiving optimal care.

With respect to individuals included in the HOMER cohort, the data were population-based, so that all individuals in BC who were being treated for HIV infection and who met the inclusion criteria for HOMER were considered for inclusion here. Thus, the results for treated individuals are generalizable to individuals initiating treatment with HAART in BC, and potentially to other settings with publically-funded health care in which similar HIV treatment guidelines are followed.

3.5. Conclusion

In this study, the direct medical costs associated with treating individuals infected with HIV were quantified in a Canadian setting. The inclusion of individuals who had never accessed HAART allowed for a comprehensive analysis of health services utilization across the spectrum of community-dwelling individuals infected with HIV. As such, the economic impact of HAART was assessed without the need for historical controls, and was not conflated with other temporal changes in HIV treatment patterns. The results of this study are consistent with

previous studies,(7, 12) suggesting that use of HAART is associated with decreases in other categories of medical costs, including hospitalizations. The implication of this finding is that increases in pharmaceutical budgets associated with expanded use of HAART may be offset by decreased need for hospitalizations amongst individuals who are adherent to HAART.

Figure 3.1: Number of subjects included from the HOMER and CHASE cohorts, Vancouver BC. Used in all analyses of health services categories other than antiretroviral therapy.

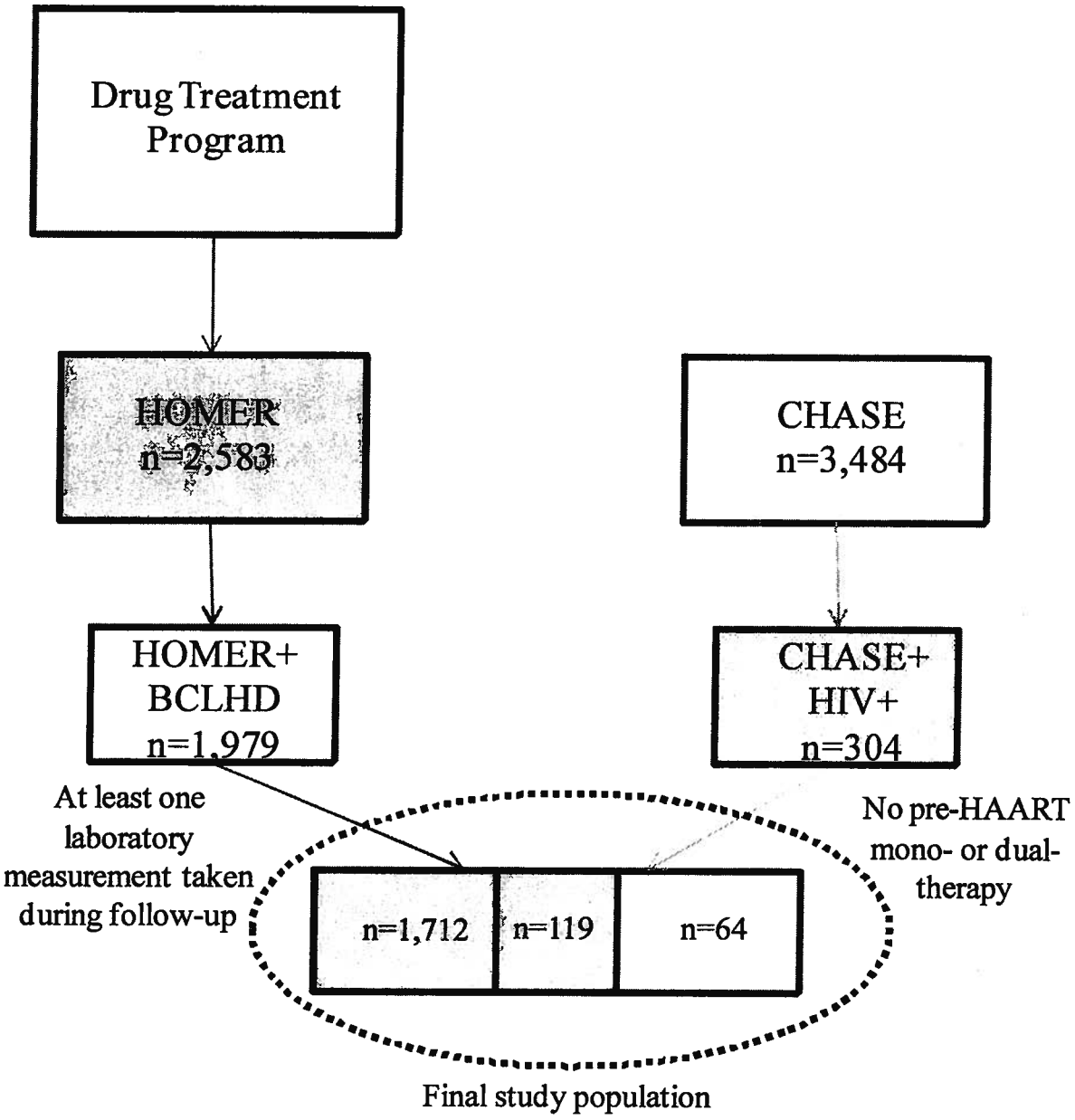

Figure 3.2: Available data and follow-up information.
Individuals included in the HOMER and/or CHASE cohorts.
∩ refers to intersection with period of linkage to the British Columbia Linked Health Database.

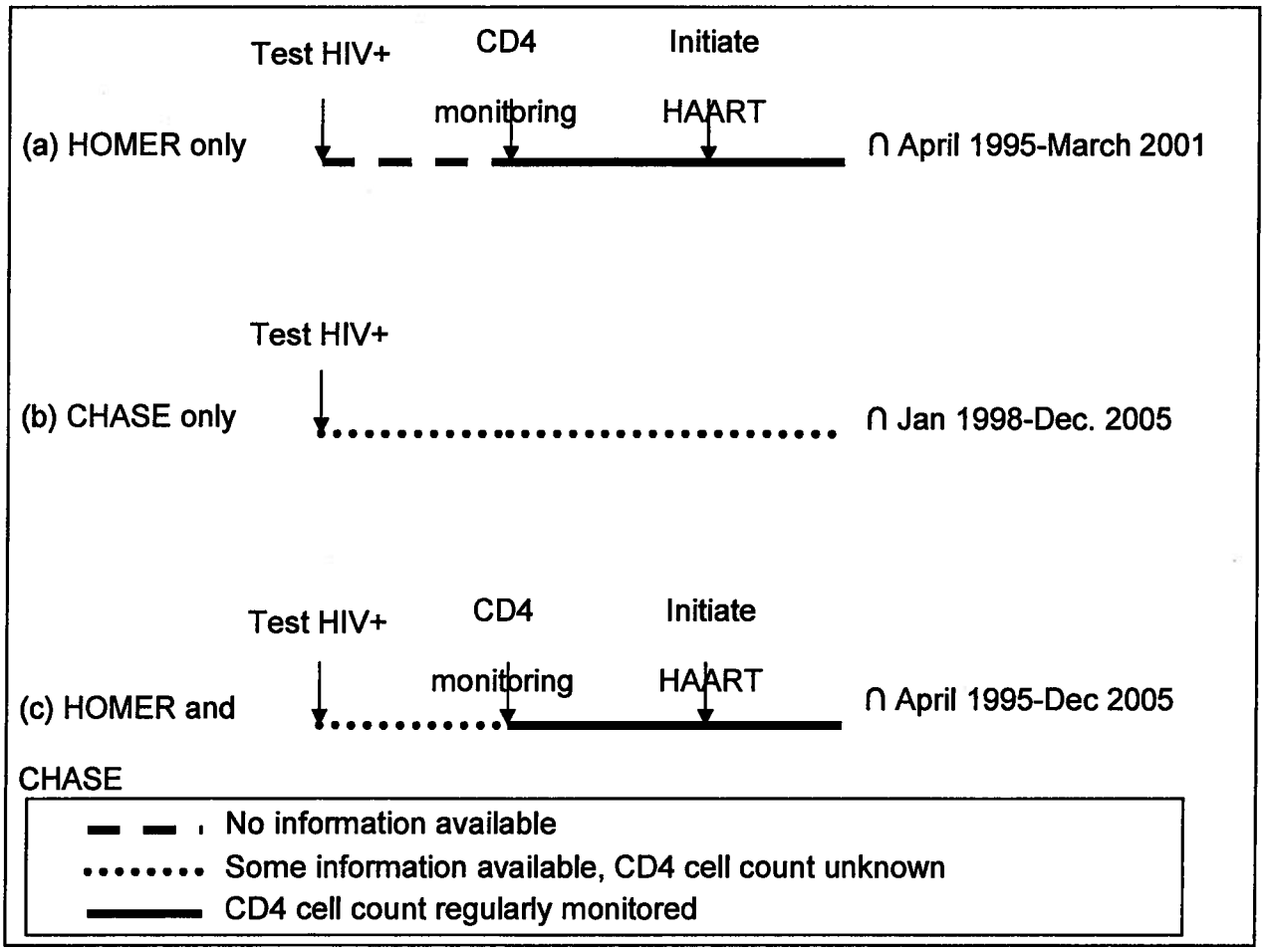

Figure 3.3: Estimated mean monthly costs.
Associated with non- highly active antiretroviral therapy (HAART) health services for individuals infected with HIV, stratified by treatment status and CD4 cell count, for (a) male non-users of injection drugs (n=1,103), (b) female non-users of injection drugs (n=180), (c) male users of injection drugs (n=441), and (d) female users of injection drugs (n=170). Treated individuals assumed to have average medication adherence observed in HOMER.

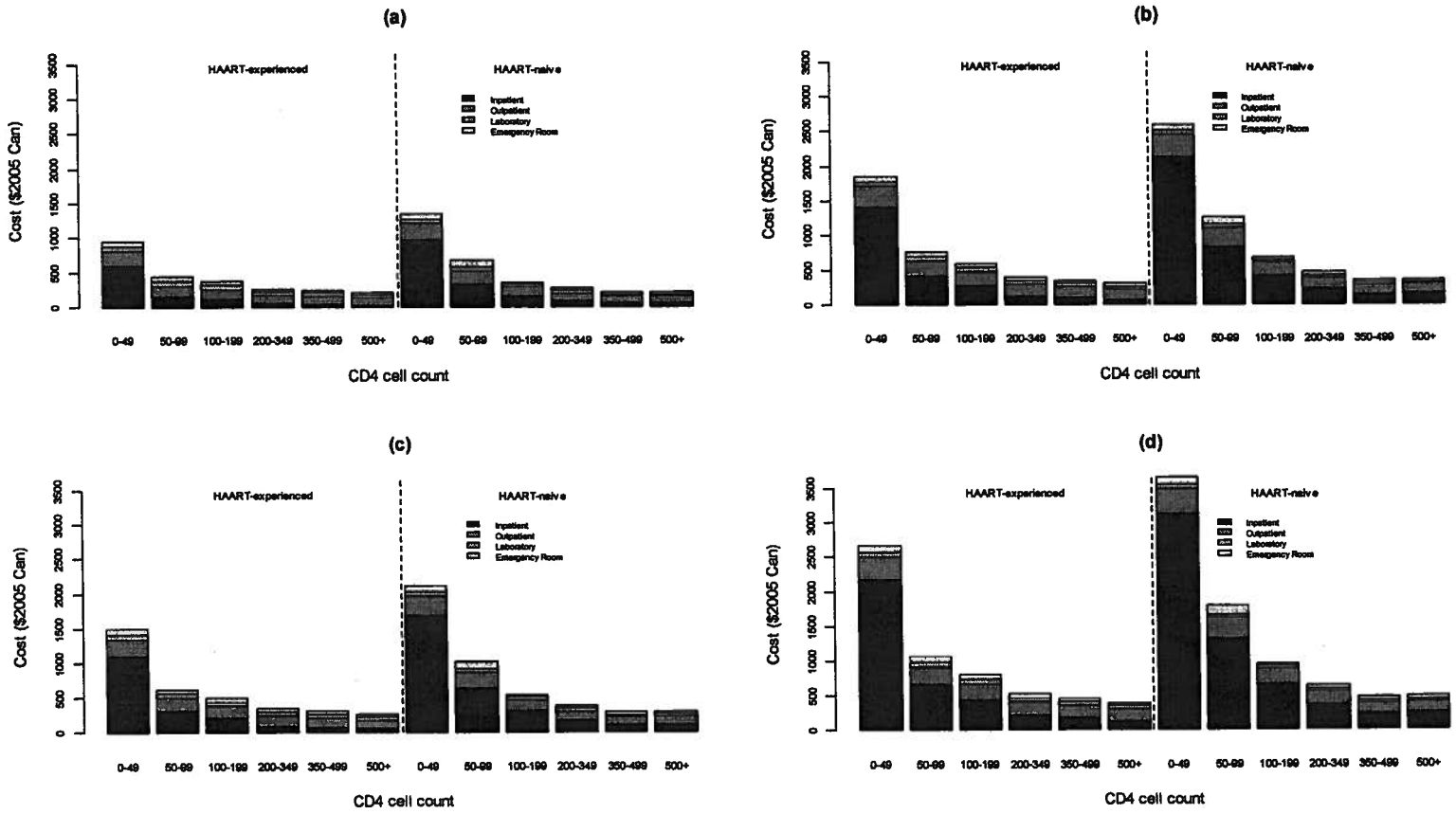

Figure 3.4: Estimated total direct medical costs.
Costs are excluding highly active antiretroviral therapy (HAART) by % medication adherence, stratified by CD4 cell count.

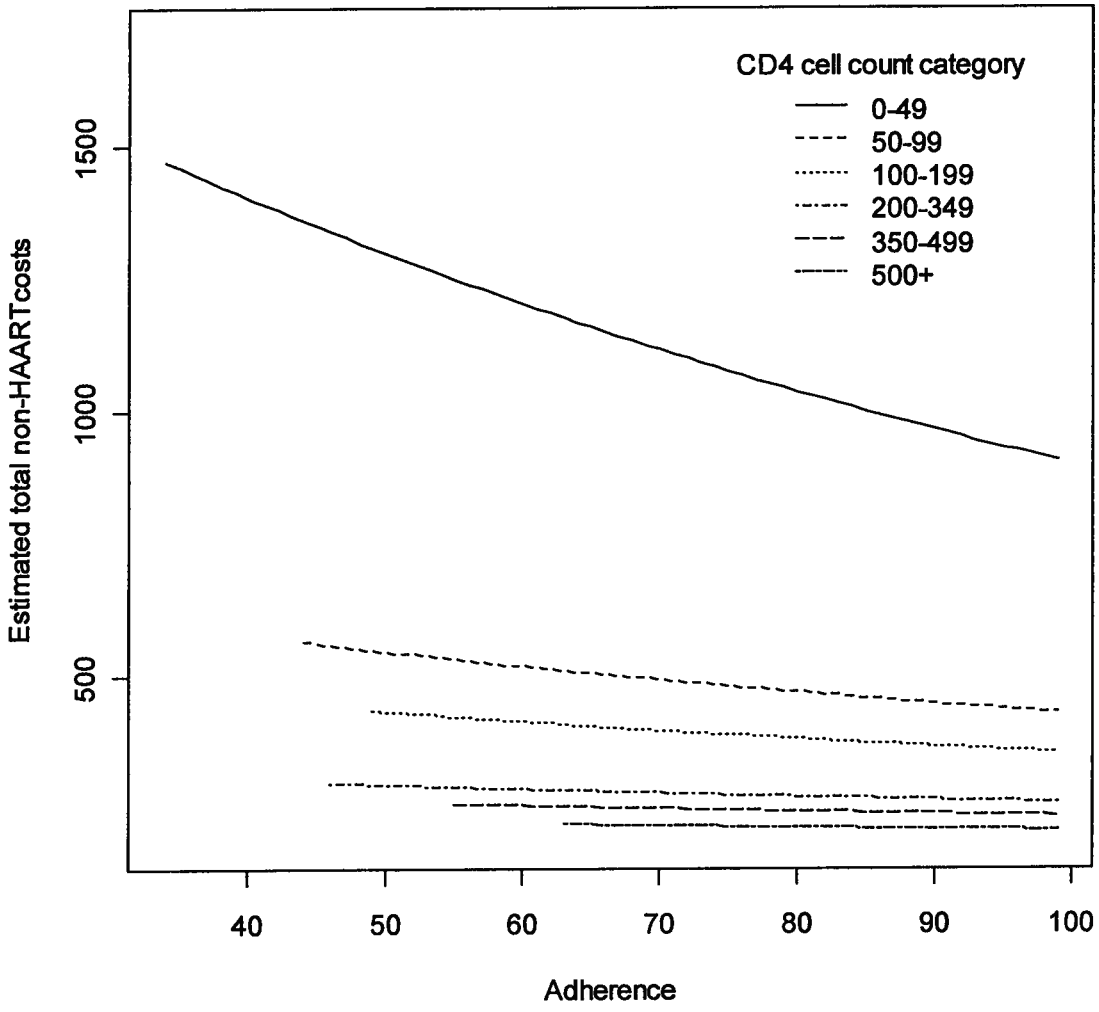

Figure 3.5: Estimated association between highly active antiretroviral therapy (HAART) cost random effects and expected hospitalization costs.
Based on a male, non-user of injection drugs, with a CD4 cell count of (a) 49, (b) 199, and (c) 499 cells/mm^3.

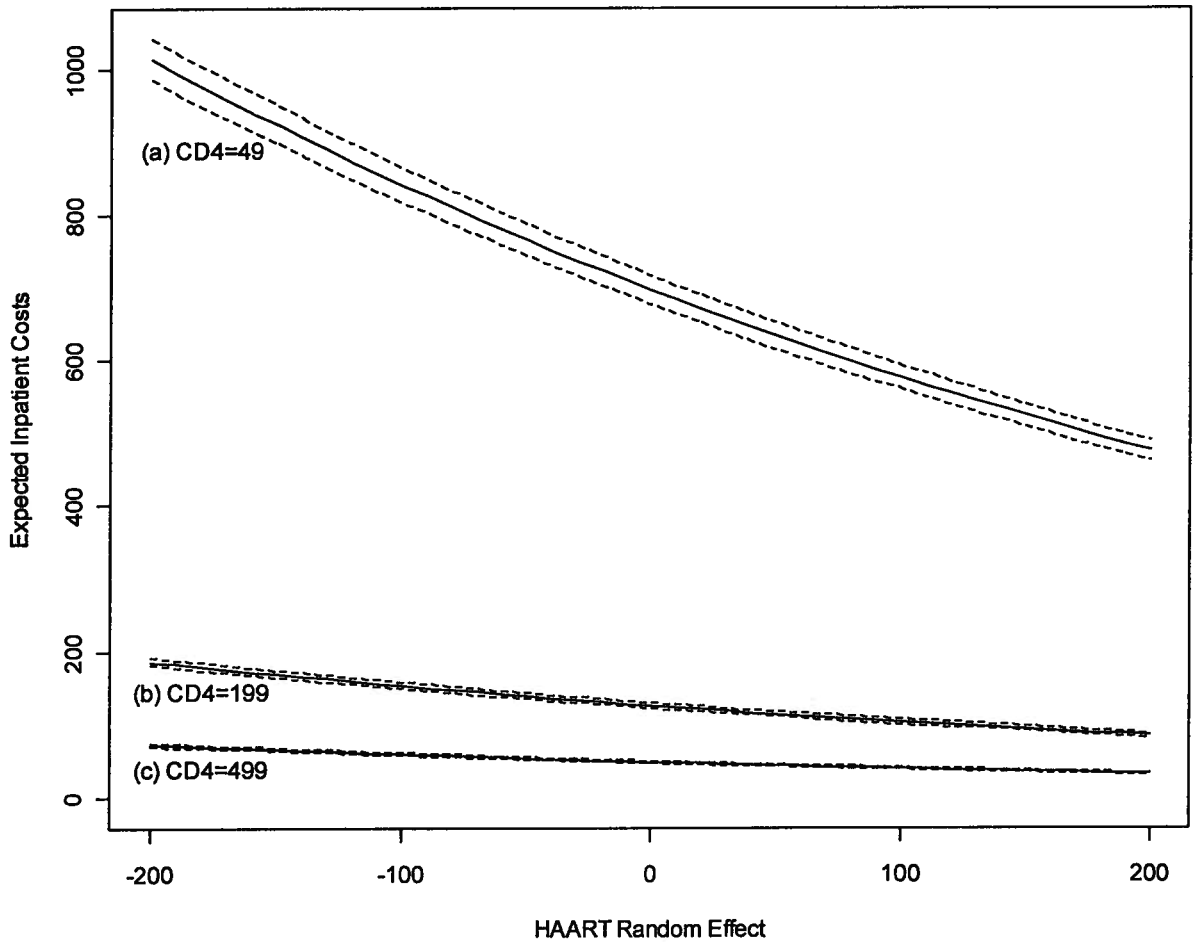

Table 3.1: Descriptive characteristics of study sample.
For (a) Inpatient, outpatient, laboratory, emergency room, and non-highly active antiretroviral therapy (HAART) medication models, and (b) HAART cost model.

(a)	HAART-experienced	HAART-naïve
Person-years contributed	4048.9	2520.1
CD4 cell count (cells/mm^3) (% person-time)		
0-49	5.5	5.5
50-99	3.7	3.7
100-199	10.5	10.5
200-349	23.6	23.6
350-500	20.0	21.6
≥ 500	31.6	25.7
No recent measurements	4.6	9.4
Sex (% person-time)		
Male	83.5	79.8
Female	16.5	20.2
Injection drug user (% person-time)		
Yes	63.8	60.8
No	36.2	39.2
Adherence (% person-time)		
0 -59%	36.1	NA
60-79%	9.0	
80-94%	13.2	
95-100	41.7	

(b)

Person-years contributed	12391
Sex (n(%))	
Male	2107 (81.6)
Female	476 (18.4)
Injection drug user (n(%))	
Yes	715 (27.7)
No	1868 (72.3)
Year of first HAART* use (n(%))	
1996-1998	927 (35.9)
1999-2001	884 (34.2)
2002-2004	772 (29.9)
Type of HAART regimen (% person-time)	
PI*	36.2
NNRTI*	63.8
Number of drugs in current HAART regimen (% person-time)	
≤ 2	34.8
3	59.9
4	3.8
≥ 5	1.5
Mean percent adherence to HAART* (standard deviation)	79.1 (29.5)

* PI = Protease inhibitor

** NNRTI=Non-nucleoside reverse transcriptase inhibitor

Table 3.2: Estimated regression coefficients and standard errors.
Two-stage random effects models describing monthly (log) costs associated with inpatient, outpatient, laboratory, emergency room, and non- highly active antiretroviral therapy (HAART) medication use, respectively. First-stage model estimates probability of non-zero cost, while second-stage model estimates (log) cost, conditional on it being non-zero.

	Estimated regression coefficient (standard error)							
	Inpatient		Outpatient		Laboratory		Emergency Room	
Variable	First-stage	Second-stage*	First-stage	Second-stage*	First-stage	Second-stage	First-stage	Second-stage*
Intercept	-1.97(0.10)	8.71(0.06)	1.03(0.08)	5.32(0.03)	-0.42(0.06)	4.82(0.03)	-2.10(0.26)	6.34(0.09)
CD4 count (cells/mm^3) (HAART*-experienced)								
0-49	0.00 (---)	0.00 (---)	0.00 (---)	0.00 (---)	0.00 (---)	0.00 (---)	0.00 (---)	0.00 (---)
50-99	1.00(0.14)	-0.10(0.10)	-0.11(0.1)	-0.16(0.05)	-0.16(0.09)	-0.03(0.04)	-0.05(0.29)	0.19(0.12)
100-199	-1.62(0.12)	-0.23(0.08)	-0.32(0.08)	-0.45(0.04)	-0.45(0.07)	-0.05(0.03)	-1.35(0.25)	-0.06(0.11)
200-349	-2.03(0.11)	-0.44(0.08)	-0.46(0.08)	-0.57(0.03)	-0.58(0.07)	-0.1(0.03)	-0.87(0.22)	0.11(0.09)
350-500	-2.49(0.12)	-0.49(0.09)	-0.65(0.08)	-0.66(0.04)	-0.68(0.07)	-0.11(0.03)	-1.05(0.23)	0.1(0.09)
≥ 500	-2.44(0.12)	-0.44(0.08)	-0.6(0.08)	-0.69(0.04)	-0.82(0.07)	-0.24(0.03)	-0.99(0.23)	0.05(0.09)
No recent measurements	-3.56(0.27)	-0.62(0.22)	-1.53(0.09)	-0.84(0.05)	-3.5(0.17)	-0.51(0.1)	-0.64(0.42)	-0.04(0.19)
(HAART*-naïve)								
0-49	0.16(0.13)	0.24 (0.09)	-0.34 (0.1)	-0.1(0.04)	-0.76(0.08)	0.65(0.04)	0.01(0.29)	0.23(0.11)
50-99	-0.84(0.27)	0.14 (0.19)	-0.45 (0.2)	-0.26(0.09)	-0.92(0.17)	0.62(0.08)	-0.04(0.58)	0.42(0.23)
100-199	-1.46(0.25)	0.01 (0.17)	-0.66 0.18)	-0.55(0.08)	-1.21(0.15)	0.6(0.07)	-1.34(0.54)	0.17(0.22)
200-349	-1.87(0.24)	-0.20 (0.17)	-0.8 (0.18)	-0.67(0.07)	-1.34(0.15)	0.55(0.07)	-0.86(0.51)	0.34(0.2)
350-500	-2.33(0.25)	-0.25 (0.18)	-0.99 (0.18)	-0.76(0.08)	-1.44(0.15)	0.54(0.07)	-1.04(0.52)	0.33(0.2)
≥ 500	-2.28(0.25)	-0.20 (0.17)	-0.94(0.18)	-0.79(0.08)	-1.58(0.15)	0.41(0.07)	-0.98(0.52)	0.28(0.2)
No recent measurements	-3.40(0.40)	-0.38 (0.31)	-1.87 (0.19)	-0.94(0.09)	-4.26(0.25)	0.14(0.14)	-0.63(0.71)	0.19(0.3)
Sex								
Male	0.00 (---)	0.00 (---)	0.00 (---)	0.00 (---)	0.00 (---)	0.00 (---)	0.00 (---)	0.00 (---)
Female	0.76(0.11)	0.04(0.05)	0.26(0.09)	0.25(0.03)	0.00(0.04)	-0.04 (0.02)	0.30(0.20)	0.02(0.04)
Injection drug user								
No	0.00 (---)	0.00 (---)	0.00 (---)	0.00 (---)	0.00 (---)	0.00 (---)	0.00 (---)	0.00 (---)
Yes	0.42(0.09)	0.07(0.05)	0.19(0.07)	0.09(0.02)	-0.02(0.03)	0.03 (0.01)	0.29(0.21)	-0.07(0.05)
Adherence (per 10% increase in refill rate)	-0.08(0.001)	-0.03 0.001)	0.09(0.001)	-0.01(0.00)	0.08(0.00)	-0.05(0.00)	-0.03(0.003)	-0.02(0.001)

* Costs log-transformed in second stage model

Table 3.3: Estimated regression coefficients and standard errors.
Random effects model describing monthly costs associated with highly active antiretroviral therapy (HAART) use.

Variable	Estimated regression coefficient (standard error)
Intercept	-23.46 (9.37)
Sex	
Male	0.00 (---)
Female	-6.28 (4.65)
Injection drug user	
No	0.00 (---)
Yes	-10.13 (3.71)
HAART Regimen	
PI*	191.59 (1.73)
NNRTI**	0.00 (---)
Number of medications in current HAART* Regimen***	290.74 (0.53)
Percent adherence to HAART	0.67 (0.06)
Year of first HAART use	
1996	0.00 (---)
1997	-22.66 (8.12)
1998	-26.38 (8.33)
1999	-24.03 (8.32)
2000	-29.34 (8.76)
2001	-29.59 (8.90)
2002	-19.40 (9.02)
2003	-43.47 (8.78)
2004	13.60 (9.09)

* PI = Protease inhibitor
** NNRTI=Non-nucleoside reverse transcriptase inhibitor
*** Does not include Ritonavir in boosted-PI regimens as an additional medication

Table 3.4: Measures of variability for statistical models.
Inpatient, outpatient, laboratory, emergency room, non- highly active antiretroviral therapy (HAART) medication and HAART models.

Cost category	Root mean square error	Mean absolute error	Residual standard deviation (cost model)	Standard deviation of random effects (use model)	Standard deviation of random effects (cost model)
Inpatient*	3321.02	3000.31	0.82	2.03	0.49
Outpatient*	281.03	136.28	0.88	1.95	0.42
Laboratory*	121.04	103.39	0.54	0.18	0.19
Emergency room*	423.31	376.56	0.49	1.48	0.17
HAART*	194.68	105.75	183.44	NA	76.63

* Costs log-transformed in second-stage model

Table 3.5: Predicted monthly HAART costs.
For various combinations of regimen type, sex, and injection drug user status. Predictions are further based on average observed adherence to HAART (79.1%), a 3-medication HAART regimen, and HAART initiation in 2004.

Demographic Category	PI*-based regimen	NNRTI**-based regimen
Male, non-user of injection drugs	1130	938
Female, non-user of injection drugs	1124	932
Male, user of injection drugs	1121	923
Female, user of injection drugs	1114	922

* PI = Protease inhibitor
** NNRTI=Non-nucleoside reverse transcriptase inhibitor

Table 3.6: Correlations between random intercepts.
For 11 statistical models describing probability of non-zero utilization and monthly cost conditional on non-zero utilization, respectively.

	Inpatient use	Outpatient use	Laboratory use	Emergency room use	Non-HAART* medication use	Inpatient cost	Outpatient cost	Laboratory cost	Emergency room cost	HAART* cost
Inpatient use	1.00	**0.23**	0.19	**0.63**	**0.29**	**0.34**	**0.53**	0.01	0.19	-0.08
Outpatient use		1.00	**0.30**	0.19	**0.42**	0.02	**0.38**	0.00	0.12	0.05
Laboratory use			1.00	**0.27**	0.11	-0.03	**0.30**	0.05	0.10	**0.21**
Emergency room use				1.00	**0.24**	0.18	**0.43**	-0.02	**0.39**	-0.03
Non-HAART* medication use					1.00	0.06	**0.31**	0.01	0.16	0.00
Inpatient cost						1.00	**0.25**	-0.04	0.10	-0.11
Outpatient cost							1.00	0.04	0.15	0.02
Laboratory cost								1.00	0.02	0.00
Emergency room cost									1.00	0.04
HAART* cost										1.00

*HAART = Highly active antiretroviral therapy

3.6. References

1. Hogg RS, Heath KV, Yip B, et al. Improved survival among HIV-infected individuals following initiation of antiretroviral therapy. Journal of the American Medical Association 1998;279:450-4.
2. Palella FJ, Delaney KM, Moorman AC, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. New England Journal of Medicine 1998;338:853-60.
3. Public Health Agency of Canada. HIV/AIDS Epi Updates 2007. Public Health Agency of Canada, 2007.
4. Stone PW, Chapman RH, Sandberg EA, et al. Measuring costs in cost-utility analyses - Variations in the literature. International Journal of Technology Assessment in Health Care 2000;16:111-24.
5. Scitovsky AA. Studying the cost of HIV-related illnesses - Reflections on the moving target. Milbank Quarterly 1989;67:318-44.
6. Levy AR, James D, Johnston KM, et al. The direct costs of HIV/AIDS care. Lancet Infectious Diseases 2006;6:171-7.
7. Crane L, Crowe R, Fine S, et al. Hospital and outpatient health services utilization among HIV-infected patients in care in 1999. Journal of Acquired Immune Deficiency Syndromes 2002;30:21-6.
8. Torti C, Casari S, Palvarini L, et al. Modifications of health resource-use in Italy after the introduction of highly active antiretroviral therapy (HAART) for human immunodeficiency virus (HIV) infection. Pharmacoeconomic implications in a population-based setting. Health Policy 2003;65:261-7.

9. Mocroft A, Monforte AD, Kirk O, et al. Changes in hospital admissions across Europe: 1995-2003. Results from the EuroSIDA study. HIV Medicine 2004;5:437-47.

10. Fleishman JA, Moore RD, Conviser R, et al. Associations between outpatient and inpatient service use among persons with HIV infection: A positive or negative relationship? Health Services Research 2008;43:76-95.

11. Beck EJ, Miners AH, Tolley K. The cost of HIV treatment and care - A global review. Pharmacoeconomics 2001;19:13-39.

12. Bozzette SA, Joyce G, McCaffrey DF, et al. Expenditures for the care of HIV-infected patients in the era of highly active antiretroviral therapy. New England Journal of Medicine 2001;344:817-23.

13. Joy R, Druyts EF, Brandson EK, et al. Impact of neighborhood-level socioeconomic status on HIV disease progression in a universal health care setting. Journal of Acquired Immune Deficiency Syndromes 2008;47:500-5.

14. Chamberlayne R, Green B, Barer ML, et al. Creating a population-based linked health database: A new resource for health services research. Canadian Journal of Public Health-Revue Canadienne De Sante Publique 1998;89:270-3.

15. Kerr T, Wood E, Grafstein E, et al. High rates of primary care and emergency department use among injection drug users in Vancouver. Journal of Public Health 2005;27:62-6.

16. Palepu A, Tyndall MW, Leon H, et al. Hospital utilization and costs in a cohort of injection drug users. Canadian Medical Association Journal 2001;165:415-20.

17. British Columbia Centre for Excellence in HIV/AIDS Drug Treatment Program. Unpublished Data 2007.

18. AIDSInfo. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. 2007.

19. Organization for Economic Development and Co-operation (OECD). Health Data 2007., 2007.

20. Richman D. Human Immunodeficiency Virus. International Medical Press, 2003.

21. Diehr P, Yanez D, Ash A, et al. Methods for analyzing health care utilization and costs. Annual Review of Public Health 1999;20:125-44.

22. British Columbia Centre for Excellence in HIV/AIDS. Therapeutic guidelines: Antiretroviral therapy for HIV-1 infected adults. 2006.

23. Duan N. Smearing estimate - A nonparametric retransformation method. Journal of the American Statistical Association 1983;78:605-10.

24. Hammer SM, Eron JJ, Reiss P, et al. Antiretroviral treatment of adult HIV infection - 2008 recommendations of the International AIDS Society USA panel. Journal of the American Medical Association 2008;300:555-70.

25. El-Sadr WM, Lundgren JD, Neaton JD, et al. CD4+count-guided interruption of antiretroviral treatment. New England Journal of Medicine 2006;355:2283-96.

26. Wood E, Hogg RS, Yip R, et al. Impact of baseline viral load and adherence on survival of HIV-infected adults with baseline CC4 cell counts >= 200 cells/mu l. AIDS 2006;20:1117-23.

27. Garattini L, Tediosi F, Di Cintio E, et al. Resource utilization and hospital cost of HIV/AIDS care in Italy in the era of highly active antiretroviral therapy. AIDS Care-Psychological and Socio-Medical Aspects of AIDS/HIV 2001;13:733-41.

28. Roberts RR, Kampe LM, Hammerman M, et al. The cost of care for patients with HIV from the provider economic perspective. AIDS Patient Care and STDs 2006;20:876-86.

29. Wood E, Montaner JSG, Bangsberg DR, et al. Expanding access to HIV antiretroviral therapy among marginalized populations in the developed world. AIDS 2003;17:2419-27.

30. Iron K, Goel V. Sex differences in the factors related to hospital utilization: Results from the 1990 Ontario Health Survey. Journal of Womens Health 1998;7:359-69.

31. Owens GM. Gender differences in health care expenditures, resource utilization, and quality of care. Journal of Managed Care Pharmacy 2008;14:S2-S6.

32. Bertakis KD, Azari R, Helms LJ, et al. Gender differences in the utilization of health care services. Journal of Family Practice 2000;49:147-52.

33. Muller C. Health Care and Gender. New York: Russell Sage Foundation, 1990.

34. Spittal PM, Craib KJP, Wood E, et al. Risk factors for elevated HIV incidence rates among female injection drug users in Vancouver. Canadian Medical Association Journal 2002;166:894-9.

35. Zierler S, Krieger N. Reframing women's risk: Social inequalities and HIV infection. Annual Review of Public Health 1997;18:401-36.

36. Bain L, Engelhardt M. Introduction to Probability and Mathematical Statistics. Duxbury, 1992.

4. ACCOUNTING FOR TIME-DEPENDENT CD4-CELL COUNT TRAJECTORIES IN ESTIMATING CUMULATIVE UTILIZATION AND COST OF HOSPITALIZATIONS IN INDIVIDUALS INFECTED WITH HIV*

4.1. Introduction

With the introduction of highly active antiretroviral therapy (HAART) for treating individuals infected with HIV, morbidity and mortality have dramatically decreased.(1, 2) A number of studies have reported an association between the introduction of HAART and a decrease in HIV-related hospitalizations.(3, 4, 5, 6) However, hospitalizations continue to comprise a substantial proportion of HIV-related direct medical costs, particularly for individuals with a CD4 cell count below 200 cells/mm^3.(7) Costs associated with hospitalizations are relatively high compared to those associated with other types of health services, meaning that even if hospitalizations occur less frequently for individuals receiving treatment with HAART, the overall costs may still present an important economic burden. In addition, following the introduction of HAART, individuals infected with HIV now experience greater life expectancy, and the distribution of long-term cumulative hospitalization costs for these individuals has not been quantified.

In Canada, an estimated 58,000 individuals were living with HIV at the end of 2005.(8) Given this level of prevalence, an estimate of the resources required for current and future HIV-related hospitalizations is potentially useful to health planners and policy makers. These resources can be expressed in terms of expected incidence of hospitalizations over time or the anticipated costs

---

* A version of this chapter will be submitted for publication. Johnston KM, Levy AR, Gustafson P, Briggs AH, Lima VD, Hogg RS, Tyndall M, Montaner JSM. Accounting for time-dependent CD4-cell count trajectories in estimating cumulative utilization and cost of hospitalizations in individuals infected with HIV

associated with these hospitalizations. The former can be used to estimate needs for personnel, space, beds, and other hospital infrastructure, while the latter can be used to project hospital budgetary requirements.

CD4 cell count is a prognostic clinical marker of HIV severity, with lower CD4 cell counts indicative of more severe disease. This measure can also be used to estimate short-term risk of developing an AIDS-defining condition.(9, 10, 11) CD4 cell count has consistently been found to display an inverse relationship with hospitalization costs.(7) Because CD4 cell count displays a dynamic trajectory over the course of an HIV infection(12) and its prognostic value is most relevant in the short term, it is preferable to account for this variable in a time-dependent fashion when modelling long-term hospitalization incidence and costs.

One method of analyzing hospitalization data is to treat hospitalizations within an individual as correlated recurrent events. Extensions of standard time-to-event methods have been developed for this class of data,(13) and these methods can easily be adapted to describe incidence of hospitalizations. Further, the costs associated with each hospitalization can be incorporated through a "marked point process,"(13) in which the cost of a hospitalization is modelled in addition to the time at which the hospitalization occurred. There is flexibility regarding the degree of dependence between the cost process and the recurrent events process. At one end of the spectrum, it can be assumed that all costs are independently and identically distributed, while at the other end of the spectrum, costs can be modelled conditional on the time of the hospitalization and other variables describing hospitalization history, e.g. number of previous hospitalizations or cumulative costs associated with previous hospitalizations.

The objective in this study was to develop a method for incorporating a time-dependent covariate into a recurrent events model in order to make inferences regarding incidence and costs of hospitalizations. We illustrate the method using data describing an observational cohort of individuals infected with HIV and receiving treatment with HAART, with CD4 cell count included as a time-dependent covariate.

4.2. Methods

4.2.1. Data

The study population was comprised of a population-based observational cohort from British Columbia (BC), Canada. The HAART Observational Medical Evaluation and Research (HOMER) cohort includes all antiretroviral-naive individuals in BC who initiated treatment with HAART after 1996. Data are available describing baseline sociodemographic characteristics, antiretroviral treatment history, and longitudinal CD4 cell count for all members of the HOMER cohort. CD4 cell count was measured on a schedule based on clinical guidelines, approximately quarterly for most individuals.(14)

A subset of individuals who enrolled in HOMER prior to 2003 was linked to the BC Linked Health Database (BCLHD), which contains hospitalization abstracts describing date and length of stay for hospitalization separations between April 1995 and March 2002. The cohort and linkage has been described in detail in Chapter 3. In this analysis, we further limited inclusion to individuals who provided at least six months follow-up on HAART (n=1,589) in order to eliminate individuals who were only ever dispensed HAART medications once and were otherwise untreated. For the subset of 84 individuals who were also included in CHASE, the more recent linkage provided hospitalization separations until December 31, 2005, and the

extended follow-up was included for these individuals. Individuals who died during the course of follow-up were considered to be right-censored at date of death, including deaths which occurred in hospital. Data describing longitudinal CD4 cell counts were available for all individuals until November 30, 2005 or date of death or last contact.

The cost of a hospitalization was estimated using an estimated cost of $610.33 per hospital day. This quantity was based on a fully-allocated costing model created for British Columbia,(15) and was updated to 2005 Canadian dollars using health-specific inflation factors.(16) This daily rate was multiplied by length of hospital stay, so that the model describing hospital costs was based on a constant multiple of length of hospital stay.

All analyses were performed using `R version 2.8.1`,(17) including the `survival`(18) and `nlme`(19) libraries.

4.2.2. Statistical models

The method described herein was based on three statistical models. The first was a model describing the incidence of hospitalizations, using a recurrent events framework, with CD4 cell count included as a time-dependent covariate. The second was a linear model with random intercepts describing the log-transformed cost of each hospitalization. The third was a quadratic random intercept model describing CD4 cell count trajectories over time. These three models were then combined to generate long-term cumulative incidence and costs of hospitalizations. For clarity and illustrative purposes, the models fit here did not include any covariates besides CD4 cell count. However, it is straightforward to extend the models to also incorporate time-constant covariates.

4.2.2.1. Incidence model

In our notation, *X(t)* describes the entire CD4 cell count history up to time *t*, *Z(t)* is a CD4 cell count at time *t*, and *N(t)* is the cumulative number of hospitalizations experienced by time *t*. In practice, CD4 cell counts were measured on an approximately quarterly basis (Chapter 2), and we assumed that CD4 count remained constant between measurements. The incidence of hospitalizations was modelled using a semi-parametric model with shared frailties:

$$\rho(t \mid X(t),U) = \int_0 U \exp(Z(s)\gamma) d\rho_0(s) \qquad \text{(Eq.4.1)}$$

for conditional mean function $\rho(t \mid X(t),U) = \mathrm{E}[N(t) \mid X(t)]$, baseline non-parametric mean function $\rho_0(t)$, and frailty term *U*. Frailty terms varied across individuals and were assumed to follow a gamma distribution with mean 1.0. A frailty term of 1.0 is interpreted as having hospitalization incidence equivalent to the population average, while frailty terms less than 1.0 indicate lower than average hospitalization incidence and frailty terms greater than 1.0 indicate greater than average hospitalization incidence. The model was fit as using the semiparametric counting process framework described by Andersen and Gill,(20) extended to include frailty terms.(13) This model can be fit in `R` using the `survival` function `coxph` by using the counting-process version of `Surv` and including the `frailty` option(21) in the survival formula (see Appendix C.1.4 of Cook et al. for model fitting details). This model choice allowed for the incorporation of multiple events per individual and the inclusion of CD4 cell count as a time-dependent covariate. Further, time spent in hospital was explicitly excluded from time considered to be at risk for a future hospitalization. Time was measured relative to first HAART prescription.

4.2.2.2. Cost model

The log-transformed cost of a hospitalization was modelled using a linear model with random intercepts of the form:

$$\log\left(C_k^{(i)} \mid Z_k^{(i)}\right) = \left(\alpha_0 + a_0^{(i)}\right) + \alpha_1 Z_k^{(i)} + \varepsilon \qquad \text{(Eq.4.2)}$$

with i indexing over individuals, $a_0^{(i)} \sim \mathrm{N}\left(0, \sigma_a^2\right)$, and $\varepsilon \sim \mathrm{N}(0, \sigma_C^2)$. Random intercepts were included to allow for individual-level variability in costs not explained by CD4 cell count. The log transformation was selected due to the tendency of cost data to display a positive skew.(22) Diagnostic plots (not shown) confirmed the appropriateness of the log-transformation for these particular cost data.

The random intercept $a_0^{(i)}$ is an adjustment to the intercept term made for each individual, accounting for the fact that, after adjusting for CD4 cell count, individuals can still have a systematic tendency to experience hospitalizations that are either more or less expensive than the population average. A positive random intercept indicates that an individual has a tendency to experience more expensive hospitalizations while a negative random intercept indicates a tendency to experience less expensive hospitalizations. Because costs were log-transformed, random intercepts are also on a log scale.

Both the incidence and cost models were also fit including covariates describing number of previous hospitalizations and cumulative number of previous hospital days, but neither of these covariates was found to be statistically significant in either model, and the associated coefficients

were all small in absolute value, suggesting that after adjusting for recent CD4 cell count, both incidence and cost of hospitalizations could be considered independently of previous hospitalization history.

4.2.2.3. CD4 cell count model

In the case of no covariates or time-constant covariates only, estimation of cumulative incidence and costs for a recurrent events process is relatively straightforward.(13) However, in the case of time-dependent covariate (e.g. CD4 cell count), estimation of these quantities for a "typical" individual requires an underlying model describing the average trajectory of the time-dependent covariate over time. Because CD4 trajectories are non-linear and display variability across individuals (Chapter 2), we used a random intercept model that was quadratic in time to describe CD4 cell count trajectories.(23) The model can be expressed as:

$$Z^{(i)}(t) = \left(\beta_0 + b_0^{(i)}\right) + \beta_1 t + \beta_2 t^2 + \varepsilon \qquad \text{(Eq.4.3)}$$

where $b_0^{(i)} \sim \mathrm{N}\left(0, \sigma_b^2\right)$ and $\varepsilon \sim \mathrm{N}(0, \sigma_z^2)$. At a given time *t*, the expected CD4 cell count for individual *i* can be expressed as $\mu_k^{(i)}(t) = \left(\beta_0 + b_0^{(i)}\right) + \beta_1 t + \beta_2 t^2$. The interpretation of random intercepts in the CD4 model is analogous to those in the cost model, except that CD4 counts were left untransformed, so random intercepts can be interpreted on the original CD4 scale. Positive random intercepts indicate an individual tends to have higher-than average CD4 cell counts over time while negative random intercepts indicate lower-than-average CD4 cell counts over time.

4.2.3. Cumulative incidence and costs

Based on the three models described above, and the assumption that CD4 cell count only changes at discrete time points $t_1, t_2, \ldots, t_m,$, the expected cumulative cost and incidence of hospitalizations for individual i at time t can be derived as:

$$\mathrm{E}\left[C^{(i)}(t) \mid U\right] = US_C \exp\left(\alpha_0 + a_0^{(i)} \mid U + \sigma_Z^2(\alpha_1 + \gamma)^2/2\right) \times \sum_{k=1}^{m^{(i)}} \exp\left(\mu_k^{(i)}(t \mid U)(\alpha_1 + \gamma)\right)\left(\rho_0(t_k) - \rho_0(t_{k-1})\right)$$

(Eq.4.4)

$$\mathrm{E}\left[N^{(i)}(t) \mid U\right] = U \exp\left(\sigma_Z^2 \gamma^2/2\right) \sum_{k=1}^{m^{(i)}} \exp\left(\mu_k^{(i)}(t \mid U)\gamma\right)\left(\rho_0(t_k) - \rho_0(t_{k-1})\right) \qquad \text{(Eq.4.5)}$$

The derivation of equations (4.4) and (4.5) is given in Appendix B. In equation (4.4), S_C refers to the smearing factor of the cost model, which is defined as the average of exponentiated residuals. The smearing factor is used to convert expected log-cost back to the untransformed expected cost scale. Under the assumption that U, a_0, and b_0 follow a multivariate normal distribution, the conditional expectations of $a_0^{(i)} \mid U$ and $\mu_k^{(i)}(t \mid U)$ are $\frac{(U-1)\sigma_{aU}}{\sigma_U^2}$ and $\beta_0 + \frac{(U-1)\sigma_{bU}}{\sigma_U^2} + \beta_1 t + \beta_2 t^2$, respectively. The covariance terms in these expressions can be approximated by the corresponding empirical quantities estimated via model fitting. In practice, the choice of discrete time points is arbitrary and may be chosen on a fine enough scale to approximate continuous change.

The variance of the estimators was estimated using 1000 non-parametric bootstrap samples. Sampling was taken at the level of the individual, and the observations associated with a sampled individual were then included in the fitting of the three statistical models. This sampling scheme was chosen due to the complex structure and observational nature of the data.(24)

4.3. Results

Descriptive characteristics of the study sample are given in Table 4.1. Individuals provided a median of 11 CD4 measurements (interquartile range [IQR] = 6-17) over a median 3.3 years follow-up (IQR=2.09-4.53). In total, 19,552 CD4 cell counts were contributed over 5,244 person-years of follow-up. The majority of follow-up time was contributed at higher CD4 cell count strata, reflective of successful treatment with HAART. The incidence of hospitalizations displayed a consistent inverse relationship with CD4 cell count strata, with an estimated incidence of 1.75 hospitalizations per person-year within the lowest CD4 cell count stratum (0-50 cells per mm^3) compared to just 0.25 hospitalizations per person-year at the highest CD4 cell count stratum ($\geq$500 cells per mm^3). The contrast between median cost per hospitalization ($3,561) and mean cost per hospitalization ($6,877) is due to the right-skewed nature of the data, with a small number of very lengthy hospitalizations. Over the course of follow-up, the data on 82 individuals were censored for analysis when they died.

Coefficients for all three statistical models are given in Table 4.2. In both the hospitalization incidence and hospitalizations cost models, the coefficient associated with CD4 cell count was negative, suggesting that individuals with lower CD4 cell count were more likely to experience hospitalizations, and, once hospitalized, they tended to require longer hospital stays.

The empirical variance and covariance terms observed between frailty terms, random intercepts from the CD4 model, and random intercepts from the cost model are given in table 4.3(a). The variance of the CD4 model random intercepts was several orders of magnitude larger than the other variables, due to the difference in scales—CD4 counts were left on the original scale while costs were log-transformed. In table 4.3(b) these covariances are translated into conditional expected values for CD4 and cost model random effects for selected values of frailties. These expectations are based on the assumption that the three variables follow a multivariate normal distribution (see Appendix B for further details). The motivation for deriving these conditional expectations was to incorporate the potential for correlation between the three processes being modelled. The outcome of primary interest in this study was expected incidence and cost of hospitalizations over time for individuals with different underlying tendencies to be hospitalized as measured by their frailty terms. By deriving conditional expectations for CD4 and cost random intercepts, we were able to incorporate the joint relationship between CD4 cell count, hospitalization cost, and frailty terms into our overall estimators for cumulative incidence and costs.

Figure 4.1 displays estimated cumulative incidence and cost of hospitalizations following initiation of HAART for an individual with a frailty term of 1.0, which corresponded to the population mean. Mean incidence and cost increased to approximately three hospitalizations and $15,000 in hospitalization costs incurred after eight years of treatment with HAART.

Figure 4.2 displays the estimated cumulative cost of hospitalizations for selected quantiles (5^{th}, 25^{th}, 50^{th}, 75^{th}, 95^{th}) of the empirical frailty distribution. Frailty terms affected the incidence and cost estimators in two ways: directly, through their multiplicative effect on hospitalization

incidence, and indirectly, through their correlation with random effects in the CD4 and cost models which are then included in the models describing hospitalization incidence and hospitalization cost, respectively. Wide variation was observed across the frailty distribution. After eight years of treatment with HAART, an individual with a frailty term corresponding to the 95th percentile of the empirical distribution was estimated to incur approximately nine hospitalizations, costing over $60,000. Over eight years, the expected cost per hospitalization, defined as cumulative cost divided by cumulative incidence, increased across frailty percentiles from approximately $5,300 in the 5th percentile to approximately $6,500 in the 95th percentile.

4.4. Discussion

In this study, we found an inverse relationship between incidence of hospitalizations and CD4 cell count. It was also observed that hospitalizations experienced by individuals with lower CD4 cell counts tended to be longer and therefore more costly. A semi-parametric frailty model was used to describe hospitalization incidence, accounting for the occurrence of multiple hospitalizations within individuals. Parametric frailty terms, assumed to follow a gamma distribution with mean 1.0, were used as a proxy for the combination of underlying health and health-seeking behaviour that leads to a particular individual's tendency to be hospitalized. We also fit a quadratic random intercept model for CD4 trajectories and a linear random intercept model for log-transformed hospitalization costs, which were used in conjunction with the hospitalization incidence model to estimate the cumulative incidence and cost of hospitalizations over time for various categories of individuals.

The results of this study could potentially be used to inform decision-making and health planning regarding the need for financial and infrastructural resources needed to accommodate HIV-

related hospitalizations over time. In addition, these results could be used as an input to an economic model, such as that described in Chapters 5 and 6 of this thesis. In those economic models, we instead used the two-stage random effects models described in Chapter 3 to incorporate hospitalization utilization and cost. The two-stage random effects model is relatively simple and averages the cost of hospitalizations over the entire follow-up period, without explicitly taking timing into account. For the purposes of the economic models described in Chapters 5 and 6 this simplification was thought to be acceptable. However, if the specific timing of individual hospitalizations was of interest, the method described in this chapter would be a more appropriate choice.

Given the correlation between the three processes modelled here (hospitalization incidence, hospitalization cost, CD4 trajectory), one option would have been to define a fully joint model, such that correlation between the processes could be explicitly considered in model fit. We instead fit the three models independently, and then accounted for correlation *post hoc* via the empirical correlation between frailties and random intercepts and conditional expectations that could be incorporated into the cumulative incidence and cost estimators. The advantage to using this method is that the estimators can be generated using widely-available statistical software, and do not require the creation of customized software programs to implement model fit, and can therefore be potentially applied in a wider variety of settings.

The expected trajectories of incidence and cost over time shown in Figure 4.2 suggest that accounting for the correlation between processes led to more informative results. Individuals with higher frailty terms were expected to experience more hospitalizations, which would be expected due to the definition of the frailty term. However, they also tended to incur higher

costs beyond what would be expected due simply to higher incidence of hospitalizations, resulting from more expensive individual hospitalizations. This was due to the fact that higher frailty terms were associated with lower CD4 cell counts (negative covariance of -6.561in Table 4.3(a)), which were in turn associated with more frequent and expensive hospitalizations Further, higher frailty terms were also associated with more expensive hospitalizations even after accounting for CD4 cell count (positive covariance of 0.054 in Table 4.3(a)). By accounting for these covariance terms in the cost and incidence estimators, we were able to fully account for the impact of a higher frailty term.

One limitation to the method suggested here for incorporating CD4 cell trajectories into the cumulative incidence and cost estimators is that the additive models described in Chapter 2 could not be used here because it was computationally infeasible to include random intercepts in an additive model with a dataset of this size. The inclusion of random intercepts was necessary to incorporate a correlated structure with the other two statistical models. In place of an additive model, we fit a quadratic random intercept model, which has been previously shown to have good predictive ability for modelling CD4 cell counts.(23) In a comparison not shown here, we found that an additive model with a term for most recent CD4 cell count had slightly better predictive ability than a quadratic random effects model, but results were comparable, and the latter model was able to adequately describe the CD4 data considered here.

A further limitation to this particular analysis was that, besides CD4 cell count, other potential sociodemographic and clinical covariates were not included in the models fit here. Results presented here can thus be interpreted as most relevant to the specific combination of sociodemographic and clinical variables observed in BC, and may not be widely generalizable.

The motivation for this exclusion was the aim to present an illustrative example, with clarity in deriving the estimators and presenting results. There is also evidence that CD4 cell count is a key predictor of hospitalization costs; several other studies have focused exclusively on CD4 cell count in analyses of health services costs.(7) A potential future application would be to incorporate additional covariates, (e.g. medication adherence and resistance, plasma viral load, age, sex, injection drug use) in order to elucidate more subtle relationships with hospitalization utilization and costs.

The method presented here is not limited to the context of using CD4 cell count trajectories to estimate incidence and cost of hospitalizations in individuals infected with HIV. In a recent study of diabetes outcomes, Clarke and co-authors predicted risk over time of seven common health outcomes in diabetes, with several time-dependent covariates.(25) However, in this study, the various outcomes were treated as independent processes and an analytical estimating function for predicted outcomes was not derived; models describing risk factor progression and outcome incidence were not integrated into a single model, and predictions were generated numerically through simulation. The estimating functions described here could potentially be applied in such a setting and can be generalized to describe cumulative incidence of a recurrent event or multi-type related events and/or the cumulative value of a marker process in the presence of a time-dependent covariate in other contexts.

Figure 4.1: Estimated (a) cost and (b) incidence functions.
With associated pointwise 95% confidence intervals for hospitalizations, accounting for dynamic CD4 cell count trajectories and assuming a frailty term of 1.0.

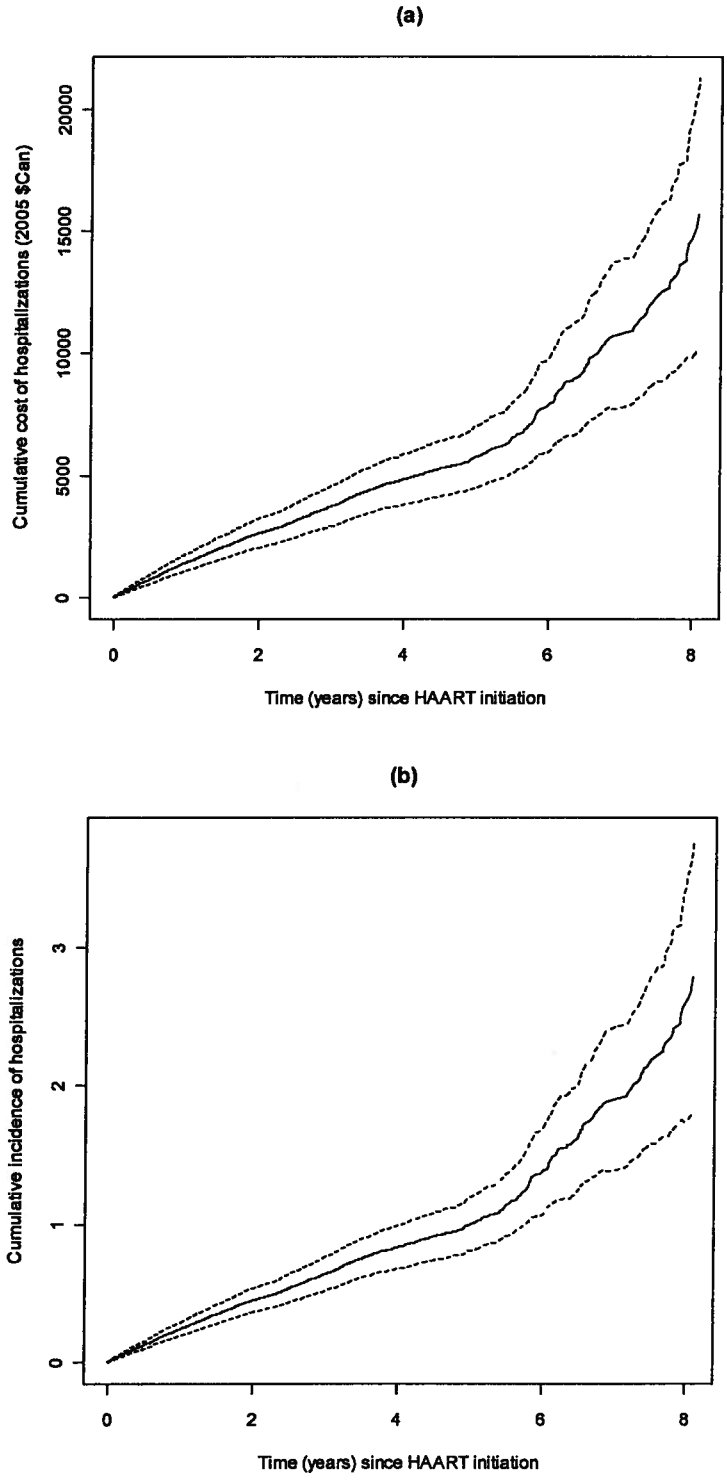

Figure 4.2: Cumulative (a) cost and (b) incidence functions for selected quantiles of the empirical frailty distribution.

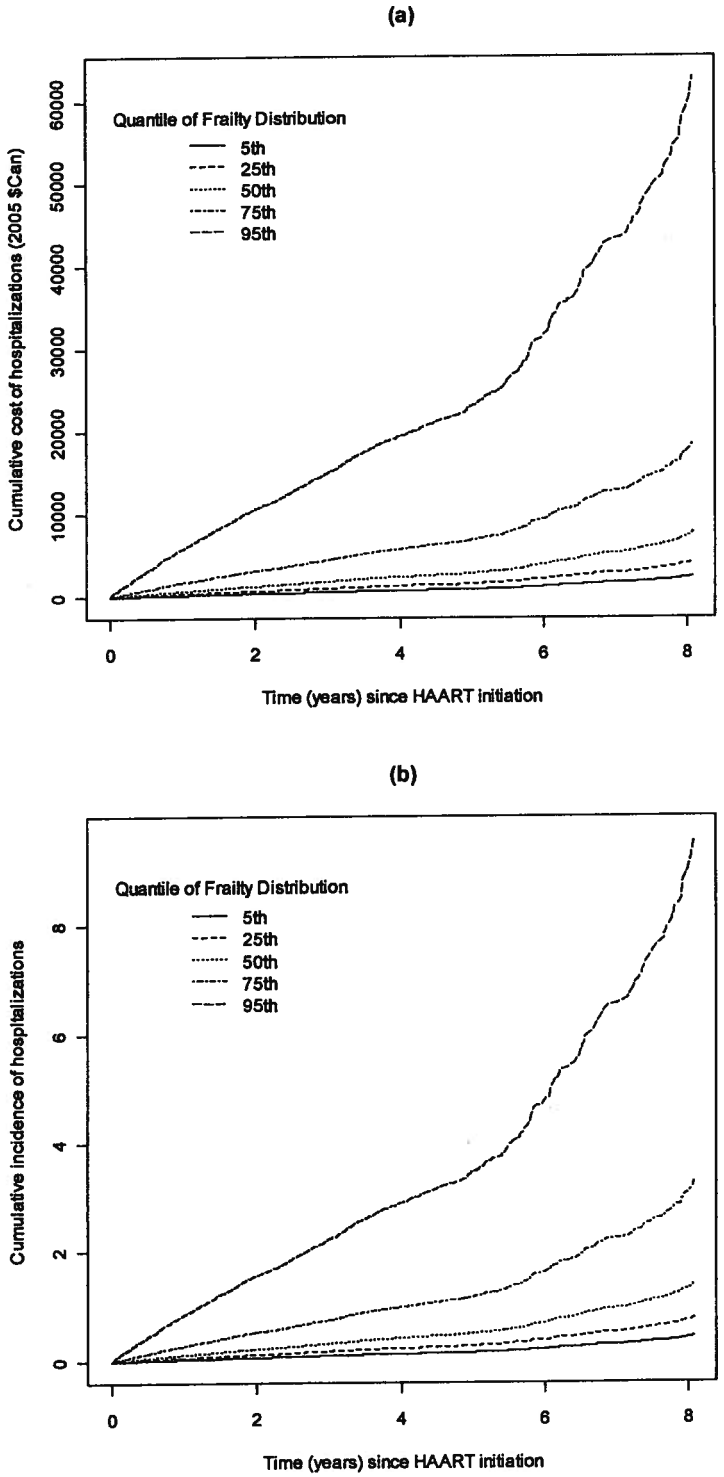

Table 4.1: Descriptive characteristics of study sample.
N=1589; IQR = Interquartile Range

Variable		
Cost per hospitalization (2005$ Canadian)	Median (IQR) Mean (95% Confidence Interval)	3561 (1424-7122) 6877 (6387-7366)
Follow-up time (years) per individual	Median (IQR)	3.32 (2.09-4.53)
Number of CD4 measurements per individual	Median (IQR)	11 (6-17)
Time between CD4 measurements (months)	Median (IQR)	3.02 (2.43-3.70)
Follow-up time contributed (years) by CD4 strata (cells/mm^3)	0-49 50-99 100-199 200-349 350-499 ≥500	229.7 244.2 647.2 1210.3 1104.2 1808.5
Hospitalization incidence per person-year by CD4 strata (cells/mm^3)	0-49 50-99 100-199 200-349 350-499 ≥500	1.75 0.72 0.60 0.43 0.32 0.25

Table 4.2: Regression coefficients and 95% confidence intervals
Fixed effects for hospitalization incidence model, hospitalization cost model, and CD4 trajectory model.

	Incidence Model	Cost Model	CD4 model
Intercept	———	8.35 (8.26 – 8.44)	352.05 (340.45 - 363.64)
CD4 coefficient (per 100 cells/mm^3)	-0.20 (-0.23 - -0.17)	-0.086 (-0.107 – 0.065)	———
Time (months)	———	———	1.96 (1.92 – 1.99)
$Time^2$ ($months^2$)	———	———	-0.00040 (-0.00039 - -0.00041)
Residual standard error	———	0.94	169.15

Table 4.3: Covariance, correlation and conditional expectation of random effects.
(a) Empirical covariance (correlation) matrix for random effects estimated within hospitalization cost model, CD4 trajectory model, and frailty terms from hospitalization incidence model; (b) Conditional expectations of random effects from CD4 and cost models for selected frailty terms based on multivariate normal distribution.

(a)

	Hospitalization cost model random effects	CD4 trajectory model random effects	Hospitalization incidence model frailties
Hospitalization cost model random effects	0.131 (1.000)	-6.561 (-0.088)	0.054 (0.210)
CD4 trajectory model random effects		44,841.30 (1.000)	-4.90 (-0.024)
Hospitalization incidence model frailties			0.91 (1.000)

(b)

Frailty term	Expected CD4 trajectory model random intercept	Expected cost model random intercept
0.16 (5^{th} percentile)	4.52	-0.05
0.28 (25^{th} percentile)	3.87	-0.04
0.51 (50^{th} percentile)	2.66	-0.03
1.17 (75^{th} percentile)	-0.91	0.01
3.36 (95^{th} percentile)	-12.71	0.14

4.5. References

1. Hogg RS, Yip B, Kully C, et al. Improved survival among HIV-infected patients after initiation of triple-drug antiretroviral regimens. Canadian Medical Association Journal 1999;160:659-65.
2. Palella FJ, Delaney KM, Moorman AC, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. New England Journal of Medicine 1998;338:853-60.
3. Fleishman JA, Moore RD, Conviser R, et al. Associations between outpatient and inpatient service use among persons with HIV infection: A positive or negative relationship? Health Services Research 2008;43:76-95.
4. Torres RA, Barr M. Impact of combination therapy for HIV infection on inpatient census. New England Journal of Medicine 1997;336:1531-2.
5. Gebo KA, Diener-West M, Moore RD. Hospitalization rates in an urban cohort after the introduction of highly active antiretroviral therapy. Journal of Acquired Immune Deficiency Syndromes 2001;27:143-52.
6. Paul S, Gilbert HM, Lande L, et al. Impact of antiretroviral therapy on decreasing hospitalization rates of HIV-infected patients in 2001. AIDS Research and Human Retroviruses 2002;18:501-6.
7. Levy AR, James D, Johnston KM, et al. The direct costs of HIV/AIDS care. Lancet Infectious Diseases 2006;6:171-7.
8. Public Health Agency of Canada. HIV/AIDS Epi Updates 2007. Public Health Agency of Canada, 2007.

9. Beral V, Coutinho R, Darbyshire J, et al. Short-term risk of AIDS according to current CD4 cell count and viral load in antiretroviral drug-naive individuals and those treated in the monotherapy era. AIDS 2004;18:51-8.

10. Mellors JW, Munoz A, Giorgi JV, et al. Plasma viral load and CD4(+) lymphocytes as prognostic markers of HIV-1 infection. Annals of Internal Medicine 1997;126:946-54.

11. Mellors JW, Margolick JB, Phair JP, et al. Prognostic value of HIV-1 RNA, CD4 cell count, and CD4 cell count slope for progression to AIDS and death in untreated HIV-1 infection. Jama-Journal of the American Medical Association 2007;297:2349-50.

12. Richman D. Human Immunodeficiency Virus. International Medical Press, 2003.

13. Cook R, Lawless J. The statistical analysis of recurrent events. Springer, 2007.

14. Hammer SM, Eron JJ, Reiss P, et al. Antiretroviral treatment of adult HIV infection - 2008 recommendations of the International AIDS Society USA panel. Journal of the American Medical Association 2008;300:555-70.

15. Palepu A, Tyndall MW, Leon H, et al. Hospital utilization and costs in a cohort of injection drug users. Canadian Medical Association Journal 2001;165:415-20.

16. Organization for Economic Development and Co-operation (OECD). Health Data 2007., 2007.

17. R Development Core Team. R: A Language and Environment for Statistical Computing. 2008.

18. Therneau T, Lumley T. survival: Survival analysis, including penalised likelihood., 2008.

19. Pinheiro J, Bates D, DebRoy S. Linear and Nonlinear Mixed Effects Models. 2008.

20. Andersen PK, Gill RD. Cox regression model for counting-processes - A large sample study. Annals of Statistics 1982;10:1100-20.

21. Therneau TM, Grambsch PM, Pankratz VS. Penalized survival models and frailty. Journal of Computational and Graphical Statistics 2003;12:156-75.

22. Diehr P, Yanez D, Ash A, et al. Methods for analyzing health care utilization and costs. Annual Review of Public Health 1999;20:125-44.

23. Taylor JMG, Law N. Does the covariance structure matter in longitudinal modelling for the prediction of future CD4 counts? Statistics in Medicine 1998;17:2381-94.

24. Wehrens R, Putter H, Buydens LMC. The bootstrap: a tutorial. Chemometrics and Intelligent Laboratory Systems 2000;54:35-52.

25. Clarke PM, Gray AM, Briggs A, et al. A model to estimate the lifetime health outcomes of patients with Type 2 diabetes: the United Kingdom Prospective Diabetes Study (UKPDS) Outcomes Model (UKPDS no. 68). Diabetologia 2004;47:1747-59.

26. Duan N. Smearing estimate - A nonparametric retransformation method. Journal of the American Statistical Association 1983;78:605-10.

27. Bain L, Engelhardt M. Introduction to Probability and Mathematical Statistics. Duxbury, 1992.

5. ESTIMATING THE LIFETIME DIRECT MEDICAL COSTS ASSOCIATED WITH HIV/AIDS IN BRITISH COLUMBIA, CANADA*

5.1. Introduction

Understanding the distribution of lifetime direct medical costs associated with treating an individual infected with HIV is important for health policy makers who allocate resources for HIV-related health services. The advent of highly active antiretroviral therapy (HAART) in the mid-1990s brought about dramatic increases in life expectancy(1) and shifted patterns of resource utilization(2) in individuals infected with HIV. Both of these changes have resulted in changes to expected costs. The survival benefits of HAART have yet to be fully realized, as life expectancy after initiating treatment with HAART exceeds its period of widespread availability.(3) Therefore, lifetime direct medical costs cannot be estimated directly, and must be extrapolated beyond the follow-up data that are currently available.(4, 5, 6)

Methods for modelling disease histories can be broadly characterized into approaches which describe a homogeneous group of individuals at the aggregate level, or approaches which incorporate heterogeneity, for example by including multiple distinct cohorts or unique disease histories at the individual level. Empirical results have shown that failure to incorporate heterogeneity into model design can lead to biased results.(7) We hypothesized *a priori* that the incorporation of individual-level heterogeneity would potentially be required to accurately model the costs associated with treating HIV. This was based on the non-linear inverse relationship

---

* A version of this chapter will be submitted for publication. Johnston KM, Levy AR, Briggs AH, Lima VD, Tyndall M, Hogg RS, Gustafson P, Montaner JSM. Estimating the lifetime direct medical costs associated with HIV/AIDS in British Columbia, Canada

that has been empirically observed between medical costs and CD4 cell count (a marker for HIV severity)(Chapter 3) combined with individual-level variability in CD4 cell count trajectories,(Chapter 2) as well as the potential for individual-level variability in direct medical costs that is not fully described by measured variables.(Chapter 3) Thus, by assuming homogeneous individuals, as would be necessary when using a standard cohort approach, it would not be possible to account for individuals who tend to incur greater healthcare costs than their health status would suggest or who spend longer periods of time with low CD4 cell counts and who have increased expenditure during those periods.

In British Columbia (BC), Canada, sociodemographic information, vital statistics, and longitudinal data describing clinical disease markers and laboratory tests are available for all individuals who are infected with HIV and accessing HAART. Further, health care is financed by through single payer (the BC Ministry of Health), and records for various health care services such as hospitalizations, physician visits, and pharmaceuticals can be linked at the individual level. All individuals have a unique health identifier, which allows for HIV-specific data to be linked to general health services data, creating a comprehensive data source for estimating the cost implications of HIV disease.

The objective in this study was to estimate the lifetime direct medical costs associated with treating an individual infected with HIV in BC. We estimated this quantity using a microsimulation approach in order to accommodate the inclusion of individual-level heterogeneity.

5.2. Methods

5.2.1. Data

The study sample was characterized by a linkage between the HAART Observational Medical Evaluation and Research (HOMER) cohort and the Community Health and Safety Evaluation (CHASE) cohort. The HOMER cohort is comprised of all antiretroviral-naive individuals in BC who initiated treatment with HAART between 1996 and 2004 (n=2,583). Individuals included in the cohort contributed longitudinal data describing HIV clinical course and resource utilization following initiation of treatment with HAART.

The CHASE cohort is a representative sample of low-income individuals living in Vancouver, BC's "Downtown East Side". The CHASE cohort was linked to the BC Centre for Disease Control HIV testing database, which provided information on HIV incident sero-status. Those individuals with confirmed HIV status were linked to centralized antiretroviral treatment records and categorized into three groups: (1) individuals who had received antiretroviral therapy and were included in both the HOMER and CHASE cohorts; (2) individuals who had received antiretroviral therapy but did not meet the inclusion criteria for HOMER; and (3) individuals who had never received antiretroviral therapy. Individuals who had received some antiretroviral therapy and who initiated antiretroviral therapy with mono- or dual-therapy prior to 1996 were excluded from all analyses. Of 3,484 individuals in the CHASE cohort, 304 had confirmed HIV-positive status, of which 183 contributed data describing HIV clinical course and resource utilization prior to the initiation of treatment and were included in this study.

5.2.2. *Costs*

The perspective taken in this analysis was that of the BC Ministry of Health, the third party payer responsible for a large majority of health care services and antiretroviral therapy in the province. Costs were estimated using a combination of charges paid and unit costs. Hospitalization costs were estimated by multiplying the length of each hospital stay in days by a daily unit cost of $610.33, which was estimated using a fully-allocated costing model developed for the province of BC.(8) For outpatient physician visits, charges paid by the Ministry of Health were available for each recorded visit, and these charges were applied directly. The unit costs associated with laboratory tests were obtained directly from a centralized testing facility, and were $60 for a plasma viral load assay, $65 for a CD4 cell count test, and $230 for a genotype based HIV drug resistance assay. The cost of an emergency room visit was estimated to be $430, based on the charges applied to non-residents of Canada.(Reference: Personal communication) For HAART medications, records describing medication name and duration dispensed were applied to unit costs(9) and standard dosing guidelines.(10) All costs were converted to 2005 Canadian dollars.(11) Costs and outcomes were discounted at rates of 0%, 3%, and 5%.

5.2.3. *Statistical models*

The majority of processes within the microsimulation were described by statistical models, described in further detail below.

Costs were estimated via two-stage random effects models, in which the first stage was a logistic regression model describing non-zero utilization in a particular time period and the second stage was a linear model of log-costs, conditional on non-zero utilization.(12) (Chapter 3) Separate models were fit for different categories of utilization, including hospitalizations, outpatient visits,

laboratory tests, emergency room visits and medications (HAART and non-HAART). All models adjusted for HAART treatment status, adherence to HAART, CD4 cell count, injection drug use status and an interaction between CD4 cell count and HAART treatment status. For individuals who died during follow-up, an indicator variable was included for the final month of life in order to account for the increased costs typically associated with late-stage palliative care. Because late-stage costs may be incurred prior to the final month of life, we also performed a sensitivity analysis in which we fit random effects models with an end-of-life indicator during the entire final year of life. Results were comparable to the model with a final month indicator only, suggesting that the final-month variable was sufficient to capture the majority of end-of-life costs.

Following treatment initiation, trajectories of CD4 and viral load were estimated using generalized additive models,(13) which were able to describe non-linear trajectories in time.(Chapter 2) For both CD4 and viral load, previous measurements were used to predict future measurements; in the case of CD4 cell count, previous viral load was also included as a predictive variable due to biological mechanisms by which viral load affects CD4 cell cout.(14) Other predictive variables included sex, injection drug use status, adherence to HAART, and the cumulative number of medication categories for which an individual had conferred resistance mutation(s).

Times until developing various resistance mutations were estimated using a Weibull survival model with gamma-distributed frailty terms.(15) A single individual could potentially develop resistance to up to four categories of antiretroviral medications, inducing an individual correlation structure in the time-to-resistance process. Within a survival model, frailty terms are

analogous to random effects, and were used here to account for that correlation. The resistance model is described in further detail in Appendix C.

Survival after treatment initiation was estimated via a Cox proportional hazards model,(15) which was chosen over a parametric form as it allows for straightforward inclusion of CD4 cell count as a time-dependent covariate via a counting process methodology.(15) The outcome in this model was all-cause mortality, including deaths due to HIV as well as those due to other causes.

The remaining processes were not described by statistical models, including: CD4 cell count and viral load trajectories during the time period between HIV infection and HAART initiation; background age-related mortality; and utility values associated with health states. Pre-treatment trajectories of CD4 and viral load could not be accurately assessed with the available data because a substantial proportion of individuals did not present for routine monitoring of these variables until they were nearing clinical indication for treatment with HAART. Further, for the majority of individuals, time of infection was not known, making treatment initiation date the most reliable choice for baseline time. Standard estimates for the untreated history of these variables were assumed to hold prior to this date.(16) Background mortality was assumed to follow the age-specific rates reported for BC for 2004.(17, 18) Although this assumption resulted in HIV-related deaths being double-counted, the proportion of population-wide deaths due to HIV were sufficiently small that removing them would not have altered rates at the number of significant figures considered. Health state utilities were stratified by CD4 cell count, based on Schackman et al.'s study of patient preferences using the SF-6D instrument.(19) This

empirical study found little variability between health states, with utility values ranging between 0.83 and 0.87, and, as such, further variability in utilities was not incorporated into the model.

5.2.4. Microsimulation structure

The integration of these statistical models and process assumptions into a comprehensive microsimulation is shown in Figure 5.1. The microsimulation began by generating baseline characteristics for all individuals, with distributions based on data observed in the HOMER and CHASE cohorts. These characteristics included demographics as well as random effects for each of the cost models and frailty terms for the resistance model. The random effects for the cost models represented an individual's tendency to utilize particular health services, above and beyond what would be described by the covariates included in the cost models (i.e. CD4 cell count, adherence, treatment status, demographics). Similarly, the frailty terms for the resistance model described an individual's tendency to develop medication resistance beyond what would be expected based on their adherence to HAART and other variables included in the resistance model. An additional variable that was generated at the outset of the microsimulation was the expected time following infection at which an individual was expected to present for CD4 monitoring. This was assumed to follow a uniform distribution over ten years, which was chosen based on the observed initial CD4 cell counts in the HOMER cohort. It was further assumed that after an individual presented for CD4 monitoring, they would initiate treatment with HAART once they were considered clinically eligible, with eligibility based on a CD4 cell count below 350 cells/mm^3.

The microsimulation then proceeded by moving forward in one-month time cycles. At each time step, CD4 cell count and viral load were updated. Treatment uptake in the upcoming month for

individuals not yet receiving HAART was determined based on their CD4 cell count and whether or not they had presented for monitoring.

Resistance mutation development was randomly determined assuming probabilities based on the Weibull frailty model. It was assumed that if an individual initiating an non-nucleoside reverse transcriptase inhibitor (NNRTI)-based regimen developed an NNRTI resistance mutation, they would switch to a protease inhibitor (PI)-based regimen, and remain on PI-based regimens for the remainder of their life, due to the increased likelihood of cross-class resistance to NNRTIs.(20) Similarly, individuals who started on a PI-based regimen were assumed to switch to an NNRTI-based regimen if they developed PI resistance, and to switch back to a PI-based regimen if they subsequently became resistant to NNRTIs. The distribution of times associated with these regimen switches were estimated based on empirical data and were accounted for in the frailty model.

Monthly costs associated with the various health services categories (hospitalizations, outpatient visits, laboratory tests, emergency room visits and medications) were then generated. The first stage random effects models were used to determine the probability of non-zero utilization for each of the categories, and random numbers were generated to determine which individuals would utilize each of the categories during the upcoming month. Conditional on non-zero utilization to a particular health services category, the second stage random effects model was used to estimate the expected log-cost associated with that category for the upcoming month.(12) The expected log-costs calculated using the second stage models were treated as point estimates associated with normally-distributed random variables. Actual log-costs were then randomly sampled from normal distributions with mean and standard deviation equal to the point estimate

and the residual standard error of the second stage model, respectively. Finally, these log-costs were transformed back to the original cost scale using the Duan smearing factor.(21) Survival throughout the upcoming interval was randomly determined based on the Cox proportional hazards model. For individuals who died during the upcoming one-month interval, the amount of time they contributed to the interval prior to death was sampled from a uniform distribution, and the final month-of-life indicator was included in the cost models to reflect the increase in expected costs associated with this interval.

5.2.5. Justification for microsimulation approach

The primary reason for using a microsimulation to estimate lifetime costs instead of a less computationally-intensive cohort-based method was the fact that with the former method it is straightforward to incorporate heterogeneity between individual-level CD4 trajectories and individual tendencies towards health-seeking behavior leading to variations in health services utilization and costs. It was hypothesized *a priori* that failure to incorporate these sources of heterogeneity could potentially lead to biased results. If the model had instead been structured as a homogeneous cohort, it would not have been possible to include this heterogeneity at the individual level. To empirically justify the use of the microsimulation approach, we removed key sources of heterogeneity from the microsimulation and re-estimated costs, in order to approximate results from a cohort-based model. In particular, we assumed a common CD4 trajectory for all individuals based on the population average and we removed the random effects from the cost models, i.e. by assuming that all individuals had a random effect of zero for all categories of health services utilization and cost. We still allowed the parameters of the cost models to vary across iterations of the probabilistic sensitivity analysis, as this source of uncertainty could also be captured within a cohort model.

5.2.6. *Assumptions*

Valid data describing CD4 cell count and survival were only available for individuals who accessed HAART, and baseline time in the Cox proportional hazards model was defined to be time of treatment initiation. Thus, the Cox proportional hazards model was only relevant for individuals following the initiation of HAART. We therefore made the assumption that individuals were not at risk for mortality until initiating treatment, meaning that model results are only applicable to individuals who accessed treatment at some point in their disease history. Because survival was described using a Cox proportional hazards model, we did not estimate a parametric form for the baseline hazard, and survival probabilities were based on a non-parametric function which only described the period for which survival data were available (approximately ten years following treatment initiation). Beyond this follow-up period, it was assumed in the base-case analysis that the monthly probability of HIV mortality remained constant. The observed mortality probability at the end of follow-up was used to estimate this quantity; this assumption was supported by the fact that mortality remained relatively constant during the latter period of observation. The microsimulation was run for a maximum of 30 years; it was assumed that there was no survival beyond this point.

5.2.7. *Uncertainty and sensitivity analysis*

Because this was a micro-level simulation, first-order uncertainty (i.e. individual-level variability within a population) was implicitly accounted for by simulating different individuals with different characteristics and different disease processes. Second-order uncertainty reflects the potential for systematic bias in results due to uncertainty in model input parameters.(22) We *a priori* identified the cost model parameters as potentially important sources of second-order

uncertainty. To quantify the impact of this uncertainty, we performed a probabilistic sensitivity analysis, in which the coefficients associated with all two-stage cost models were sampled from multivariate normal distributions. In addition, CD4 and viral load trajectories were sampled from normal distributions with the mean determined by the corresponding generalized additive model and standard deviation based on the residual variance of the generalized additive model. This incorporated variability into the trajectories that was not described by the assumed statistical model.

In addition, several one-way sensitivity analyses were performed. We examined the impact of changing the discount rate to 0% and 5%, respectively. We also examined several scenarios regarding survival probabilities extrapolated beyond the observed follow-up of the data source: (1) that the monthly mortality probability decreased by 50% after the follow-up period; (2) that the monthly mortality period increased by 50% after the follow-up period; and (3) that the maximum lifetime was 40 years following infection, compared to the base-case assumption of 30 years.

Another one-way sensitivity analysis was performed to estimate an upper bound on antiretroviral costs. In the base-case analysis, antiretroviral costs were estimated by a statistical model fit using observed dispensing data, which implicitly accounted for actual medication adherence, i.e. medications that were prescribed but not dispensed were not included in cost estimates. While this is accurate from the perspective of a third-party payer, we also wished to estimate a hypothetical lifetime cost based on the assumption of 100% medication adherence. This sensitivity analysis was based on a regimen of boosted atazanavir and tenofovir with lamivudine,

which costs $1,438/month in BC. This particular regimen was chosen because it is a relatively high-cost combination that is currently widely used in treatment-naive individuals in BC.

5.2.8. Internal validation

As a measure of internal validation, the process used to generate costs within the microsimulation was also used to generate estimated costs for the individuals who contributed data to the health services utilization and cost models. For these individuals, we randomly generated costs for each person-time interval they contributed, based on observed values of the variables considered within the microsimulation. We then summed and log-transformed both observed costs and randomly-generated costs across person-time intervals for each individual, and compared the aggregate measures. These measures were compared both graphically and using a linear regression model.

5.2.9. Computational details

All programming and statistical analyses were performed in `R` version `2.8.1`.(23) The microsimulation was designed as a series of matrix manipulations, which allowed for all individuals and all iterations of the probabilistic sensitivity analysis to be computed simultaneously. As a result, the only "loop" required was that associated with the 1-month time cycles. All microsimulation runs included 500 simulated individuals and 500 probabilistic sensitivity analysis iterations, for a total of 250,000 simulated life histories per run. A common random number stream was used across iterations of the probabilistic sensitivity analysis and across scenario and one-way sensitivity analyses. This increased efficiency by reducing individual-level variability across comparisons.(24)

5.3. Results

Model input parameters for demography and treatment-related variables are given in Table 5.1. Individuals generated within the microsimulation were assigned characteristics randomly from these distributions, which were based on distributions observed in the HOMER cohort. Sex and injection drug use history were included as covariates in the statistical models describing monthly direct medical costs, CD4 cell count and plasma viral load models. Medication adherence was also included in all these statistical models, as well as the medication resistance time-to-event model. PI- vs. NNRTI-based HAART was an indicator variable in the medication cost and resistance models, and choice of initial regimen was assumed to have implications for an individual's future resistance and treatment profile. Age was not found to be an important predictor in any of the statistical models fit, so it was only incorporated into the lifetime simulation through background (non-HIV-related) mortality. The distributions for all parameters were chosen based on empirical examination of the HOMER cohort.

The non-linear relationships between CD4 cell count and the monthly direct medical costs associated with hospitalizations, physician visits, laboratory tests, and emergency room utilization, estimated using two-stage random-effects models,(12) are shown in Figure 5.2. Factors associated with increased costs included HAART treatment status, female sex, and injection drug user status. The shaded area indicates the increased costs associated with CD4 cell counts below 200 cells/mm^3. In all plots, the initial stepwise decline describes the time between seroconversion and HAART initiation, assumed to be homogeneous across individuals, while the subsequent "random walk" represents post-HAART trajectories characterized by individual variability.

In the base-case analysis, monthly antiretroviral costs (not included in Figure 5.2) were estimated to range between $1,114 and $1,130 across demographic categories for a PI-based regimen and $922 and $938 for a NNRTI-based regimen.(Chapter 3) The estimated costs associated with the final month of life ranged from $15,164 and $24,559 across demographic categories.

The empirical covariance matrix of random effects estimated in the health services utilization and cost models is given in Table 5.2. Note that due to data transformations the utilization random effects are reported on a logit scale, while all non-antiretroviral cost random effects are reported on a log-cost scale. Antiretroviral costs were left untransformed, so these random effects are reported on a cost scale. These random effects represent individuals' tendencies to utilize health services beyond what would be expected based on their demographics and current CD4 cell count.

Randomly-generated CD4 cell count trajectories resulting from the microsimulation are shown in Figure 5.3, with periods with CD4 cell counts below 200 cells/mm^3 highlighted in the shaded area. The analogous common trajectory assumed for the cohort approach is shown in Figure 5.4. All individuals displayed here spent at least one month with CD4 cell counts below 200 cells/mm^3, in which direct medical costs are increased (Figure 5.2). These individual trajectories were consistent with the empirical data used to fit the statistical models.
In contrast, the population average CD4 trajectory remained consistently above the 200 cells/mm^3 threshold throughout a simulated lifetime (Figure 5.4). The discrepancy between these trajectories results from the fact that, *on average*, at any given time point, the majority of individuals have CD4 cell counts well above 200 cells/mm^3 (Figure 5.4). However, at the

individual level, the majority of individuals experience at least transient periods with lower CD4 cell counts (Figure 5.3), and this phenomena is only captured by incorporating individual-level variability into projections. In practice, this was achieved by incorporating the residual standard error of the CD4 model into the individualized trajectories. Thus, although both the individual and average trajectories were based on the same underlying statistical models, assuming the population average trajectory applied to all individuals failed to capture observed months with low CD4 counts and corresponding high medical costs.

Using the microsimulation approach, the estimated lifetime cost discounted at 3% was $267,822 (95% confidence interval: $257,500 – $278,049) (Table 5.3). Undiscounted lifetime cost was estimated to be $408,396 (95% confidence interval: $390,859 – $425,274). Estimated discounted lifetime costs were robust to varying assumptions about long-term survival, with differences from baseline ranging between 6% ($252,196) and 11% ($299,681). Assuming all individuals were fully compliant increased estimated discounted lifetime costs by 19% (to $319,596). Across scenario analyses, estimated life-years ranged between 22.4 years and 28.3 years following infection, while estimated quality-adjusted life years (QALYs) ranged between 11.7 and 20.7 QALYs. The majority of variation in life-years is explained by differing assumptions regarding long-term survival, while the variation in QALYs is explained largely by different choices of discount rates. Note that the variability reported in Table 5.3 refers to population-level variability observed across iterations of the probabilistic sensitivity analysis. At the individual level, simulated lifetime costs discounted at 3% ranged between $2,531 and $1.2 million.

Figure 5.5 compares the base-case distribution of lifetime costs to the corresponding distribution when key sources of heterogeneity were removed. Consistent with our *a priori* hypothesis, we observed that estimated lifetime costs costs were systematically lower when we did not incorporate individualized CD4 trajectories or random effects.

The comparison between observed and randomly generated log-costs, which was used to internally validate the method for using the two-stage utilization and cost models within the microsimulation, is shown in Figure 5.6. The solid line in the figure is the identity line with intercept 0 and slope 1. The data points show a clear positive relationship and are randomly distributed about the identity line, suggesting that the method used to generate costs provided an unbiased estimate of actual costs. Further, a linear regression of randomly generated log-costs against actual log-costs , constrained to have an intercept of 0, yielded an estimated slope of 1.004 and an R^2 coefficient of 0.993.

5.4. Discussion

In this study, we estimated the discounted lifetime direct medical cost of treating a person initiating HAART for HIV infection to be $267,822 (95% confidence interval: $257,500-$278,049). This value changed less than 20% when varying assumptions about survival and adherence. To our knowledge, this is the first study to estimate this parameter in a Canadian setting during the HAART era. Alternative data-analytic methods to the microsimulation approach have been suggested for estimating lifetime cost estimation, including an adaptation of Kaplan-Meier analysis for censored cost data (25) and a phase-of-illness approach.(26) Such methods were not appropriate in this particular setting because they are best suited to a data source in which full lifetime estimates are available for a majority of patients. In contrast, for

HIV during the HAART era, a large majority of patients who have initiated HAART since 1996 are still alive, necessitating the use of a simulation-based methodology to extrapolate assumptions regarding lifetime inference.

A strength of the microsimulation was that the majority of underlying processes were incorporated using population-based individual-level data, which reduced the need for assumptions based on expert opinion or drawn from the literature. Based on the quality and scope of the available data, the model is likely to be applicable other Canadian settings with similar HIV treatment guidelines and health care systems.

The results were strengthened by the robustness observed across assumptions regarding long-term survival. Because improvements in life expectancy currently exceed the availability of HAART,(3) long-term life-expectancy represents a key assumption that must be made when estimating lifetime medical costs. As more data become available over time describing long-term survival on HAART, the microsimulation can be updated to include survival models with longer follow-up. However, extensive sensitivity analyses indicated that the estimates derived here are likely to be within 20% of the actual mean value.

The study was also subject to several limitations related to both the methodological approach and the input data. The complexity of the model limited the number of probabilistic sensitivity analysis iterations that could be reasonably undertaken. While the 500 iterations (resulting in 250,000 simulated individuals) in each scenario analysis were sufficient to quantify the impact of second order uncertainty, a potential improvement could be to incorporate an optimized sampling scheme (e.g. latin hypercube sampling) in place of the Monte Carlo approach.(27)

Another limitation was the fact that only limited data were available for individuals who did not initiate HAART or who initiated HAART late in the course of disease. Because of this limitation, we assumed a common untreated course for CD4 cell count and plasma viral load prior to HAART initiation,(16) and individual-level variability was not incorporated into these pre-HAART trajectories. We anticipate this particular limitation to have had a limited impact on results; because treatment is initiated at a CD4 cell count of 350 cells/mm^3, untreated trajectories by definition describe CD4 cell counts above this threshold, which are associated with little cost variability (Figure 5.2).

A limitation due to the observational nature of the input data was the potential for informative missing data, e.g. if individuals with lower CD4 cell counts were less likely to present for routine laboratory monitoring. However, health services utilization and cost data were extracted from a comprehensive administrative source, so the ability of the microsimulation to validate observed costs (Figure 5.6) suggests that missing CD4 data did not present a major source of bias. Data on date of death may have been missing for individuals who died outside the province of BC. We made the assumption that the number of such deaths was likely to be negligible, and, further, that individuals who died out-of-province would not be systematically different with respect to disease course or health services utilization.

The comparison between the full microsimulation and the modified microsimulation with key heterogeneity sources removed provided insight into cost drivers in HIV disease. With cost model parameters and sociodemographic characteristics following common distributions across the two approaches, the different assumptions regarding CD4 trajectories and individualized random effects represented the primary source of different outcomes. We found that the

inclusion of individualized trajectories and random effects consistently led to systematically higher cost estimates. There are intuitive explanations for the importance of both sources of heterogeneity on lifetime cost estimates. With respect to CD4 trajectories, the relationship between CD4 cell count and mean monthly direct medical costs is non-linear (see Figure 5.2), with similar costs across the highest CD4 cell count strata and large increases in costs when CD4 cell count falls below 100 cells/mm^3. This pattern was observed to hold across all sociodemographic categories. Therefore, time spent in the lowest CD4 cell count strata has a large effect on lifetime medical costs. During the modern HAART era, on average, CD4 cell count is expected to remain in the highest strata, reflecting the success of HAART in improving immunologic function. However, at the individual level, some individuals do experience periods of reduced CD4 cell count, resulting from treatment failure, medication resistance, or suboptimal adherence.(Chapter 2) By incorporating individual-level trajectories based on actual observed trajectories, such periods could be realistically incorporated. With respect to random effects, it is important to note that while they are constrained to be centered about 0 (population average),(28) the majority of cost models were fit to log-transformed data to account for right skew. Thus, when model estimates were back-transformed to the original scale via exponentiating and applying the smearing factor, positive random effects tended to have a larger impact relative to negative random effects. This reflects the fact that while the majority of individuals are clustered together with relatively low monthly costs, a small proportion of individuals have very high monthly costs. By incorporating random effects, this right-skewed cost distribution can be explicitly included in the model.

The motivation for comparing results based on heterogeneous CD4 trajectories and random effects to those based on homogeneous values was to justify the use of a microsimulation model

as opposed to a cohort based model. A recent study reported cost-effectiveness results in HIV to be similar when comparing microsimulation and cohort approaches, with the microsimulation displaying slightly higher predictive ability.(29) The advantage to a cohort is that it would have had fewer data requirements and have been less computationally intensive, allowing for easier implementation and a larger number of probabilistic sensitivity analysis iterations. While the results based on the homogeneous approach were still generated using a microsimulation model, the approach represented the "disabling" of key microsimulation features, in order to approximate the results expected from a cohort model. Because the key difference between the two approaches was characterized by the inclusion of individual-level heterogeneity, it would have potentially been possible to design a cohort model that could approximate the microsimulation results. Here we assumed a single homogeneous CD4 trajectory based on the observed population average and did not incorporate any variability in random effects. Alternatively, a model could be designed assuming several distinct cohorts, each characterized by a different combination of CD4 trajectory and random effects. While such a model may avoid the systematic underestimation of costs observed when using the population average approach considered here, there is little to guide the number of cohorts that would be required. Further, as the number of distinct CD4 trajectories and random effects values increase, the number of required cohorts increases in a multiplicative fashion, and may become unrealistically large. In contrast, the full microsimulation approach makes efficient use of all available data.

Figure 5.1: Microsimulation structure

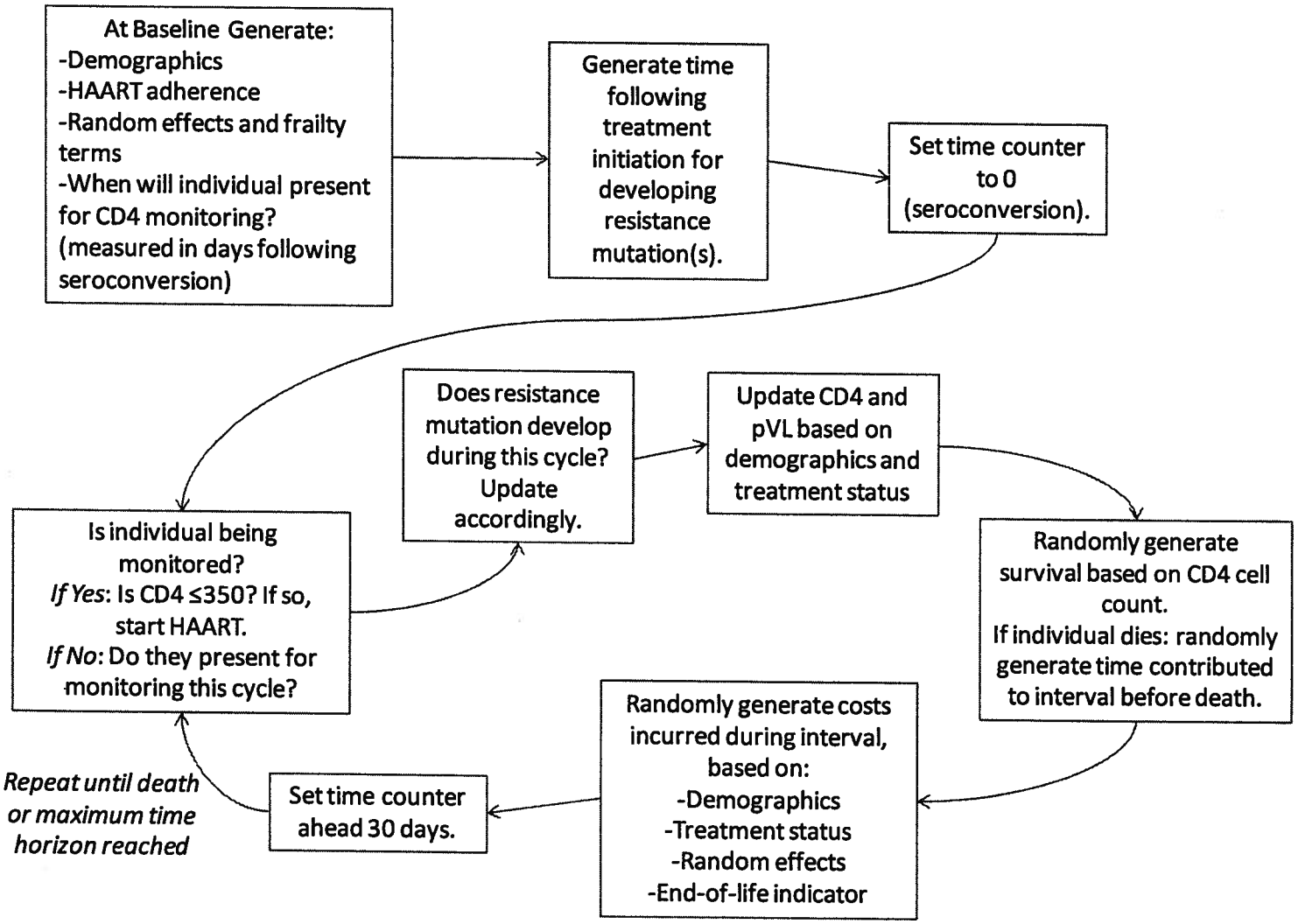

Figure 5.2: Estimated monthly costs (2005 CDN$)
Hospitalizations, physician visits, laboratory tests, and emergency room utilization by CD4 cell count strata.

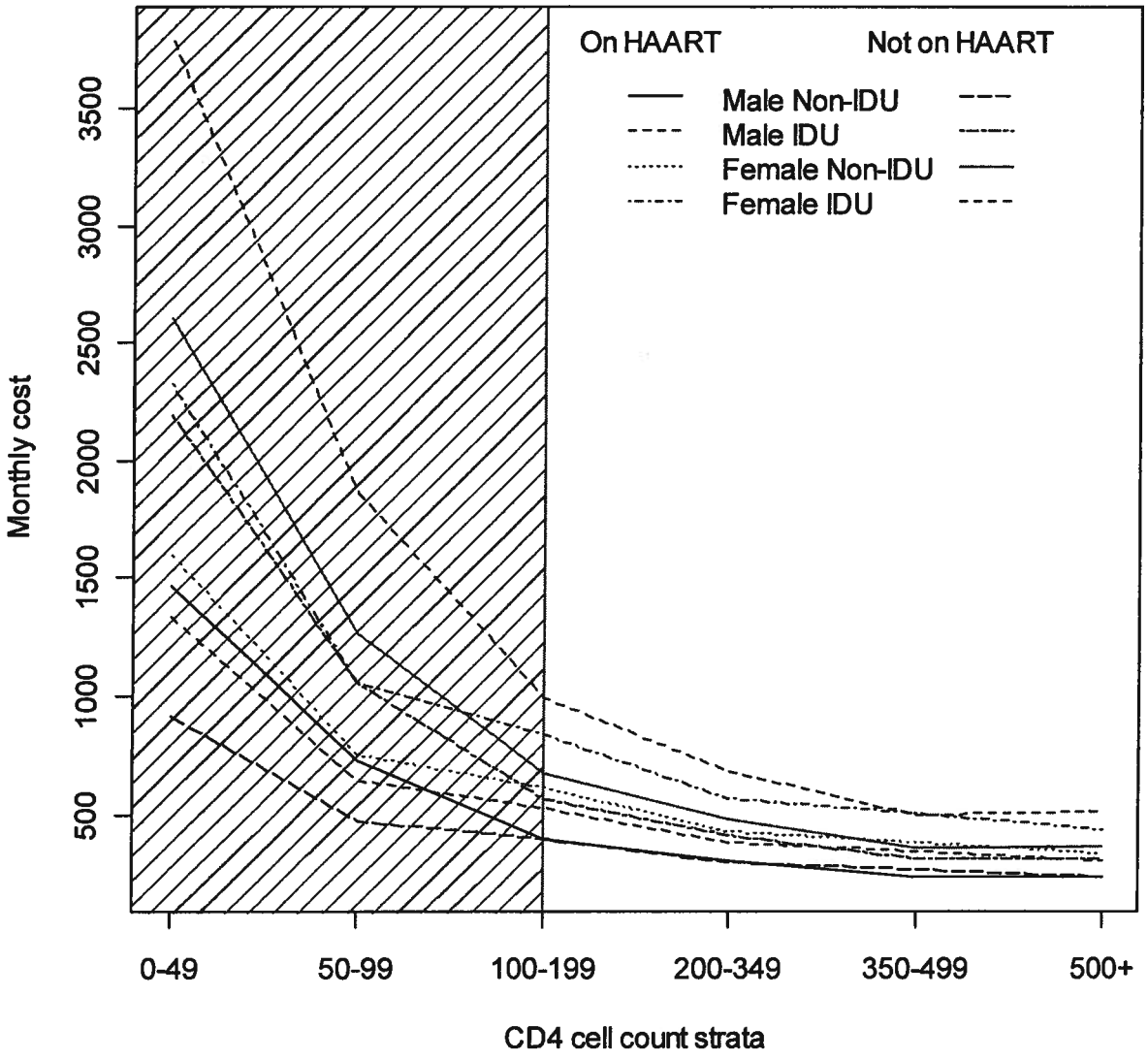

Figure 5.3: Example CD4 trajectories from microsimulation.
CD4 cell counts below 200 cells/mm^3 shaded.

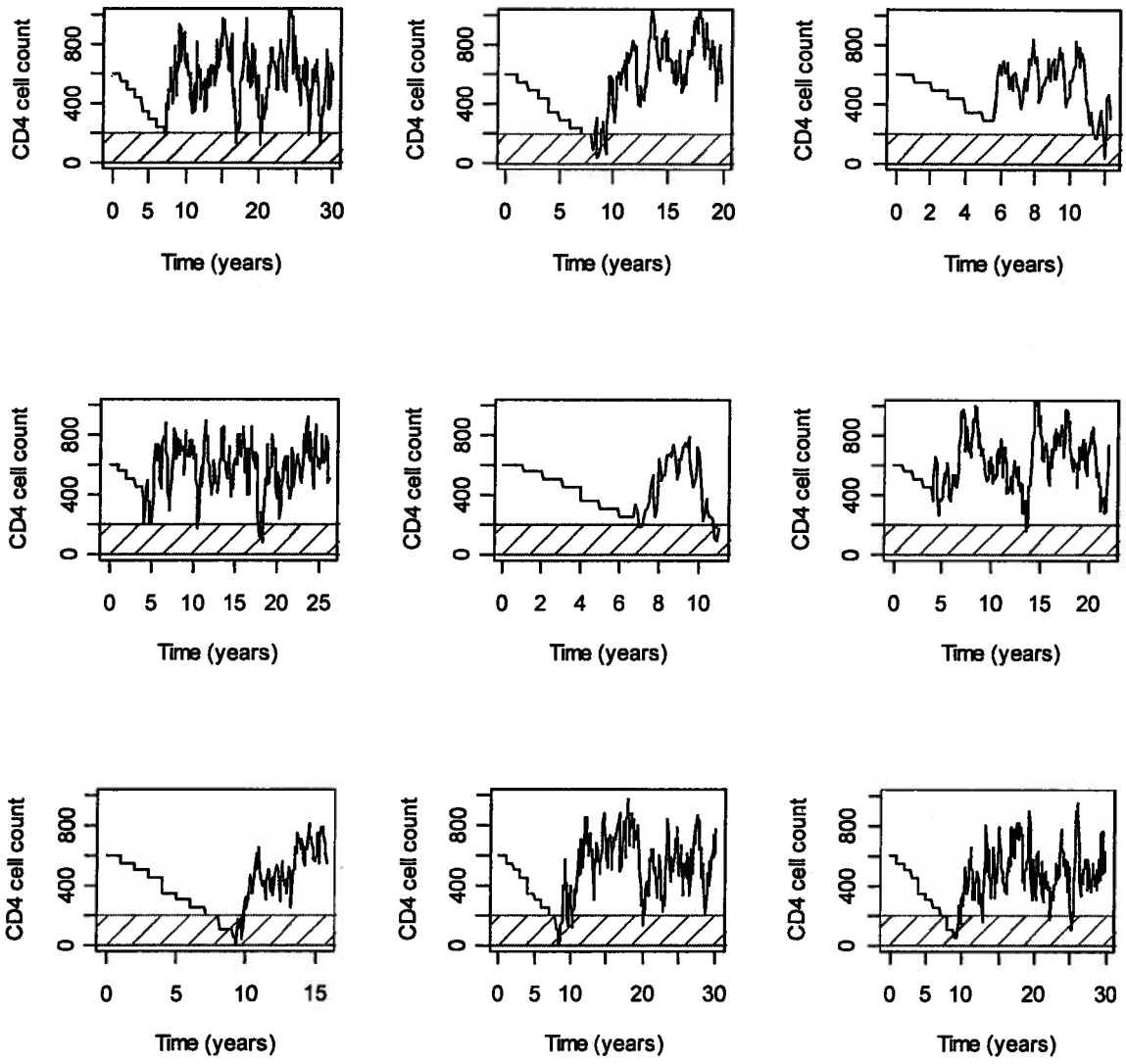

Figure 5.4: Common CD4 trajectory from cohort approach.
CD4 cell counts below 200 cells/mm^3 shaded.

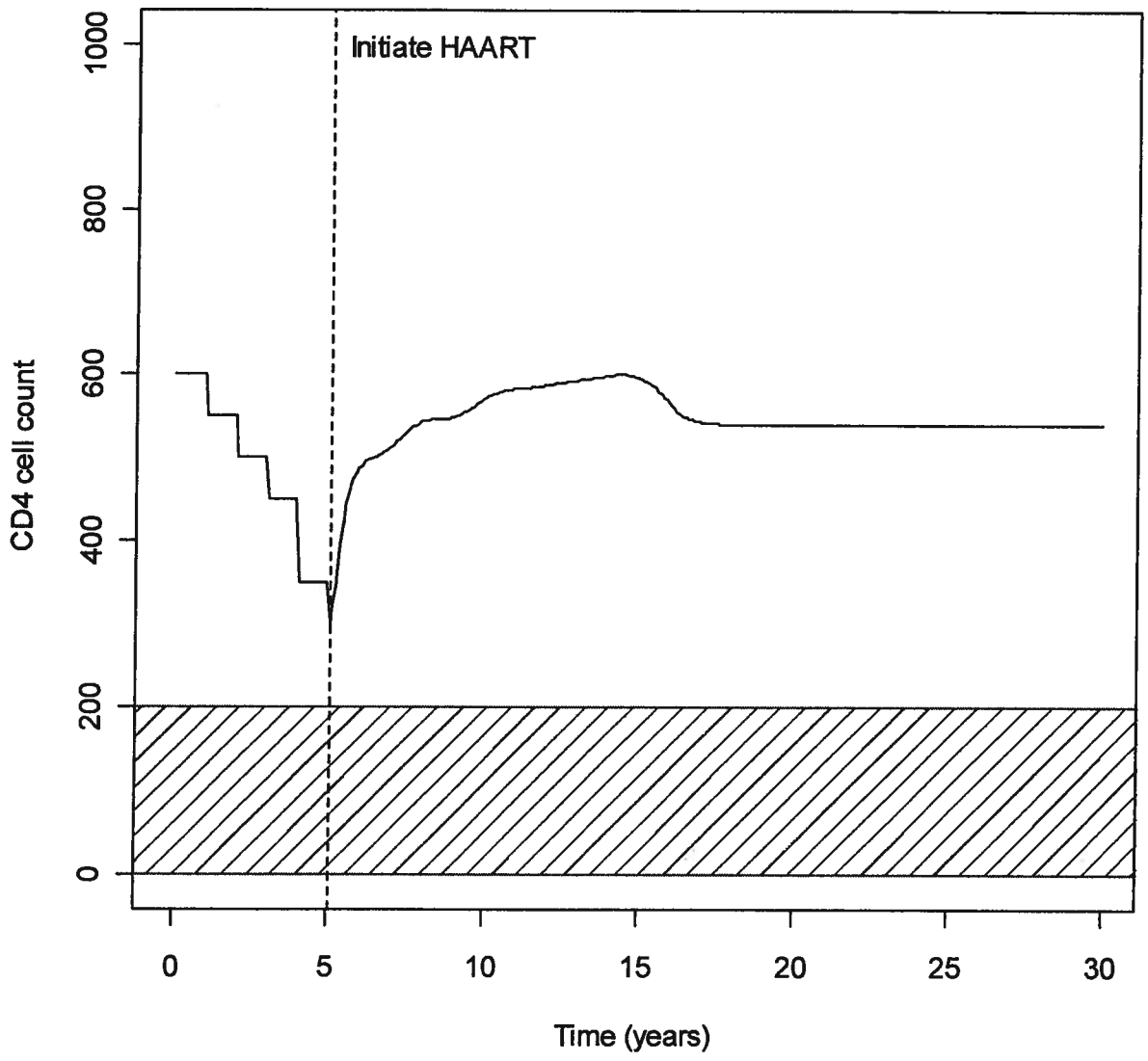

Figure 5.5: Distribution of lifetime costs.
Estimated within base-case microsimulation approach and comparison with key sources of heterogeneity removed.

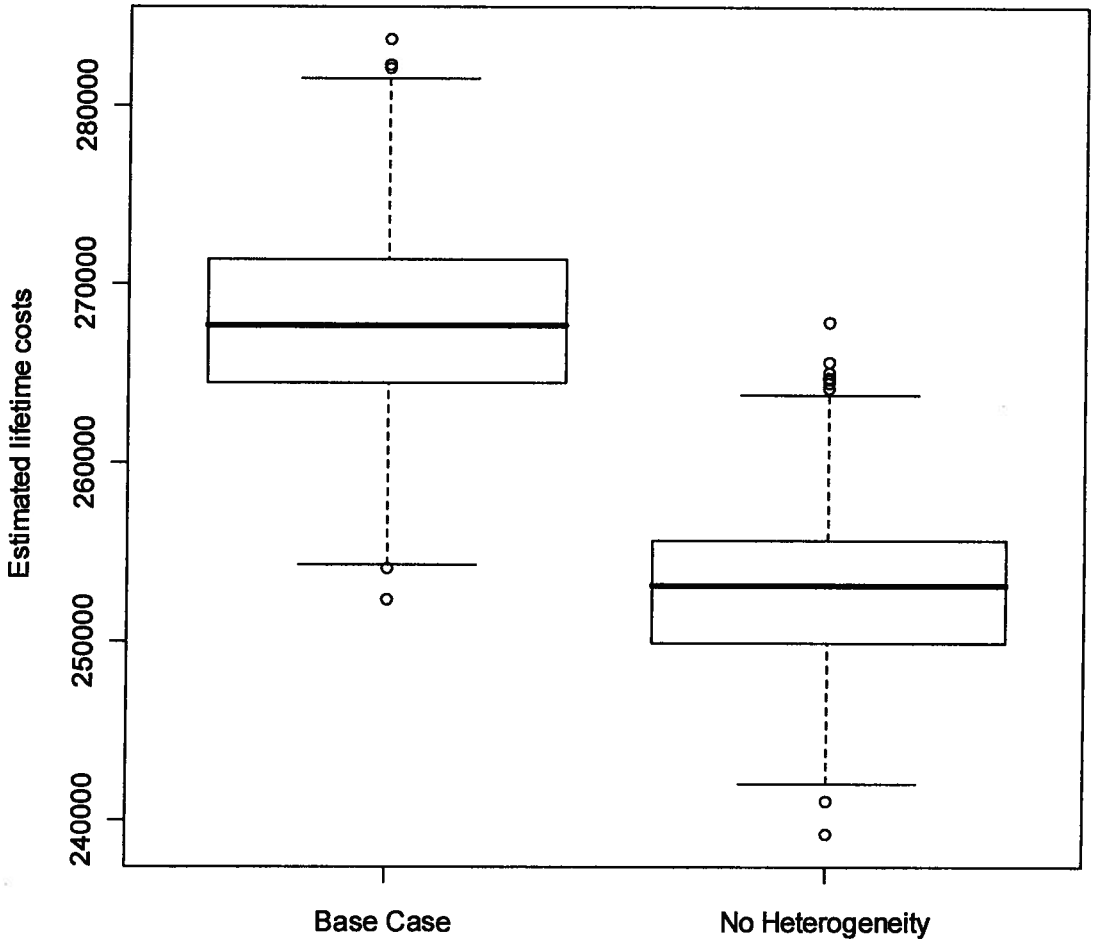

Figure 5.6: Observed vs. randomly generated log-costs used for internal validation purposes. Line shown has an intercept of 0 and a slope of 1.

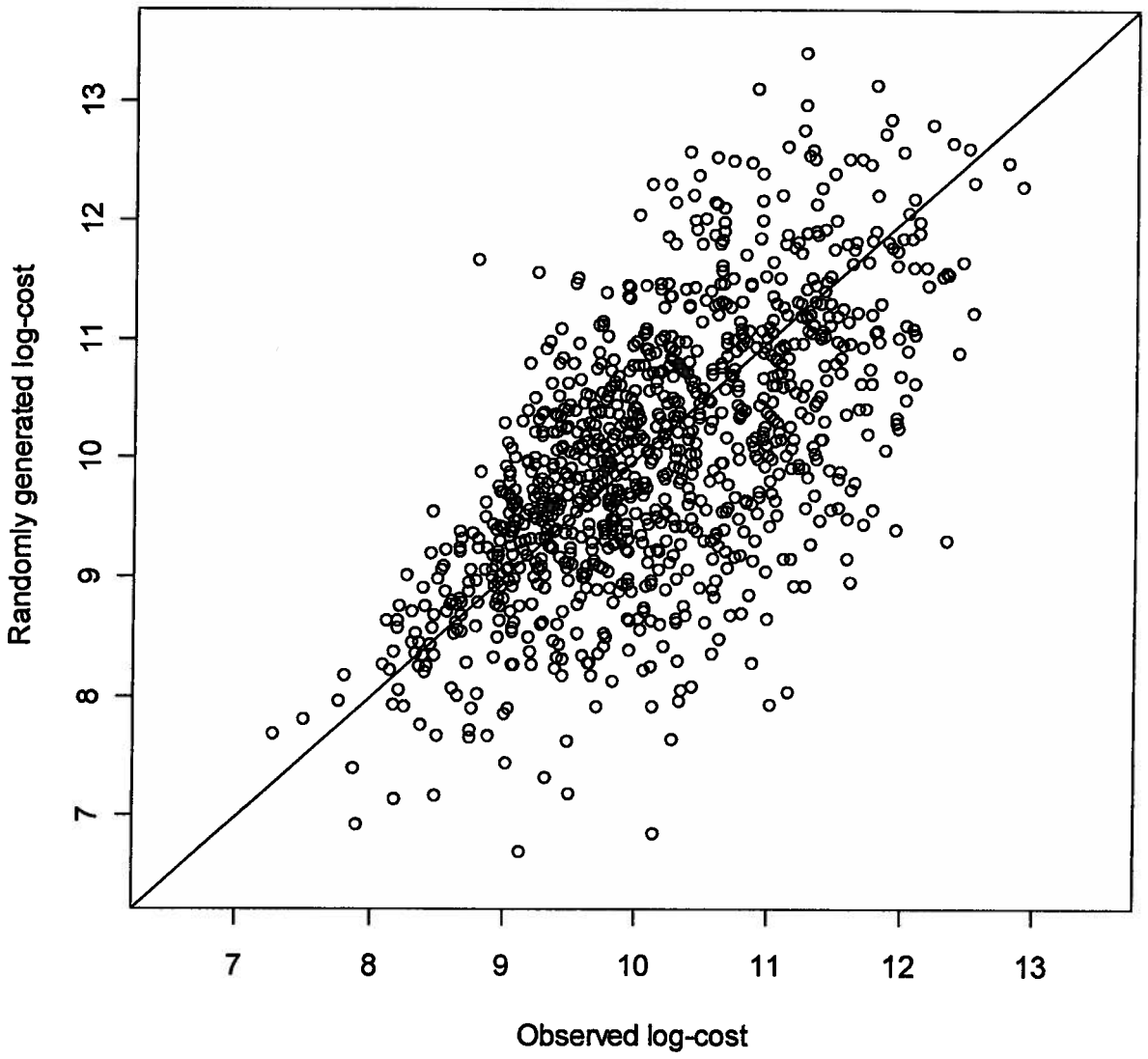

Table 5.1: Observed demographic and treatment characteristics in the HOMER cohort.
N=2,583; used to populate the microsimulation model of lifetime costs of treating persons infected with HIV in BC.(HAART=Highly Active Antiretroviral Therapy; PI=Protease Inhibitor; NNRTI=Non-nucleoside reverse transcriptase inhibitor.)

Variable		Distributional assumptions for microsimulation
Sex (%) Male Female	 81.6 18.4	Binomial
Injection drug use (%) No Yes	 72.3 27.7	Binomial
Initial HAART regimen (%) PI-based NNRTI-based	 50.0 50.0	Binomial
Age at treatment initiation (years) (mean (standard deviation))	 37.2 (3.1)	Normal
Adherence to HAART (refill %) (mean (standard deviation))	 79.1 (29.5)	Beta with parameters estimated by method of moments

Table 5.2: Empirical covariance matrix of random effects.
Estimated in the health services utilization and cost models. (ER=Emergency Room; HAART=Highly Active Antiretroviral Therapy)

	Inpatient Utilization	Outpatient Utilization	Laboratory Utilization	ER Utilization	Non-HAART Medication Utilization	Inpatient Cost	Outpatient Cost	Laboratory Cost	ER Cost	Non-HAART Medication Cost	HAART Cost
Inpatient Utilization	1.0435	0.2021	0.0622	0.7116	0.4267	0.0996	0.1962	0.0010	0.0205	0.2147	-4.6753
Outpatient Utilization		0.8526	0.0849	0.2212	0.4238	0.0027	0.1136	-0.0005	0.0157	0.2078	2.4391
Laboratory Utilization			0.0943	0.0937	0.0812	-0.0025	0.0340	0.0020	0.0041	0.0409	3.8435
ER Utilization				1.0592	0.5452	0.0658	0.1587	-0.0035	0.0384	0.2157	-1.8038
Non-HAART medication Utilization					10.0928	0.0216	0.1089	0.0164	0.0710	0.3361	1.5397
Inpatient Cost						0.1086	0.0282	-0.0023	0.0038	0.0150	-2.0708
Outpatient Cost							0.1358	0.0018	0.0051	0.0958	0.4764
Laboratory Cost								0.0199	0.0002	-0.0061	0.0314
ER Cost									0.0107	0.0053	0.2900
Non-HAART medication Cost										0.7831	2.5914
HAART Cost											3594.43

Table 5.3: Estimated lifetime costs, life-years, and quality-adjusted life years.
Base-case analysis, scenario analyses, and 1-way sensitivity analyses. Base-case analysis assumes treatment initiation threshold of 350 cells/mm^3, discounting rate of 3%, and 30-year time horizon.

Scenario	Lifetime cost (95% confidence interval) 2005 CDN$	Undiscounted Life-years (95% confidence interval)	Quality-adjusted life-years (95% confidence interval)
Base-case (30-year time horizon; constant mortality after follow-up).	267,822 (257,500 – 278,049)	23.9 (23.2 – 24.6)	14.4 (14.0 - 14.7)
0% Discount rate	408,396 (390,859 – 425,274)	As for base-case	20.7 (20.1 - 21.3)
5% Discount rate	209,192 (201,714 – 216,940)	As for base-case	11.7 (11.4 - 11.9)
Reduced mortality rates after end of follow-up	276,805 (266,001-287618)	24.8 (24.1-25.5)	14.8 (14.1-15.1)
Increased mortality rates after end of follow-up	252,196 (242,103–262,423)	22.4 (21.7-23.0)	13.7 (13.4-14.0)
40-year time horizon	299,681 (286,850–311,344)	28.3 (27.2-29.3)	15.7 (15.3-16.2)
100% adherence to boosted atazanavir and tenofovir with lamivudine ($1438/month)	319,596 (309,150 – 330,037)	23.9 (23.2 – 24.6)	14.4 (14.0 - 14.7)

5.5. References

1. Palella FJ, Delaney KM, Moorman AC, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. New England Journal of Medicine 1998;338:853-60.
2. Levy AR, James D, Johnston KM, et al. The direct costs of HIV/AIDS care. Lancet Infectious Diseases 2006;6:171-7.
3. King JT, Justice AC, Roberts MS, et al. Long-term HIV/AIDS survival estimation in the highly active antiretroviral therapy era. Medical Decision Making 2003;23:9-20.
4. Yazdanpanah Y, Goldie SJ, Losina E, et al. Lifetime cost of HIV care in France during the era of highly active antiretroviral therapy. Antiviral Therapy 2002;7:257-66.
5. Freedberg KA, Losina E, Weinstein MC, et al. The cost effectiveness of combination antiretroviral therapy for HIV disease. New England Journal of Medicine 2001;344:824-31.
6. Schackman BR, Gebo KA, Walensky RP, et al. The lifetime cost of current human immunodeficiency virus care in the United States. Medical Care 2006;44:990-7.
7. Kuntz KM, Goldie SJ. Assessing the sensitivity of decision-analytic results to unobserved markers of risk: Defining the effects of heterogeneity bias. Medical Decision Making 2002;22:218-27.
8. Palepu A, Tyndall MW, Leon H, et al. Hospital utilization and costs in a cohort of injection drug users. Canadian Medical Association Journal 2001;165:415-20.
9. British Columbia Centre for Excellence in HIV/AIDS Drug Treatment Program. Unpublished Data 2007.

10. AIDSInfo. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. 2007.

11. Organization for Economic Development and Co-operation (OECD). Health Data 2007., 2007.

12. Diehr P, Yanez D, Ash A, et al. Methods for analyzing health care utilization and costs. Annual Review of Public Health 1999;20:125-44.

13. Wood S. Generalized Additive Models: An Introduction with R. Chapman & Hall/CRC, 2006.

14. McCune JM. The dynamics of CD4(+) T-cell depletion in HIV disease. Nature 2001;410:974-9.

15. Klein JP, Moeschberger ML. Survival analysis: techniques for censored and truncated data. Springer, 2003.

16. Richman D. Human Immunodeficiency Virus. International Medical Press, 2003.

17. BC Stats. Population Estimates — British Columbia. 2004.

18. Statistics Canada. Deaths by single year of age and geography — Both sexes. 2004.

19. Schackman BR, Goldie SJ, Freedberg KA, et al. Comparison of health state utilities using community and patient preference weights derived from a survey of patients with HIV/AIDS. Medical Decision Making 2002;22:27-38.

20. Bangsberg DR, Moss AR, Deeks SG. Paradoxes of adherence and drug resistance to HIV antiretroviral therapy. Journal of Antimicrobial Chemotherapy 2004;53:696-9.

21. Duan N. Smearing estimate - A nonparametric retransformation method. Journal of the American Statistical Association 1983;78:605-10.

22. Briggs AH. Handling uncertainty in cost-effectiveness models. Pharmacoeconomics 2000;17:479-500.

23. R: A Language and Environment for Statistical Computing. 2006.

24. Shechter SM, Schaefer AJ, Braithwaite RS, et al. Increasing the efficiency of Monte Carlo cohort simulations with variance reduction techniques. Medical Decision Making 2006;26:550-3.

25. Lin DY. Estimating medical costs from incomplete follow-up data. Biometrics 1997; 53:419-434.

26. Yabroff KR, Warren JL, Schrag D. Comparison of Approaches for Estimating Incidence Costs of Care for Colorectal Cancer Patients. Medical Care 2009; 47:S56-S63.

27. Oakley JE, O'Hagan A. Probabilistic sensitivity analysis of complex models: a Bayesian approach. Journal of the Royal Statistical Society Series B-Statistical Methodology 2004;66:751-69.

28. Diggle P, Heagerty P, Liang, K, Zeger S. Analysis of longitudinal data. Oxford University Press, 2002.

29. Simpson KN, Strassburger A, Jones WJ, et al. Comparison of Markov Model and Discrete-Event Simulation Techniques for HIV. Pharmacoeconomics 2009; 27:159-165.

6. COST-EFFECTIVENESS OF INCREASING ACCESS TO HIGHLY ACTIVE ANTIRETROVIRAL THERAPY FOR TREATING HIV/AIDS*

6.1. Introduction

More than 25 years after the first diagnosed cases of HIV/AIDS, the disease continues to present a significant global health burden. Currently, an estimated 33.0 million individuals are estimated to be living with HIV, with 2.7 million new infections and 2 million deaths estimated to have occurred in 2007.(1) In some countries in sub-Saharan Africa, adult prevalence is greater than 20%.(1) Although treatment with antiretroviral medications has been successful in reducing morbidity and mortality,(2) it does not eradicate the infection within infected individuals,(3) making the prevention of new infections a priority in curbing the growth of the epidemic.

It has been suggested that increasing coverage with highly active antiretroviral therapy (HAART) amongst persons infected with HIV, who have a medical indication for HAART, may help to prevent new infections.(4) The mechanism is thought to be through a reduction in HIV-1 viral load among treated individuals,(5) which reduces the likelihood of HIV transmission to uninfected persons.(6, 7) A mathematical model describing transmission of HIV in sub-Saharan Africa has been used to show that universal screening coupled with immediate HAART for those testing positive could result in dramatic decreases in HIV incidence and long-term treatment costs in this setting.(8)

---

* A version of the chapter will be submitted for publication. Johnston KM, Levy AR, Montaner JS, Lima VD, Briggs AH, Hogg RS, Tyndall M, Gustafson P. Cost-Effectiveness of increasing access to highly active antiretroviral therapy for treating HIV/AIDS.

Given the relatively high lifetime cost of treating individuals infected with HIV,(9; Chapter 5) an effective prevention strategy may be a cost-effective measure.(4) Simplified costing analyses, which included only the costs of HIV testing and HAART have supported this empirically.(8) However, there are a number of expenses associated with increased use of HAART that must be accounted for in addition to the cost associated with providing lifelong therapies to more individuals. Additional expenses include the provision of medical services to individuals infected with HIV who are now living longer due to the increase in life expectancy conferred by HAART, and thus requiring medical care for an extended period in time.(10)

In addition, a potential limitation to using increased treatment with HAART as a complementary form of prevention is that it does not address infections transmitted during primary infection, which is the period immediately following seroconversion. Individuals in the primary infection phase typically experience transient, high levels of viral load.(11) Further, they may not be aware of their HIV status, in which case no modification of risk behaviour would be expected, making primary infection a plausible period for further transmission.(12) Because individuals experiencing primary infection are not likely to be captured by a program to increase use of HAART within existing clinical guidelines, this source of transmission would not be affected by such a program.

To fully quantify the strengths, limitations, and cost implications of the proposed strategy to increase use of HAART, a formal economic evaluation is required. The objective of such an evaluation is to account for all relevant costs associated with implementing the program, and to

compare these costs with the cost reductions and health benefits expected to result from increased use of HAART, in both the short- and long-term.

In British Columbia (BC), Canada, the HIV epidemic can be characterized as being mature and relatively stable, with comparatively consistent annual incidence of new positive tests over the past ten years.(13) Using the proposed updated definition of a concentrated epidemic to be one in which transmission is primarily restricted to particular HIV-vulnerable groups,(14) BC can be classified as having a concentrated epidemic with the majority of infections occurring within the subpopulations of men who have sex with men, sex trade workers, and injection drug users.(15) The objective of this study was to perform a comprehensive economic evaluation of a program to increase the uptake of HAART in BC. A secondary objective was to track the proportion of new infections transmitted by individuals in the primary infection phase, and to compare this proportion across the hypothetical strategies regarding HAART coverage.

6.2. Methods

6.2.1. Model structure

The model utilized in this study was designed as an integration of an existing microsimulation model of the clinical and economic course of HIV with a series of difference equations describing HIV transmission in discrete time steps.

The microsimulation is described in detail in Chapter 5. Table 6.1 describes the processes included in the microsimulation and the statistical methods used to model them. All of these processes were modelled using statistical analyses of individual-level data that were available for

all individuals in BC receiving treatment with HAART.(16) In particular, CD4 and viral load trajectories following the initiation of HAART were described by non-linear statistical models,(17) and direct medical costs were described by random effects models that incorporated both utilization and level of use of particular health services.(18) The microsimulation described individual disease trajectories, so that all processes were generated separately to account for random variability between individuals. Based on BC data, within the microsimulation it was assumed that 81.6% of individuals were male, 72.3% were injection drug users, and average medication adherence (defined as the proportion of months receiving HAART) was 79.1%.

One addition to the previously-described microsimulation was the incorporation of untreated life histories, which were necessary to account for the implications of allowing a proportion of individuals to remain untreated. Individuals who did not access treatment were assumed to experience CD4 cell count trajectories associated with the natural untreated history of HIV.(11) Following primary infection, untreated viral load trajectories were characterized by a "set-point" that was randomly generated with a mean of 4 log-copies,(19) and an assumed increase of 0.5 log-copies per year in the absence of treatment. HIV-related mortality in the untreated group was based on data collected prior to the introduction of HAART.(20)

The disease transmission model was based on a model created previously,(21) updated to be integrated with the microsimulation. We considered a population of susceptible individuals at high risk for acquiring HIV infection, and a population of individuals infected with HIV at varying disease stages. The baseline infected population was assumed to consist of 8000 individuals, of whom 50% were assumed to have accessed treatment at some point in their

disease history. The population size was chosen through a process of empirical calibration, as it resulted in model-predicted numbers of new infections and new treatment initiations that were consistent with observed data.(21)

Simulated time moved forward in one-month intervals, during which individual disease histories and resource utilization profiles were updated within the microsimulation, and new infections were estimated using a series of difference equations. Discrete-time difference equations were used in place of continuous-time differential equations because the model was structured to "stop" at each one-month time interval and calculate new infections based on the distribution of clinical variables determined by the microsimulation. The difference equations described a process in which susceptible individuals became infected at a rate determined by the respective numbers of susceptible and infected individuals, the distribution of viral load amongst infected individuals, the baseline level of risk behaviour observed in the population, and any decrease in risk behaviour due to increased viral load. Mathematically, the equations are expressed:

$$S_{t+30} = S_t - \left(S_t \beta_0 \sum_{j=0}^{4} I_t^{(j)} \gamma^{(j)} (1 - \delta^{(j)}) / N \right) + \lambda$$

$$I_{t+30} = I_t + \left(S_t \beta_0 \sum_{j=0}^{4} I_t^{(j)} \gamma^{(j)} (1 - \delta^{(j)}) / N \right) - \mu_t$$

In this notation, t is measured in days, S_t refers to the number of susceptible individuals at time t, $I_t^{(j)}$ refers to the number of individuals in viral load category j at time t, β_0 is the baseline population force of infectivity, $\gamma^{(j)}$ is the increase in infectivity associated with viral load category j, $\delta^{(j)}$ is a dampening of the increased infectivity associated viral load category j due to a decrease in risk behaviour, N is the total number of individuals in the population (susceptible and infected), λ is the net migration into the susceptible population accounting for migration in and

out and mortality (note that λ could be positive or negative, depending on migration patterns), and μ_t is the mortality rate of infected individuals at time t, which varies according to the disease stage distribution at time t. The model structure for a given time period is displayed in Figure 6.1.

The viral load categories considered were: primary infection during the 60 days following seroconversion ($I^{(0)}$); viral load below 3 log copies per mL ($I^{(1)}$), viral load in the interval [3,4) log copies per mL ($I^{(2)}$), viral load in the interval [4,5) log copies per mL ($I^{(3)}$), and viral load greater than or equal to 5 log copies per mL ($I^{(4)}$). Category $I^{(1)}$ was taken as the baseline category, so $\gamma^{(1)}$ and $\delta^{(1)}$ were both defined to be 1, meaning that individuals in this viral load category represent the baseline level of infectivity and risk behaviour. For increasing viral load categories, it was assumed that an increase of 1 log copy/mL was associated with a 2.45-fold increase in infectivity(6) and a 10% reduction in risk behaviour, based on the assumption that viral load could be used as an acceptable proxy for morbidity. This decrease in risk behaviour was varied in the sensitivity analysis. It was further assumed that during the period of primary infection, viral load reached a level of 6 log copies/mL with no reduction in risk behaviour.

Newly-infected individuals were randomly assigned to one of two categories: the first category of individuals would remain untreated throughout the course of their disease, while the second category would at some point access treatment with HAART. Individuals assigned to the second category were then randomly assigned a point in their disease history (between zero and ten years following infection) at which they would present for routine clinical monitoring. These individuals were assumed to initiate treatment with HAART once they had presented for

monitoring and their CD4 cell count first dropped below 350 cells/mm^3. This procedure allowed us to account for the fact that some individuals may not present for treatment until their CD4 cell count has dropped well below the threshold for clinical eligibility.

A 1000-iteration probabilistic sensitivity analysis(22) was performed to assess the impact of uncertainty in model parameters on cost-effectiveness results. Nested within each iteration of the sensitivity analysis, individual life histories were generated to account for heterogeneity in clinical and economic processes. Parameter values used in the difference equations and associated distributions for the probabilistic sensitivity analysis are given in Table 6.2.

6.2.2. Cost and utility estimates

Costs incurred by infected individuals included costs associated with physician visits, hospitalizations, HAART, non-HAART medications, laboratory tests, and emergency room visits. These costs were generated randomly based on statistical models, adjusted for treatment status, medication adherence, and CD4 cell count.(18) In the base-case analysis, HAART costs were based on actual costs incurred by infected individuals in BC, which includes some individuals with sub-optimal adherence. We also performed a sensitivity analysis, assuming 100% adherence to an arbitrary regimen of boosted atazanavir and tenofovir with lamivudine, which is a commonly-used first-line regimen in BC and costs $1,438/month.

Costs incurred by susceptible individuals were assumed to be $4,133/year, the average Canadian health care cost reported for 2005.(23) To account for health-related quality of life, life years were converted to quality-adjusted life years by applying utilities to time spent in various health

states. The utility associated with a health state is a number between zero and one which is used to weight time spent in the state, with one representing perfect health and zero representing death.(24) Health state utilities for infected individuals were assumed to follow published estimates,(25) while a utility of 0.87 was assumed for the susceptible state. All costs were converted to 2005 Canadian dollars, using health care-specific cost inflators.(26)

6.2.3. Scenario analyses

We considered two scenarios regarding treatment with HAART. In both scenarios, it was assumed that the infected population at baseline was based on a historical treatment rate of 50%. In the first scenario it was assumed that this treatment rate would remain at 50% for the entire simulated time period. In the second scenario it was assumed that treatment uptake would immediately increase to 75% and that 75% of newly infected individuals would continue to seek treatment throughout the simulated time period.

6.2.4. Model outcomes

The primary outcome compared between the two scenarios was incremental cost-effectiveness, defined as the ratio between incremental costs and incremental quality-adjusted life years across scenarios. Cost-effectiveness was assessed using the incremental net benefit approach,(27) based on the commonly used willingness-to-pay threshold of $50,000 per quality-adjusted life year.(28) The net benefit is a synthetic economic quantity, equal to the product of the willingness-to-pay threshold and the difference in effectiveness, less the difference in costs, simultaneously incorporating costs, survival, and quality of life. For a given willingness-to-pay threshold, a positive incremental net benefit based on that threshold indicates that the alternative

intervention is cost-effective. Net benefit was considered at both the overall population level, in which costs and quality-adjusted life years were aggregated across all infected and susceptible individuals over time, and the patient-centred level, in which the only costs and quality-adjusted life years considered were those accrued by the subgroup of individuals who were infected at baseline. Patient-centred net benefit reflects the direct health benefits due to expanding HAART use, while overall population net benefit further incorporates the benefits due to reducing new infections. Uncertainty in the model was expressed using empirical 95% confidence bounds for the incremental net benefit(29) and a cost-effectiveness acceptability curve.(30)

A second outcome used to evaluate the two scenarios was the number of new HIV infections expected to occur under each scenario. New infections were separated into those resulting from individuals during the primary infection stage of their disease versus those resulting from individuals who were past the primary infection stage.

6.3. Results

Figure 6.2 displays the estimated cumulative incremental net benefit over time for the HAART expansion scenario relative to a continuing coverage rate of 50%. The incremental net benefit reported here is based on a willingness-to-pay threshold of $50,000 per quality-adjusted life year; an incremental net benefit greater than zero indicates that the intervention is considered cost-effective at this threshold.(27) Using the $50,000 threshold, the HAART expansion scenario was estimated to reach cost-effectiveness—indicated by a positive incremental net benefit—within four years. The cumulative net benefit was further estimated to increase steadily over time, reaching approximately $900 million (95% Confidence Interval [CI]: $493 million -

$1.45 billion) after 30 years. Uncertainty in results due to uncertainty in input parameters was assessed through the probabilistic sensitivity analysis, with 95% confidence bounds included in Figure 6.2. Within four years of implementation of the HAART expansion program, the lower 95% confidence bound for the net benefit became positive. As an alternative to the net benefit approach, the results of the probabilistic sensitivity analysis are also shown in a cost-effectiveness acceptability curve in Figure 6.3. For a willingness-to-pay threshold of $20,000, over 80% of iterations indicated cost-effectiveness of the HAART expansion scenario, while over 90% of iterations indicated cost-effectiveness for a willingness-to-pay threshold of $50,000.

When we assumed increased costs of $1,438 for monthly HAART costs, the scale of cumulative net benefit was reduced, reaching a plateau of approximately $760 million over thirty years. However, a similar pattern of cost-effectiveness was observed, with the HAART expansion scenario reaching cost-effectiveness within seven years and remaining cost-effective throughout the 30-year duration of simulated time (results not shown).

The overall population and patient-centered incremental net benefits are shown in Figure 6.4 (note that the overall population curve is equivalent to the point estimate shown in Figure 6.2). For the first five years of expanded HAART, it is estimated that overall population net benefit is comprised almost entirely of net benefits accrued by the initial group of individuals infected at baseline. The general trend continues for approximately ten years, with the overall population net benefits gradually increasing beyond the patient-centered net benefits. However, beyond ten years, there is an increasing divergence, with the overall population net benefits being

increasingly explained by averted infections. At the end of the 30 year simulated period, patient-centered net benefits account for approximately half of overall population net benefits.

Cumulative new infections predicted under each scenario are shown in Figure 6.5. The HAART expansion scenario was consistently associated with fewer new infections. During the 30-year simulated time period, the HAART expansion scenario was predicted to yield 26% fewer new infections. In the mature epidemic simulated here, a relatively small proportion of new infections was estimated to result from the period of primary infection.

6.4. Discussion

In this study, we demonstrated the potential cost-effectiveness of expanding the use of HAART—as defined by increasing the proportion of HIV-infected individuals with CD4 cell counts below 350 cells/mm^3 receiving HAART from 50% to 75%—in BC. Increased treatment was found to reduce the incidence of new infections, and, despite the up-front acquesition costs associated with an increase in HAART use, it was estimated that the strategy became cost-effective within four years. Although HAART expansion would not reduce the probability of transmission during the primary infection phase, this limitation was not predicted to substantially decrease the overall effectiveness of the strategy.

Several other mathematical models have been used to assess the impact of HAART on HIV transmission. Relative to the results presented here, some models have resulted in similar or more dramatic effects on HIV incidence(8, 21, 31), while others have resulted in more conservative estimates.(32, 33) All investigators have reported some reduction in the incidence

of new infections when treatment rates are increased. When evaluating infectious disease prevention strategies, the current phase of the particular epidemic being modelled is an important factor,(34) so some discordance between studies undertaken in different settings is to be expected. To our knowledge, ours is the first study to incorporate a comprehensive economic evaluation in this context.

The microsimulation component of the model, which described the clinical course and economic implications of HIV disease, was based on high-quality population-based data for BC, and is thus expected to be accurate for BC and other areas with similar health care systems. Because this component of the model was based on observational data, it describes actual clinical and virological outcomes observed in practice, rather than idealized outcomes based on the assumption of optimal medication adherence.

Another strength of the model was that it addressed multiple sources of complexity of HIV disease through the integration of the microsimulation with the transmission model. Because HIV is an infectious disease, a dynamic modelling method, such as the transmission model used here, is required to quantify the impact of a prevention program. In addition, due to the relatively long clinical course of HIV, it also displays properties of a chronic disease, and lifetime direct medical costs may vary substantially across individuals.(Chapter 5) Microsimulation methods based on statistical models have been well developed within the health economics literature for describing the costs associated with a chronic disease.(35) These methods provide a framework for quantifying the impact of HAART use at the individual level. By combining the two modelling techniques, we were able to exploit the strengths of two

complementary methods: one for addressing the infectious nature of HIV transmission, and one for addressing the chronic nature and individual-level variability associated with HIV clinical processes and medical costs.

A limitation of the study was the paucity of data available for assigning parameters to the disease transmission component of the model. Parameters for which no empirical data were available were chosen using a combination of expert opinion and calibration to historical incidence and treatment rates.(21) In order to reduce model complexity, we assumed a single baseline infectivity parameter (β_0), and did not differentiate between different routes of transmission (e.g. men who have sex with men, injection drug use, sex trade work). Thus, β_0 can be viewed as a implicitly incorporating the relative sizes of different risk groups, and the level of risk behaviour and risk of transmission associated with each risk group. It is unlikely that β_0 is actually static, as the relative size of risk groups and the behaviour of individuals change over time. However, we made the simplifying assumption that the β_0 parameter would remain constant over the simulated time period. We do not expect that this parameter uncertainty in β_0 had a major impact on model results. Within this study, it was not our objective to exactly replicate transmission patterns observed in BC or predict the specific number of infections to be expected in upcoming years. Rather, we wished to evaluate the difference in outcome between two treatment strategies in the context of a mature, concentrated epidemic consistent with that observed in BC. Because our primary outcomes were incremental in nature and all comparisons used the same parameter estimates, it was unlikely that any incorrect assumptions regarding specific parameter values would have biased overall results in either direction. Further, we allowed parameters to vary across a relatively wide range of plausible values in the probabilistic sensitivity analysis (see

Table 2). The results of this sensitivity analysis were consistent with the base-case analysis regarding cost-effectiveness over time, using a willingness-to-pay threshold of $50,000 per quality-adjusted life year.

In order to not overestimate the net benefit associated with the HAART expansion scenario, wherever possible, we made conservative assumptions to make increased uptake of HAART appear *less* effective. Firstly, we made the conservative choice of 0.87 as the utility value associated with individuals susceptible for HIV infection. It has been suggested that assigning a utility of 1.0 ("perfect health") to the absence of a particular health condition ignores the likely presence of other chronic and acute health conditions, and can result in overly optimistic cost-effectiveness results.(36, 37) In the absence of empirical data describing the utility associated with individuals susceptible to HIV infection, we selected the highest utility reported for individuals infected with HIV as a lower bound for this utility.(25) Choosing the lowest plausible utility value for susceptible individuals biases results against a prevention program. If a utility of 1.0 is used instead, the cumulative net benefit after 30 years increases to 1.1 billion dollars. A second conservative assumption was the choice of $4,133/year as the direct medical cost associated with individuals susceptible for HIV infection. This figure was reported by the Canadian Institute of Health Information as the Canadian average for 2005.(23) However, it was higher than the annual costs associated with HIV infection observed in the highest CD4 cell stratum.(Chapter 3) This discrepancy may be due to younger age or lower health seeking behaviour in the population of interest relative to the Canadian average. Regardless of the reason, the result within the model was that medical costs tended to *decrease* immediately following HIV infection. This tendency will also bias results against a prevention program.

Finally, for all parameters related to primary infection, we assumed values on the upper end of the plausible range. We assumed that the phase would last 60 days, that viral load would remain at 6 log-copies/mL throughout the phase, and that there would be no decrease in risk behaviour. Transmission during the primary phase was assumed to be unaffected by the program to increase uptake of HAART.

6.5. Conclusion

In this study, an intervention to increase the HAART treatment rate from 50% to 75% of HIV-infected individuals with CD4 count below 350 cells/mm^3 in BC was demonstrated to be a cost-effective strategy. This result was obtained under a number of conservative assumptions which were chosen to bias results against the intervention. Due to these assumptions, it is plausible that the actual net benefit associated with the intervention is even higher than that reported here. These cost-effectiveness results are consistent with public health objectives: all individuals who are eligible for an established life-saving treatment should receive it.

Figure 6.1: Schematic of model structure.

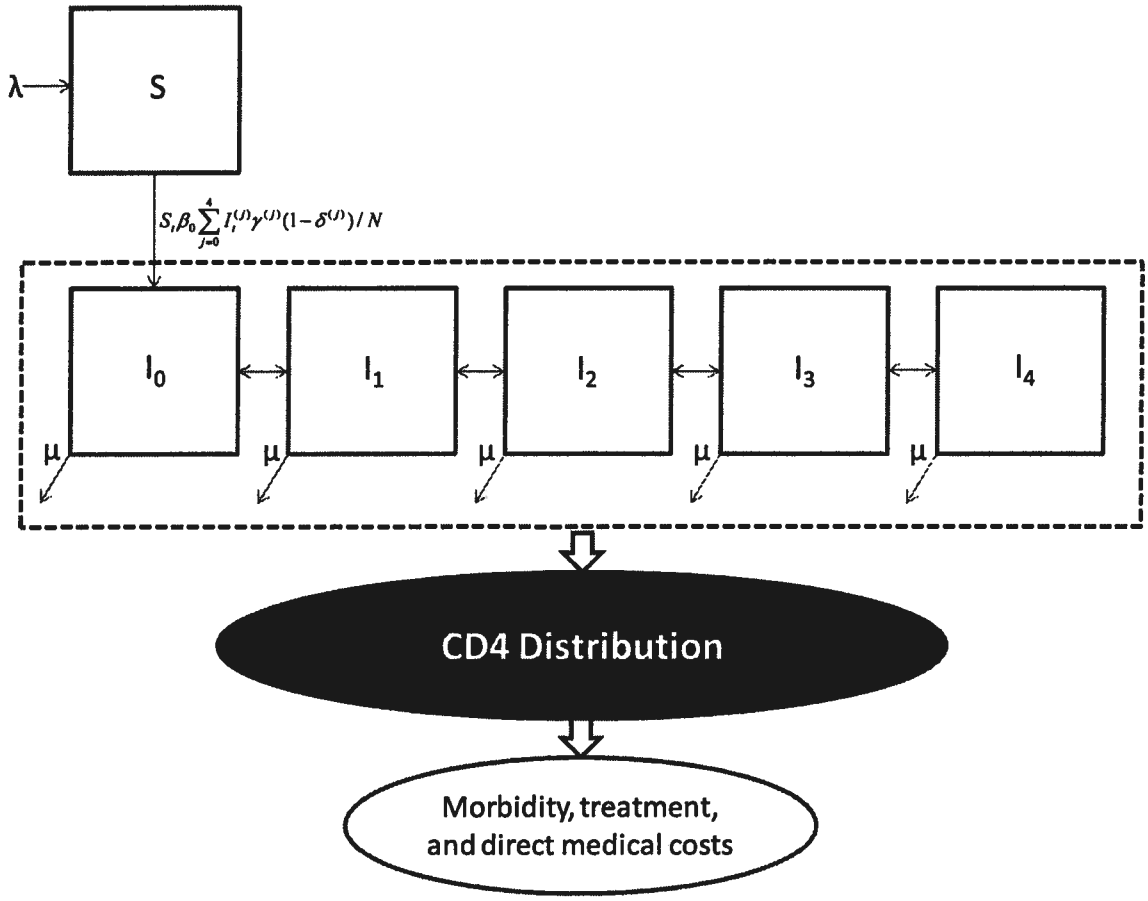

Figure 6.2: Incremental net benefit.
Increasing uptake of highly active antiretroviral therapy (HAART) from 50% to 75% over a simulated time period of 30 years, based on a willingness-to-pay thresholds of $50,000 per quality-adjusted life year. Solid line represents mean of empirical distribution of probabilistic sensitivity analysis results, while dashed lines represent a 95% confidence interval.

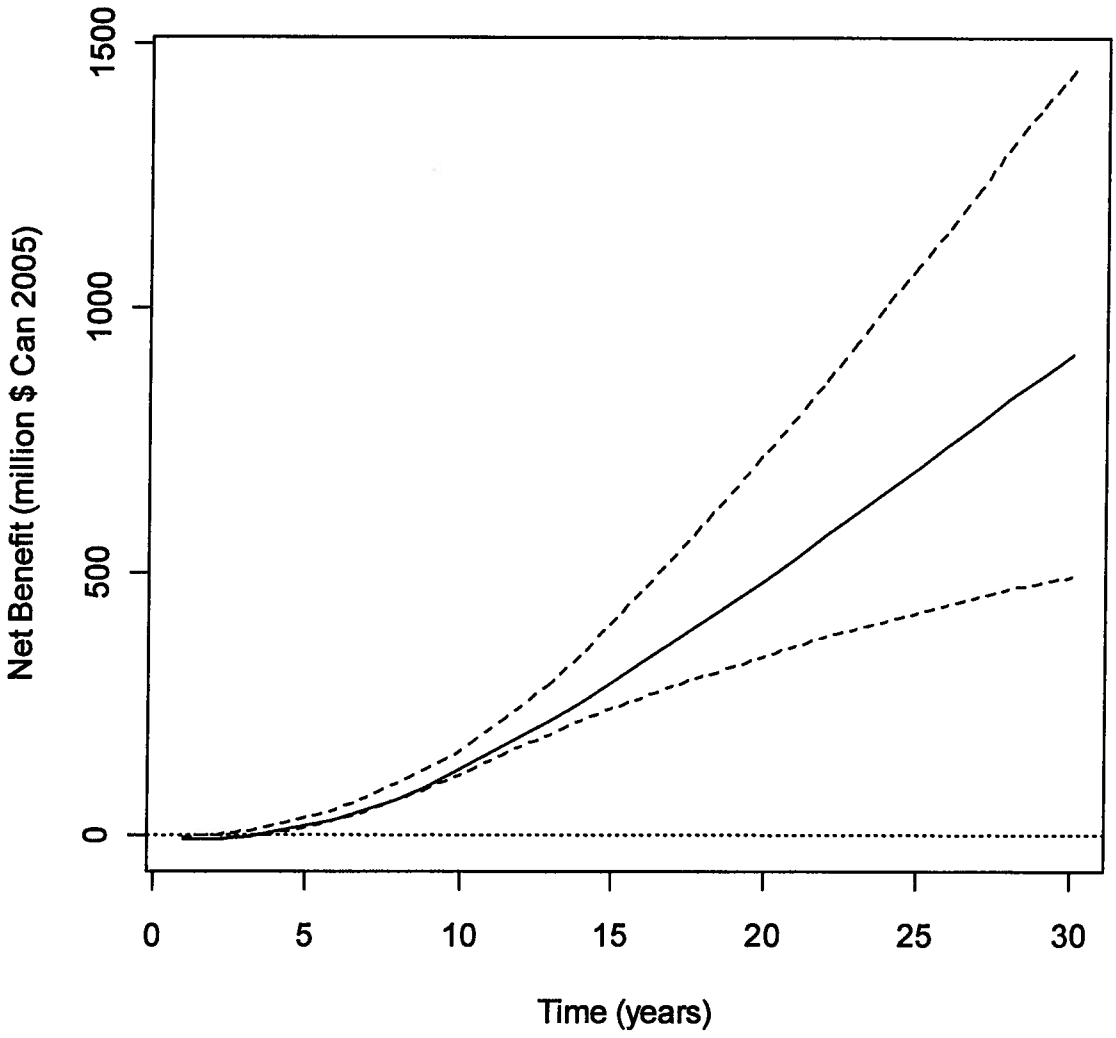

Figure 6.3: Cost-effectiveness acceptability curve.
Increasing uptake of highly active antiretroviral therapy (HAART) from 50% to 75% over a simulated time period of 30 years showing proportion of probabilistic sensitivity analysis iterations that were cost-effective for willingness-to-pay thresholds between 0 and $80,000.

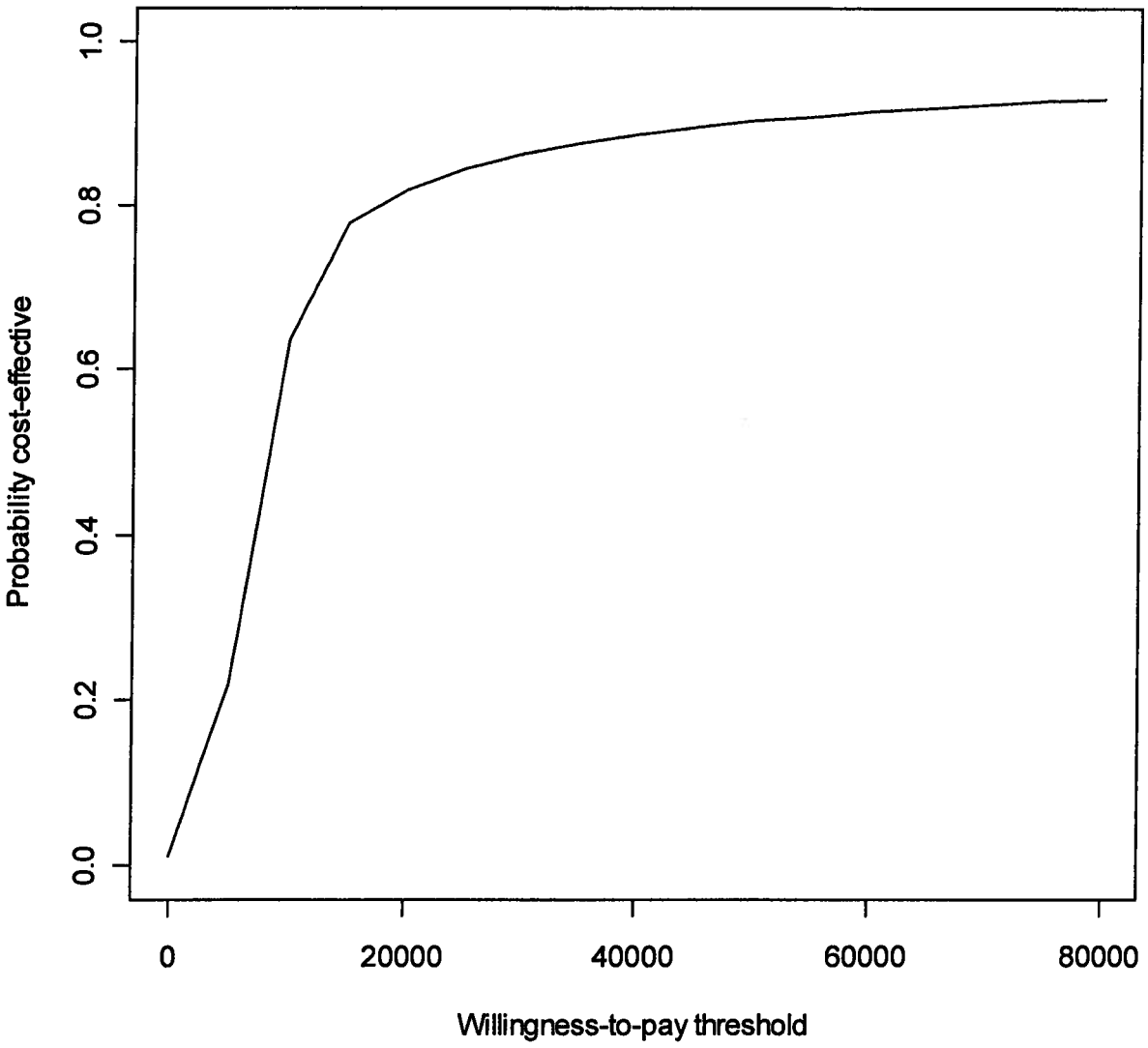

Figure 6.4: Overall population and patient-centered incremental net benefit.
Increasing uptake of highly active antiretroviral therapy (HAART) from 50% to 75% over a simulated time period of 30 years, based on a willingness-to-pay threshold of $50,000 per quality-adjusted life year.

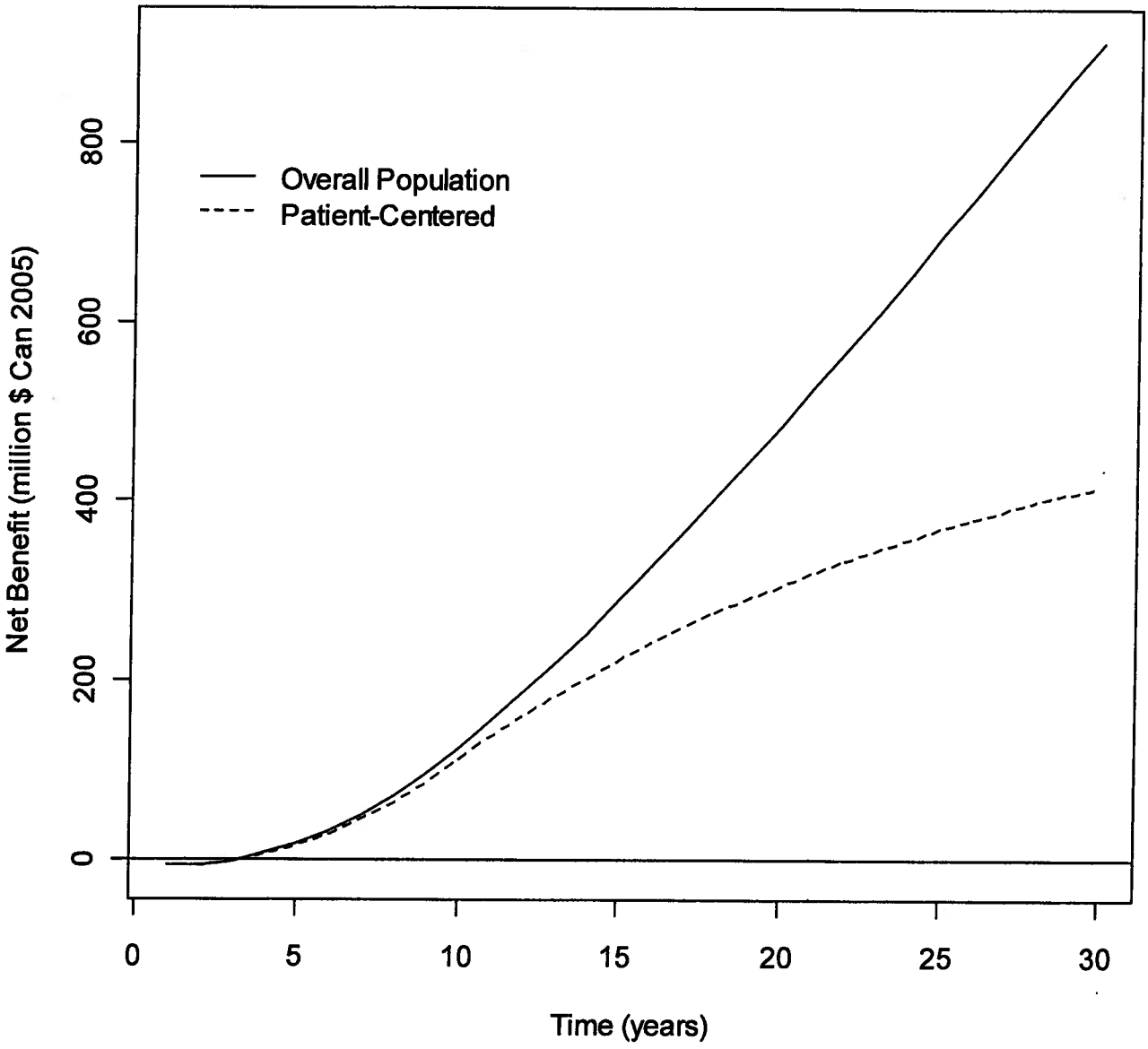

Figure 6.5: Estimated cumulative number of new infections over time.
Scenario1: increasing uptake of highly active antiretroviral therapy (HAART) from 50% to 75%; and Scenario 2: uptake of HAART consistent at 50% over a simulated time period of 30 years. Infections transmitted during primary infection phase identified.

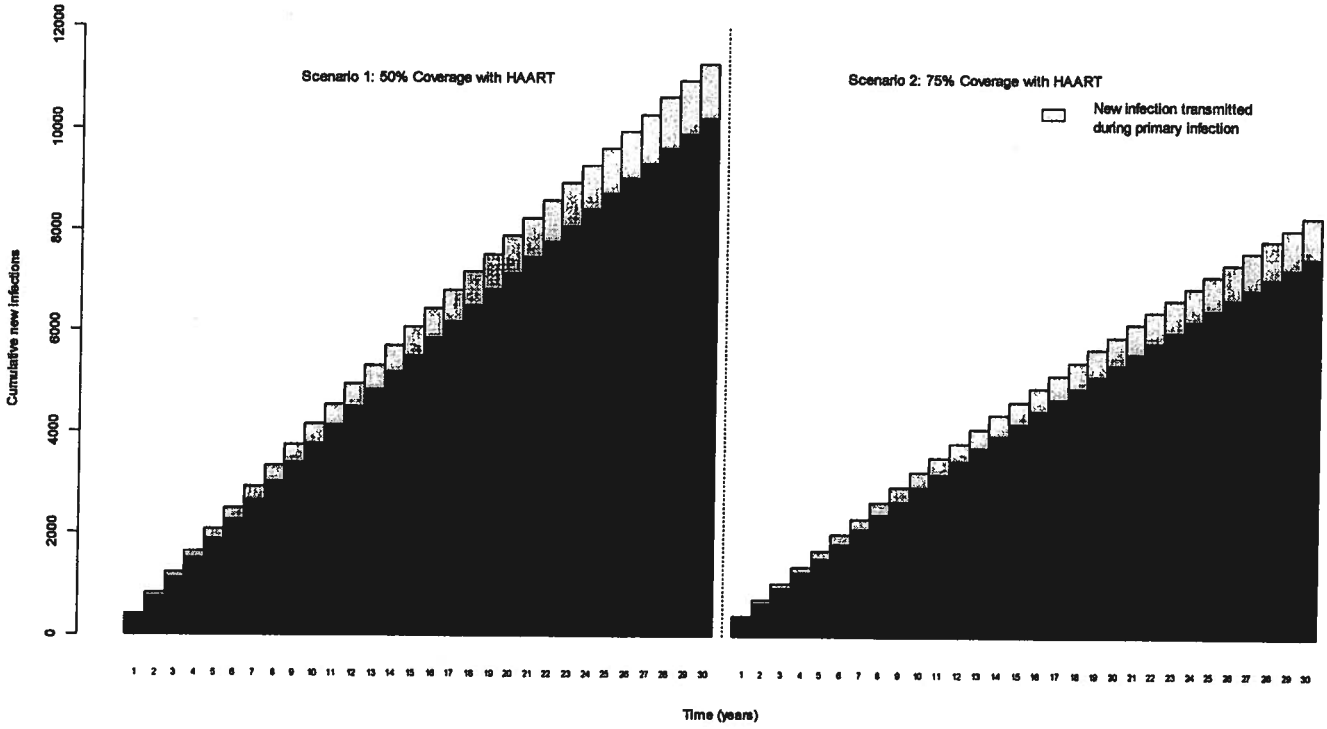

Table 6.1: Statistical methods used to describe processes included in microsimulation.
All models fit using individual-level data from British Columbia.

Process	Statistical methods used
CD4 cell count trajectories	• Non-linear statistical model(17)
Plasma viral load trajectories	• Non-linear statistical model(17)
Medication resistance*	• Weibull time-to-event model with frailty terms due to correlation between resistance to different classes of medication(38)
Accumulation of direct medical costs	• Series of two-stage random effects models(18)
HIV-Related mortality following highly active antiretroviral therapy (HAART) initiation	• Cox proportional hazards model incorporating CD4 cell count as a time-dependent covariate(38)

* Defined as developing a first resistance mutation to each of four categories of antiretrovirals: protease inhibitors, non-nucleoside reverse transcriptase inhibitors, lamivudine, and nucleoside reverse transcriptase inhibitors (excluding lamivudine).

Table 6.2: Parameters associated with transmission model.

Parameter	Description	Point estimate	Distribution used in probabilistic sensitivity analysis	Source(s)
S_0	Susceptible population at baseline	15,000	Normal (σ=2,000)	• Assumption
β_0	Baseline force of infectivity (associated with viral load of 3 log copies/mL)	0.02	Beta (α=15.66, β=767.34)	• Empirical calibration(21) • Beta distribution parameters estimated using method of moments assuming standard deviation of 0.005.
$\gamma^{(j)}$	Increase in infectivity associated with viral load category *j* relative to baseline	2.45 per log increase	Log-normal (σ=0.15)	• Quinn et al.(6)
$\delta^{(j)}$	Decrease in risk behaviour associated with viral load category *j* relative to baseline	0.10 per log increase (excluding primary infection)	Beta (α=0.8, β=7.2)	• Assumption • Beta distribution parameters estimated using method of moments assuming standard deviation of 0.10.
λ	Net annual migration into susceptible population (accounting for mortality)	500	Normal(σ=100)	• Empirical calibration
μ_t	Mortality rate at time *t*.	Dependent on disease stage and age distribution and treatment history of simulated population.	N/A	• Untreated: Babiker et al.(20) • Treated: Cox proportional hazards analysis of British Columbia data. • Non-HIV related: British Columbia age-specific mortality. (39, 40)
Viral load set point	Baseline viral load (assumed to increase by 0.5 log-copies per year)	4 log-copies	Normal(σ=0.5)	• Fraser et al.(19) • British Columbia data

6.6. References

1. UNAIDS. Report on the global AIDS epidemic. Geneva, 2008.
2. Palella FJ, Delaney KM, Moorman AC, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. New England Journal of Medicine 1998;338:853-60.
3. Pomerantz RJ. Residual HIV-1 disease in the era of highly active antiretroviral therapy. New England Journal of Medicine 1999;340:1672-4.
4. Montaner JSG, Hogg R, Wood E, et al. The case for expanding access to highly active antiretroviral therapy to curb the growth of the HIV epidemic. Lancet 2006;368:531-6.
5. Hogg RS, Rhone SA, Yip B, et al. Antiviral effect of double and triple drug combinations amongst HIV-infected adults: lessons from the implementation of viral load-driven antiretroviral therapy. AIDS 1998;12:279-84.
6. Quinn TC, Wawer MJ, Sewankambo N, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1. New England Journal of Medicine 2000;342:921-9.
7. Cohen MS, Gay C, Kashuba ADM, et al. Narrative review: Antiretroviral therapy to prevent the sexual transmission of HIV-1. Annals of Internal Medicine 2007;146:591-U63.
8. Granich R, Gilks C, Dye C, et al. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet 2009;373:48-57.

9. Levy AR, James D, Johnston KM, et al. The direct costs of HIV/AIDS care. Lancet Infectious Diseases 2006;6:171-7.

10. Hogg RS, Yip B, Kully C, et al. Improved survival among HIV-infected patients after initiation of triple-drug antiretroviral regimens. Canadian Medical Association Journal 1999;160:659-65.

11. Richman D. Human Immunodeficiency Virus. International Medical Press, 2003.

12. Hollingsworth T, Anderson R, Fraser C. HIV-1 Transmission, by Stage of Infection. Journal of Infectious Diseases 2008;198:687-93.

13. British Columbia Centre for Disease Control. STI/HIV Prevention and Control Annual Report. 2008.

14. Wilson D. HIV Epidemiology: A Review of Recent Trends and Lessons. Global HIV/AIDS Program The World Bank 2006.

15. Public Health Agency of Canada. HIV/AIDS Epi Updates 2007. Public Health Agency of Canada, 2007.

16. Wood E, Hogg RS, Yip B, et al. Is there a baseline CD4 cell count that precludes a survival response to modern antiretroviral therapy? Aids 2003;17:711-20.

17. Wood S. Generalized Additive Models: An Introduction with R. Chapman & Hall/CRC, 2006.

18. Diehr P, Yanez D, Ash A, et al. Methods for analyzing health care utilization and costs. Annual Review of Public Health 1999;20:125-44.

19. Fraser C, Hollingsworth TD, Chapman R, et al. Variation in HIV-1 set-point viral load: Epidemiological analysis and an evolutionary hypothesis. Proceedings of the National Academy of Sciences of the United States of America 2007;104:17441-6.

20. Babiker A, Darby S, De Angelis D, et al. Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: a collaborative re-analysis. Lancet 2000;355:1131-7.

21. Lima VD, Johnston K, Hogg RS, et al. Expanded access to highly active antiretroviral therapy: A potentially powerful strategy to curb the growth of the HIV epidemic. Journal of Infectious Diseases 2008;198:59-67.

22. Briggs AH. Handling uncertainty in cost-effectiveness models. Pharmacoeconomics 2000;17:479-500.

23. Canadian Institute for Health Information. National Health Expenditure Trends: 1975-2006. Ottawa, 2006.

24. Torrance GW. Measurement of health state utilities for economic appraisal - A review. Journal of Health Economics 1986;5:1-30.

25. Schackman BR, Goldie SJ, Freedberg KA, et al. Comparison of health state utilities using community and patient preference weights derived from a survey of patients with HIV/AIDS. Medical Decision Making 2002;22:27-38.

26. Organization for Economic Development and Co-operation (OECD). Health Data 2007., 2007.

27. Zethraeus N, Johannesson M, Jonsson B, et al. Advantages of using the net-benefit approach for analysing uncertainty in economic evaluation studies. Pharmacoeconomics 2003;21:39-48.

28. Laupacis A, Feeny D, Detsky AS, et al. How attractive does a new technology have to be to warrant adoption and utilization - Tentative guidelines for using clinical and economic evaluations. Canadian Medical Association Journal 1992;146:473-81.

29. Koerkamp BG, Hunink MGM, Stijnen T, et al. Limitations of acceptability curves for presenting uncertainty in cost-effectiveness analysis. Medical Decision Making 2007;27:101-11.

30. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves - facts, fallacies and frequently asked questions. Health Economics 2004;13:405-15.

31. Velasco-Hernandez JX, Gershengorn HB, Blower SM. Could widespread use of combination antiretroviral therapy eradicate HIV epidemics? Lancet Infectious Diseases 2002;2:487-93.

32. Baggaley RF, Garnett GP, Ferguson NM. Modelling the impact of antiretroviral use in resource-poor settings. PLoS Medicine 2006;3:493-504.

33. Salomon JA, Hogan DR. Evaluating the impact of antiretroviral therapy on HIV transmission. AIDS 2008;22:S149-S59.

34. Wasserheit JN, Aral SO. The dynamic topology of sexually transmitted disease epidemics: Implications for prevention strategies. Journal of Infectious Diseases 1996;174:S201-S13.

35. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. Health Economics 2006;15:1295-310.

36. Fryback DG, Lawrence WF. Dollars may not buy as many QALYs as we think: A problem with defining quality-of-life adjustments. Medical Decision Making 1997;17:276-84.

37. King JT, Styn MA, Tsevat J, et al. "Perfect health" versus "disease free": The impact of anchor point choice on the measurement of preferences and the calculation of disease-specific disutilities. Medical Decision Making 2003;23:212-25.

38. Klein JP, Moeschberger ML. Survival analysis: techniques for censored and truncated data. Springer, 2003.

39. BC Stats. Population Estimates — British Columbia. 2004.

40. Statistics Canada. Deaths by single year of age and geography — Both sexes. 2004.

7. DISCUSSION

7.1. Study Results and Implications

7.1.1. Summary of findings

The studies included as Chapters 2-6 in this dissertation describe a number of clinical, economic, and policy-related findings:

- The use of generalized additive models to quantify non-linear trajectories for plasma viral load and CD4 cell count during the highly active antiretroviral therapy (HAART) era (Chapter 2). The non-linear statistical technique was shown to be an effective method for incorporating individual-level variability in trajectories as well as describing broad population trends.
- The development of alternative methods for statistical modelling of health services utilization and costs associated with HIV during the HAART era (Chapters 3 and 4). Key findings from these chapters include: a consistent inverse trend observed between CD4 cell count and direct medical costs; an inverse relationship between expenditure on HAART and expenditure on hospitalizations; and pairwise correlations between CD4 cell count trajectories, hospitalization incidence, and hospitalization costs. In addition, Chapter 3 describes statistical models that were used to generate estimates of monthly costs and associated variability for incorporation into a microsimulation model of HIV.
- The incorporation of the statistical models from Chapters 2 and 3, respectively, into a microsimulation model for generating stochastic estimates of the lifetime direct medical costs associated with HIV in British Columbia (BC) during the HAART era (Chapter 5). The base-case estimate of lifetime direct medical costs was approximately $268,000 (95% Confidence Interval [CI]: $258,000-$278,000) (2005 CDN$). The inclusion of individual-level heterogeneity was found to be important; when individualized CD4 cell

counts and random effects were removed, lifetime costs were estimated to be systematically lower.

- The integration of the microsimulation model with a disease transmission model, in order to evaluate the economic implications of expanding access to HAART in BC from 50% to 75% coverage of individuals with CD4 cell counts below 350 cells/mm^3, accounting for both the individual-level therapeutic effects and the population-level transmission effects of HAART (Chapter 6). Under a number of conservative assumptions, the HAART expansion strategy was found to be cost-effective at a standard willingness-to-pay-threshold of $50,000 per quality-adjusted-life-year. For a conservative threshold of $20,000 per quality-adjusted-life-year, the cost-effectiveness acceptability curve indicated a probability of cost-effectiveness greater than 80%.

7.1.2. Cost effectiveness of expanded access to HAART

The results of Chapter 6, which illustrate the cost-effectiveness of expanding access to HAART in BC, represent a culmination of the models described in Chapters 2, 3, and 5, combined with a set of calibrated difference equations introduced to incorporate disease transmission. The existing models incorporated into the cost-effectiveness analysis included population-based statistical models describing CD4 and viral load trajectories, health services utilization and expenditure, medication resistance, and survival, as well as a microsimulation model which combined all these elements to generate estimates of lifetime costs. Thus, the cost-effectiveness of a strategy to expand access to HAART in BC from 50% to 75% of individuals with CD4 cell count below 350 cells/mm^3 represents the key finding of the dissertation. The results of the cost-effectiveness analysis were based on all available clinical and economic data, and should therefore provide reliable accuracy within the province of BC.

A key element in determining cost-effectiveness of HAART expansion is the assumed impact of treatment on transmission of new infections, as disease prevention represents a key area of cost savings. The transmission component of the model described here generated results that were similar to a published transmission model for BC.(1) This is to be expected, since the two transmission models were based on similar assumptions and input data, with minimal differences in methodology. Both BC models were consistent with a recently published model describing the effects of early testing and immediate treatment with HAART on HIV transmission in sub-Saharan Africa.(2) While the model of sub-Saharan Africa predicted more dramatic decreases in transmission of new infections, it was also based on stronger assumptions regarding treatment coverage and CD4 count threshold for initiating therapy.

7.1.3. International variability in HIV direct medical costs

A motivating reason for generating BC-specific estimates of lifetime direct medical costs given the availability of existing estimates(3, 4) was the potential for international variability in this parameter. International variability across unit costs for elements of direct medical expenditure has been demonstrated,(5) and this variability will be compounded when accruing costs over the course of a lifetime. Consistent with this hypothesis, the estimate of lifetime direct medical costs in BC was lower than corresponding estimates for the United States (US)(3) and France.(4) When the three cost estimates were converted to a common currency and year (2005 $Canadian), the BC estimate of $268,000 is 14% (95% CI: 11%-17%) lower than the French estimate and 45% (95% CI: 43%-47%) lower than the US estimate.

There are at least two potential explanations for these cost discrepancies. The first is the difference in monthly antiretroviral costs. The costs for the same medications can vary across countries, with the US typically paying the highest prices.(6) An international comparison of monthly direct medical costs in HIV found that France and the US reported the highest monthly antiretroviral costs (relative to Canada, Italy, and England).(5) Schackman et al. reported the monthly antiretroviral costs assumed within the US lifetime cost model; when converted to 2005 $CDN, they ranged between $1,440 and $4,760. This is substantially higher than the range assumed within the BC model described here ($922-$1,130). This difference does not necessarily indicate a bias in either model—the US costs were based on reported wholesale prices paid by Medicaid while the BC costs were based on analysis of population-based dispensing data—but rather is indicative of actual cost differences between the US and BC, highlighting the need for jurisdiction-specific estimates. The second potential reason for the discrepancy between lifetime cost estimates for France and BC is the fact that the estimate for France was based on survival estimates from the post-HAART era, but health services utilization and cost estimates calculated between 1994 and 1998. This includes several years during which HAART was not readily available, and, as a result, cost estimates for hospitalizations may have been biased upwards due to increased hospitalization rates prior to HAART availability.

7.1.4. Importance of individual-level variability

A recurring theme throughout the studies described in this dissertation was the importance of capturing individual-level heterogeneity in clinical and economic variables. This heterogeneity was incorporated in multiple ways, targeted to the method of analysis within the individual studies. In Chapter 2, substantial individual-level variability in CD4 and viral load was demonstrated. This variability was captured within the statistical models by using a transition

structure(7) in which previous measurements were used as a covariate for predicting future measurements. The effectiveness of this approach was demonstrated by comparing actual CD4 and viral load trajectories to predicted trajectories (see Figures 2.5 and 2.6). Of particular note was the fact that predicted trajectories were able to empirically capture short-term fluctuations that could not be predicted given current understanding of the biology of HIV.

In Chapter 3, random effects were used to capture individual-level variability in tendency to utilize various categories of health services, after adjusting for demographics and CD4 cell count. These random effects described individual tendencies in health and health-seeking behaviour that could not be explained by measured clinical or demographic variables, and were used to quantify the pairwise correlations between categories of health services utilization and cost. In particular, the random effects associated with HAART and hospitalizations, respectively, were used to estimate the relationship between expenditure on HAART and expected expenditure on hospitalizations. In Chapter 4, which focussed on trajectories of hospitalization incidence and costs given time-varying CD4 cell count, random effects were used similarly. The empirical correlations between hospitalization incidence, hospitalization cost, and CD4 cell count trajectories were incorporated into cumulative incidence and cost estimators. This allowed the cumulative estimators to account for the tendency of individuals who experienced lower-than-average CD4 cell counts to also have high hospitalization incidence and cost, even after adjusting for CD4 counts, and overall incidence and cost could be estimated accordingly.

The microsimulation in Chapter 5 was based on the statistical models of Chapters 2 and 3, and, as such, made use of the methods described therein for incorporating individual-level variability.

The CD4 and viral load models were used to generate individualized trajectories of these variables, with most recent measurements used to predict current measurements. The residual standard deviation of the models was used to further incorporate random error that was not described by the models. The random effects estimated in Chapter 3 were also incorporated into the microsimulation. Each simulated individual was assigned a separate random effect for each category of health services utilization and cost. These values were randomly drawn from the empirical distributions observed when fitting the models, accounting for correlations between random effects. The inclusion of these random effects in the microsimulation enabled us to allow individuals to accrue costs that were systematically above or below the population average after adjusting for measured covariates, and for costs to cluster together across categories of health services. It was found that if the above-described sources of heterogeneity were not included in the microsimulation, lifetime cost estimates were biased towards lower values.

7.2. Study Strengths

An important strength of the studies described in this dissertation was the high-quality population-level data that were available for producing estimates. The three main data sources utilized were the HAART Observational Medical Evaluation and Research (HOMER) cohort, the Community Health and Safety Evaluation (CHASE) cohort, and the BC Linked Health Database (BCLHD), each of which brought specific strengths to individual studies. The data sources are described briefly below and in detail in Chapter 3.

The HOMER cohort is comprised of all antiretroviral-naïve individuals in BC who initiated treatment with HAART after 1996. Available data included baseline cross-sectional variables describing age, sex, and injection drug use status; longitudinal results of laboratory tests (CD4

cell count, plasma viral load, medication resistance); and longitudinal antiretroviral dispensing records. Access to the HOMER cohort provided a rich source of CD4 cell count and viral load data which allowed for fitting the non-linear models in Chapter 2. Another highlight of the HOMER data was the fact that antiretroviral dispensing data allowed for monthly HAART costs to be observed on actual observed medication adherence, as opposed to hypothetical pricing of specific regimens, which requires explicit assumptions regarding medication adherence. A subset of individuals from HOMER was linked to the BCLHD (described in further detail below), which contains comprehensive health services utilization records.

The CHASE cohort is a representative sample of low-income individuals residing in Vancouver's downtown east side. A linkage to the BC Centre for Disease Control HIV testing database allowed for the identification of individuals with a confirmed HIV diagnosis. This HIV-positive cohort subset was then linked to all BC antiretroviral treatment records, such that individuals with positive HIV serostatus who had never accessed treatment with HAART could be identified. Such individuals would not likely be recorded in routinely collected administrative health databases whose records require contact with a specialized treatment centre. While the CHASE data source did not provide access to clinical information such as CD4 cell count or viral load, it did provide a linkage to the BCLHD for all cohort members. As such, we were able to capture the utilization patterns that were reflective of the actual community of individuals living with HIV, rather than the specialized subset of individuals who are receiving optimal treatment for their disease.

In BC, the universal health insurance scheme requires individuals to have a unique identifier, allowing individual-level linkages between hospitalizations, physician visits, and

pharmaceuticals. The BCLHD contains comprehensive health services utilization records,(8) and provides information on utilization patterns for all individuals in BC. The linkages between HOMER, CHASE, and the BCLHD resulted in a rich data source describing sociodemographics, clinical progression, and health resource utilization for a wide cross-section of individuals infected with HIV.

A further strength of the dissertation was the range of analytical methods that were applied across studies. As a result of the high-quality individual-level data that were available for these studies, specialized statistical methods could be specifically tailored to the objectives of interest within each study.

The decision and mathematical models were built and validated based on suggested best practices for modelling studies.(9) In particular: results of the lifetime cost model were internally validated against observed costing data; results of the mathematical model were calibrated against observed incidence and treatment data to aid in parameter assignment; and a cross-validation exercise was performed to explain differences in results between the lifetime cost model and similar models created for other countries. The incorporation of statistical models was also performed in compliance with suggested guidelines, with models based on defensible biostatistical and epidemiological techniques and described in detail.(9)

7.3. Study Limitations

7.3.1. Data limitations

There were also limitations associated with the data sources utilized here. One such limitation was the observational nature of the CD4 cell count and plasma viral load data. These

measurements were not collected as part of a trial protocol, but rather were recorded as individuals presented for routine clinical monitoring and management. While the interquartile ranges of intermeasurement times, given in Chapter 2, suggested relatively little variablility, with the majority of intermeasurement times between one and four months, there was still potential for some individuals to provide measurements either more or less often than this standard frequency. This may have resulted in a phenomenon similar to selection bias:(10) although all individuals who initiated treatment with HAART were included in the sample, those individuals who contributed the most measurements were most influential on model fit, as they contributed a relatively large proportion of the data.(11) If individuals who were sickest—and therefore expected to have low CD4 cell counts and high viral loads—tended to provide more measurements due to more frequent monitoring of severe disease, then model-based estimates would be biased towards low CD4 cell counts and high viral loads. Conversely, if frequent monitoring were associated with high levels of health-seeking behaviour, it would be expected that individuals who provided the most measurements were also maintaining their health in other ways, and would therefore tend to have high CD4 cell counts and low viral load, and model-based estimates would be biased in this direction. In practice, it is possible that both phenomena were occurring simultaneously, and it is difficult to ascertain the net impact on model results. However, the predictive ability of the models, as displayed in Figures 2.5 and 2.6 suggested that whatever biases were present did not have a material impact on model fit.

A second data-related limitation was the fact that all longitudinal data sources were specific to BC, and did not contain information describing migration out-of-province, creating the potential for information bias with respect to misspecification of outcome measures.(10) If an individual moved outside of BC, their health services utilization would be incorrectly recorded as zero from

that point forward, biasing monthly cost estimates downwards. This bias would not continue throughout the entire follow-up period, as such individuals would not be contributing CD4 counts if they were living out-of-province, and the monthly cost models adjusted for time periods during which a CD4 measurement had not been recorded for one year. We made the assumption that a relatively small proportion of individuals moved out of province,(12) and, further, that such individuals did not have a systematic tendency to incur higher-than-average direct medical costs. If that had been the case, the bias to overall cost estimates would have been more substantial, as the most "expensive" individuals would be incorrectly assigned to incur zero medical costs.

A third limitation to the data source was the fact that the linkage between HOMER and the BCLHD was only current to March 2001. While this provided approximately five years follow-up during the post-HAART era, it is possible that health services utilization patterns have changed since this time due to advances in HAART regimens.(13) The CHASE-BCLHD linkage extended to December 2005, but of the 1,895 individuals included in the health services utilization analyses described in Chapter 3, just 11% were included in the CHASE cohort. The remaining 89% were only members of the HOMER cohort and provided BCLHD follow-up until 2001 only.

A fourth potential limitation was that, as with any analysis of observational data, there is the potential for statistical models to be misspecified and for confounding variables to have been left out of the analysis. If this were the case, some of the exposure/outcome variable relationships estimated within the statistical models may be spurious, due to the presence of one or more unmeasured variable(s) related to both exposure and outcome that induced an observed

relationship between the two.(10) However, in this dissertation, the primary function of the statistical models was to serve as predictive tools to provide input parameters to the decision and mathematical models, which lessened the potential impact of confounding. For example, in Chapter 3, positive relationships between female sex and health care utilization and cost were consistently found. While there are some plausible biological reasons for these relationships (notably, pregnancy and childbirth), they are not likely to explain the entire observed relationship.(14) An alternative hypothesis is that health-seeking behaviour was a confounding variable, such that women were more likely to exhibit greater health-seeking behaviour and individuals with greater health-seeking behaviour were more likely to utilize more health care services. If the objective of a study were to discern the causal relationship between sex and health care utilization, this would represent a major potential limitation. However, from the perspective of predicting long-term costs, it is acceptable to assume that the observed relationship is correct. When projecting lifetime costs, it was not necessary to understand the nuances of *why* sex is related to utilization, only to be confident that a consistent relationship was observed in the data.

7.4. Future Research

One potential area for future research would be to simply repeat the analyses and models described in this dissertation with updated data. In particular, an updated linkage to the BCLHD would provide a more accurate description of health resource utilization during the modern HAART era. Further, as more time accrues following the introduction of HAART, analysis of survival time will be more accurate, allowing for greater confidence in long-term predictions.

A second area for future research would be to revisit the statistical models described in Chapters 2 and 3 from the perspective of generating stand-alone epidemiologic studies. The current versions of these studies were designed with the primary objective of providing required inputs to the microsimulation and transmission models. Specific extensions that can be made would be to perform a more rigorous evaluation of the generalized additive modelling technique described in Chapter 2, specifically in comparison to other predictive models that have been applied to CD4 cell count data, such as quadratic random effects models.(15) In Chapter 3, the correlations between different categories of health services utilization were primarily derived for use as inputs into the microsimulation model. The underlying reasons for particular correlations and resulting implications were not explored in detail. A follow-up study could use the same data and analytical methods to evaluate potential theoretical models of health and health-seeking behaviour that could explain the observed relationships.

In addition, further data collection could inform refinements to model design and parameter estimation. The majority of BC-specific parameters within the transmission model were estimated via a process of empirical calibration, in which parameters were adjusted until model-predicted estimates of new infections and number of individuals initiating treatment matched observed numbers. Alternatively, a data collection study could be performed to empirically confirm assumptions regarding average level of risk behaviour and its relationship with disease severity, the size of the population at high risk for contracting HIV, and annual migration in and out of the HIV-infected and high risk populations.

An additional study requiring the collection of primary data would be to more accurately estimate dynamic treatment patterns following initiation of HAART, and assess the implications

of treatment interruption on morbidity, independent of CD4 cell count. The observational administrative data sources were not sufficiently detailed to address this issue. Medication dispensing records could not accurately describe time-dependent medication adherence. For example, if an individual was only dispensed 75% of a yearly supply of medication, it was not possible to discern whether they routinely took 75% of their medications throughout the year, or if they maintained 100% adherence for 75% of the year and experienced a treatment interruption for the remainder. If treatment interruption is to be evaluated, more detailed records are required. A study designed to address this question could sample a subset of individuals infected with HIV who could fill out detailed daily treatment logs. This time-dependent treatment information could be either linked to the BCLHD as before, or to self-reported data describing health services utilization and morbidity.

7.5. Contributions to Original Knowledge

7.5.1. Methodological contributions and implications

A novel methodology was presented for incorporating a time-dependent covariate into a recurrent events model with a marked point process (Chapter 4). In this particular example, CD4 was the time-dependent covariate, hospitalizations represented the recurrent event, and the marked point process associated with each event was the cost of the hospitalization. However, the methodology could also be applied in alternative contexts. One such extension would be to estimate cumulative hospitalization costs in a different disease area with an analogous clinical marker that varies over time.

The primary methodological contribution of this dissertation was the integration of a detailed microsimulation model with a population disease transmission model, resulting in a single

comprehensive model that could be used to evaluate both individual-level and population-level policies (Chapter 6). The microsimulation model described in Chapter 5 is an efficient method for describing the therapeutic effects of HAART at the individual level, while the compartmental model described in Chapter 6 represents an established method for describing disease transmission at the population level. It would have been possible to incorporate disease transmission directly into the microsimulation, but this would have required describing individual-level behaviour in greater detail than that required to satisfy the objectives of the study. Conversely, the disease progression and economic processes considered within the microsimulation could have potentially been incorporated directly into a mathematical model of disease transmission, but it is not clear how the statistical models which populated model parameters could have been seamlessly incorporated into a mathematical model as they were into the microsimulation. By instead combining the two forms of modelling into an integrated model, the cost-effectiveness of HAART expansion could be addressed in a more efficient manner, with each element of the process described by the most appropriate modelling technique.

Previous economic models incorporating HIV transmission have tended to primarily focus on either the chronic nature or the infectious nature of HIV, with the remaining processes addressed in a simplified manner. To date, when mathematical models have been used to describe HIV transmission, simplifying assumptions have been made regarding direct medical costs, with homogeneous costs based on population averages assumed to apply uniformly to all infected individuals.(2, 16) Conversely, models which have utilized a more complex decision model structure to describe the clinical course of HIV and related economic implications have been based on simplifying assumptions regarding the dynamic nature of HIV transmission.(17, 18)

To our knowledge, the integrated model described here is the first economic model of HIV transmission to incorporate the complexity of both processes.

7.5.2. Policy contributions and implications

The results of this dissertation provide two contributions to health policy decisions in BC. Firstly, the lifetime direct medical cost estimates represent a formal combination of all available clinical and health services utilization data, and can be used to plan for current and future allocation of health resources in BC (Chapter 5). Secondly, increased treatment with HAART was shown to be associated with reduced expenditure for other health services (Chapters 3 and 6). This reduced expenditure is due to a reduced need for acute care services such as hospitalizations in individuals being optimally treated with HAART (Chapter 3) as well as a decrease in new infections due to reduced infectivity in individuals experiencing viral load suppression after HAART initiation (Chapter 6). These findings provide both empirical and theoretical support for the proposition(19) that, in addition to providing public health benefits, expanding treatment with HAART for individuals infected with HIV in BC is attractive for pragmatic economic reasons as well.

7.6. References

1. Lima VD, Johnston K, Hogg RS, et al. Expanded access to highly active antiretroviral therapy: A potentially powerful strategy to curb the growth of the HIV epidemic. Journal of Infectious Diseases 2008;198:59-67.
2. Granich R, Gilks C, Dye C, et al. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet 2009;373:48-57.
3. Schackman BR, Gebo KA, Walensky RP, et al. The lifetime cost of current human immunodeficiency virus care in the United States. Medical Care 2006;44:990-7.
4. Yazdanpanah Y, Goldie SJ, Losina E, et al. Lifetime cost of HIV care in France during the era of highly active antiretroviral therapy. Antiviral Therapy 2002;7:257-66.
5. Levy AR, James D, Johnston KM, et al. The direct costs of HIV/AIDS care. Lancet Infectious Diseases 2006;6:171-7.
6. Wagner JL, McCarthy E. International differences in drug prices. Annual Review of Public Health 2004;25:475-95.
7. Diggle P, Heagerty P, Liang, K, Zeger S. Analysis of longitudinal data. Oxford University Press, 2002.
8. Chamberlayne R, Green B, Barer ML, et al. Creating a population-based linked health database: A new resource for health services research. Canadian Journal of Public Health-Revue Canadienne De Sante Publique 1998;89:270-3.
9. Weinstein MC, O'Brien B, Hornberger J, et al. Principles of Good Practice for Decision Analytic Modeling in Health-Care Evaluation: Report of the ISPOR Task Force on Good Research Practices—Modeling Studies. Value in Health 2003; 6:9-17.

10. Szklo M, Nieto F. Epidemiology - Beyond the basics. Jones and Bartlett, 2007.

11. Touloumi G, Pocock SJ, Babiker AG, et al. Impact of missing data due to selective dropouts in cohort studies and clinical trials. Epidemiology 2002;13:347-55.

12. Rachlis BS, Hogg RS, Wood E, et al. Factors associated with geographic migration among a cohort of injection drug users. Health & Place 2008;14:536-43.

13. Phillips AN, Lepri AC, Lampe F, et al. When should antiretroviral therapy be started for HIV infection? Interpreting the evidence from observational studies. AIDS 2003;17:1863-9.

14. Muller C. Health Care and Gender. New York: Russell Sage Foundation, 1990.

15. Taylor JMG, Law N. Does the covariance structure matter in longitudinal modelling for the prediction of future CD4 counts? Statistics in Medicine 1998;17:2381-94.

16. Pinkerton S, Holtgrave D. A method for evaluating the economic efficiency of HIV behavioral risk reduction interventions. AIDS and Behavior 1998;2:189-201.

17. Sanders GD, Bayoumi AM, Sundaram V, et al. Cost-effectiveness of screening for HIV in the era of highly active antiretroviral therapy. New England Journal of Medicine 2005;352:570-85.

18. Paltiel AD, Walensky RP, Schackman BR, et al. Expanded HIV screening in the United States: Effect on clinical outcomes, HIV transmission, and costs. Annals of Internal Medicine 2006;145:797-806.

19. Montaner JSG, Hogg R, Wood E, et al. The case for expanding access to highly active antiretroviral therapy to curb the growth of the HIV epidemic. Lancet 2006;368:531-6.

APPENDIX A: DERIVATION OF CONDITIONAL DISTRIBUTION OF HAART COSTS

Let $I^{(1)}$ be an indicator variable denoting non-zero inpatient utilization in a particular month, $I^{(2)}$ be a random variable describing inpatient costs during a month with non-zero utilization, H be a random variable describing monthly HAART costs, and $\boldsymbol{x}_{ij}$ be a vector of covariates for individual i at time point j.

For individual i at time point j, the two-stage random effects model for hospitalization utilization and log-transformed costs can be written:

$$\text{logit}\left[\text{P}\left(I_{ij}^{(1)}=1 \mid X_{ij}=x_{ij}, U_i=u_i\right)\right]=\boldsymbol{x}_{ij}'\beta^{(1)}+u_i$$
$$\text{E}\left[\log\left(I_{ij}^{(2)} \mid X_{ij}=x_{ij}, V_i=v_i\right)\right]=\boldsymbol{x}_{ij}'\beta^{(2)}+v_i$$

where $U_i \sim N(0,\sigma_U^2)$, and $V_i \sim N(0,\sigma_V^2)$. The model for monthly HAART costs can be written:

$$\text{E}\left[H_{ij} \mid X_{ij}=x_{ij}, W_i=w_i\right]=\boldsymbol{x}_{ij}'\beta^{(3)}+w_i$$

where $W_i \sim N(0,\sigma_W^2)$. The pairwise correlations between U, V, and W can be estimated empirically using the values observed in Table 3.6. If we assume further that the variables follow a multivariate normal distribution, then, for a given value W=w, the conditional distribution of U and V is bivariate normal with the following moments:(36)

$$\boldsymbol{\mu} = \begin{pmatrix} \rho_{UW}\dfrac{\sigma_U}{\sigma_W}w \\ \rho_{VW}\dfrac{\sigma_V}{\sigma_W}w \end{pmatrix} \quad \Sigma = \begin{bmatrix} \sigma_U^2\left(1-\rho_{UW}^2\right) & \sigma_U\sigma_V\left(\rho_{UV}-\rho_{UW}\rho_{VW}\right) \\ \sigma_U\sigma_V\left(\rho_{UV}-\rho_{UW}\rho_{VW}\right) & \sigma_V^2\left(1-\rho_{VW}^2\right) \end{bmatrix}$$

Equation (3.1) implies a non-linear relationship between U, V, and $C^{(l)}$, such that a theoretical joint distribution for $U,V|W$ cannot be used to derive a theoretical distribution for $C^{(l)}|W$. We therefore simulated 10,000 realizations of $U,V|W$ from their known bivariate distribution, and used these values to obtain an approximate empirical confidence interval for $C^{(l)}|W$.

APPENDIX B: DERIVATION OF THE CUMULATIVE COST FUNCTION

Let $X^{(i)}(t)$ denote the history of covariate $X^{(i)}$ for individual i at time t. Assume that $X^{(i)}$ changes at discrete time points, with $X^{(i)}(t) = \left(Z_1^{(i)}, Z_2^{(i)}, \ldots, Z_{m^{(i)}}^{(i)}\right)$ over time intervals

$$\left\{[t_0, t_1), [t_1, t_2), \ldots, [t_{m^{(i)}-1}, t_{m^{(i)}} = t)\right\},$$

where Z_k is the covariate value during interval $[t_{k-1}, t_k)$. Let $C_{kj}^{(i)}$ denote the cost associated with the j^{th} event for individual i during time interval $[t_{k-1}, t_k)$, and $C^{(i)}(t)$ the cumulative cost at time t. Let $N^{(i)}(t)$ denote the cumulative number of events at time t, with $n_k^{(i)}$ events occurring during interval $[t_{k-1}, t_k)$, so that $C^{(i)}(t)$ can be written

$$\sum_{k=1}^{m^{(i)}} \sum_{j=1}^{n_k^{(i)}} C_{kj}^{(i)}$$

Assume that $Z_k^{(i)} \mid t \sim N\left(\mu_k^{(i)}(t), \sigma_Z^2\right)$ with $\mu_k^{(i)}(t) = \left(\beta_0 + b_0^{(i)}\right) + \beta_1 t + \beta_2 t^2$ and $b_0^{(i)} \sim \mathrm{N}\left(0, \sigma_b^2\right)$.

Assume that the cost of an event is described by the linear random intercept model

$$\log\left(C_k^{(i)} \mid Z_k^{(i)}\right) = \left(\alpha_0 + a_0^{(i)}\right) + \alpha_1 Z_k^{(i)} + e_k^{(i)}$$

with $a_0^{(i)} \sim \mathrm{N}\left(0, \sigma_a^2\right)$, and $e_k^{(i)} \sim \mathrm{N}(0, \sigma_C^2)$. Let S_C refer to the smearing factor, defined to be the mean of exponentiated residuals, which can be used to transform expectations from the log-scale back to the original cost scale,(26) i.e.:

$$\mathrm{E}\left[C_k^{(i)} \mid Z_k^{(i)}\right] = S_C \times \mathrm{E}\left[\log\left(C_k^{(i)} \mid Z_k^{(i)}\right)\right]$$

Assume that, for a frailty term U, $N(t)$ is described by a semi-parametric recurrent-events model with mean function:

$$\rho(t \mid X(t),U) = \int_0^t U \exp(Z(s)\gamma) d\rho_0(s)$$

Consider the covariance matrix of random intercepts a_0, b_0, and frailty terms U:

$$\Sigma = \begin{bmatrix} \sigma_a^2 & \sigma_{ab} & \sigma_{aU} \\ & \sigma_b^2 & \sigma_{bU} \\ & & \sigma_U^2 \end{bmatrix}$$

Under the assumption that a_0, b_0, and U are distributed multivariate normally, for a given frailty term U, the conditional expectation of a_0 and b_0 is:(27)

$$\begin{pmatrix} \mu_{a_0} \mid U \\ \mu_{b_0} \mid U \end{pmatrix} = \frac{(U-1)}{\sigma_U^2} \begin{pmatrix} \sigma_{aU} \\ \sigma_{bU} \end{pmatrix}$$

The cumulative cost function for individual i with frailty term U can be derived as follows:

$$\begin{aligned}
\mathrm{E}\left[C^{(i)}(t) \mid U\right] &= \mathrm{E}_{X^{(i)}}\left[\mathrm{E}\left(C^{(i)}(t) \mid X^{(i)}(t),U\right)\right] \\
&= \mathrm{E}_{Z^{(i)}}\left[\mathrm{E}\left(\sum_{k=1}^{m^{(i)}} \sum_{j=1}^{n_k^{(i)}} C_{kj}^{(i)} \mid Z_k^{(i)}\right) \mid U\right] \\
&= \mathrm{E}_{Z^{(i)}}\left[\sum_{k=1}^{m^{(i)}} \mathrm{E}\left(C_{kj}^{(i)} \mid Z_k^{(i)}\right) \mathrm{E}\left(N_k^{(i)} \mid Z_k^{(i)}\right) \mid U\right] \\
&= \mathrm{E}_{Z^{(i)}}\left[\sum_{k=1}^{m^{(i)}} S_C \exp\left(\left(\alpha_0 + a_0^{(i)}\right) + \alpha_1 Z_k^{(i)}\right)\left(\rho_0(t_k) - \rho_0(t_{k-1})\right) U \exp\left(Z_k^{(i)}\gamma\right) \mid U\right] \\
&= U \sum_{k=1}^{m^{(i)}} \left(\rho_0(t_k) - \rho_0(t_{k-1})\right) \mathrm{E}_{Z^{(i)}}\left[\exp\left(\alpha_0 + a_0^{(i)} + (\alpha_1 + \gamma) Z_k^{(i)}\right) \mid U\right] \\
&= U \exp\left(\alpha_0 + a_0^{(i)} \mid U + \sigma_Z^2 (\alpha_1 + \gamma)^2 / 2\right) \sum_{k=1}^{m^{(i)}} \exp\left(\mu_k^{(i)}(t \mid U)(\alpha_1 + \gamma)\right)\left(\rho_0(t_k) - \rho_0(t_{k-1})\right)
\end{aligned}$$

The final expectation is derived from the fact that $\exp\left((\alpha_1 + \gamma) Z_k^{(i)}\right)$ is a moment generating function and $Z_k^{(i)}$ is assumed to be normally distributed. Based on the conditional expectations

derived above, the conditional random effects associated with the CD4 and cost models can be approximated as follows:

$$\begin{aligned} \mathrm{E}\left[a_0^{(i)} \mid U\right] &= \frac{(U-1)\sigma_{aU}}{\sigma_U^2} \\ &= \frac{(U-1)\,\mathrm{corr}(a_0, U)\sigma_a}{\sigma_U} \end{aligned}$$

and

$$\begin{aligned} \mu_k^{(i)}(t \mid U) &= \beta_0 + \frac{(U-1)\sigma_{bU}}{\sigma_U^2} + \beta_1 t + \beta_2 t^2 \\ &= \beta_0 + \frac{(U-1)\,\mathrm{corr}(b_0, U)\sigma_b}{\sigma_U} + \beta_1 t + \beta_2 t^2 \end{aligned}$$

The baseline rate function $\rho_0(t)$ can be estimated non-parametrically via the generalized Nelson-Aalen estimator. The cumulative incidence function can be derived analogously, which results in:

$$\mathrm{E}\left[N^{(i)}(t) \mid U\right] = U \exp\left(\sigma_z^2 \gamma^2 / 2\right) \sum_{k=1}^{m^{(i)}} \exp\left(\mu_k^{(i)}(t \mid U)\gamma\right)\left(\rho_0(t_k) - \rho_0(t_{k-1})\right).$$

APPENDIX C: TIME UNTIL GAINING AND LOSING RESISTANCE MUTATIONS FOR INDIVIDUALS RECEIVING HAART

Summary

The objective of this analysis was to estimate time until gaining a resistance mutation to each of four medication categories (3TC, other nucleoside reverse transciptase inhibitors (NRTIs), non-nucleoside reverse transciptase inhibitors (NNRTIs), and protease inhibitors (PIs)). Of particular interest was the relationship between medication adherence and gaining resistance mutations, and the interactions between adherence levels and medication categories.

Methods

Data

Data were obtained from the HAART Observational Medical Research cohort (n=2583), which consists of all antiretroviral-naïve individuals in British Columbia who initiated treatment with HAART between 1996 and 2004. Information is available describing baseline sociodemographic and clinical measurements. This dataset was linked to longitudinal resistance testing results.

Medications were separated into four categories: 3TC, other NRTIs, NNRTIs, and PIs. 3TC was categorized separately to other NRTIs because of the high likelihood of developing a resistance mutation to this particular medication and the lack of cross-resistance to other NRTIs.(1) For each medication category, time was measured relative to first exposure to the category. Individuals who had never received a medication of that category, or who had resistance detected to the category prior to first exposure were excluded from the analysis for that particular medication category, but were eligible to contribute to analyses for other categories.

Estimating time at risk

Time at risk for a particular resistance mutation was defined to be cumulative number of days spent taking the medication prior to developing resistance, which was approximated by number of days during which medication was dispensed. Gaps in exposure were not counted towards time spent at risk for gaining a resistance mutation. It was assumed that no new mutations developed during times when viral load was consistently suppressed below 1000 copies/mL.(1)

Exact time of developing a resistance mutation was not known, and could be considered interval-censored, bounded below by the date of the last negative test and above by the date of the first positive test. This interval could be further narrowed due to periods of time with consistently suppressed viral load following a negative resistance test.

Explanatory variables

Explanatory variables considered were sex, calendar year of first antiretroviral use, self-reported injection drug user status, baseline CD4 cell count, baseline plasma viral load, medication category, medication adherence, and interactions between medication and adherence categories. Medication adherence was defined as the percentage of months in which antiretroviral medications were dispensed during the first year of therapy.

Statistical methods

A parametric survival model was fit to describe time until gaining resistance mutations. Based on diagnostics and model flexibility, a Weibull distribution was assumed for the hazard rate. The interval-censored nature of the data was explicitly accounted for. Frailty terms were

included to account for multiple observations from a single individual. These frailties were assumed to follow a gamma distribution with mean 1.

Results

Descriptive characteristics of the study sample are given in Table C.1.

Figure C.1 displays Kaplan-Meier curves describing time until developing a resistance mutation, stratified by medication adherence and medication category, respectively. The probability of resistance decreased with increasing adherence, particularly the 95-100% adherence category (p-value for trend <0.0001). Resistance occurred most frequently to 3TC and non-nucleoside reverse transcriptase inhibitors (p-value for trend <0.0001). Both these trends are consistent with previously published results.(1, 2)

The Kaplan-Meier results are intended for describing broad trends only, because the interval-censored nature of the data and the potential for correlation within an individual were not accounted for. The parametric frailty model was fit to incorporate these elements into the analysis.

Estimated coefficients estimated using the parametric frailty model are given in Table C.2, while corresponding hazard ratios are given in Table C.3. Note that for a Weibull model with scale term σ, the relationship between the coefficient γ and the hazard ratio associated with a particular variable is expressed as:(3)

$$\alpha = \exp(-\gamma / \sigma)$$

The frailty terms were constrained to have a mean of 1.0. The estimated variance associated with the gamma distribution was 0.38.

The interaction terms estimated using the Weibull model are consistent with the hypothesis that the relationship between medication adherence and risk of resistance differs according to drug category.(2) For example, while the estimated risk of resistance was lowest at 95-100% adherence for all drug categories, this reduction was less pronounced for NNRTIs. Thus, at this level of adherence, risk of resistance to NNRTIs was estimated to be more than seven times greater than risk of resistance to 3TC, the medication category with the next greatest risk.

Figure C.1: Kaplan-Meier curves describing time until gaining resistance mutation, stratified by (a) medication adherence, and (b) medication category.

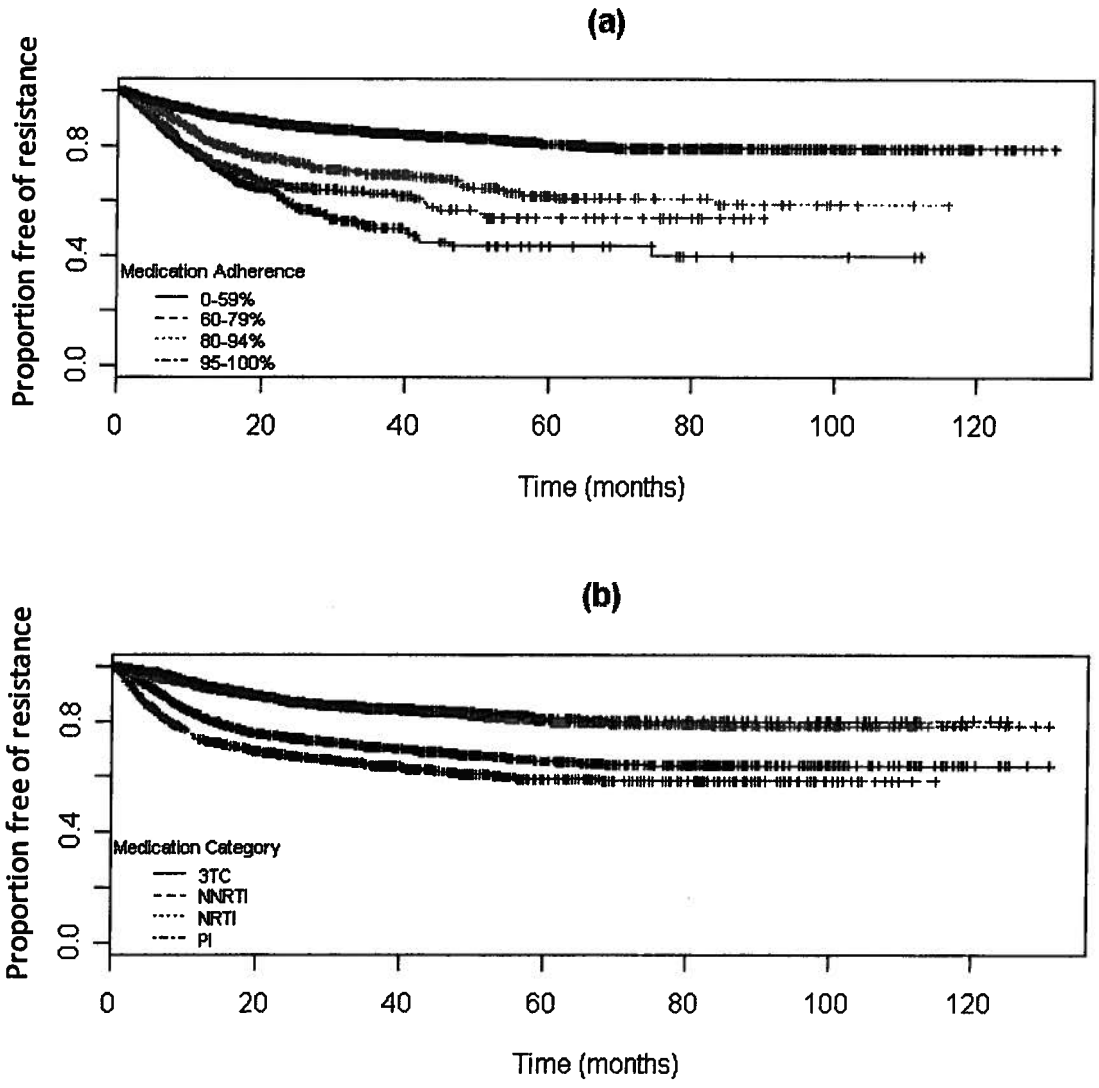

Table C.1: Descriptive characteristics of study sample

Variable	
Sex (%)	
Male	81.6
Female	18.4
Injection drug user (%)	
No	72.3
Yes	27.7
Year of first antiretroviral use (%)	
1996-1998	35.9
1999-2001	34.2
2002-2004	32.9
Mean baseline CD4 cell count (cells/mm^3) (standard deviation (SD))	237.9 (198.7)
Mean baseline plasma viral load (log copies/mL) (SD)	5.00 (0.85)
Medication Adherence (%)	
<60%	23.7
60-79%	9.7
80-94%	12.9
≥95%	53.8
Medication Category (number of individuals exposed)	
3TC	2514
NRTI	2583
NNRTI	1642
PI	1891
Medication Category (total person-years follow-up)	
3TC	4474
NRTI	5173
NNRTI	2550
PI	2837

Table C.2: Regression coefficients estimated using Weibull survival model.

Variable	Estimate (standard error)
Intercept	8.57 (0.54)
Sex	
Male	--
Female	-0.02 (0.17)
Injection drug user	
No	--
Yes	-0.39 (0.14)
Year of first antiretroviral use (continuous)	0.08 (0.03)
Baseline CD4 cell count (cells/mm^3)	0.001 (0.0003)
Baseline plasma viral load (log copies/mL)	-0.35 (0.09)
Medication category	
3TC	--
NRTI	1.1 8 (0.12)
NNRTI	-0.87 (0.11)
PI	1.58 (0.16)
Medication category	
<60%	--
60-79%	0.29 (0.23)
80-94%	0.76 (0.22)
≥95%	1.80 (0.18)
Medication/Adherence interactions	
NRTI*60-79% adherence	0.24 (0.19)
NRTI*80-94% adherence	0.50 (0.19)
NRTI*≥95% adherence	-0.32 (0.14)
NNRTI*60-79% adherence	-0.11 (0.18)
NNRTI*80-94% adherence	0.60 (0.18)
NNRTI*≥95% adherence	0.64 (0.14)
PI*60-79% adherence	0.04 (0.26)
PI*80-94% adherence	-0.39 (0.22)
PI*≥95% adherence	-0.55 (0.19)
Log (scale)	-0.51 (0.03)

Table C.3: Estimated hazard ratios estimated using Weibull survival model.

Variable	Estimate
Sex	
Male	1.00
Female	1.04
Injection drug user	
No	1.00
Yes	1.93
Year of first antiretroviral use (continuous)	0.87
Baseline CD4 cell count (per 100 cells/mm^3)	0.84
Baseline plasma viral load (log copies/mL)	1.80
Medication category	
3TC	
<60% adherence	1.00
60-79% adherence	1.00
80-94% adherence	1.00
≥95% adherence	1.00
NRTI	
<60% adherence	0.14
60-79% adherence	0.09
80-94% adherence	0.06
≥95% adherence	0.24
NNRTI	
<60% adherence	4.28
60-79% adherence	2.87
80-94% adherence	1.85
≥95% adherence	7.41
PI	
<60% adherence	0.07
60-79% adherence	0.05
80-94% adherence	0.03
≥95% adherence	0.12
Medication adherence	
<60%	
3TC	1.00
NRTI	1.00
NNRTI	1.00
PI	1.00
60-79%	
3TC	0.62
NRTI	0.41
NNRTI	0.75
PI	0.58
80-94%	
3TC	0.28
NRTI	0.12
NNRTI	0.10
PI	0.54
≥95%	
3TC	0.05
NRTI	0.09
NNRTI	0.17
PI	0.12

References

1. Harrigan PR, Hogg RS, Dong WWY, et al. Predictors of HIV drug-resistance mutations in a large antiretroviral-naive cohort initiating triple antiretroviral therapy. Journal of Infectious Diseases 2005;191:339-47.
2. Bangsberg DR, Moss AR, Deeks SG. Paradoxes of adherence and drug resistance to HIV antiretroviral therapy. Journal of Antimicrobial Chemotherapy 2004;53:696-9.
3. Klein JP, Moeschberger ML. Survival analysis: techniques for censored and truncated data. Springer, 2003.

APPENDIX D: ETHICS APPROVAL CERTIFICATE

PROVIDENCE HEALTH CARE
Research Institute

UBC-Providence Health Care
Research Institute
Office of Research Services
11th Floor Hornby Site - SPH
c/o 1081 Burrard St.
Vancouver, BC V6Z 1Y6
Tel: (604) 806-8567
Fax: (604) 806-8568

ETHICS CERTIFICATE OF EXPEDITED APPROVAL

PRINCIPAL INVESTIGATOR:	DEPARTMENT:	UBC-PHC REB NUMBER:
Adrian Levy	UBC/Medicine, Faculty of/Health Care & Epidemiology	H07-01506

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:

Institution	Site
Providence Health Care	St. Paul's Hospital

Other locations where the research will be conducted:
N/A

COINVESTIGATOR(S):
Paul A. Gustafson
Karissa M. Johnston
Mark W. Tyndall
Julio S.G. Montaner

SPONSORING AGENCIES:
Canadian Institutes of Health Research (CIHR)

PROJECT TITLE:
A new method of integrating epidemiological and health services modelling techniques for studying infectious diseases: An example using HIV/AIDS

THE CURRENT UBC-PHC REB APPROVAL FOR THIS STUDY EXPIRES: September 7, 2008

The UBC-PHC Research Ethics Board Chair or Associate Chair, has reviewed the above described research project, including associated documentation noted below, and finds the research project aceeptable on ethical grounds for research involving human subjects and hereby grants approval.

DOCUMENTS INCLUDED IN THIS APPROVAL:		APPROVAL DATE: September 7, 2007
Document Name	**Version**	**Date**
Protocol: A new method of integrating epidemiological and health services modelling techniques for studying infectious diseases: An example using HIV/AIDS	N/A	June 15, 2007

CERTIFICATION:

1. The membership of the UBC-PHC REB complies with the membership requirements for research ethics boards defined in Part C Division 5 of the Food and Drug Regulations of Canada.
2. The UBC-PHC REB carries out its functions in a manner fully consistent with Good Clinical Practices.
3. The UBC-PHC REB has reviewed and approved the research project named on this Certificate of Approval

including any associated consent form and taken the action noted above. This research project is to be conducted by the principal investigator named above at the specified research site(s). This review of the UBC-PHC REB have been documented in writing.

Approval of the UBC-PHC Research Ethics Board or Associate Chair, verified by the signature of one of the following:

Dr. I. Fedoroff,
Chair

Dr. J. Kernahan,
Associate Chair