Title
Causal Inference Methods and Their Application To HIV Observational Study Data

Author
Anglemyer, Andrew Thomas

Publication Date
2010

Peer reviewed|Thesis/dissertation

Causal Inference Methods and Their Application To HIV Observational Study Data

By

Andrew Thomas Anglemyer

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Epidemiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor John Colford, Chair
Professor Alan Hubbard
Professor Malcolm Potts

Spring 2010

Causal Inference Methods and Their Application To HIV Observational Study Data

Andrew Thomas Anglemyer

Abstract

Causal Inference Methods and Their Application To HIV Observational Study Data

by

Andrew Thomas Anglemyer

Doctor of Philosophy in Epidemiology

University of California, Berkeley

Professor John M. Colford, Jr., Chair

Worldwide, particularly in areas with no treatment availability or antenatal programs, approximately 1600 children are diagnosed with human immunodeficiency virus (HIV) every day,(5) and over 300,000 deaths from acquired immune deficiency syndrom (AIDS) among children occur annually worldwide.(6) The Centers for Disease Prevention and Control (CDC) estimated a total of 142 children less than 13 years old were infected with HIV perinatally in 2005,(7) while the World Health Organization (WHO) estimates 2 million children (0-14 years) globally living with HIV (1.8 million living in Sub-Saharan Africa alone).(8) Epidemiologists and biostatisticians are actively trying to estimate the causal effects of highly active antiretroviral therapy (HAART) in order to establish which treatments are best and when to they should be initiated. This proves to be a challenging task for several reasons including the unique dynamics of pediatric HIV populations and the lack of randomized evidence. However, with an abundance of observational data, analytical approaches designed to help researchers establish causal effects from observational studies have been developed—referred to within the present studies as causal inference techniques. In this dissertation, I performed a systematic review of studies that used so-called causal inference methods (i.e. propensity scores, instrumental variables, marginal structural models, and structural equation models) in the context of HIV/AIDS research and assessed the interpretability and content of the identified studies. I use empirical examples from a marginal structural model (MSM) analysis and instrumental variable (IV) analysis using Pediatric Spectrum of Disease (PSD) surveillance program data. Specifically, I estimate the causal effect of triple therapy (e.g. HAART) on time to C diagnosis, and time to C diagnosis/death among HIV-infected children and perform an adapted instrumental variable analysis in order to estimate the causal effect of HAART on the hazard of AIDS events or death.

The systematic review revealed that approximately 43% of all studies using causal inference methods on HIV/AIDS data were published in 2007 and 2008. Studies using MSMs were less likely to discuss specific model selection than studies using any other causal inference method (OR=0.26; 95% CI 0.08-0.72). Using a g-comp approach, where I define $\Psi_1 (p_0)(t_k) \equiv P(T_a > t_k)$ as all treated and $\Psi_0 (p_0)(t_k) \equiv P(T_a > t_k)$ as all

untreated, the causal effect of HAART suggested that among children who initiated therapy within 6 months of birth the effect in delaying a C diagnosis, $\Psi HZ(p_0)(t_k)$= -0.466 (95% CI -1.20-0.565), is seemingly stronger than children who initiated therapy within 12 months of birth ($\Psi HZ(p_0)(t_k)$= -0.321 (95% CI -1.151-0.300)). Additionally, though not statistically significant, the effect of triple therapy initiated within the first 12 months of life on time to C diagnosis is potentially greater among *symptomatic* children (12 Months: $\Psi HZ_{symptomatic}$ (p0)(t_{36}): -0.587 (95% CI -1.217-0.480)) than among *asymptomatic* children (12 Months: $\Psi HZ_{asymptomatic}$ (p0)(t_{36}): -0.106 (95% CI -1.054-0.739)). The instrumental variable analysis yielded the naïve rate ratio comparing an early-defined IV (1997 cut-off) non-HAART era with the HAART era—estimated at RR_{ITT}=2.17 (95% CI: 1.34-3.52). As a result of HAART use misclassification by calendar era, an instrumental variable estimator was used, yielding a RR_{IV} = 3.91 (95% CI 2.41, 6.34), 80% higher than the naïve result.

Regardless of year of publication, all HIV studies are deficient by varying degrees in all assessed areas. Researchers using causal inference methods should describe their methods in a more transparent and interpretable way so that the results may reach a wider audience. Together Chapter 3 and 4 use causal inference methods to not only help establish the effectiveness of HAART on preventing advanced disease and/or mortality, but they also attempt to address the need to establish optimal timing of treatment for treatment guidelines. The overarching benefits of these methods are that they define a parameter of interest not dependent on a particular model assumption (semi-parametric), and they define explicit identifiability assumptions under which these estimators produce estimates of so-called causal association, which are related to distributions of counterfactuals.

for kara.

Contents

List of Tables

List of Figures

Acknowledgements

This dissertation would not have been completed without the help of people I have been lucky enough to have in my life. My advisor and mentor, Jack Colford, has been an ever-present support in my academic endeavors. He graciously took me under his wing several years ago and I am confident that I would not be where I am now without his kindness and motivation. Even with nearly 20 epidemiology graduate students working with him, Alan Hubbard has found the time and patience to guide me in statistical methods throughout this adventure. It is his encouragement that has reassured me that I am not alone in the eternally complex world of biostatistics. At the most frustrating times in my coding process, Ori Stitelman came to the rescue—showing me that sometimes the easiest way to think about things is the best way. Six years ago I began studying public health issues at Berkeley; one of the first classes I took was International Health co-taught by Malcolm Potts. He saw in me an inspired student who truly wants to help the public health community. Working with him in the Maternal and Child Health department provided me the opportunity to refine where and how I wanted pursue my future in epidemiology. And in 2005, I was in Russia for a public health internship, which fell through abruptly after arriving. With late notice, Wayne Enanoria and Tomas Aragon at the Center for Infectious Disease Preparedness provided me the opportunity to work with their center for an internship. It is with this opportunity that I not only discovered my love of epidemiology, but I was also introduced to Jack Colford. Bonnie Maldonado and Amy Sturt at Stanford University rescued me when my first project fell through. They provided me with the data, which were used for my analyses contained within this dissertation. Though the world of epidemiology is foreign to my immediate family, each and every one of my family members has been supportive from the beginning. Their thoughts and prayers have lifted the weight of expectations from this project. In epidemiology and biostatistics we are often lost in the numbers; I would like to remind the reader that behind every data point is a child's experience. And Kara--she has been with me through two project proposals and numerous struggles and long nights throughout this journey. She has been my inspiration and I could not ask for a better support system and love. Her beauty will save the world.

Chapter 1

Introduction

1.1 Introduction

The goal of this dissertation is to describe the methodological and research options for epidemiologists who are working with observational data dealing with human immunodeficiency virus/acquired immune deficiency syndrome (HIV/AIDS). As the use of causal inference research approaches becomes more prominent, the application of these epidemiological methods should involve just as much, if not more, care to ensure quality and validity as with traditional analytical techniques. I will use pediatric observational study data to highlight different causal inference approaches. Treatment guidelines for HIV-infected children are country and institution-dependent; it is with these causal inference approaches that I hope to contribute to the pediatric HIV research field to help establish the best HIV treatment scenario and schedule. The analyses heretofore aim to not only support findings from randomized trials but to also estimate the impact of HIV interventions on a population level. Results from these analyses may not only presumably help reduce the individual burden of disease through treatment initiation guidance, but could also impact future treatment costs and requirements within countries with a high incidence of pediatric HIV.

1.2 Specific aims

This dissertation has two specific aims:

1. Review the most common causal inference methods used to analyze HIV observational data

2. Apply techniques reviewed under Aim 1 to:

    a) Estimate the effect of highly active antiretroviral therapy on time to AIDS or death among HIV-infected children attending referral clinics from 1988 to 2009.

    b) Estimate the population effect of highly active antiretroviral therapy on the hazard of AIDS or death among HIV-infected children attending referral clinics from 1988 to 2009.

My second aim is an application of two causal inference methods on data from the Pediatric Spectrum of Disease (PSD) multicenter active surveillance program specifically for children who have been exposed to HIV perinatally from 1988 to present. The outcomes that are my primary focus are: a C diagnosis (a definition of severe HIV disease progression); C diagnosis or death; and death.

1.3 Structure of the dissertation

This dissertation is structured to include the following chapters:

- In Chapter 1 I highlight the need for wider, more appropriate uses of causal inference techniques applied to HIV/AIDS data. I also discuss how these methods could influence treatment guidelines if adopted more widely to reach an international audience.

- Chapter 2 highlights Specific Aim 1. I briefly introduce four different causal inference techniques--propensity scores; instrumental variables, marginal structural models, and structural equation models—and describe the temporal trends in publications and quality assessments of the described techniques within HIV/AIDS settings.

- Chapter 3 addresses Specific Aim 2. I use marginal structural models as estimated by g-comp to estimate the causal effect of highly active antiretroviral therapy on time to a C diagnosis or death among an HIV-positive pediatric population.

- Chapter 4 applies an adapted instrumental variable analysis to estimate the effect of highly active antiretroviral therapy on the hazard of a C diagnosis or death among an HIV-positive pediatric population.

- Chapter 5 is a synthesis of my results from Chapters 2-4 and presents possible directions for future research.

- The appendices include reference materials for Chapters 2-4 (A) and the manuscripts of Chapters 2-4 submitted for publication (B).

1.4 Prevention, treatment, and study of HIV/AIDS

Though the World Health Organization (WHO) advocates for a four-tiered approach for preventing mother-to-child-transmission (MTCT),(1) the public health response has been predominantly compartmentalized within one or two aspects separately. The recommended approach to reducing MTCT worldwide is by focusing on: 1) prevention of HIV in women, most importantly young women; 2) prevention of unintended pregnancies among HIV positive women; 3) prevention of vertical transmission; and 4) support for the mother and family.(1) As I discuss, highlight, and implement new epidemiological techniques for estimating the effects of specific treatments for pediatric HIV, it is important to keep in mind that all of these cases could have been prevented by not only HIV prevention services but also sexual and reproductive health services. A schematic illustrating the different points of intervention for prevention of MTCT is shown below. The public health approach has focused primarily on identifying, through testing, the HIV positive pregnant women to ensure adequate MTCT prevention efforts during and after pregnancy. However, the case could be made that an integrated focus on sexual and reproductive health services with HIV/AIDS prevention services would have a wider-catchment area for HIV prevention.(2) For example, in Figure 1, Women 2-4 would have remained uninfected had they had access to sexual and reproductive health services (e.g. condoms). Furthermore, this same access would have prevented any unwanted pregnancies, and thus would have prevented any vertical transmission. Woman 1's use of reproductive health services not only prevented her own HIV infection, but also vertical infections by not becoming pregnant. To illustrate that sexual and reproductive services are not only helpful in protecting a woman from HIV, but are also a necessary component to prevent MTCT, Woman 3's access to reproductive services prevented any unwanted pregnancies. The HIV status of children of Women 2 and 4 could be addressed through a combination of antenatal and HIV services during pregnancy, delivery, and early infancy. As such, services that focus primarily on the pregnant woman are further downstream in the infection chain and must focus on the woman *and* infant in order to prevention transmission.

Figure 1.1 Stages of HIV Vertical Transmission and Provision of Prevention Services

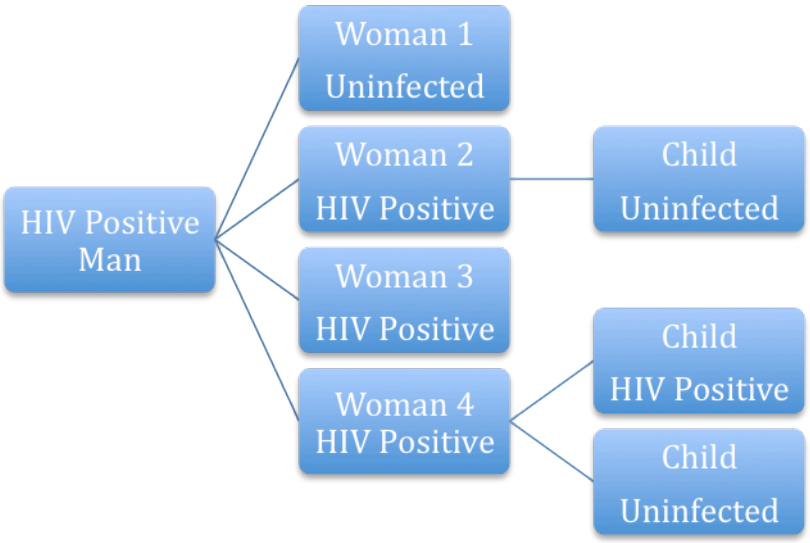
Transmission preventable by *reproductive services* by either preventing initial infection to women or by preventing future pregnancies among infected women

{----------}

STAGE 1 STAGE 2 STAGE 3

Preventable by treating HIV+ pregnant women, administering prophy-laxis to child and through safe feeding practices

{----------}



Prevention strategies focusing on the child have to cope with three avenues of infection via MTCT: in utero; at delivery; and through breastfeeding. It is estimated that the probability of infection in utero and at time of delivery is approximately 15-30%, though if breastfeeding from 18-24 months the overall probability of vertical transmission increases to about 30-45%.(3)

Factors affecting the risks at each of these stages are highlighted in Table 1.1. One of the greatest risks for vertical transmission is the severity of HIV disease among the mothers. Maternal disease severity is most commonly described by the viral load, the amount of HIV contained within her blood, and is often an indication of either recent infection or advanced AIDS. As seen in Table 1.1, maternal viral load is an ever-present MTCT risk from pregnancy through breastfeeding.

Table 1.1 Maternal and neonatal factors that may increase the risk of HIV transmission

Pregnancy	Labour and Delivery	Breastfeeding
▪ High maternal viral load (new infection or advanced AIDS) ▪ Viral, bacterial, or parasitic placental infections, such as malaria ▪ Sexually transmitted infections (STIs)	▪ High maternal viral load (new infection or advanced AIDS) ▪ Rupture of membranes for more than 4 hours[1] ▪ Invasive delivery procedures that increase contact with mother's infected blood or body fluids (e.g. episiotomy, artificial rupture of membranes) ▪ Chorioamnionitis (from untreated STI or other infection) ▪ Preterm delivery ▪ Low birthweight	▪ High maternal viral load (new infection or advanced AIDS) ▪ Duration of breastfeeding ▪ Mixed feeding (giving water, other liquids, or solid foods in addition to breastfeeding) ▪ Breast abscesses, nipple fissures, mastitis ▪ Oral disease in the baby (e.g. thrush or sores)
[1] Studies have found there is an increased rate of HIV transmission after a mother's membranes have been ruptured for more than 4 hours. The longer the membranes are ruptured, the higher the risk of HIV transmission.		

Adapted from the WHO/CDC Prevention of Mother-to-Child Transmission of HIV Generic Training Package (2008).(4)

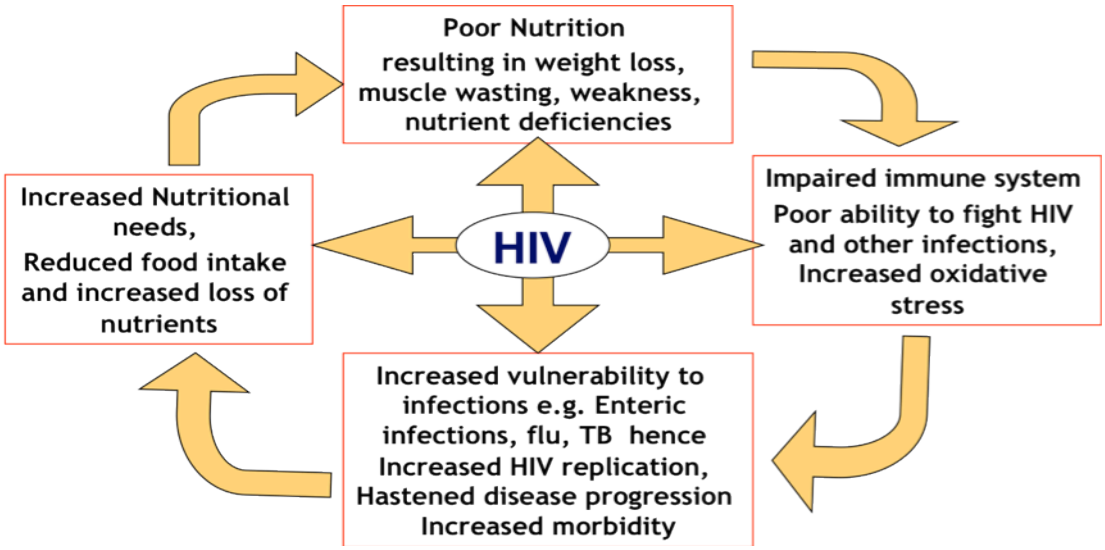
As illustrated in Figure 1.1, prevention efforts have multiple avenues of approach to prevent MTCT. Before pregnancy, reproductive health services provide options for safer sex and family planning alternatives to prevent unintended pregnancies. Additionally, sexual health services provide testing and treatment options for sexually transmitted infections. Testing for HIV and counseling are available at antenatal clinics and during prenatal care. And, providing antiretroviral therapy not only benefits the woman's health, but also in turn reduces the risk of vertical transmission. All these approaches in concert with counseling about healthy feeding options will likely reduce the risk of infection for a child born to an HIV positive mother.

Despite all the various sectors of prevention, United Nations AIDS (UNAIDS) estimates that worldwide, particularly in areas with no treatment availability or antenatal programs, approximately 1600 children are diagnosed with HIV every day.(5) Furthermore, over 300,000 deaths among infected children occur annually worldwide.(6) The influence prevention strategies have had on stopping MTCT is ever apparent when the numbers of pediatric infections in the United States and Sub-Saharan Africa are contrasted. The Centers for Disease Prevention and Control (CDC) estimated a total of 142 children less than 13 years old were infected with HIV perinatally in 2005,(7) while the WHO estimates 2 million children (0-14 years) globally living with HIV (1.8 million living in Sub-Saharan Africa alone).(8)

The factors affecting HIV progression in children are multi-dimensional. Perhaps one of the most influential aspects to HIV progression, aside from access to antiretroviral therapy (ART) treatment, is nutrition. As illustrated in Figure 1.2, the relationships

between HIV, nutrition, coinfections, and baseline immune function are synergistic.(9) Some research suggests that HIV-infected newborns with fewer t-cells are more likely to progress quicker than HIV-infected newborns with more t-cells.(10) The presence of coinfections, such as tuberculosis, has also been independently linked to HIV progression in children.(11; 12) Much like the WHO's recommended multi-tiered approach to prevention of MTCT, HIV treatment should also include multiple aspects to adequately slow disease progression.

Figure 1.2 Relationships Between HIV, Nutrition, Coinfection, and Baseline Immune Function



Source: Adapted from RCQHC and FANTA 2003.(9)

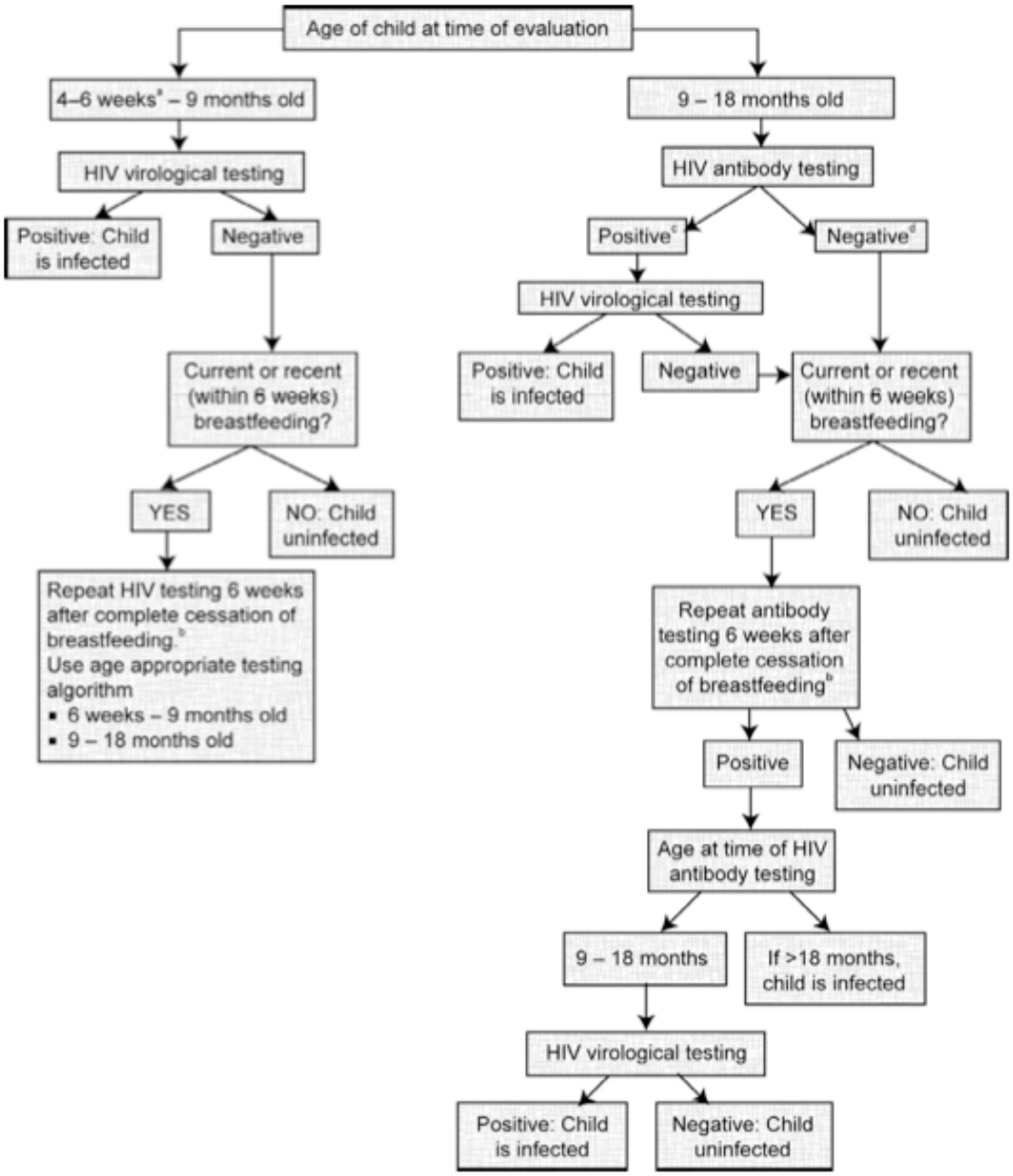
Children born to HIV positive mothers are at a unique risk for not only HIV infection, but also immune health in general. Infants born to HIV+ mothers are a particularly vulnerable population for myriad reasons. Pregnant women infected with HIV are more likely to give birth to a low birth weight (LBW) infant if they are not taking antiretroviral medications.(13) Without access to appropriate drugs and care, a reality in the developing world, these fetuses are likely to seroconvert, as well. In turn, a LBW child vertically infected with HIV has two significant threats to his immune function—LBW and HIV.

Children with a compromised immune system, as a result of HIV infection or other factors such as impaired intrauterine growth, are especially vulnerable to poor immune responses. Cytotoxic T cells, mostly CD8s are the protagonists in cell-mediated illnesses, while helper T cells, e.g. CD4s, secrete cytokines to activate more cytotoxic cells for damage. A possible consequence of impaired immune response is a condition referred to as immune reconstitution inflammatory syndrome (IRIS).(14) This situation occurs whenever an individual's immune system is stabilized after long-term suppression (e.g. HIV infection). Indeed, reconstituting the lymphocyte structure and production among

HIV+ individuals will raise levels of CD4, which in turn creates the opportunity for more infections like tuberculosis or other cell mediated illnesses by activating more cytotoxic cells. IRIS is also seen in infants born LBW with low lymphocyte counts. After therapy their counts raise, as does their risk for opportunistic infections.

Diagnosing infants early is essential to providing adequate care, but young pediatric populations are more difficult to diagnose than adults. The WHO has put together an algorithm for diagnosing HIV in infants (see Figure 1.3).(15) Essentially the status of children under 18 months old is only confirmed through more expensive HIV virological testing (e.g. polymerase chain reaction-PCR); even with a negative virological test, if the child has breastfed recently he could still seroconvert in the near future. The HIV rapid test (e.g. enzyme-linked immunosorbent assay-ELISA) most commonly used among adult populations is not useful in infants under 9 months because maternal antibodies may still be present and can affect the test's results. The costs of the virological tests can pose a barrier for adequate diagnoses in resource-limited settings.

Figure 1.3 HIV Diagnosis in Infants and Young Children Under 18 Months



[a] A positive virological test at any age indicates HIV infection: NAT is recommended beginning at 4–6 weeks of age to maximize sensitivity.

[b] If a child experiences HIV-related symptoms, regardless of prior test results, repeat test even if child has not stopped breastfeeding.

[c] If virological testing is not available, repeat antibody testing 6 weeks after complete cessation of breastfeeding. If the child is less than 18 months of age at time of the repeat test, the child should be test again at 18 months of age or older (as per national guidelines).

[d] A negative antibody testing for a child 9 to 18 months of age, who is not breastfeeding, can be used to exclude HIV infection.

Just as identifying HIV positive children calls for different approaches than those that would be employed for adult populations, studies of treatment effectiveness in adults may not be reflective of its effectiveness in pediatric populations. Results from adult studies that have explored highly active antiretroviral therapy (HAART) and its effect on mortality are likely not reflective of its effect on mortality if applied to pediatric populations.(16) Typically CD4 counts and viral load are clinical parameters used to assess disease progression in adult populations, but these same parameters have a wide variability in children.(17-19) Additionally, in children HIV affects neuro-cognitive development, growth, and an immune system that is not yet fully mature.(20-22) Though adult observational studies have shown that patients who start HAART early (higher CD4 counts) have better clinical outcomes than adults who start at CD4 counts below 200 cells/μL, these results are likely not generalizeable to the pediatric HIV infected community.

Deciding when to begin HIV treatment among children is an ongoing debate that has life-long implications for the children and economic consequences for governments. Until recently(23) there was little agreement between the three primary HIV/AIDS health-governing bodies about the optimal time for treatment initiation for children under 12 months. Table 1.2 contrasts the treatment guidelines for HIV-positive children.(24) A decision to recommend treatment for all HIV positive children regardless of symptoms was recently made by all three institutions based on recent randomized controlled trial (RCT) evidence suggesting reduced mortality risk for early treated children versus deferred treated children. Generally, the WHO is more conservative in its recommendations as they have a lower CD4% or CD4 count treatment initiation threshold than Paediatric European Network for Treatment of AIDS (PENTA) or the CDC in most pediatric age groups.

Table 1.2 Contrasts of Pediatric Treatment Initiation Guidelines by Governing Body

		PENTA 2008	CDC 2008	WHO 2009
0-11 months	Clinical Immunological Virological	Treat all	Treat all	Treat all
12-35 months	Clinical	Treat CDC Stage B or C/WHO stage 3 or 4	Treat CDC Stage B or C	WHO stage 4 or severe 3
	Immunological	Treat < 25% or < 1000	Treat < 25%	Treat < 20% or < 750
	Virological	Consider > 100,000 copies/mL	Consider > 100,000 copies/mL	
36-59 months	Clinical	Treat CDC Stage B or C/WHO stage 3 or 4	Treat CDC Stage B or C	WHO stage 4 or severe 3
	Immunological	Treat < 20% or < 500	Treat < 25%	Treat < 20% or < 750
	Virological	Consider > 100,000 copies/mL	Consider > 100,000 copies/mL	
60+ months	Clinical	Treat CDC Stage B or C/WHO stage 3 or 4	Treat CDC Stage B or C	WHO stage 4 or severe 3
	Immunological	Treat < 350	Treat < 350	Treat < 15% or < 200
	Virological	Consider > 100,000 copies/mL	Consider > 100,000 copies/mL	

Adapted from PENTA Guidelines.(24)

1.5 Alternative analytic approaches

One of the key reasons why it has been so difficult to establish the optimal timing for treatment initiation is because of a lack of randomized evidence. The effects of ARVs in preventing severe disease and mortality have been well-established for years in observational and randomized studies. In turn, it is often difficult to ethically justify randomizing some children to treatment and others to none or deferred treatment. A couple of RCTs have actually estimated the effect of delaying treatment (i.e. HAART) to see if there was any benefit.(25; 26) The CHERI study estimated the effect of early HAART versus delayed HAART on mortality among HIV positive infants.(26) This trial was prematurely terminated because of an unbalanced, disproportionate number of deaths in the delayed treatment arm. A small feasibility RCT study was also conducted which aimed to estimate the impact of delaying HAART.(25) The study population for this small feasibility trial was comprised of HIV positive children from 1 year to 12 years old, thus not addressing the issue of delaying treatment among infants. The results from the larger trial will likely not be available until 2011. Observational study data results supporting these findings are not widely available either.(27; 28) In one of the few observational studies exploring when to initiate treatment among infants, Chiappini et al compared children who were treated with HAART early with children who were deferred treatment and found that the early treated group had significantly lower viral load than deferred treatment group and that they were less likely to progress to a C diagnosis.(27) Similarly, Newell et al studied a cohort of HIV infected children and concluded that initiating ART in the first 5 months of life and the use of HAART were both predictive of an improved CD4 z-score 6 months after treatment initiation.(28)

Despite a lack of randomized evidence, researchers have developed techniques in an attempt to assert causation from observational study data. The counterfactual framework was developed in 1990 by Rubin to help illustrate the need for and the use of causal inference methods.(29) The theory was later introduced more into mainstream epidemiology by Greenland in 2000.(30) In general, the counterfactual is the idea that an outcome would not have occurred had, counter to fact, a previous event not occurred. The goal of counterfactuals is to achieve a pseudo-randomized population as one would have in an RCT. Much like the scenario when a person's treatment assignment through randomization in an RCT is not dependent on information connected to outcome studied, the counterfactual framework creates the situation in which person A would have had equal opportunity to be treated to x.

Studying HIV/AIDS observational data can be complex and multifaceted, posing numerous challenges for biostatisticians and epidemiologists alike. Researchers are often charged with estimating the true impact of treatments to adequately prevent disease spread or treat patients with HIV/AIDS. However, to estimate the true effects, nuisance factors need to be addressed and/or removed, often requiring special techniques.

Common approaches to estimate treatment effects are often biased because of model selection or unmeasured confounding. For example, one of the traditional methods for controlling the biases introduced by measured confounders are multivariable regression techniques. However, regression techniques are only as good as the measured confounder data.(31) Moreover, model building can become cumbersome even with comprehensive confounder data collected; often investigators will settle for model interpretability over adequate adjustment for bias.(31) Furthermore, time-dependent confounding is often addressed by using methods such as extended Cox regression models in which one would add an interaction term made up of the time-dependent variable of interest and some function of time.(32) Under two specific conditions, these conventional methods fail: C1) when there is a time-dependent risk factor of an event that also predicts subsequent treatment; and C2) when previous treatment history predicts a future risk factor.(33) Causal inference methods have been developed in order to help reconcile these problems commonly seen in HIV epidemiological research. These concepts are reintroduced later in Chapters 2-4.

The application of causal inference techniques to HIV/AIDS observational study data has increased steadily over the last 20 years. In Chapter 2, I performed a systematic review of the literature to explore the temporal trends of four causal inference methods used with HIV/AIDS data—including instrumental variables, propensity scores, marginal structural models, and structural equation models--the frequency of publications by authors and journals, assess the transparency and quality of the methods employed, and explore the networks of affiliated institutions. I identified 70 papers for all methods; approximately forty-one percent of all HIV/AIDS studies using the listed causal inference methods were published in 2007 and 2008. Specifically, approximately half (47%) of the studies using marginal structural models were published in 2007 and 2008.

Though the application of causal inference techniques has become more popular in the epidemiologic literature, their use with pediatric HIV/AIDS data is infrequent. Through 2008, only three studies have applied causal inference methods to pediatric HIV data, all three used marginal structural models.(34-36) Fox et al used marginal structural models the estimate the effect of maternal CD4 count on child mortality after adjusting for breastfeeding and low birth weight. The researchers found that a child whose mother's CD4 was less than 200 had an increased 18 month mortality risk when compared to children whose mother's CD4 was greater than 500 (HR=3.0 95% 1.2-7.9).(34) The other two causal inference applications through 2008 were both Patel et al.(35; 36) The researchers used marginal structural models to estimate the effects of HAART on CD4 cell percentage in a pediatric cohort study. After 5 years of follow up, they found that the initiation of HAART increased CD4 percentage by 2.34% (95% 1.35%-3.33%) after adjusting for confounding by disease severity.(35) With the same study population after 10 years of follow up, the authors estimated the effect of HAART on mortality and found that marginal hazard ratio comparing HAART and non-HAART use was 0.24 (95% 0.11-0.51) after adjusting for disease severity.(36) In Chapter 3 to estimate the causal effect of HAART on time to AIDS progression or death, I have estimated a marginal structural (association of marginal counterfactual means with changes in uniformly applied treatment) using g-computation. Similar to Patel et al's approach, in this analysis I

avoided numerous pitfalls commonly seen with traditional techniques, including, but not limited to, and unmeasured biases and model misspecification. In Chapter 4, I use calendar year in an adapted instrumental variable analysis to estimate the causal effect of HAART on the hazard of C diagnosis or death. This analysis further adjusted for misclassification of HAART use and any covariate that may have been related to calendar year and outcome. This approach, never before used on an HIV positive pediatric population, is particularly appropriate when trying to estimate the effect of a treatment on a population level. Together Chapter 3 and 4 not only help establish the effectiveness of HAART on preventing advanced disease and/or mortality, but they also attempt to address the need to establish optimal timing of treatment for treatment guidelines.

Bibliography

1. World Health Organization. Guidance on Global Scale-up of the prevention of mother-to-child transmission of HIV. [Internet]. 2007.Available from: www.who.int/hiv/mtct/PMTCT_enWEBNov26.pdf

2. Leach-Lemens C. Preventing unintended pregnancies in women living with HIV in resource-poor settings. HIV and AIDS Treatment in Practice 2010 Mar;(155):1-12.

3. De Cock KM, Fowler MG, Mercier E, de Vincenzi I, Saba J, Hoff E, Alnwick DJ, Rogers M, Shaffer N. Prevention of mother-to-child HIV transmission in resource-poor countries: translating research into policy and practice. Jama 2000;283(9):1175.

4. World Health Organization. WHO/CDC Prevention of Mother-To-Child-Transmission of HIV Generic Training Package. 2008;

5. HIV/AIDS Update [Internet]. [January 2010];Available from: http://www.unaids.org/

6. UNAIDS/World Health Organization. HIV/AIDS Update [Internet]. [January 2010].Available from: http://data.unaid-s.org/pub/EPISlides/2007/2007_epiupdate_en.pdf

7. Centers for Disease Control and Prevention. HIV/AIDS Surveillance Report, 2005. Atlanta: US Department of Health and Human Services; 2007.

8. World Health Organization. HIV and AIDS Estimates and data, 2007 and 2001. World Health Organization; 2008.

9. Regional Centre for Quality of Health Care (RCQHC), Food and Nutrition Technical Assistance (FANTA). Developing and Applying National Guidelines on Nutrition and HIV/AIDS. [Internet]. 2003;Available from: www.fantaproject.org

10. Petru A. 2008;

11. Putong N, Pitisuttithum P, Supanaranon W. Mycobacterium tuberculosis infection among HIV/AIDS patients in thailand: clinical manifestations and outcomes. Southeast Asian Journal of Tropical Medicine and Public Health [2002];33:346-351.

12. Singh A, Bairy I, Shivananda P. Spectrum of opportunistic infections in AIDS case. Indian Journal of Medical Science 2003;57:16-21.

13. Schulte J, Dominguez K, Sukalac T. Declines in low birth weight and preterm birth among infants who were born to HIV infected women during an era of increased use of maternal antiretroviral drugs: pediatric spectrum of HIV disease, 1989-2004. Pediatrics 2007;119:e900-3906.

14. Shelburne S, Visnegarwala F, Darcourt J. Incidence and risk factors for immune reconstitution inflammatory syndrome during highly active antiretroviral therapy. AIDS [date unknown];19:299-406.

15. World Health Organization. Testing and counseling for prevention of mother-to-child transmission of HIV (TC for PMTCT): Reference Guide. [Internet]. 2006;Available from: http://www.womenchildrenhiv.org/wchiv?page=vc-10-00

16. Keating MJ, Flinn I, Jain V, Binet JL, Hillmen P, Byrd J, Albitar M, Brettman L, Santabarbara P, Wacker B, others. Therapeutic role of alemtuzumab (Campath-1H) in patients who have failed fludarabine: results of a large international study. Blood 2002;99(10):3554.

17. McIntosh K, Shevitz A, Zaknun D, et al. Age and Time related changes in extracellular viral load in children vertically infected by human immunodeficiency virus. The Pediatric Infectious Disease Journal. 1996; 15(12): 1087-1091.

18. Palumbo PE, Kwok S, Waters S, Wesley Y, Lewis D, McKinney N, Bardeguez A, Connor EM, Oleske JM. Viral measurement by polymerase chain reaction-based assays in human immunodeficiency virus-infected infants. The Journal of pediatrics 1995;126(4):592–595.

19. Shearer WT, Quinn TC, LaRussa P, Lew JF, Mofenson L, Almy S, Rich K, Handelsman E, Diaz C, Pagano M, others. Viral load and disease progression in infants infected with human immunodeficiency virus type 1. The New England journal of medicine 1997;336(19):1337.

20. Burns D, Mofenson L. Paediatric HIV-1 Infection. The Lancet 1999;354((Supplement II)):1.

21. de Martino M, Tovo PA, Galli L, Gabiano C, Chiarelli F, Zappa M, Gattinara GC, Bassetti D, Giacomet V, Chiappini E, others. Puberty in perinatal HIV-1 infection: a multicentre longitudinal study of 212 children. Aids 2001;15(12):1527.

22. Macmillan C, Magder L, Browers P, et al. Head growth and neurodevelopment of infants born to HIV-1-infected drug-using women. Neurology 2001;57:1402-1411.

23. World Health Organization. Report of the WHO Technical Reference Group, Paediatric HIV/ART Care Guideline Group Meeting. Geneva, Switzerland: WHO Headquarters; 2008.

24. PENTA Steering Committee. PENTA 2009 guidelines for the use of antiretroviral therapy in paediatric HIV-1 infection. HIV Medicine 2009;10(10):591-613.

25. Ananworanich J, Kosalaraksa P, Siangphoe U, Engchanil C, Pancharoen C, Lumbiganon P, Intasan J, Apateerapong W, Chuenyam T, Ubolyam S, others. A feasibility study of immediate versus deferred antiretroviral therapy in children with HIV infection. AIDS Research and Therapy 2008;5(1):24.

26. Violari A, Cotton MF, Gibb DM, Babiker AG, Steyn J, Madhi SA, Jean-Philippe P, McIntyre JA. Early antiretroviral therapy and mortality among HIV-infected infants. New England Journal of Medicine 2008;359(21):2233.

27. Chiappini E, Galli L, Tovo P. Five year follow up of children with perinatal HIV 1 infection receiving early highly active antiretroviral therapy. BMC Infectious Diseases 2009;9:140.

28. Membership I, In VC, Falsey AR, Walsh EE, Capellan J, Gravenstein S, Zambon M, Yau E, Gorse GJ, Edelman R, others. CD4 cell response to antiretroviral therapy in children with vertically acquired HIV infection: is it associated with age at initiation? The Journal of Infectious Diseases 2006;193(7):954–962.

29. Rubin D. Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies. 1990;

30. Greenland S. An introduction to instrumental variables for epidemiologists. International Journal of Epidemiology 2000;29(4):722-729.

31. Shah B, Laupacis A, Hux J, Austin P. Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review. Journal of Clinical Epidemiology 2005;58(6):550-559.

32. Kleinbam D, Klein M. Survival Analysis. Second Edition. Springer-Verlag; 2005.

33. Hernan MA, Brumback B, Robins J. Marginal structural models to estimate joint causal effect of nonrandomized treatments. Journal of the American Statistical Association 2001;96(454):440-448.

34. Fox MP, Brooks DR, Kuhn L, Aldrovandi G, Sinkala M, Kankasa C, Horsburgh R, Thea DM. Role of breastfeeding cessation in mediating the relationship between maternal HIV disease stage and increased child mortality among HIV-exposed uninfected children. International Journal of Epidemiology 2009;38(2):569.

35. Patel K, Hernán M, Williams P, Seeger J, McIntosh K, Dyke R, Seage III G, Pediatric AIDS Clinical Trials Group 219/219C Study Team. Long-term effect of highly active antiretroviral therapy on CD4 cell evoloution among children and adolescents infected with HIV: 5 years and counting. CLIN INFECT DIS 2008;46(11):1751-1760.

36. Patel K, Hernan MA, Williams PL, Seeger JD, McIntosh K, Dyke RBV, Seage III GR, Pediatric AIDS Clinical Trials Group 219/219C Study Team. Long-Term Effectiveness of Highly Active Antiretroviral Therapy on the Survival of Children and Adolescents with HIV Infection: A 10-Year Follow-Up Study. CLIN INFECT DIS 2008;46(4):507-515.

Chapter 2

Review of Causal Inference Methods Used in HIV/AIDS Epidemiologic Studies

2.1 ABSTRACT

Background

To adequately prevent disease spread or treat sufferers of HIV/AIDS, epidemiologists and biostatisticians are charged with finding the true impact of interventions or treatments. The gold standard for testing clinical effectiveness is the randomized controlled trial (RCT). Randomization of treatment assignment is a requisite tool in controlling confounding in RCTs, both known and unknown.(1; 2) When the randomization assumption is satisfied, the estimated effects derived from a study are said to be causal. However, for multiple reasons RCTs cannot often be employed. Using the counterfactual framework more recently reintroduced by Rubin,(3) researchers attempt to define and estimate the causal effects of an intervention or treatment. To help to determine whether a specific causal effect is identifiability from the data (and to even define what it is with regards to a specific intervention on a graph), directed acyclic graphs (DAGs) are often used. These in essence define a non-parametric structural equation model.(4)

Traditional methods for controlling biases, like regression techniques, are only as good as the measured confounder data and how true the model is.(5) Causal inference methods including propensity scores, instrumental variables, marginal structural models, and structural equations are alternative techniques with causal effect interpretations.

The propensity score is defined as a subject's conditional probability of treatment or exposure (arbitrarily defined as one of two possible levels), given the observed potential confounders.(5) The principle of random allocation to treatments in RCTs, instrumental variables (IV) are variables that only affect the outcome through their effect on the treatment or exposure alone (e.g., random treatment assignment).(6) A method proposed by Robins as early as 1997 and later emerging as a significant step forward in so-called causal inference methods is the parameter one estimates from a marginal structural model (MSM).(7) This methods to estimate an MSM can accomdate the presence of time-dependent covariates, often simultaneously intermediate variables and confounders, in the estimation of the marginal ("causal") association of history of exposure or treatment.(8) Often used in psychology research, structural equation modeling (SEM) involves a network of independence assumptions and equations which typically includes a parametric model for a DAG which includes measured and latent variables.(9) This review aims to establish the temporal trends of these causal inference methods used with HIV/AIDS data, the frequency of publications by authors and journals, assess the transparency and quality of the methods employed, and explore the networks of affiliated institutions.

Methods

We performed a systematic review of studies that used causal inference methods in the context of HIV/AIDS research. The initial search strategy collected all publications through December 2008 using Medline.

The interpretability of the study design to measure the causal effect of antiretroviral therapy and an HIV-related outcome was evaluated for each study based on the following 4 fields: 1) traditional interpretability; 2) discussion of statistical analysis-specific assumptions; 3) discussion of confounding measures; and 4) model or instrument selection.

Results

The literature search yielded 70 papers which satisfied the eligibility criteria and were included in this review. Approximately forty-one percent of all HIV/AIDS studies using the listed causal inference methods were published in 2007 and 2008. Approximately half (47%) of the studies using marginal structural models were published in 2007 and 2008.(10-25)

Nearly two-thirds (63.0%) of all studies made comparisons between causal inference results and results using traditional methods; two-thirds (68.5%) of all studies referenced any causal inference method-specific assumptions. Over two-thirds (67.1%) of all studies discussed in detail the instrument or (treatment) model selection.

Over a third of all HIV/AIDS studies using propensity scores (36.8%) did not relate the causal inference results with results using traditional methods or failed to show the benefit of applying these techniques to the study data. Furthermore, propensity score studies were not likely to discuss any causal inference method-specific assumptions (42.1%). Over half of MSM studies (52.8%) stated the method by which the treatment model was selected. Studies using MSMs are more likely to relate causal inference results to those generated from more traditional methods and discuss model assumptions, than studies using any other causal inference method (OR=4.87; 95% CI 1.77-14.71, OR=5.87; 95% CI 1.98-20.23, respectively). Studies using MSMs are less likely to discuss specific model selection than studies using any other causal inference method (OR=0.26; 95% CI 0.08-0.72).

Among MSM studies published before 2007, the largest network is associated with Harvard University with 9 affiliated publications. In 2007 or 2008, the largest networks were among Johns Hopkins and University of California-San Francisco each with 7 publications, and University of California-Berkeley with 6 publications.

Discussion

We have found an increasing trend in appearance of most of these methods as nearly forty percent of studies using one of the listed methods were published in 2007 or 2008. As some of the most technical methods are published more often in journals such as *AIDS* and *American Journal of Epidemiology*, the readership may begin to employ the methods within their own research. Hernan, Robins, and Cole have authored more HIV/AIDS studies using causal inference methods than any other researchers. As a result, their respective affiliated institutions, Johns Hopkins and Harvard University, have some of the largest networks for MSMs. It should be noted, however, that there was a geographical and institutional shift in network size for MSM studies from pre-2007 to 2007-2008. All HIV studies using causal inference methods are deficient by varying

degrees in traditional interpretability, assumption discussion, covariate and confounding discussion, and model and instrument selection discussion. Perhaps the fact that most of the MSM studies make comparisons to traditional methods partly describes their increase in publication frequency.

2.2 Background

As researchers and scientists across the globe continue to work toward reducing the threat of human immunodeficiency virus (HIV) and acquired immune deficiency syndrome (AIDS), they must also constantly react to a changing research environment. Just as the unforgiving mutations of the virus make vaccine development a daunting task for virologists and immunologists, the complexities of HIV/AIDS epidemiology are multifaceted and ever-changing for biostatisticians and epidemiologists. To adequately prevent disease spread or treat patients with HIV/AIDS, epidemiologists and biostatisticians are charged with finding the true impact of interventions or treatments. Unfortunately, the true effects are often obscured by nuisance factors and special analysis techniques are needed to remove these nuisance factors. The analytic techniques that estimate a true causal relationship are referred to as causal inference methods. In this review of causal inference methods used with HIV/AIDS data, we will explore the temporal trends of their appearance in the literature, the frequency of publications by authors and journals, assess the transparency and quality of the methods employed, and explore the networks of affiliated institutions.

Researchers are often tasked with establishing the causal effects of an exposure or treatment on a disease or other outcome. In the perfect world, we would be able to see how X (treatment) affects Y (outcome) by treating everyone to X and following up to see if the Y occurs. Then, we would re-run the experiment under the exact same conditions and have everyone who was once before treated with X now be untreated. The difference we would see in Y between the two scenarios could be causally attributed to the changes in exposure to X. This scenario is difficult to achieve in human studies as other factors can interfere with the X-Y relationship and the exact same conditions are nearly impossible to guarantee. In an attempt to simulate these conditions, the accepted gold standard for testing clinical effectiveness is the randomized controlled trial (RCT). However, often an RCT cannot be performed because of ethical or plausibility reasons and researchers are left with data from observational studies.

2.2.1 Counterfactual Framework

When trying to assert causation, it is often helpful to consider the contribution the counterfactual framework has had on causal inference. The counterfactual is the idea that

an outcome would not have occurred had, counter to fact, a previous event not occurred. Neyman introduced this framework in 1923 and more recently Rubin reintroduced it in 1990(3); it was later made more accessible to the epidemiology community through Greenland's contributions in 1999(26) and 2000.(27) Using studies of HIV or AIDS as a backdrop, suppose we were interested in studying the effects of antiretroviral therapy (ART) ($X=x_{(x=1 \text{ received})}$ or $x_{(x=0 \text{ not received})}$) on death ($Y=y_{(y=1 \text{ yes})}$ or $y_{(y=0 \text{ no})}$). If person A received ART (x=1), he would be followed up to determine whether he would be alive at the end of the follow-up period. Then, this same person A during the same time period would subsequently be untreated with x (x=0) and followed up to determine whether he would survive. In turn, for each subject there can be two treatment scenarios and two possible outcomes: (x=1)-(y=1); (x=1)-(y=0); (x=0)-(y=1); (x=0)-(y=0). However, we can only observe person A's true treatment scenario and outcome and the three remaining unrealized possibilities are known as the counterfactuals. Causal effect methods commonly found in HIV/AIDS observational studies are heavily dependent on the counterfactual framework. Furthermore, the HIV/AIDS counterfactual example can be extended to a practical longitudinal example. Similar to the previous example, let a ϵ {0,1} indicate ART receipt. Under this scenario, person A has two possible outcomes—Y(1) when receiving ART and Y(0) not receiving ART. As such, the causal effect within that individual is the difference between the two outcomes. Furthermore, the average causal effect across the whole population is E{Y(1)-Y(0)}. These potential outcomes are often depicted in terms of a linear model termed a Marginal Structural Model(8):

$$E\{Y(a)\} = \alpha + \beta a \qquad [1]$$

To reconcile the realistic circumstance in which we only have one observable scenario, we can create the situation in which person A would have had equal opportunity to be treated to x. Randomized controlled trials achieve this very goal by ensuring in study design that person A's treatment allocation is not dependent on information connected to outcome Y, also known as the randomization assumption.(28)

2.2.2 Randomization Assumption

Random allocation or randomization of treatment assignment is a requisite tool in controlling both known and unknown confounding in RCTs.(1; 2) When the randomization assumption is satisfied, the estimated effects derived from a study are said to be an estimate of so-called causal associations—confounding is absent.

Randomization reduces the residual or within-group variance and ultimately minimizes the bias from both measured and unmeasured confounding factors.(29) Subsequently, its use allows for a clearer establishment of the intervention's or treatment's effect.(1) It is generally understood that instituting a treatment randomization scheme implies that there are assumed, *a priori*, uncontrolled effects from unknown variables. Furthermore, equal randomization is the most statistically efficient ratio because for any given total sample

size it maximizes statistical power.(30) And, if one were interested in estimating the difference in means, the power of t-test is affected in the case of large treatment imbalances, (e.g. 70% in one arm versus 30% in another arm).(30)

2.2.3 Introduction to Directed Acyclic Graphs

Directed acyclic graphs are nonparametric causal models which illustrate the relationships between variables involved in a study. By definition, variables, or nodes, represented in DAGs cannot both affect and be affected by another variable at the same time.(31) In theory it would be possible to depict variables over time as so-called feedback loops represent a coarsening of a DAG that more finely represents cause and effect over time. In the HIV example above, assume patients living in certain neighborhoods were more likely to die from HIV/AIDS and also were less likely to be treated with ART. Figure 2.1 demonstrates this confounding scenario. In this situation, zip code would be a confounder of the ART-death relationship. In fact, zip code not only directly affects death, but also indirectly affects death through ART. Thus, zip code-death forms a directed path, as does also the zip code-ART-death path.

Backdoor Pathways in DAGs
In Figure 2.1, if one were interested in the direct effect of ART on death, one would have to explore whether the presence of a backdoor path is confirmed. That is to say, in the present example is there a path from ART, against the flow, to another variable that is associated with death? In fact, zip code represents a variable that should be addressed in the analysis of ART's direct effect on death to ensure that the randomization assumption does not fail. It should be noted, however, when looking at ART's direct effect on death, the pathway death-nutrition-ART does not represent a backdoor path because the path would have to go against the flow of not only the first path from nutrition but also the path from ART-nutrition.

Variable Adjustment in DAGs
DAGs are particularly useful in determining which variables are important in the analysis. In the Figure 2.1 representing simple confounding, zip code's effect on that relationship will have to be removed to realize an unbiased direct effect of ART on death. Essentially, we are setting zip code at a fixed level such that it has no influence on the ART-death relationship. Furthermore, assume that one of the major side effects of initiating ART is a loss of appetite. In this scenario, when person A initiates ART he would have worse nutrition than those not starting ART. Furthermore, a poor diet in immune-compromised individuals leads to poor survival outcomes. In this respect, the direct effect of ART on survival is modified by nutrition. Again, to estimate an unbiased direct effect of ART on death, apart from appetite, this effect modification will need to be addressed.

In contrast, calendar year is neither a confounder nor an effect modifier. As such, calendar year is an external time-dependent variable that cannot be affected by the severity of disease (an indication for ART initiation). Moreover, Figures 2.1 assumes

that calendar year is independent of death given ART use. This external variable will be introduced again later in the discussion of instrumental variables.

2.2.4 Traditional Methods for Confounding Adjustment

The traditional methods for controlling the biases introduced by measured confounders include using multivariable regression techniques. However, regression models are parametric and depend heavily on how true the model is. In the best-case scenario with a presumably correctly identified coefficient, the covariates within the model are arbitrarily selected, thus biasing the coefficient estimate and resulting in model misspecification. Moreover, model building can become cumbersome even with comprehensive confounder data collected; often investigators will settle for parameter interpretability over adequate adjustment for bias.

2.2.5 Causal Inference Approaches for Confounding Adjustment

Causal inference methods have been developed in order to help reconcile these problems commonly seen in HIV epidemiological research and observational studies with high-dimensional studies in general. We will review 5 causal inference methods that have relevance to addressing common data and statistical problems found in HIV/AIDS studies: propensity score matching, instrumental variable approaches, estimation of marginal structural models and structural equation models, and accelerated failure time models.

Randomized Controlled Trials and Causal Inference

Unlike an RCT with a properly performed randomization procedure, observational studies are inherently affected by confounding bias. While an RCT randomizes study participants to receive treatments, the physicians or the participants themselves select receipt of treatments in an observational study. Therefore, in the latter scenario, an argument for causality is difficult because the effect seen could either be a result of the treatment or it could be a result arising from the reason for selecting the treatment.(6) In the presence of noncomplicance, data from RCTs are most often analyzed using the intention to treat (ITT) principle. Under this rule, once person A is randomized to a treatment X, he should be included in any future analyses comparing treatment assignment arms as if he actually received treatment X despite actual receipt of treatment X. In this respect, ITT is actually determining the effect of assigning person A to treatment X. If randomization failed or treatment assignments were rarely followed, non-compliance adjustments are needed to ensure the results are a true reflection of the causal effect of treatment X.

Propensity Scores

In the absence of treatment assignment, observational studies are inherently more prone to confounding bias than controlled trials. Propensity scores (PS) were developed in 1983 by Rosenbaum and Rubin to control for known confounding bias and yet preserve parameter interpretability. The propensity score is defined as a subject's conditional probability of treatment or exposure as opposed to another treatment or exposure, given the observed potential confounders.(5) To estimate the propensity score, a logistic regression can be used with the confounders as the independent variables and the treatment as the dependent variable. As such, two subjects with the same propensity score will have an equal estimated probability of treatment. Propensity scoring emulates the randomization procedure of RCTs, as the distribution of confounders between treatment groups is similar. It is important to note that the propensity model selection should consider the best balance of confounders between treatment arms.(5) It should be noted that propensity score adjustment assumes (like all but instrumental variable methods) that there is no unmeasured confounding to bias the results.

In 1999, McLaughlin et al published what is thought to be the first HIV/AIDS study using propensity score methods.(32) The researchers used propensity scores to adjust for confounding in the estimated effect of zidovudine (ZDV) and pneumocystis carinii pneumonia (PCP) prophylaxis on hospitalizations and death. Specifically, the authors used propensity scores for ZDV and PCP prophylaxis for each subject in order to adjust the estimated effect of ZDV and PCP prophylaxis on hospitalizations and mortality. Propensity scores were entered into the proportional hazards models as covariates. Outcomes were measured within proportional hazards regression models; the authors found that the adjusted relative risk of death associated with ZDV was 36% (95% CI 0.2-0.4). among treated patients when compared to untreated patients. The adjusted relative risk of death among subjects treated with PCP prophylaxis was 49% (95% CI 0.3-0.8) when compared to subject untreated with PCP prophylaxis.

Instrumental Variables

The use of instrumental variables (IV) can be dated back over a half-century when they have been found in econometric theory.(33) Like the principle of random allocation to treatments in RCTs, IVs are variables that only affect the outcome through their effect on the treatment or exposure alone.(6) Moreover, referring once again to the counterfactual framework and the randomization assumption, all counterfactual observations are independent of the process of treatment allocation. Additionally, the variation in the identified instrument is assumed to be substantial enough to cause variation in the treatment. Instrumental variable methods can actually control confounding, measured or unmeasured, under the specific assumption that the outcome is conditionally independent given the variable of interest. Though this assumption may not be directly testable, the use of DAGs helps justify the use of a specific IV. In Figure 2.1, the implicit assumption is that calendar year is only related to the outcome (death) through the exposure (ART). Thus, given the assumptions of the graph, calendar year presumably makes for a possible instrumental variable. As a result, the researcher will be able to estimate how much the variation in ART that is explained by the calendar year affects death.

The earliest application of IVs analyzing HIV/AIDS data and meeting my inclusion criteria was in 2001.(34) In a study of population effectiveness of ART in reducing AIDS diagnoses in an HIV positive population, Tarwater et al used calendar year as an external time-dependent variable.(34) This allowed the researchers to account for different infection durations. Specifically, the authors compared the incidence of AIDS in subjects who had the same disease duration in different eras of ART. The relative hazard in the no therapy era was 1.52 (95% CI 0.93-2.49) and 0.30 (95% CI 0.18-0.51) in the HAART era.

Marginal Structural Models

A method proposed by Robins as early as 1992 and later emerging as a significant step forward in causal inference methododology are proposed estiamtors for marginal structural model (MSM).(35) Basic methods can be divided in estimation equation approaches (the so-called inverse probability of treatment weighting-IPTW-and its double-robust extension) and graphical computation (G-comp) approaches. Much like all causal inference approaches, MSMs assume no unmeasured confounding. Additonally, another important assumption is the experimental treatment assignment (ETA) assumption,(36) which is a scenario when at each time point, there is a covariate level l_k at which all patients either receive or do not receive the identical treatment.

In 2000, Hernan et al used MSM to estimate the causal effect of ZDV on survival of HIV-positive men.(37) In an attempt to eliminate bias from time-dependent confounding, the authors used weights to obtain IPTW partial likelihood estimates. Specifically, the marginal structural Cox model yielded a mortality rate ratio of 0.7 (95% CI 0.6-1.0) for ZDV use after adjusting for CD4 count and other time-dependent covariates.

Source of Inference in Marginal Structural Models

Some options for producing confidence intervals for the MSM parameters include robust or sandwich estimators. Bootstrapping also helps to make a case for a probability-based inference about an effect based on an estimated effect using a population-based sample.(7; 38) Rather than making assumptions about the population, the researcher can make conclusions about a population's characteristics using the data from the sampled population.

Structural Equation Models

Often used in psychology research, structural equation modeling (SEM) involves a network of independence assumptions and equations.(9) In this network of equations, each variable may only appear as a dependent variable once, but may appear in any equation as a causal variable. As such, the network of equations allows the researchers to see how each dependent variable changes as its causal variables change. The response variables in this network of equations are referred to as endogenous; all others are exogenous.(9)

In 1991, van der Welde et al published the first SEM analysis within an HIV/AIDS study population. The authors found that when studying AIDS_related health behavior, it may

be beneficial to explore not only one's motivation to protect one's self but also to include other variables such as social norms and previous behavior attributes.

2.3 Methods

2.3.1 Literature Search and Study eligibility

I performed a systematic review of studies that used causal inference methods in the context of HIV/AIDS research. The initial search strategy collected all publications through December 2008 using Medline. Articles containing the textwords "propensity", or "instrumental AND variable", or "marginal structural model", or "structural equation model" and indexed to include "HIV" or "AIDS" textwords were selected with MeSH subject headings "Acquired Immunodeficiency Syndrome", "AIDS-Associated Nephropathy", "AIDS Dementia Complex", "AIDS Serodiagnosis", "AIDS-Related Opportunistic Infections", "AIDS-Related Complex", "AIDS Vaccines", "HIV Seropositivity", "HIV Long-Term Survivors", "HIV", or "HIV Infections". Once the eligible articles were identified, a cross-reference search using Web of Science was performed. All studies citing the included Medline publications were included for the initial review. Following identification of eligible cross-referenced publications from Web of Science, a final search was performed on the bibliographies of methods-based researchers who are either first or senior authors of more than 2 eligible publications. These authors include: Cole S, Hernán M, van der Laan M, Petersen M, and Robins J. The eligibility criteria are summarized in Table 2.1. All analyses were performed using R.(39)

2.3.2 Qualitative Study Assessment

The interpretability of the study design to measure the causal effect of antiretroviral therapy and an HIV-related outcome was evaluated for each study based on the following 4 fields: 1) traditional interpretability; 2) discussion of statistical analysis-specific assumptions; 3) discussion of confounding measures; and 4) model or instrument selection. The rubric used to assess these fields is illustrated in Table 2.2.

Traditional Interpretability

To assess the interpretability of results from HIV/AIDS studies that employ causal inference methods, epidemiology and biostatistics literacy criteria were established. Our aim was to explore the frequency HIV/AIDS researchers compare their causal inference results to results they would have achieved had they used traditional methods. With studies using propensity scores, we were also exploring whether the researchers demonstrated the benefit of using propensity adjustment by showing the distribution of

covariates before and after adjustment. With the remaining methods, contrasts of results from traditional and causal inference methods were evaluated.

Statistical Analysis Assumptions

All statistical methods have important assumptions that should either be tested or at least discussed in papers using them, particularly if the methods are more contemporary or highly specific. Often studies will not list the methods-specific assumptions, but the researchers will rather acknowledge the assumptions through their discussion of methods or limitations. For example, in a study using propensity scores, the assumption of no unmeasured confounding may not be explicitly identified, but the researchers may discuss controlling for all known confounders. Instrumental variables are assumed to be independent of the outcome and assumed to have enough variance to induce variance in the treatment. Marginal structural models make several assumptions, to which a study employing these methods should at least make a reference. Among these assumptions are the ETA assumption described previously, no unmeasured confounding or randomization assumption, and appropriate model specification.(8) In structural equation models, exogenous variables are assumed independent.(9)

Confounding Measures

As estimating the causal effect of x on y while minimizing bias is the main objective of causal inference methods, it is important for studies to identify specifically which covariates may bias this causal effect. A general discussion of confounder adjustment may not help future studies trying to expound on the study's results.

Model or Instrument Selection

It is not enough to simply employ these methods to control for confounding and reduce bias. The technique by which the researcher selects his model or instrument is just as important as the researcher recognizing *a priori* the necessity to perform the causal inference. In studies using propensity scores, a case for the specific variables used to model the propensity of treatment is helpful to understand their influence on the estimated effect and control of bias. Similarly, the variables included in the treatment model for marginal structural models should not only be identified but also justified. Some of the model selection procedures may be as basic as an acknowledged prior knowledge or research, but may also include stepwise addition or deletion techniques, Akaike information criterion, or super learning applications like Deletion/Substitution/ Addition (DSA) algorithm. A discussion about the ways the authors selected the included variables (e.g. based on prior studies) is enough to satisfy this criterion. A justification for using a specific instrument in studies using instrumental variables is necessary to understand its influence, or lack thereof, on the estimated causal effect.

2.4 Results

2.4.1 Literature Search

The initial search on Medline, Web of Science, and selected bibliographies yielded 1535 potential studies, of which 932 were later removed based on publication date, duplication, title, or abstract details. The remaining 603 papers were reviewed for eligibility and 70 satisfied the eligibility criteria and were included in this review. The selection flow is detailed in Figure 2.2.

2.4.2 Temporal Trends

Though some of the included causal inference methods were developed by the onset of the HIV epidemic, the highest concentration of these methods in the HIV/AIDS literature was not apparent for about two decades. Propensity scores were introduced near the early stages of the HIV epidemic, however the first appearance of these methods in the HIV/AIDS literature was in 1999 (McLaughlin et al).(32) It was another four years before the next publication using propensity scores appeared in 2003.(40) The trend in propensity score publications is markedly increasing as about 42% of all HIV studies using these methods were published in 2007 and 2008.(41-48) The appearance of instrumental variables in HIV/AIDS literature occurred in 2001 (Tarwater et al).(34) Two years later the next instrumental variables publication appeared in the literature (Bhattacharya et al).(49) Though only seven instrumental variable studies are known to have been published before 2009,(6; 24; 34; 49-52) four of the seven were published in the last four years.(6; 24; 51; 52) In 2000 the first MSM study appeared in the literature (Hernan et al),(37) and every subsequent year saw an increase in publications using these methods on HIV/AIDS data. Approximately half (47%) of the studies using marginal structural models were published in 2007 and 2008.(10-14; 16-25; 53) The use of structural equations on HIV/AIDS data has its origin in 1991 (Van der Velde et al).(54) Only one other publication using this method would appear in the 1990s,(55) while the majority (54.6%) was published since the end of 2005.(56-61) Approximately forty-three percent of all HIV/AIDS studies using the listed causal inference methods were published in 2007 and 2008. Temporal trends in the appearance of causal inference methods in HIV/AIDS publications are described in Table 2.3 and illustrated in Figure 2.3.

2.4.3 Causal Inference Study Epidemiology and Publication Characteristics

Figure 2.4 is a dot-chart showing the frequency of causal inference publications analyzing HIV/AIDS data among all affiliated researchers with at least two included papers. Hernan has the most affiliated publications with 12 included studies; Cole had 10 and Robins had 9 associated publications. Figure 2.5 is a dot-chart illustrating the frequency of publications among first authors. Petersen had the most first authorships with 5 publications, followed by Hernan with 4 and Cole with 3 publications. About one in six (17.1%) of all included studies was principally authored by one of these three researchers. It should be noted, however, that all of the publications that were principally authored by these researchers employed marginal structural models. Robins had the most senior authorships with 7 publications, followed by van der Laan with 3 publications. See Figure 2.6 for a dot-chart illustrating the frequency of publications among senior authors.

Over one-third (37.1%) of all the studies were published in *AIDS, American Journal of Epidemiology*, *Journal of Acquired Immune Deficiency Syndromes*, or *Statistics in Medicine* (see Figure 2.7). The most common exposure or outcome was antiretroviral therapy or HIV disease progression (64.3%).

1. Epidemiology of Studies Using Propensity Scores

All studies using propensity scores had unique first authors and all but two (*AIDS*) were published in unique journals. Antiretroviral therapy or HIV disease progression were the most common exposure or outcome studied (47.4%). High-risk behaviors were the second most common exposure or outcome studied (21.1%) in publications using propensity scores (see Table 2.4).

2. Epidemiology of Studies Using Instrumental Variables

All studies using instrumental variables had unique first authors and all but two (American Journal of Epidemiology) were published in unique journals (see Table 2.5). Additionally, all but one of these studies explored the impact of ART on disease progression or high-risk behavior. Bhattacharya et al explored the relationship between type of insurance (private versus public) and HIV-related mortality.(49)

3. Epidemiology of Studies Using Marginal Structural Models

Among publications using marginal structural models, several researchers were the primary authors on more than one publication (see Table 2.6). Peterson was the first author of five HIV/AIDS studies using marginal structural models,(13-16; 23) Hernan was the primary author on four HIV/AIDS studies,(37; 62-64) and Cole(11; 65; 66) was the primary author for three studies. Lopez-Gatell,(12; 20) Brumback,(40; 67) and Patel(21; 22) were the primary authors for two studies each. Over a third (36.1%) of all HIV/AIDS studies using MSMs were published in *AIDS* or *American Journal of*

Epidemiology. Antiretroviral therapy or HIV disease progression was the most common exposure or outcome studied (80.6%).

4. Epidemiology of Studies Using Structural Equation Models

Studies using structural equation models were authored by unique authors and generally published in different journals, though *AIDS and Behavior* had two publications (see Table 2.7). All publications using structural equation models explored either ART adherence or high-risk behaviors.

2.4.4 Summary of Study Quality Assessment Results

Application of the quality assessment tool for these studies revealed that the most common weaknesses were traditional interpretability, a discussion of model or instrument selection, and a discussion of assumptions, though these results are highly method-dependent. The results are described in detail in Tables 2.8-2.11. Figure 2.8 illustrates the proportion of causal inference studies which satisfied the specific study assessment criteria.

Nearly two-thirds (63.0%) of all studies made comparisons between causal inference results and results using traditional methods; two-thirds (68.5%) of all HIV/AIDS studies referenced any causal inference method-specific assumptions. Nearly all studies (95.9%) had a discussion about the type of confounding being controlled for using the causal inference method and listed the specific potential confounders. Over two-thirds (67.1%) of all studies discussed in detail the instrument or (treatment) model selection.

1. Assessment of Studies Using Propensity Scores

Over a third of all HIV/AIDS studies using propensity scores (36.8%) did not relate the causal inference results with results using traditional methods or failed to show the benefit of applying these techniques to the study data. Furthermore, propensity score studies were not likely to discuss any causal inference method-specific assumptions (42.1%). As the primary goal of using propensity scores is to control for confounding and selection bias, nearly all (89.5%) discussed specifically which potential confounders they were controlling with propensity scores. Over two-thirds of propensity score studies discussed the procedure of selecting the treatment model.

2. Assessment of Studies Using Instrumental Variables

Studies using instrumental variables were likely to have causal effects compared to traditional methods (71.4%). Additionally, studies using instrumental variables were just as likely to discuss the method-specific assumptions made as specific instrument selection (85.7%). All instrumental variable studies discussed in detail the specific confounding being controlled for in the study.

3. Assessment of Studies Using Structural Equation Models

No studies using structural equation models made comparisons of their results with results from traditional methods. Nearly half (45.5%) of these studies discussed their inherent assumptions. All studies using structural equations discussed their specific mediating variables and model selection.

4. Assessment of Studies Using Marginal Structural Models

Studies using marginal structural models were likely to have causal effects compared to traditional methods (80.6%). Additionally, studies using marginal structural models were likely to discuss the method-specific assumptions made (86.1%). Nearly all MSM studies discussed in detail the confounding variables they aimed to control (97.2%). Just over half of MSM studies (52.8%) stated the method by which the treatment model was selected.

2.4.5 Temporal Trends and Year of Publication

Approximately seventy percent of all studies analyzed data based in the United States. The temporal trends suggest that these causal inference methods are being applied to non-US data more often in 2007 and 2008 than in previous years (OR=1.55; 95% CI 0.56-4.29), though this relationship is not statistically significant. Among propensity score studies specifically, the most recent publications seem to be more likely from non-US data than in previous years of publication (OR=2.67; 95% CI 0.40-20.25). A similar trend is found among studies using MSMs published in 2007 or 2008 as these studies are analyzing non-US data less often than studies published before 2007 (OR=1.17; 95% CI 0.26-5.17).

2.4.6 Traditional Interpretability and Year of Publication

Over all studies, the year of publication does not seem to affect a study's probability of making comparisons between causal effects and effects from traditional methods (OR=0.69; 95% CI 0.26-1.81). Propensity score studies published in calendar years 2007 or 2008 appeared to be just as likely to have a traditional interpretability component as papers published before 2007 (OR=0.95; 95% CI 0.14-6.73), though counts are too small to attain statistical significance. Counts are too small for any reasonable estimates to be made from method-specific studies using instrumental variables or structural equations. No study using structural equations made any comparisons with traditional methods. However, among studies using MSMs, year of publication does not seem to have any effect on a study's likelihood for making comparisons with traditional methods (OR=0.28; 95% CI 0.04-1.55).

2.4.7 Discussion of Assumptions and Year of Publication

Among all HIV/AIDS studies using a causal inference method, the year of publication does not appear to have an effect on the probability of those studies discussing method-specific assumptions (OR=1.22; 95% CI 0.45-3.43). Studies using propensity scores published in 2007 or 2008 seem to be more likely to reference method-specific assumptions than propensity score studies published before 2007 (OR=1.75; 95% CI 0.27-11.92), though counts are too small to attain statistical significance. Calendar year of publication of studies using MSMs does not appear to affect the likelihood of those studies discussing MSM-specific assumptions (OR=1.41; 95% CI 0.21-11.83).

2.4.8 Discussion of Confounding and Year of Publication

Calendar year of publication of all causal inference studies does not seem to have an effect on whether the study will address specific confounding controlled for (OR=0.35; 95% CI 0.02-3.86). Specifically, calendar year of publication of studies using propensity scores does not seem to have an effect on whether the study will address specific confounding controlled for (OR=0.70; 95% CI 0.02-19.75). Nearly all MSM studies discussed their specific confounding concerns.

2.4.9 Discussion of Treatment Model/ Instrument Selection and Year of Publication

Calendar year of publication of all causal inference studies does not seem to have an effect on whether the study has a discussion about model or instrument selection (OR=0.82; 95% CI 0.30-2.20). Specifically, calendar year of publication of studies using propensity scores does not seem to have an effect on whether the study will address model selection (OR=1.71; 95% CI 0.24-15.75). Calendar year of publication of studies using MSMs does not appear to affect the likelihood of those studies discussing model selection (OR=0.65; 95% CI 0.17-2.41).

2.4.10 Comparison of Studies Using MSMs Vs Any Other Causal Inference Method

Studies using MSMs are just as likely to be using non-US data as studies using any other causal inference method (OR=0.80; 95% CI 0.29-2.18). Studies using MSMs are more likely to relate causal inference results to those generated from more traditional methods than studies using any other causal inference method (OR=4.87; 95% CI 1.77-14.71). Studies using MSMs are more likely to discuss specific model assumptions than studies using any other causal inference method (OR=5.87; 95% CI 1.98-20.23). Studies using MSMs are seemingly more likely to discuss specific confounding than studies using any other causal inference method (OR=2.00; 95% CI 0.18-44.18). Studies using MSMs are less likely to discuss specific model selection than studies using any other causal inference method (OR=0.26; 95% CI 0.08-0.72).

2.4.11 MSM-Specific Study Assessment

Among studies using MSMs, whether the authors stated specifically how they estimated their confidence intervals or standard errors for the causal effect of interest was investigated. Approximately eighty-six percent of all MSM studies reported their source of inference (see Table 2.12). Specifically, approximately twenty-two percent used the bootstrapping method to estimate their standard errors or confidence intervals, fourteen percent used the "sandwich" method, eight percent used generalized estimating equations, and forty-two percent used a non-specific robust method.

There are three common methods of estimating the parameters in marginal structural models: G-comp, double robust, or IPTW. Though all three of these methods control for confounding, albeit in different ways, only two studies used any method other than IPTW to estimate the MSM parameters.(15; 16) Both of these studies used all three methods to compare results.

2.4.12 Network of Publishing Institutions

The network of institutions represented by all the listed authors is illustrated in Figures 2.9a-2.12b. Table 2.13 is a supplemental table numerically showing the frequency of publications from each affiliated institution (with a minimum of 2). Institution abbreviations are detailed in Appendix A.1.

Networks of Marginal Structural Model Studies
Among MSM studies published before 2007 (Figure 2.9a), the largest network is associated with Harvard University with 9 affiliated publications. The next largest networks are associated with University of Washington with 5 papers, and University of California-San Francisco and Johns Hopkins University each with 4 affiliated publications. Among MSM studies published in 2007 or 2008 (Figure 2.9b), the largest

networks were among Johns Hopkins and University of California-San Francisco each with 7 publications, followed by University of California-Berkeley with 6 publications.

Networks of Propensity Score Studies
Among propensity scores studies published before 2007 (Figure 2.10a), the largest network is associated with University of California-Los Angeles with 4 publications, followed by Johns Hopkins University with 3 affiliated publications. The networks for propensity score papers published in 2007 or 2008 are shown in Figure 2.10b. The Centers for Disease Control and Prevention has the largest network with 3 affiliated publications among these publications.

Networks of Instrumental Variables Studies
Among studies using instrumental variables published before 2007 (Figure 2.11a), the largest network is only two; University of Pittsburgh, Northwestern University, Johns Hopkins University, University of California-Los Angeles, and RAND all have two affiliated publications before 2007. Among instrumental variables studies published in 2007 or 2008 (Figure 2.11b), no network is larger than 1.

Networks of Structural Equation Model Studies
Among studies using structural equations published before 2007 (Figure 2.12a), the largest network is only two—the University of Miami. Among structural equation model studies published in 2007 or 2008 (Figure 2.12b), no network is larger than 1.

2.5 Discussion

We have performed a systematic review of the HIV/AIDS literature for studies using causal inference methods including propensity scores, instrumental variables, marginal structural models, and structural equations. We have found an increasing trend in appearance of most of these methods as over forty percent of studies using one of the listed methods were published in 2007 or 2008. Compared to all other methods, publications using MSMs had the highest proportion published in 2007 or 2008 (47.2%), followed by propensity score studies (42.1%). HIV/AIDS studies using instrumental variables and structural equation models have not seen the resurgence that other studies using other causal inference methods have.

The journals in which the studies were published may have some impact on the method's future use. Likely due to method-specific technical issues and readership, some methods are more often found in statistical or economics journals (e.g. IVs). Moreover, IVs have their origins in economics and have yet to be adopted as a common technique in epidemiology.(33) Figure 2.7 illustrates the frequency of appearance of causal inference publications in specific journals. The highest frequencies are found in the journals *AIDS*, *American Journal of Epidemiology*, *Journal of Acquired Immune Deficiency Syndrome*, and *Statistics in Medicine*. As some of the most technical methods are published more often in journals such as *AIDS* and *American Journal of Epidemiology*, the readership may begin to employ the methods within their own research.

Some authors and affiliated institutions have contributed greatly to the dissemination of causal inference methods used for HIV data. As described in Figure 8, Hernan, Robins, and Cole have authored more HIV/AIDS studies using causal inference methods than any other researcher. As a result, their respective affiliated institutions, Johns Hopkins and Harvard University have some of the largest networks. It should be noted, however, that there was a geographical and institutional shift in network size for MSM studies from pre-2007 to 2007-2008 as Johns Hopkins, University of California-San Francisco, University of California-Berkeley, and University of California-Los Angeles had the largest networks most recently, while Harvard University was the most prolific producer of MSM studies prior to 2007.

The study assessments found that, regardless of year of publication, all HIV studies are deficient by varying degrees in traditional interpretability, assumption discussion, covariate and confounding discussion, and model and instrument selection discussion. Over all studies, traditional interpretability was the most common deficiency, but this is likely due to no study using structural equation models making any comparisons to traditional methods. Though not seen in our review, as a method has been used long enough, a traditional results comparison may be less important in the eyes of the researcher as limited space may be predicated on other results. Studies using IVs were most deficient in having a traditional interpretability component. While this is likely a reflection of the difficulty in making such comparisons with these methods, it remains an important aspect of promoting the use of a method.

Often some assumptions are not testable, but acknowledging them in the analysis should be policy to ensure study validity. Some assumptions, like the experimental treatment assignment (ETA) for MSMs, are often testable(36) and should be adequately described. Publications using propensity scores were most deficient in discussing inherent methods-specific assumptions.

The most deficient area for studies using MSMs was a discussion of their treatment model selection. Not only is this key for other researchers to understand fully how to implement MSMs, but it is also particularly important to ensure unbiased results. If incorrect covariates are used in building the treatment model, biased estimators are possible.(68) In particular, the inclusion of variables which predict only treatment, i.e. not confounders, can affect the estimator's performance. Regarding confounders and other covariates, most studies, regardless of causal inference method, described or listed the variables the method hoped to control for. In turn, future studies of similar research questions may be able to control for similar confounding.

Most studies using MSMs reported robust or sandwich confidence intervals. By explicitly stating the source of their confidence intervals, researchers once again have the opportunity for not only making a case for a true causal effect, but allow other researchers to learn from their methods.

Studies using MSMs are more likely to discuss their results as they relate to conventional methods than studies using any other causal inference method. Moreover, MSM-based HIV publications are more likely to discuss the (treatment) model selection than the treatment or instrument model in any other causal inference method. Though in comparison to the most technically difficult methods MSMs are not terribly difficult to understand, perhaps the fact that most of the MSM studies make comparisons to traditional methods partly describes their increase in publication frequency.

There are several limitations concerning this review that should be considered. Firstly, though every attempt was made to capture all relevant studies, some may have been missed. A problem encountered when performing this review was that often studies will employ a method that is technically an included causal inference method, but the researchers fail to identify the method as specifically one of the included methods. We did not include these studies, as it would not have been reasonable to capture all of these studies using unidentified causal inference methods. In turn, we could have ended up with a skewed sample of only the studies with an identified causal inference method found on Medline and studies with an unidentified causal inference method found by our cross-referencing search. Secondly, as inherent in any review, we have to consider publication and author bias as a potential issue. If a prominent researcher is one of the authors on a causal inference paper, it may be more likely to be published. As the techniques and methods increase in complexity, many technical experts are required to collaborate. As a result, multiple institutions will likely have representation as either the problem necessitates or as the method becomes more recognized.(69) Though publication bias is likely not the case regarding these publications and their concentration of authorships and institutions, some researchers believe that a looming threat to publication privilege is, for example, the department chair who wants his name on every publication from within his department.(70) Thirdly, non-health related journals, like economics journals, might have more studies with HIV data. These results may not be generalizeable to studies of other diseases. In fact, the trends found in HIV studies may not be indicative of the trends among *all* epidemiologic studies. Some of the study assessments may not be as high as they could be because authors would reference previous work on the data in an attempt to avoid discussing technical details. Journal limitations may prevent authors from discussing further. It should be noted that a possible limitation to our dot-chart describing the contributions of senior authors may be more of a misrepresentation issue. That is to say, sometimes the last author is the researcher who either contributed the least amount to the study or simply received a gift authorship with no specific research role.(71) Moreover, in some cases as some prominent journals have limits on the number of authors allowed, the author occupying the senior author spot may not be the senior author at all. In fact, recently researchers found a relationship between journals with authorship limitations and the number of databases accessing those particular journals,(72) which would have limited our ability to find these studies in our search.

Tables

Table 2.1 Causal Inference Study Eligibility Criteria

1) Prospective or retrospective study published in peer-reviewed journal before calendar year 2009

2) At least 1 of the following endpoints was ascertained:
   i. Incident AIDS
   ii. Incident HIV
   iii. Stage of HIV disease
   iv. Response of HIV/AIDS to therapy
   v. HIV/AIDS disease progression
   vi. Death

Or, behavioral or clinical risks were assessed within an exclusively HIV positive population.

3) One of the causal inference techniques described above explicitly stated and used in analysis (propensity scores, instrumental variables, MSM, structural equation model)

4) Results from application of method to dataset (not a subset of data for illustrative purposes) are published and have not been published previously elsewhere

Table 2.2: Quality Assessment Rubric Tool Applied To All Studies

Traditional Interpretability	**Is the method described as appropriate when compared to other methods? Or, are results from both causal inference and traditional methods given? (Alternatively, are before- and after-propensity score analyses done?)**
Discussion of Statistical Analysis-Specific Assumptions	**Are assumptions discussed generally and as they apply to the study data?**
Discussion of Confounding Measures	**Are the specific confounders the causal inference method aims to control or the type of (potential) confounding adjusted for discussed? This may also include mediating factors found in SEMs.**
Model and Instrument Selection	**Is the model or instrument selection technique discussed?**

Table 2.3 Trends of Causal Inference Methods Used in Studies of HIV and HIV Risk

Statistical method	Total	No. of articles	No. of articles	No. of articles	No. of articles	No. of articles
		Pre-00	2001-02	2003-04	2005-06	2007-08
Propensity scoring[±,1]	19 (26.0%)	1 (5.3%)	0 (0.0%)	3 (15.8%)	7 (36.8%)	8 (42.1%)
Instrumental Variables[±,2]	7 (9.6%)	0 (0.0%)	1 (14.3%)	2 (28.6%)	2 (28.6%)	2 (28.6%)
Marginal Structural Models[±]	36 (49.3%)	1 (2.8%)	3 (8.3%)	7 (19.4%)	8 (22.2%)	17 (47.2%)
Structural Equation Models[±]	11 (15.1%)	3 (27.3%)	0 (0.0%)	2 (18.2%)	2 (18.2%)	4 (36.4%)
Total[3]	73	5 (6.8%)	4 (5.5%)	14 (19.2%)	19 (26.0%)	31 (42.5%)

[±]Proportions reported are among studies using that statistical method
1. One of these publications is also included in Marginal Structural Models
2. Two of these publications are also included in Marginal Structural Models
3. Sum of total is more than sum of included studies due to papers using multiple methods

Table 2.4: Propensity Score-specific study epidemiology grouped by exposure or outcome keywords

	Lead Author	Journal	Exposure	Primary outcome measure
Propensity Scoring				
Antiretroviral Therapy (ART) or Disease Progression	Sanguan-wongse(41) (2008)	Journal of Acquired Immune Deficiency Syndromes	ART	Time to death
	Potard(42) (2007)	Antiviral Therapy	HAART	HIV RNA, CD4
	Braithwaite(43) (2007)	AIDS	ART initiations	Adherence, change in HIV RNA, CD4, HIV RNA suppression
	Nosyk(73) (2006)	BioMed Infectious Diseases	HAART	Hospital readmission
	Merito(74) (2006)	European Journal of Health Economics	Time of ART initiation	Cost effectiveness, incident AIDS, death
	Liu(75) (2006)	AIDS Research and Therapy	HAART	Quality of life

(Continued on next page…)

Table 2.4 Propensity Score-specific study epidemiology grouped by exposure or outcome keywords (continued)

	Lead Author	Journal	Exposure	Primary outcome measure
Antiretroviral Therapy (ART) or Disease Progression	Liu(76) (2006)	Quality of Life Research	HIV and use of HAART	Quality of life
	Chu(77) (2005)	American Journal of Epidemiology	Hormonal contraception	CD4 cell count, viral load
	Brumback(40) (2003)	Biometrics	ART	CD4 cell count
High-Risk Behaviors	Zule(48) (2008)	Drug and Alcohol Dependence	Needle sharing	HIV, HCV
	El-Bassel(78) (2005)	Social Science and Medicine	High-risk sex; intimate partner violence	Intimate partner violence, sexual risk related factors
	Wenzel(79) (2004)	Prevention Medicine	Shelters vs low income housing	Physical or sexual abuse, substance abuse, HIV risk behavior
	Rotheram-Borus(80) (2003)	Prevention Science	HIV prevention programs	High risk behavior, drug use
Health Care or Prevention Programs	Mahal(47) (2008)	AIDS	HIV/AIDS	Healthcare utilization, spending, lost income
	Gangopad-hyay(81) (2005)	Sexually Transmitted Diseases	STI/HIV programs	STDs
	McLaughlin(32) (1999)	International Journal of Quality in Health Care	Primary source of and access to care	Time to treatment, hospitalization, death
Other	Anuwatnon-thakate(46) (2008)	PLoS ONE	Directly observed therapy	TB outcomes
	Tai(45) (2007)	Journal of Infectious Diseases	Pregnancy	AIDS defining illness, death
	Albalak(44) (2007)	Archives of Internal Medicine	Time	TB/HIV coinfection

Table 2.5 Instrumental Variable-specific study epidemiology grouped by exposure or outcome keywords

	Lead Author	Journal	Exposure	Primary outcome measure
Instrumental Variables				
ART or Disease Progression	Shiels(24) (2008)	Journal of Acquired Immune Deficiency Syndromes	HAART	AIDS-defining cancers
	Bond(6) (2007)	Statistics in Medicine	Nelfinavir	HIV RNA Concentration
	Lakdawalla(51) (2006)	The Quarterly Journal of Economics	HAART	Risky sexual activity
	Cain(52) (2006)	American Journal of Epidemiology	HAART	Multiple AIDS defining illnesses
	Hogan(50) (2004)	Statistical Methods in Medical Research	HAART	CD4 Cell Count
	Bhattacharya(49) (2003)	Journal of Health Economics	Public vs private insurance	HIV-related mortality
	Tarwater(34) (2001)	American Journal of Epidemiology	ART	Time to AIDS

Table 2.6 Marginal Structural Model-specific study epidemiology grouped by exposure or outcome keywords

	Lead Author	Journal	Exposure	Primary outcome measure
Marginal Structural Models				
ART or Disease Progression	Shiels(24) (2008)	Journal of Acquired Immune Deficiency Syndromes	HAART	AIDS-defining cancers
	Petersen(23) (2008)	AIDS	Time until switching ART	All cause mortality
	Patel(21) (2008)	Clinical Infectious Diseases	HAART, HAART with PI, HAART with NON RTI	CD4 cell %
	Patel(22) (2008)	Clinical Infectious Diseases	HAART	CD4 cell count and survival

(Continued on next page…)

Table 2.6 Marginal Structural Model-specific study epidemiology grouped by exposure or outcome keywords (Continued)

	Lead Author	Journal	Exposure	Primary outcome measure
ART or Disease Progression (continued)	Fox(25) (2008)	International Journal of Epidemiology	Maternal CD4 count	HIV exposed, uninfected
	Fairall(19) (2008)	Archives of Internal Medicine	HAART	Time to death
	De Beaudrap(17) (2008)	AIDS Research and Human Retroviruses	ART	Tx discontinuation, time to death, time to progression, CD4, adverse effects, virological response
	Dolev(18) (2008)	AIDS	HAART	Skin or anogenital warts
	Petersen(15) (2007)	Clinical Infectious Diseases	Pillbox organizers	Adherence, HIV RNA level
	Petersen(16) (2007)	AIDS	Boosted single or boosted double PI ART	Viral suppression
	Petersen(13) (2007)	American Journal of Epidemiology	Time until switching ART	Future CD4 cell counts
	Petersen(14) (2007)	Statistics in Medicine	Time until switching ART	CD4 cell count
	Cole(11) (2007)	American Journal of Epidemiology	HAART	HIV RNA concentration
	Perez(53) (2007)	Gaceta Sanitaria	HAART	Time to AIDS and death
	Hogg(82) (2006)	PLoS Medicine	Resistance to non-nucleoside RTI	Death
	Hernan(64) (2006)	Basic and Clinical Pharmacology and Toxicology	HAART	AIDS or death
	De Luca(83) (2006)	Antiviral Therapy	Lopinavir/ritonavir or efavirenz plus 2 nucleoside analogues	Viral failure, CD4 recovery, and clinical progression
	Cole(66) (2005)	American Journal of Epidemiology	HAART	CD4 cell count

(Continued on next page…)

Table 2.6 Marginal Structural Model-specific study epidemiology grouped by exposure or outcome keywords (Continued)

	Lead Author	Journal	Exposure	Primary outcome measure
ART or Disease Progression (continued)	Sterne(84) (2005)	Lancet	HAART	AIDS or death
	Hogan(50) (2004)	Statistical Methods in Medical Research	HAART	CD4 cell count
	Casper(85) (2004)	Journal of Acquired Immune Deficiency Syndrome	HAART, CD4, oral inflammation	HHV-8 shedding
	Barron(86) (2004)	AIDS	Discontinuation of ART	Death
	Brumback(67) (2004)	Statistics in Medicine	Zidovudine	CD4 cell count
	Brumback(40) (2003)	Biometrics	ART	CD4 cell count and viral load
	Ko(87) (2003)	Biometrics	HAART	CD4 cell count
	Cole(65) (2003)	American Journal of Epidemiology	HAART	Time to AIDS and death
	Hernan(63) (2002)	Statistics in Medicine	Zidovudine	CD4 cell count
	Hernan(62) (2001)	Journal of the American Statistical Association	Zidovudine and prophylaxis therapy	Survival
	Hernan(37) (2000)	Epidemiology	Zidovudine	survival

(Continued on next page…)

Table 2.6 Marginal Structural Model-specific study epidemiology grouped by exposure or outcome keywords (Continued)

	Lead Author	Journal	Exposure	Primary outcome measure
Other	Lopez-Gatell(20) (2008)	AIDS	Incident TB	Death
	Brown(10) (2007)	AIDS	Incident and prevalent HSV-2	HIV acquisition
	Lopez-Gatell(12) (2007)	American Journal of Epidemiology	Incident TB	Death
	Bachmann(88) (2006)	AIDS	AIDS death, illness, poverty	Monthly adult equivalent income and expenditure, illness episodes, death
	Brookhart(89) (2006)	Computational Statistics and Data Analysis	Drinking water patterns	Prevalent GI illness
	Wang(90) (2005)	AIDS	HIV infection	Overdose mortality
	Eisenberg(91) (2002)	Epidemiology and Infection	Drinking water patterns	Diarrhea

Table 2.7 Structural Equation Models-specific study epidemiology grouped by exposure or outcome keywords

	Lead Author	Journal	Exposure	Primary outcome measure
Structural Equation Models				
Adherence and Behavior	Rice(61) (2008)	Journal of Adolescence	Social network influences	HIV risk behaviors
	Bull(59) (2008)	AIDS and Behavior	Internet-based intervention	Change in proportion engaged in protected sex
	Cha(60) (2008)	International Journal of Nursing Studies	Social support, depression, self-efficacy	ART adherence
	Sodergard(58) (2007)	Patient Education and Counseling	Attitudes toward meds, goals, right support system	Adherence-behavior
	Naar-King(57) (2006)	AIDS Care	Stage of change, social support	Alcohol and drug abuse
	Llabre(56) (2006)	AIDS Patient Care and STDs	HAART adherence	Validity in predicting HIV viral load
	Prado(92) (2004)	AIDS and Behavior	Stressors	Religious involvement
	Lim(93) (2003)	Journal of Occupational Health Psychology	Knowledge, homophobia, fear	Organizational outcomes
	Sengupta(94) (2000)	Journal of Acquired Immune Deficiency Syndromes	Distrust	Willingness to participate in AIDS research
	Kraft(55) (1995)	Social Science and Medicine	Attitudes and information	Restrictive AIDS policies
	Van der Velde(54) (1991)	Journal of Behavioral Medicine	Protection motivation theory, conflict theory	AIDS related health behaviors

Table 2.8: Results of Study Quality Assessment of Propensity Scores Publications

Lead Author	Data Country of Origin	Traditional Interpretability	Assumptions Discussed	Confounding Or Mediating Variables Discussed	Model/ Instrument Selection Discussed
Zule(48) (2008)	US	Yes	No	Yes	Yes
Sanguan-wongse(41) (2008)	Thailand	Yes	Yes	Yes	Yes
Mahal(47) (2008)	Nigeria	No	Yes	Yes	Yes
Anuwatnon-thakate(46) (2008)	Thailand	Yes	Yes	Yes	Yes
Tai(45) (2007)	US	Yes	No	Yes	Yes
Potard(42) (2007)	France	No	Yes	Yes	Yes
Braithwaite(43) (2007)	US	Yes	No	Yes	No
Albalak(44) (2007)	US	No	No	No	No
Nosyk(73) (2006)	Canada	Yes	No	Yes	No
Merito(74) (2006)	Italy	Yes	No	Yes	Yes
Liu(76) (2006)	US	No	No	Yes	Yes
Liu(75) (2006)	US	Yes	Yes	Yes	Yes
Gangopadhyay(81) (2005)	India	No	No	Yes	No
El-Bassel(78) (2005)	US	No	Yes	Yes	Yes
Chu(77) (2005)	US	Yes	Yes	Yes	Yes
Wenzel(79) (2004)	US	Yes	No	Yes	Yes
Rotheram-Borus(80) (2003)	US	Yes	No	No	No
Brumback(40) (2003)	US	Yes	Yes	Yes	Yes
McLaughlin(32) (1999)	US	No	No	Yes	No

Table 2.9: Results of Study Quality Assessment of Instrumental Variable Publications

Lead Author	Data Country of Origin	Traditional Interpretability	Assumptions Discussed	Confounding Or Mediating Variables Discussed	Model/ Instrument Selection Discussed
Instrumental Variables					
Shiels(24) (2008)	US	No	Yes	Yes	Yes
Bond(6) (2007)	Europe	Yes	Yes	Yes	Yes
Lakdawalla(51) (2006)	US	Yes	Yes	Yes	Yes
Cain(52) (2006)	US	No	Yes	Yes	Yes
Hogan(50) (2004)	US	Yes	Yes	Yes	Yes
Bhattacharya(49) (2003)	US	Yes	Yes	Yes	Yes
Tarwater(34) (2001)	US	No	No	Yes	No

Table 2.10: Results of Study Quality Assessment of Marginal Structural Model Publications

Lead Author	Data Country of Origin	Traditional Interpretability	Assumptions Discussed	Confounding Or Mediating Variables Discussed	Model/ Instrument Selection Discussed
Marginal Structural Models					
Shiels(24) (2008)	US	No	No	Yes	No
Peterson(23) (2008)	US	No	Yes	Yes	Yes
Patel.2(21) (2008)	US	Yes	Yes	Yes	No
Patel.1(22) (2008)	US	Yes	Yes	Yes	No
Lopez-Gatell(20) (2008)	US	Yes	Yes	Yes	No
Fox(25) (2008)	Zambia	Yes	Yes	No	Yes
Fairall(19) (2008)	South Africa	Yes	Yes	Yes	No
Dolev(18) (2008)	US	No	No	Yes	No
De Beaudrap(17) (2008)	Senegal	No	Yes	Yes	Yes

Table 2.10: Results of Study Quality Assessment of Marginal Structural Model Publications (continued)

Lead Author	Data Country of Origin	Traditional Interpretability	Assumptions Discussed	Confounding Or Mediating Variables Discussed	Model/ Instrument Selection Discussed
Peterson.1(13) (2007)	US	Yes	Yes	Yes	Yes
Peterson.2(14) (2007)	US	No	Yes	Yes	Yes
Peterson.3(16) (2007)	US	Yes	Yes	Yes	Yes
Peterson.4(15) (2007)	US	Yes	Yes	Yes	Yes
Lopez-Gatell(12) (2007)	US	Yes	Yes	Yes	No
Cole(11) (2007)	US	Yes	Yes	Yes	No
Brown(10) (2007)	Zim-babwe/ Uganda	Yes	Yes	Yes	Yes
Perez(53) (2007)	Spain	Yes	Yes	Yes	No
Hogg(82) (2006)	Canada	Yes	Yes	Yes	Yes
Hernan(64) (2006)	France	Yes	Yes	Yes	Yes
De Luca(83) (2006)	Italy	Yes	No	Yes	No
Brookhart(89) (2006)	US	Yes	Yes	Yes	Yes
Bachmann(88) (2006)	South Africa	Yes	No	Yes	No
Wang(90) (2005)	US	Yes	Yes	Yes	Yes
Sterne(84) (2005)	Switzer-land	Yes	Yes	Yes	No
Cole(66) (2005)	US	Yes	Yes	Yes	Yes
Hogan(50) (2004)	US	Yes	Yes	Yes	Yes
Casper(85) (2004)	US	No	No	Yes	Yes
Brumback(67) (2004)	US	No	Yes	Yes	No
Barron(86) (2004)	US	Yes	Yes	Yes	No

Table 2.10: Results of Study Quality Assessment of Marginal Structural Model Publications (continued)

Lead Author	Data Country of Origin	Traditional Interpretability	Assumptions Discussed	Confounding Or Mediating Variables Discussed	Model/ Instrument Selection Discussed
Ko(87) (2003)	US	Yes	Yes	Yes	Yes
Cole(65) (2003)	US	Yes	Yes	Yes	Yes
Brumback(40) (2003)	US	Yes	Yes	Yes	Yes
Hernan(63) (2002)	US	Yes	Yes	Yes	No
Eisenberg(91) (2002)	US	Yes	Yes	Yes	Yes
Hernan(62) (2001)	US	Yes	Yes	Yes	No
Hernan(37) (2000)	US	Yes	Yes	Yes	Yes

Table 2.11: Results of Study Quality Assessment of Structural Equation Models Publications

Lead Author	Data Country of Origin	Traditional Interpretability	Assumptions Discussed	Confounding Or Mediating Variables Discussed	Model/ Instrument Selection Discussed
Structural Equation Models					
Rice(61) (2008)	US	No	No	Yes	Yes
Cha(60) (2008)	US	No	No	Yes	Yes
Bull(59) (2008)	US	No	Yes	Yes	Yes
Sodergard(58) (2007)	Sweden	No	No	Yes	Yes
Naar-King(57) (2006)	US	No	No	Yes	Yes
Llabre(56) (2006)	US	No	Yes	Yes	Yes
Prado(92) (2004)	US	No	No	Yes	Yes

Table 2.11: Results of Study Quality Assessment of Structural Equation Models Publications (continued)

Lead Author	Data Country of Origin	Traditional Interpretability	Assumptions Discussed	Confounding Or Mediating Variables Discussed	Model/ Instrument Selection Discussed
Structural Equation Models					
Lim(93) (2003)	Singapore	No	No	Yes	Yes
Sengupta(94) (2000)	US	No	Yes	Yes	Yes
Kraft(55) (1995)	Norway	No	Yes	Yes	Yes
Van der Velde(54) (1991)	Nether-lands	No	Yes	Yes	Yes

Table 2.12 Source of Inference Among Studies Using Marginal Structural Models

Author (Year)	Source of Inference	Author (Year)	Source of Inference	Author (Year)	Source of Inference	Author (Year)	Source of Inference
Hernan[50] (2000)	GEE	Casper[95] (2004)	robust	Hogg[92] (2006)	unknown	Debeau-drap[16] (2008)	robust
Hernan[33] (2001)	robust	Hogan[31] (2004)	sandwich	Perez[64] (2007)	robust	Dolev[17] (2008)	unknown
Eisen-berg [101] (2002)	sandwich	Cole[75] (2005)	robust	Brown[9] (2007)	GEE	Fairall[18] (2008)	robust
Hernan[73] (2002)	sandwich	Sterne[94] (2005)	robust	Cole[10] (2007)	sandwich	Fox[24] (2008)	robust
Brum-back[52] (2003)	sandwich	Wang[100] (2005)	robust	Lopez-Gatell[11] (2007)	bootstrap	Lopez-Gatell[19] (2008)	robust
Cole[76] (2003)	bootstrap	Bach-mann[98] (2006)	robust	Peter-sen.1[12] (2007)	robust	Patel[21] (2008)	unknown
Ko[97] (2003)	robust	Brook-hart[99] (2006)	bootstrap	Peter-sen.2[13] (2007)	bootstrap	Patel.2[20] (2008)	unknown
Barron[96] (2004)	robust	De Luca[93] (2006)	unknown	Peter-sen.3[15] (2007)	bootstrap	Peter-sen[22] (2008)	bootstrap
Brum-back[77] (2004)	bootstrap	Hernan[74] (2006)	robust	Petersen.4 (2007)[14]	bootstrap	Shiels[23] (2008)	GEE

Table 2.13. Supplemental Table for Figures 13a-16b (Network of Institutions and Affiliated Authors of HIV-Related Studies Using Causal Inference Methods)[1, 2]

	MSMs		Propensity Scores		Instrumental Variables		Structural Equations		All Methods	
Institution	Before 2007	2007-2008	Before 2007	2007-2008	Before 2007	2007-2008	Before 2007	2007-2008	Before 2007	2007-2008
Harvard	9	4	1	1	0	0	0	0	11	4
Johns Hopkins	4	7	3	0	2	1	0	0	9	8
UCLA	3	4	4	1	2	1	0	1	9	7
UCSF	3	7	1	0	0	0	1	0	5	7
Pittsburgh	1	2	1	1	2	1	0	1	4	5
UC Berkeley	2	6	0	0	0	0	0	0	2	6
Cook County	3	2	2	0	0	0	0	0	5	2
UW	5	0	1	0	0	0	0	0	6	0
Northwestern	1	2	1	0	2	1	0	0	4	3
SUNY Brooklyn	2	2	1	0	0	0	0	0	3	2
Tulane	0	2	0	0	0	0	0	0	0	2
Montefiore	1	1	2	0	0	0	0	0	3	1
Georgetown	1	1	2	0	0	0	0	0	3	1
CDC	1	0	0	2	0	0	0	1	1	3
TMOPH	0	0	0	2	0	0	0	0	0	2
Phuket	0	0	0	2	0	0	0	0	0	2
Bangkok Metropolitan Health Admin.	0	0	0	2	0	0	0	0	0	2
Office of Disease Prevention and Control	0	0	0	2	0	0	0	0	0	2
Bamrashadura	0	0	0	2	0	0	0	0	0	2
Chiang Rai Provincial PH Office	0	0	0	2	0	0	0	0	0	2
Research Institute of Tuberculosis	0	0	0	2	0	0	0	0	0	2
Stanford	0	1	1	0	1	0	0	0	2	1
RAND	0	0	1	0	2	0	0	0	3	0
Columbia	0	1	1	0	0	0	0	0	1	1

Table 2.13. Supplemental Table for Figures 12a-15b (Networks of Institutions and Affiliated Authors of HIV-Related Studies Using Causal Inference Methods)[1, 2] (continued)

	MSMs		Propensity Scores		Instrumental Variables		Structural Equations		All Methods	
Institution	Before 2007	2007-2008	Before 2007	2007-2008	Before 2007	2007-2008	Before 2007	2007-2008	Before 2007	2007-2008
Saint Paul's	1	0	1	0	0	0	0	0	2	0
UBC	1	0	1	0	0	0	0	0	2	0
INSERM	1	0	0	1	0	0	0	0	1	1
Brown	2	0	0	0	1	0	0	0	3	0
USC	1	2	0	0	0	0	0	0	1	2
UNC Chapel Hill	0	1	0	0	0	0	1	0	1	1
University of Miami	0	0	0	0	0	0	2	0	2	0
University of East Anglia	1	1	0	0	0	0	0	0	1	1
University of the Free State	1	1	0	0	0	0	0	0	1	1
Lincoln Medical, NY	1	1	0	0	0	0	0	0	1	1
Kaiser	0	2	0	0	0	0	0	0	0	2

1. Institutional abbreviations are described in Appendix A.1.
2. Included are institutions with a minimum of 2 publications using one of the causal inference methods

Figures

Figure 2.1. DAG Representing Theoretical Antiretroviral Therapy and Death Relationship

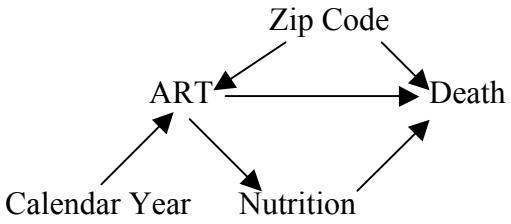


Figure 2.2. ***Study Selection Flow***

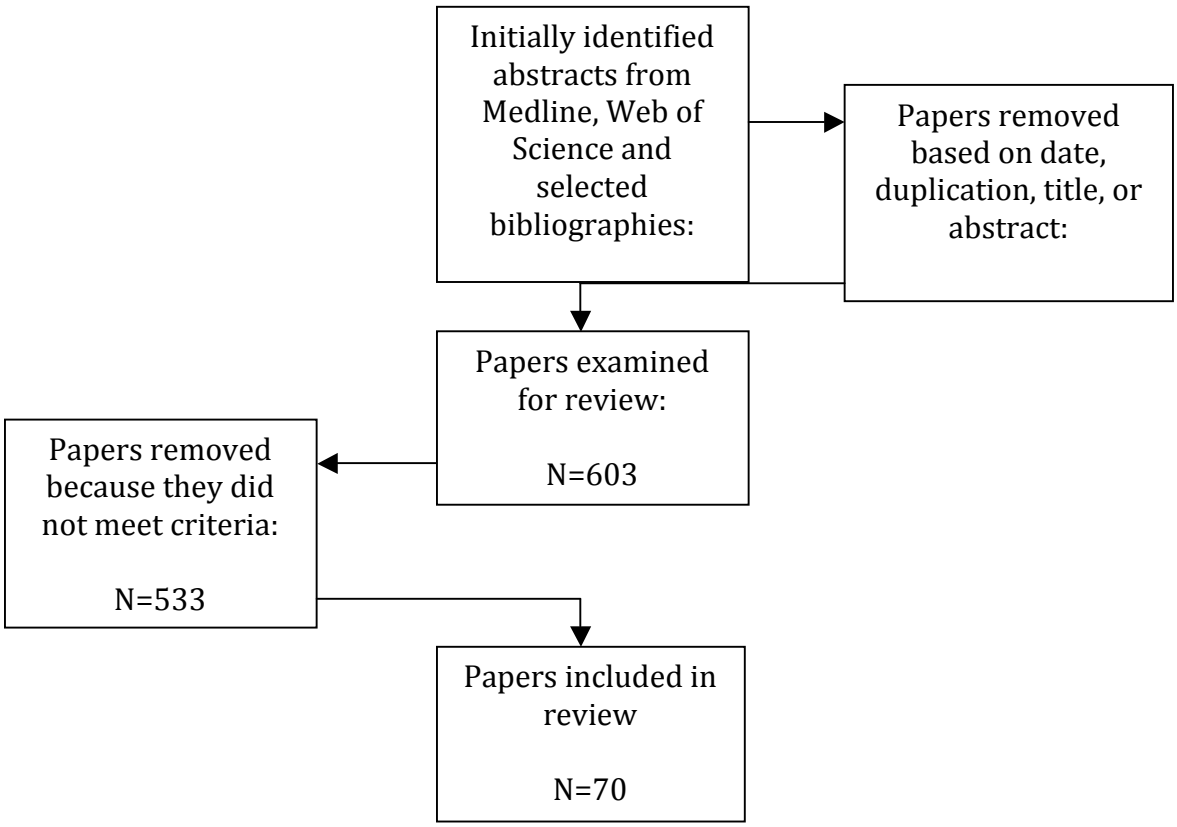


Figure 2.3. Histograms Showing Temporal Trends in Appearance of Causal Inference Methods in HIV/AIDS Publications

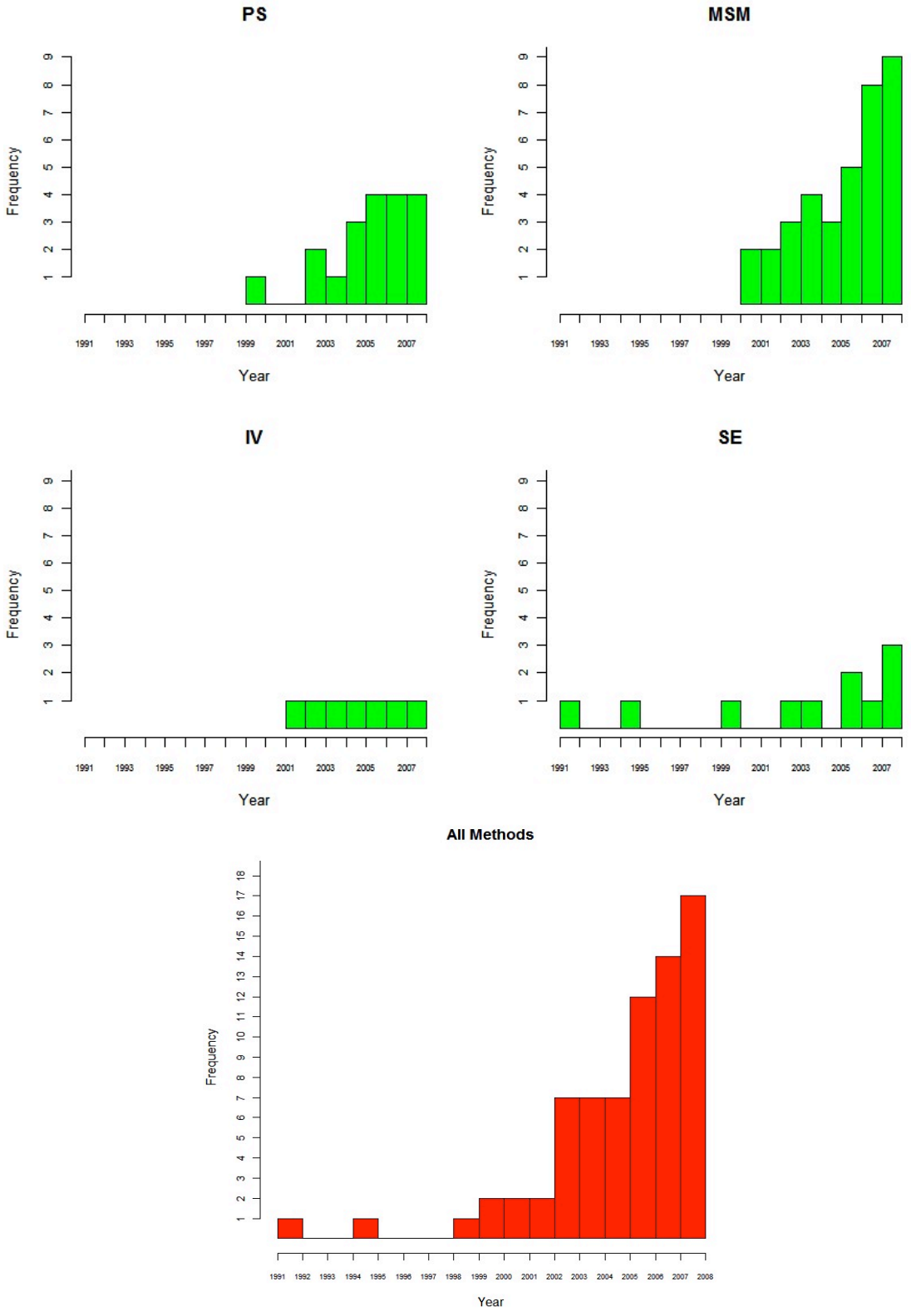


Figure 2.4 Dot Chart of Frequency of Publications (with a minimum of 2) By All Associated Authors Using Causal Inference Methods With HIV/AIDS

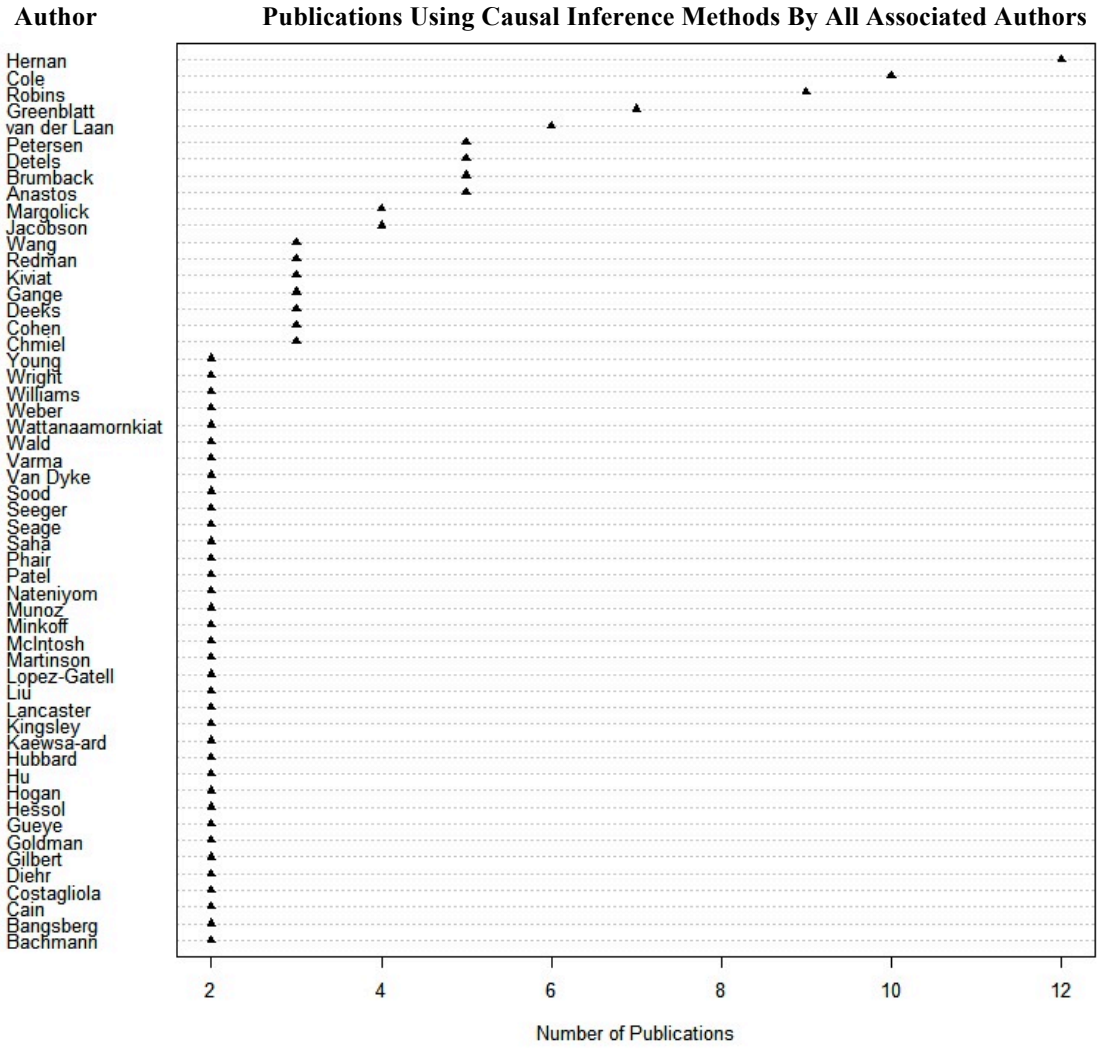


Figure 2.5 Dot Chart of Frequency of Publications By First Authors Using Causal Inference Methods With HIV/AIDS

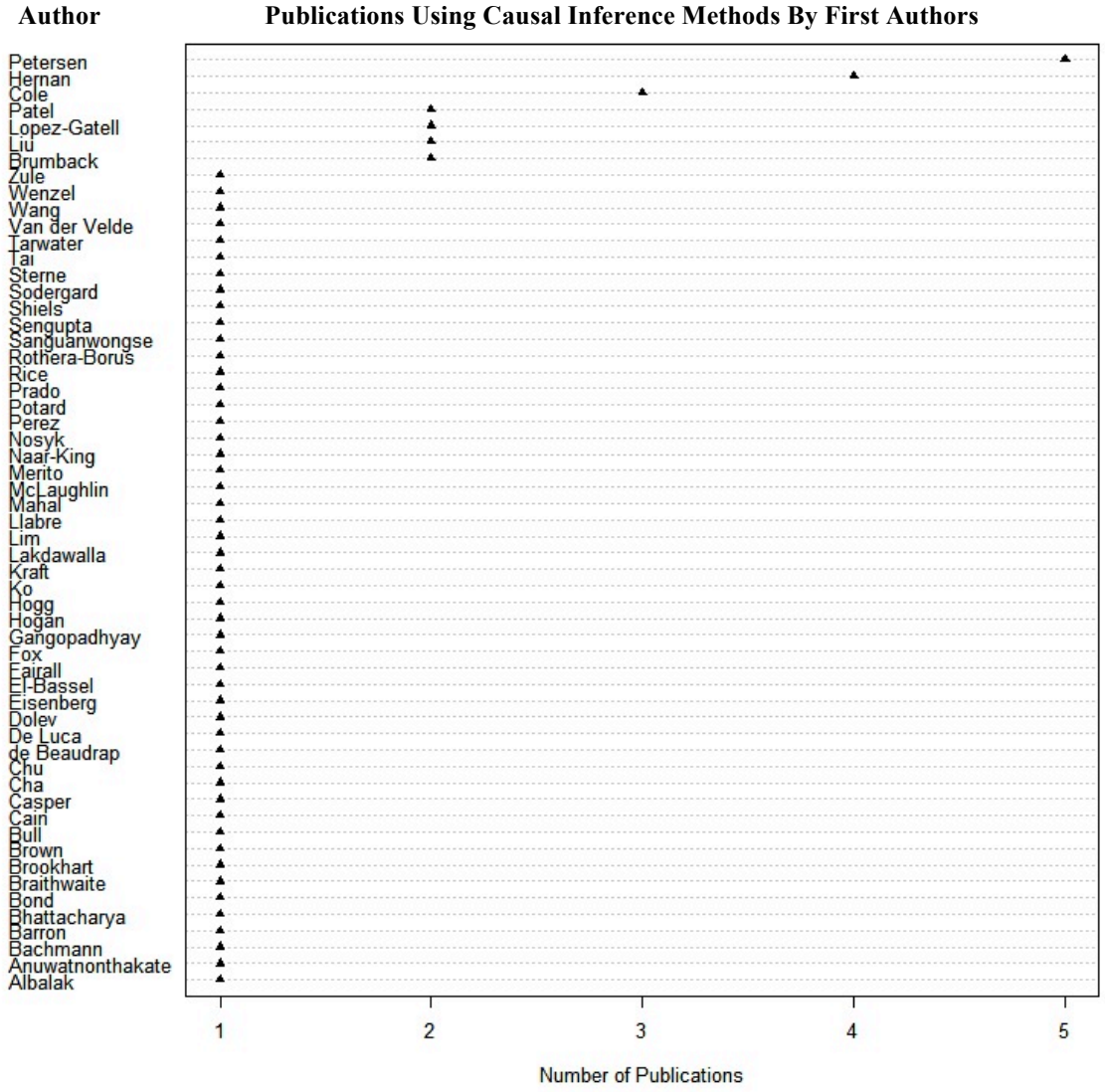


Figure 2.6 Dot Chart of Frequency of Publications By Senior Authors (Last Author) Using Causal Inference Methods With HIV/AIDS

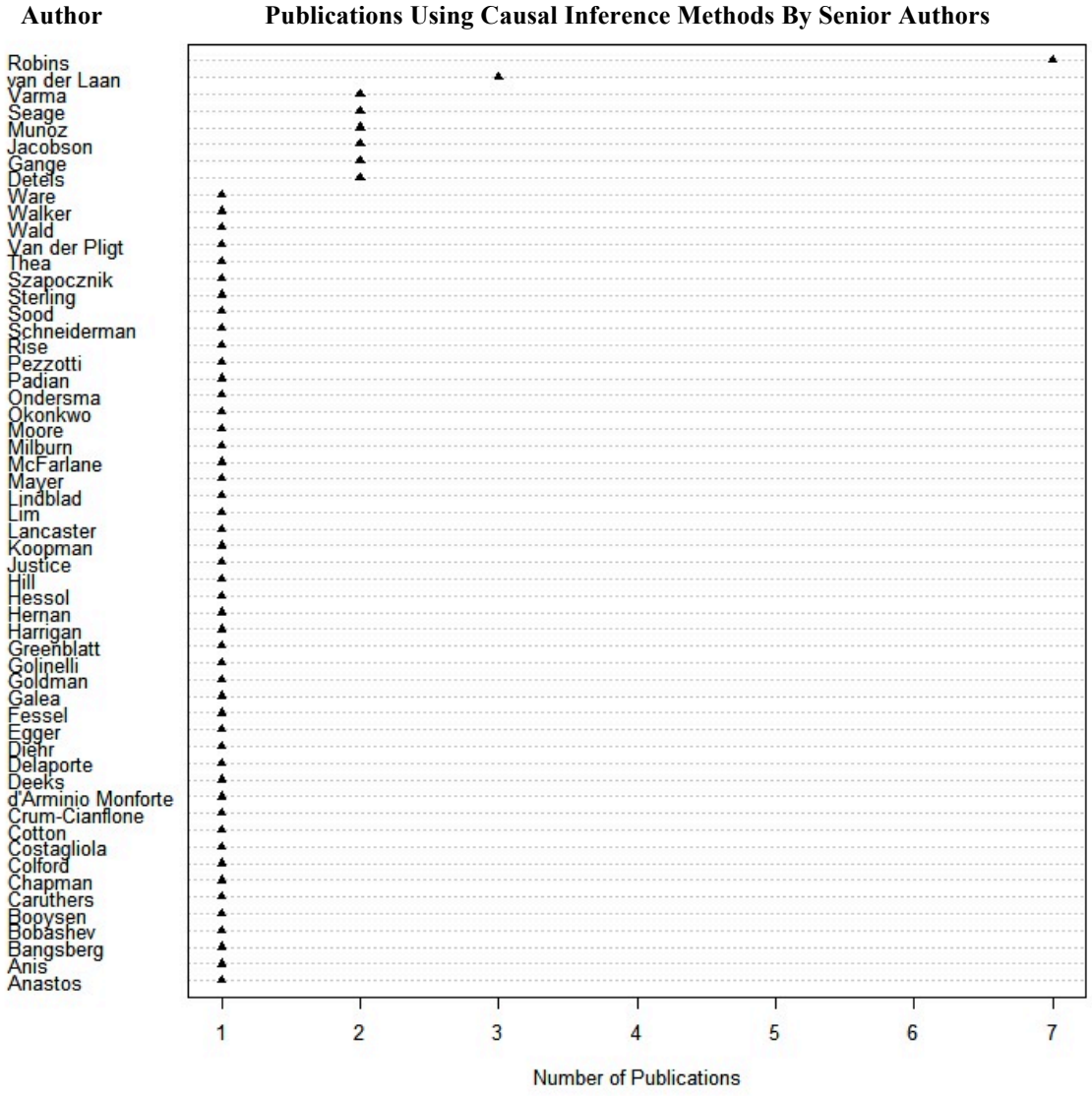


Figure 2.7. Dot Chart of Frequency of Appearance of Publications Using Causal Inference Methods With HIV/AIDS Data By Journal

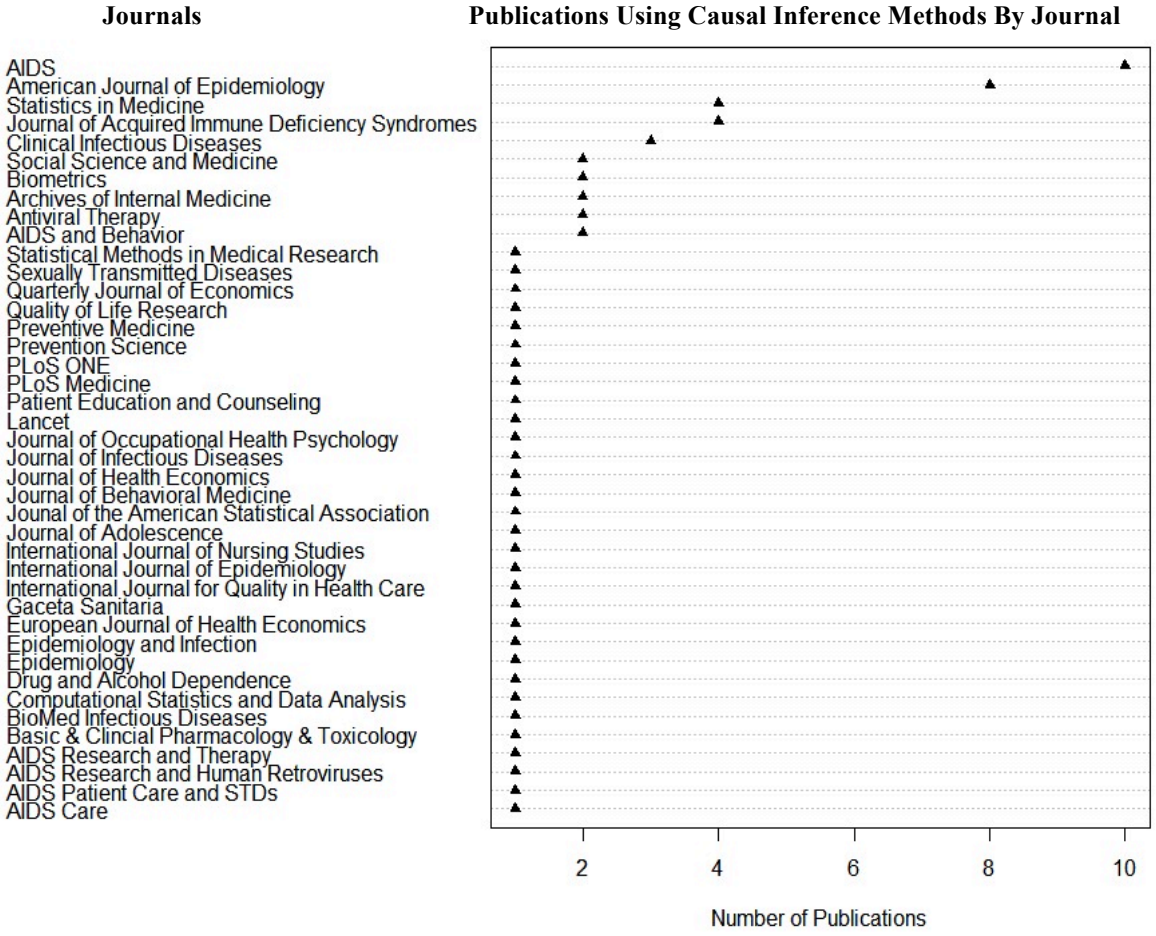


Figure 2.8. Bar-plots of Results of Study Quality Assessments By Method

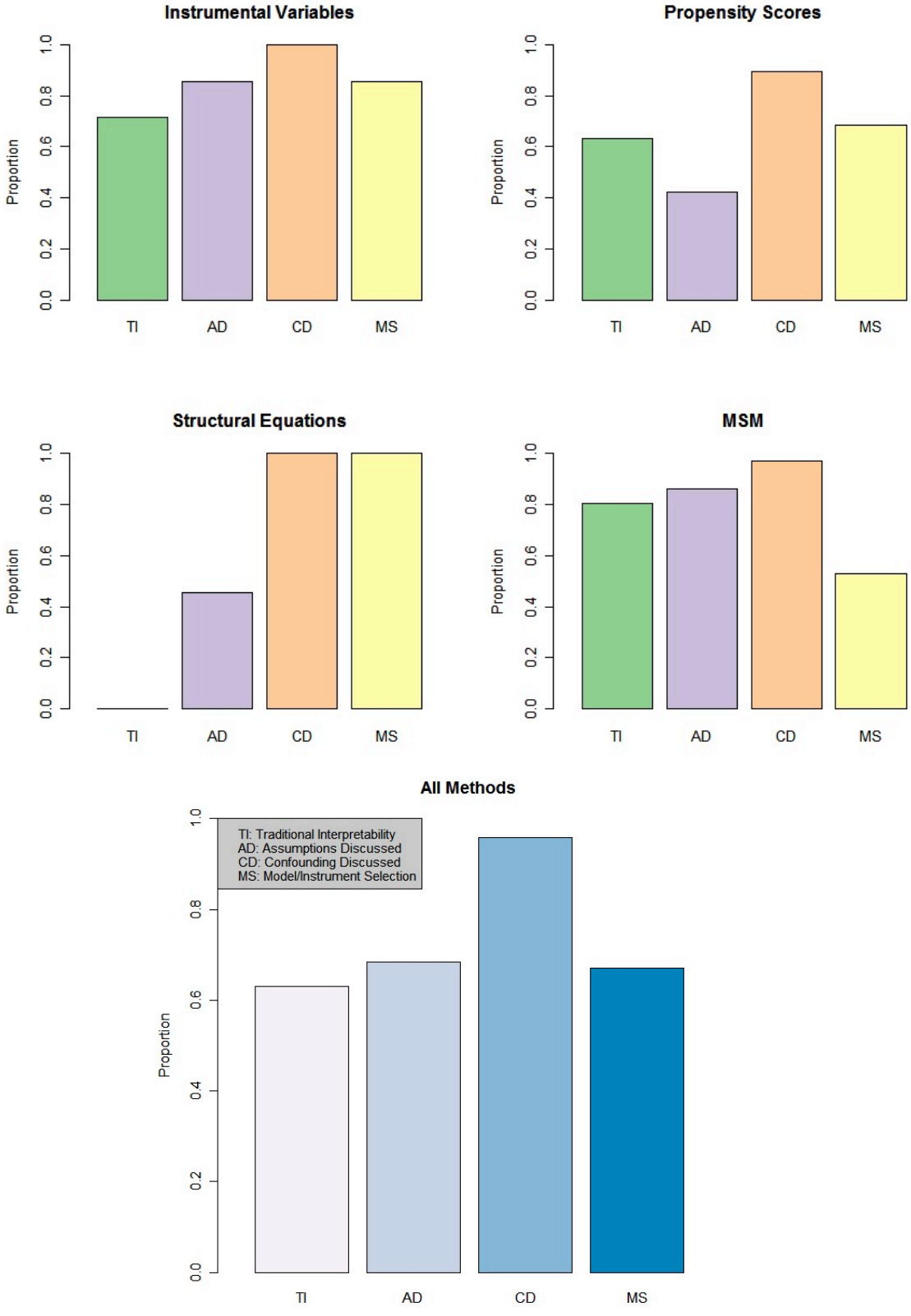


Figure 2.9a. Network of Institutions and Affiliated Authors of HIV-Related Studies Published Before 2007 Using Marginal Structural Models

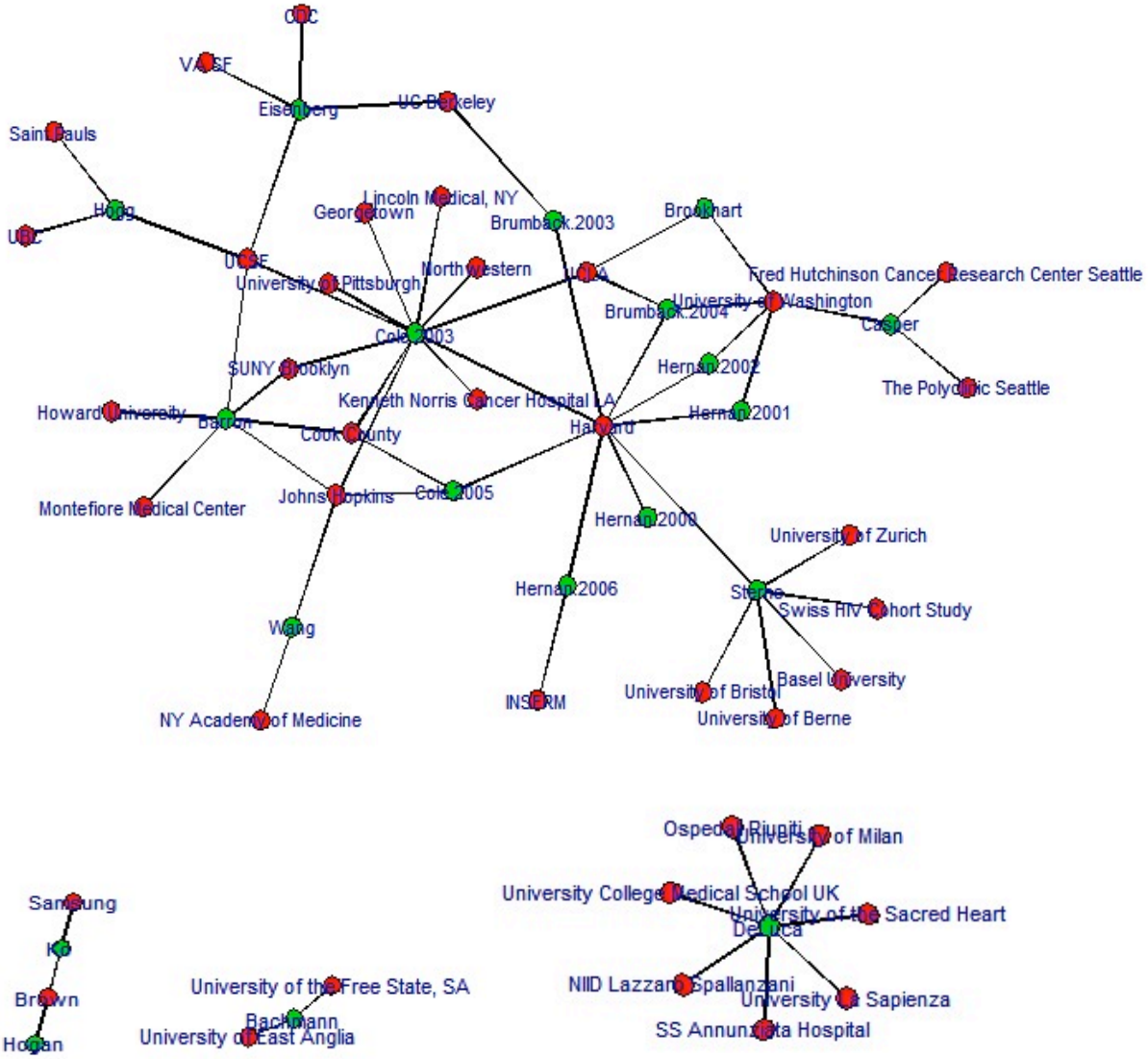


Figure 2.9b. Network of Institutions and Affiliated Authors of HIV-Related Studies Published In 2007-08 Using Marginal Structural Models

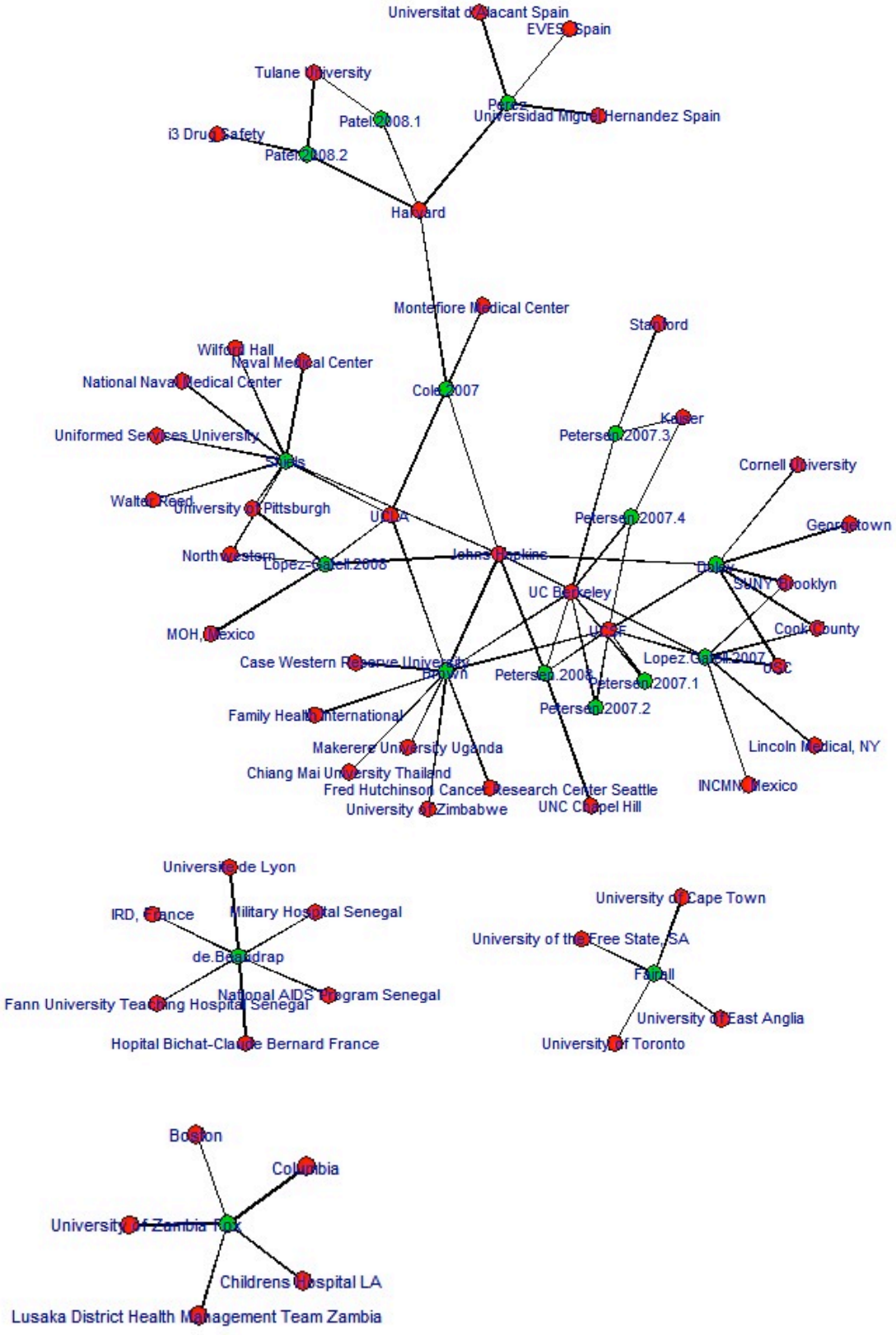


Figure 2.10a. Network of Institutions and Affiliated Authors of HIV-Related Studies Published Before 2007 Using Propensity Scores

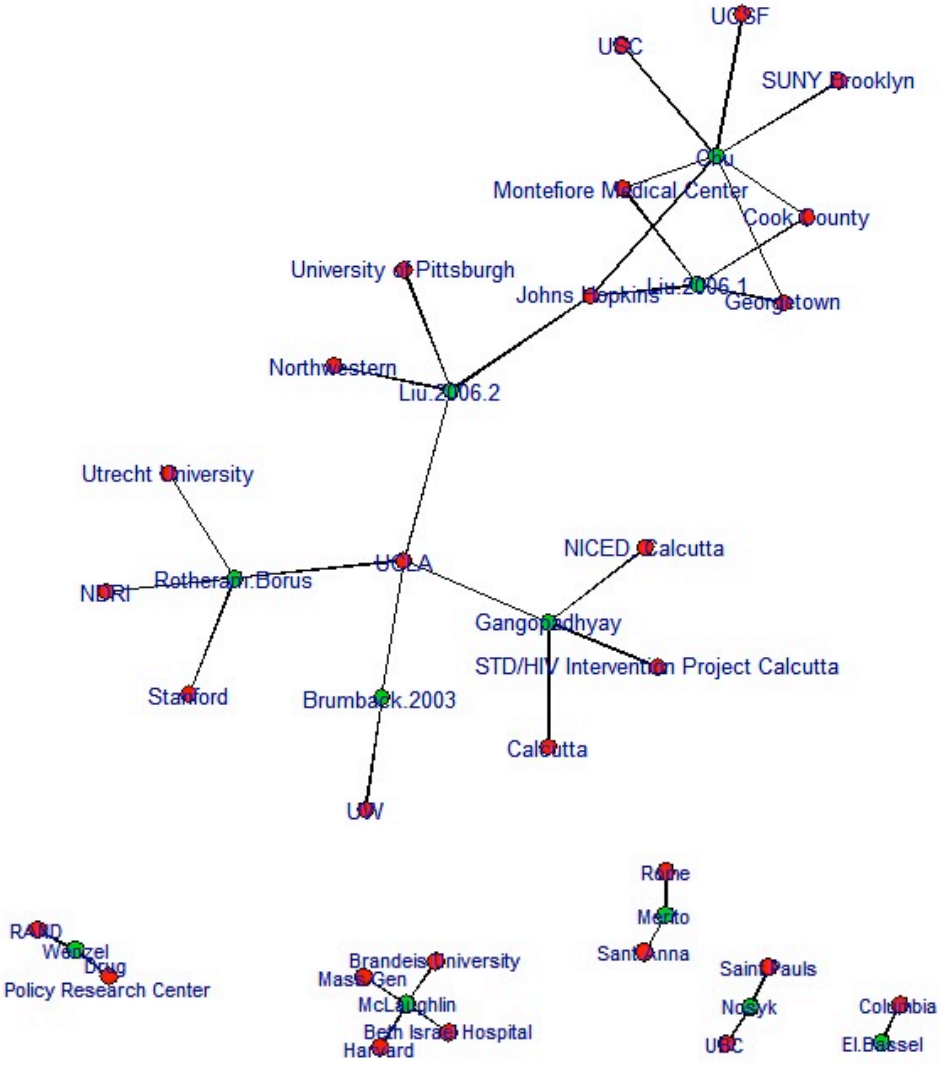


Figure 2.10b. Network of Institutions and Affiliated Authors of HIV-Related Studies Published In 2007 Or 2008 Using Propensity Scores

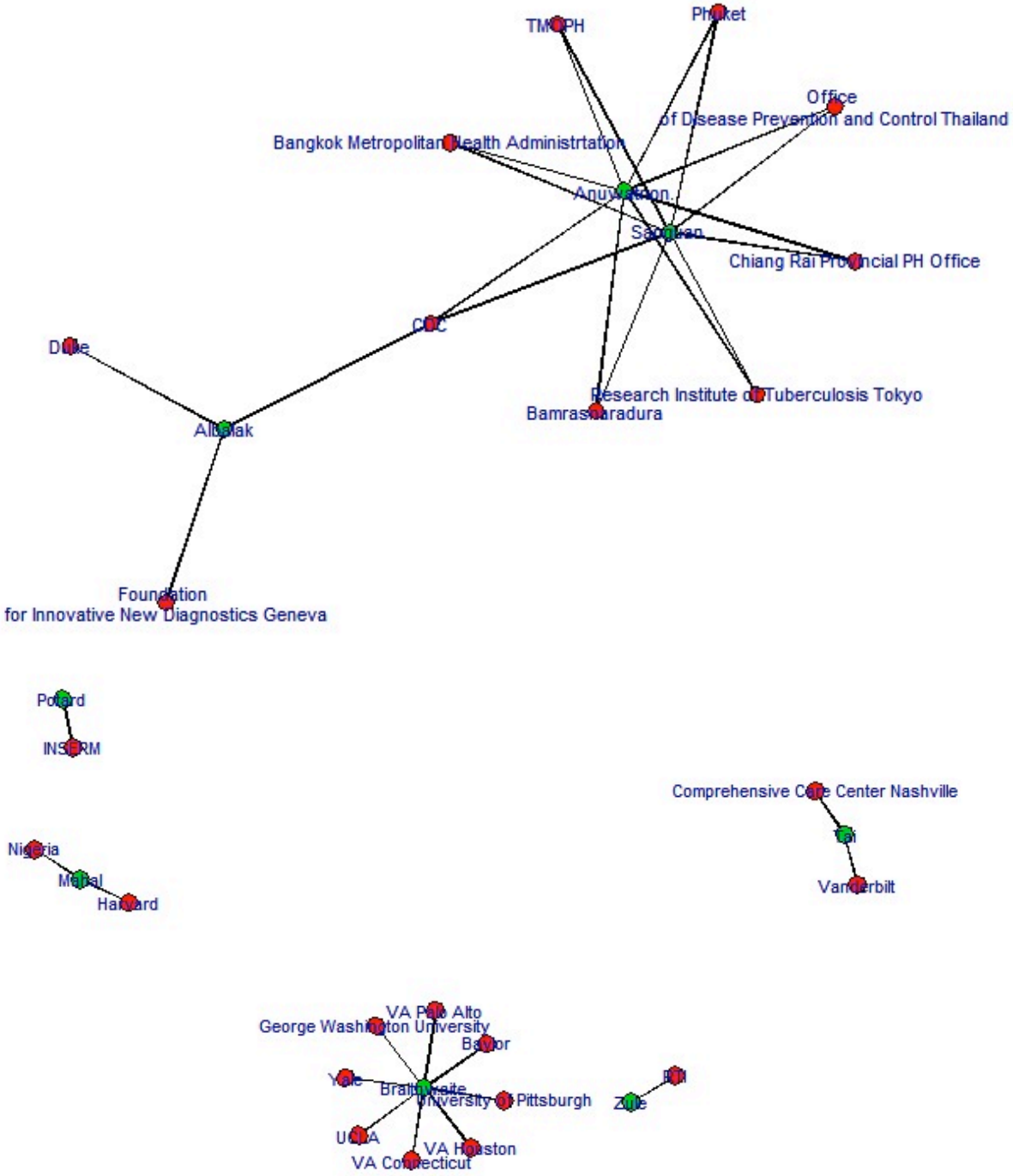


Figure 2.11a. Network of Institutions and Affiliated Authors of HIV-Related Studies Published Before 2007 Using Instrumental Variables

Figure 2.11b. Network of Institutions and Affiliated Authors of HIV-Related Studies Published in 2007 or 2008 Using Instrumental Variables

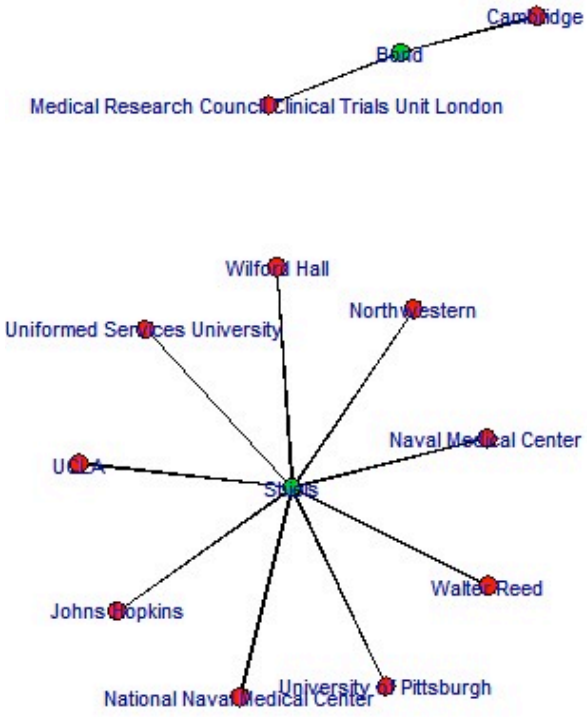


Figure 2.12a. Network of Institutions and Affiliated Authors of HIV-Related Studies Published Before 2007 Using Structural Equations

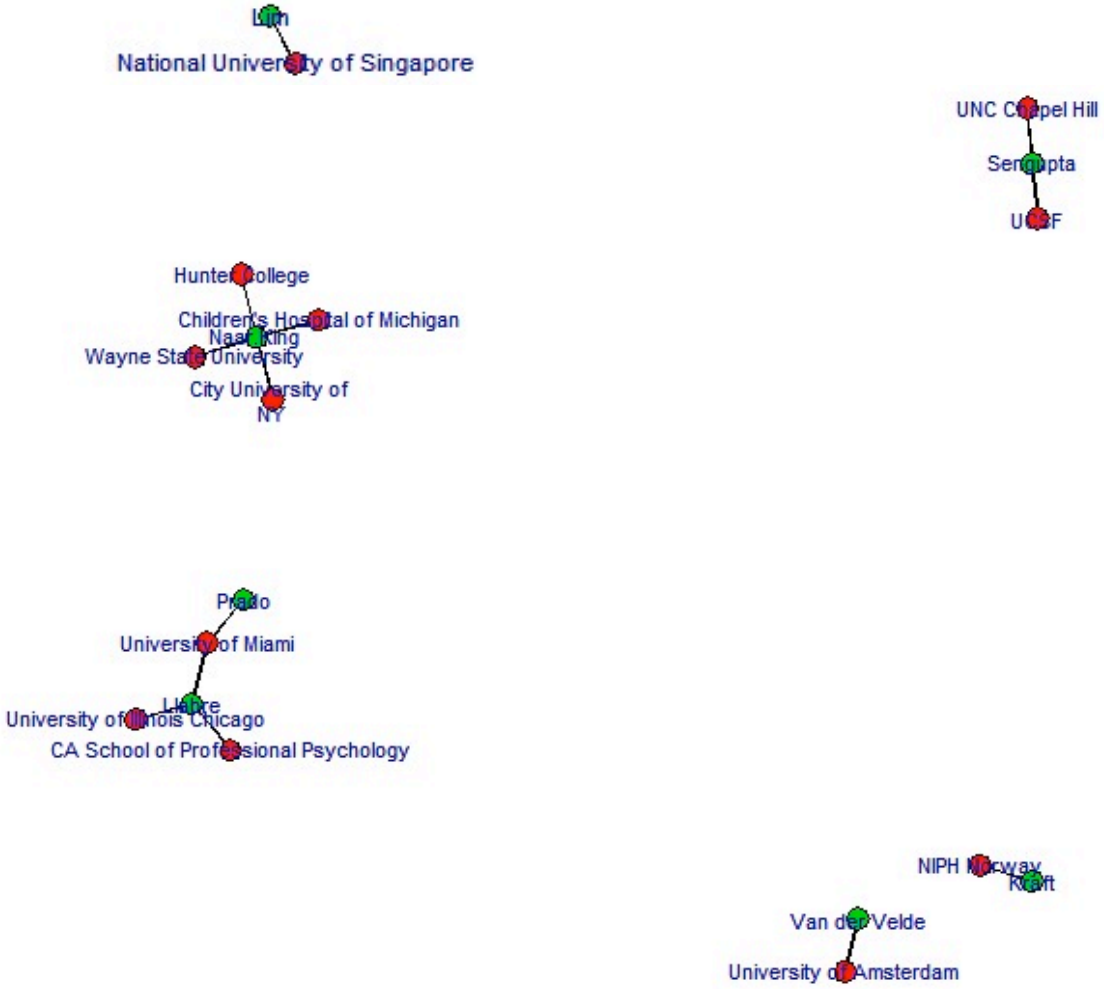
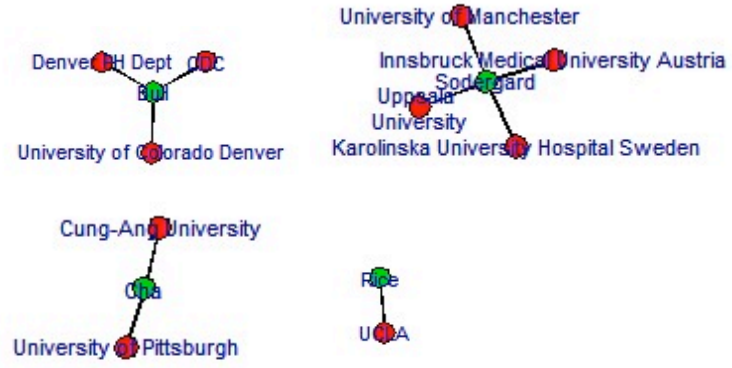


Figure 2.12b. Network of Institutions and Affiliated Authors of HIV-Related Studies Published in 2007 or 2008 Using Structural Equations



Bibliography

1. Jadad A. Randomized Controlled Trials: A User's Guide. First Edition. BJM Publishing Group; [date unknown].

2. Schulz K, Grimes D. Generation of allocation sequences in randomised trials: chance, not choice. Lancet 2002;359(9305):515-519.

3. Rubin D. Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies. 1990;

4. Pearl J. Causality: models, reasoning, and inference. Cambridge University Press; 2000.

5. Shah B, Laupacis A, Hux J, Austin P. Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review. Journal of Clinical Epidemiology 2005;58(6):550-559.

6. Bond SJ, White IR, Sarah Walker A. Instrumental variables and interactions in the causal analysis of a complex clinical trial. Statist. Med. 2007;26(7):1473-1496.

7. Robins JM. Marginal structural models. 1998;

8. Robins JM, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000;11(5):550–560.

9. Greenland S, Brumback B. An overview of relations among causal modelling methods. International journal of epidemiology 2002;31(5):1030.

10. Brown JM, Wald A, Hubbard A, Rungruengthanakit K, Chipato T, Rugpao S, Mmiro F, Celentano DD, Salata RS, Morrison CS, others. Incident and prevalent herpes simplex virus type 2 infection increases risk of HIV acquisition among women in Uganda and Zimbabwe. Aids 2007;21(12):1515.

11. Cole SR, Hernan MA, Anastos K, Jamieson BD, Robins JM. Determining the effect of highly active antiretroviral therapy on changes in human immunodeficiency virus type 1 RNA viral load using a marginal structural left-censored mean model. American journal of epidemiology 2007;

12. Lopez-Gatell H, Cole SR, Hessol NA, French AL, Greenblatt RM, Landesman S, Preston-Martin S, Anastos K. Effect of tuberculosis on the survival of women infected with human immunodeficiency virus. American journal of epidemiology 2007;165(10):1134.

13. Petersen ML, Deeks SG, Martin JN, van der Laan MJ. History-adjusted marginal structural models for estimating time-varying effect modification. American journal of epidemiology 2007;166(9):985.

14. Petersen ML, Deeks SG, van der Laan MJ. Individualized treatment rules: Generating candidate clinical trials. Statist. Med. 2007;26(25):4578-4601.

15. Petersen ML, Wang Y, van der Laan MJ, Guzman D, Riley E, Bangsberg DR. Pillbox Organizers Are Associated with Improved Adherence to HIV Antiretroviral Therapy and Viral Suppression: A Marginal Structural Model Analysis. CLIN INFECT DIS 2007;45(7):908-915.

16. Petersen ML, Wang Y, van der Laan MJ, Rhee SY, Shafer RW, Fessel WJ. Virologic efficacy of boosted double versus boosted single protease inhibitor therapy. AIDS 2007;21(12):1547–54.

17. de Beaudrap P, Etard J, Guèye FN, Guèye M, Landman R, Girard P, Sow PS, Ndoye I, Delaporte E, for the ANRS 1215/1290 Study Group. Long-Term Efficacy and Tolerance of Efavirenz- and Nevirapine-Containing Regimens in Adult HIV Type 1 Senegalese Patients. AIDS Research and Human Retroviruses 2008;24(6):753-760.

18. Dolev JC, Maurer T, Springer G, Glesby MJ, Minkoff H, Connell C, Young M, Schowalter K, Cox C, Hessol NA. Incidence and risk factors for verrucae in women. EPIDEMIOLOGY AND SOCIAL AIDS. 22 (10): 1213-1219, June 19, 2008. [date unknown];

19. Fairall LR, Bachmann MO, Louwagie G, van Vuuren C, Chikobvu P, Steyn D, Staniland GH, Timmerman V, Msimanga M, Seebregts CJ, others. Effectiveness of antiretroviral treatment in a South African program: a cohort study. Archives of Internal Medicine 2008;168(1):86.

20. López-Gatell H, Cole SR, Margolick JB, Witt MD, Martinson J, Phair JP, Jacobson LP. Effect of tuberculosis on the survival of HIV-infected men in a country with low tuberculosis incidence. AIDS 2008;22(14):1869.

21. Patel K, Hernán M, Williams P, Seeger J, McIntosh K, Dyke R, Seage III G, Pediatric AIDS Clinical Trials Group 219/219C Study Team. Long-term effect of highly active antiretroviral therapy on CD4 cell evoloution among children and adolescents infected with HIV: 5 years and counting. CLIN INFECT DIS 2008;46(11):1751-1760.

22. Patel K, Hernan MA, Williams PL, Seeger JD, McIntosh K, Dyke RBV, Seage III GR, Pediatric AIDS Clinical Trials Group 219/219C Study Team. Long-Term Effectiveness of Highly Active Antiretroviral Therapy on the Survival of Children and Adolescents with HIV Infection: A 10-Year Follow-Up Study. CLIN INFECT DIS 2008;46(4):507-515.

23. Petersen ML, van der Laan MJ, Napravnik S, Eron JJ, Moore RD, Deeks SG. Long-term consequences of the delay between virologic failure of highly active antiretroviral therapy and regimen modification. AIDS 2008;22(16):2097.

24. Shiels MS, Cole SR, Wegner S, Armenian H, Chmiel JS, Ganesan A, Marconi VC, Martinez-Maza O, Martinson J, Weintrob A, others. Effect of HAART on incident cancer and noncancer AIDS events among male HIV seroconverters. JAIDS Journal of Acquired Immune Deficiency Syndromes 2008;48(4):485.

25. Fox MP, Brooks DR, Kuhn L, Aldrovandi G, Sinkala M, Kankasa C, Horsburgh R, Thea DM. Role of breastfeeding cessation in mediating the relationship between maternal HIV disease stage and increased child mortality among HIV-exposed uninfected children. International Journal of Epidemiology 2009;38(2):569.

26. Robins J. Identifiability and exchangeability for direct and indirect effects. 1992;

27. Greenland S. An introduction to instrumental variables for epidemiologists. International Journal of Epidemiology 2000;29(4):722-729.

28. Jewell N. Statistics for Epidemiology. Chapman and Hall; 2004.

29. Bluhm R. From hierarchy to network: a richer view of evidence for evidence-based medicine. Perspect Biol Med 2005;48(4):535-47.

30. Lachin J. Statistical properties of randomization in clinical trials. Control Clin Trials 1988;9(4):289-311.

31. Hernan MA, Hernandez-Diaz S, Robins JM. A Structural Approach to Selection Bias. Epidemiology 2004;15(5):615-625.

32. McLaughlin TJ, Soumerai SB, Weinrib D, Aupont O, Cotton D. The association between primary source of ambulatory care and access to a outcomes of treatment among AIDS patients. International Journal for Quality in Health Care 1999;11(4):293.

33. Riersol E. Confluence analysis by means of lag moments and other methods of confluence analysis. Econometrica 1941;9:1-24.

34. Tarwater PM, Mellors J, Gore ME, Margolick JB, Phair J, Detels R, Munoz A. Methods to assess population effectiveness of therapies in human immunodeficiency virus incident and prevalent cohorts. American journal of epidemiology 2001;154(7):675.

35. Robins JM, Rotnitzky A. Recovery of information and adjustment for dependent censoring using surrogate markers. AIDS Epidemiology, Methodology Issues 1992;

36. Tager IB, Haight T, Sternfeld B, Yu Z, van Der Laan M. Effects of Physical Activity and Body Composition on Functional Limitation in the Elderly. Epidemiology 2004;15(4):479-493.

37. Hernan M, Brumback B, Robins J. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. Epidemiology 2000;11(5):561–570.

38. Robins J. Marginal structural models versus structural nested models as tools for causal inference. In: Statistical Models in Epidemiology: The Environment and Clinical Trials. New York, NY: Springer-Verlag; 1999 p. 95-134.

39. R-Project. www.r-project.org: [date unknown].

40. Brumback B, Greenland S, Redman M, Kiviat N, Diehr P. The intensity-score approach to adjusting for confounding. Biometrics 2003;59(2):274–285.

41. Sanguanwongse N, Cain KP, Suriya P, Nateniyom S, Yamada N, Wattanaamornkiat W, Sumnapan S, Sattayawuthipong W, Kaewsa-ard S, Ingkaseth S, others. Antiretroviral therapy for HIV-infected tuberculosis patients saves lives but needs to be used more frequently in Thailand. JAIDS Journal of Acquired Immune Deficiency Syndromes 2008;48(2):181.

42. Potard V, Rey D, Mokhtari S. First-line highly active antiretroviral regimens in 2001-2002 in the French Hospital Database on HIV: combination prescribed and biological outcomes. Antiviral Therapy 2007;12:317-324.

43. Braithwaite RS, Kozal MJ, Chang CC, Roberts MS, Fultz SL, Goetz MB, Gibert C, Rodriguez-Barradas M, Mole L, Justice AC. Adherence, virological and immunological outcomes for HIV-infected veterans starting combination antiretroviral therapies. Aids 2007;21(12):1579.

44. Albalak R, O'Brien R, Kammerer S. Trends in tuberculosis/human immunodeficiency virus comorbidity, United States, 1993-2004. Archives of Internal Medicine 2007;167(22):2443-2452.

45. Tai J, Udoji M, Barkanic G, Byrne D, Rebeiro P, Byram B, Kheshti A, Carter J, Graves C, Raffanti S, Sterling T. Pregnancy and HIV Disease Progression during the Era of Highly Active Antiretroviral Therapy. J INFECT DIS 2007;196(7):1044-1052.

46. Anuwatnonthakate A, Limsomboon P, Nateniyom S, Wattanaamornkiat W, Komsakorn S, Moolphate S, Chiengsorn N, Kaewsa-ard S, Sombat P, Siangphoe U, Mock PA, Varma JK. Directly Observed Therapy and Improved Tuberculosis Treatment Outcomes in Thailand. PLoS ONE 2008;3(8):e3089.

47. Mahal A, Canning D, Odumosu K, Okonkwo P. Assessing the economic impact of HIV/AIDS on Nigerian households: a propensity score matching approach. AIDS 2008;22:S95.

48. Zule WA, Bobashev G. High dead-space syringes and the risk of HIV and HCV infection among injecting drug users. Drug and Alcohol Dependence 2009;100(3):204–213.

49. Bhattacharya J, Goldman D, Sood N. The link between public and private insurance and HIV-related mortality. Journal of Health Economics 2003;22(6):1105–1122.

50. Hogan JW, Lancaster T. Instrumental variables and inverse probability weighting for causal inference from longitudinal observational studies. Statistical Methods in Medical Research 2004;13(1):17-48.

51. Lakdawalla D, Sood N, Goldman D. HIV Breakthroughs and Risky Sexual Behavior. The Quarterly Journal of Economics 2006;121(3):1063–1102.

52. Cain LE, Cole SR, Chmiel JS, Margolick JB, Rinaldo Jr CR, Detels R. Effect of highly active antiretroviral therapy on multiple AIDS-defining illnesses among male HIV seroconverters. American journal of epidemiology 2006;163(4):310.

53. METODOLÓGICA N. Aplicación de modelos estructurales marginales para estimar los efectos de la terapia antirretroviral en 5 cohortes de seroconvertores al virus de la inmunodeficiencia humana Marginal structural models application to estimate the effects of antiretroviral therapy in 5 cohorts of HIV seroconverters. Gac Sanit 2007;21(1):76–83.

54. Velde FW, Pligt J. AIDS-related health behavior: Coping, protection motivation, and previous behavior. Journal of behavioral medicine 1991;14(5):429–451.

55. Kraft P, Rise J. Prediction of attitudes towards restrictive AIDS policies: a structural equation modeling approach. Social Science and Medicine 1995;40(5):711.

56. Llabre MM, Weaver KE, Durán RE, Antoni MH, McPherson-Baker S, Schneiderman N. A measurement model of medication adherence to highly active antiretroviral therapy and its relation to viral load in HIV-positive adults. AIDS Patient Care & STDs 2006;20(10):701–711.

57. Naar-King S, Wright K, Parsons J, Frey M, Templin T, Ondersma S. Transtheoretical Model and substance use in HIV-positive youth. AIDS Care 2006;18(7):839-845.

58. Södergaard B, Höfer S, Halvarsson M, Sönnerborg A, Tully MP, Lindblad \K. A structural equation modeling approach to the concepts of adherence and readiness in antiretroviral treatment. Patient Education and Counseling 2007;67(1-2):108–116.

59. Bull S, Pratte K, Whitesell N, Rietmeijer C, McFarlane M. Effects of an Internet-Based Intervention for HIV Prevention: The Youthnet Trials. AIDS Behav 2008;13(3):474-487.

60. Cha ES, Erlen JA, Kim KH, Sereika SM, Caruthers D. Mediating roles of medication–taking self-efficacy and depressive symptoms on self-reported medication adherence in persons with HIV: A questionnaire survey. International journal of nursing studies 2008;45(8):1175.

61. Rice E, Stein J, Milburn N. Countervailing social network influences on problem behaviors among homeless youth. Journal of Adolescence 2008;31:625-639.

62. Hernan MA, Brumback B, Robins J. Marginal structural models to estimate joint causal effect of nonrandomized treatments. Journal of the American Statistical Association 2001;96(454):440-448.

63. Hernan MA, Brumback BA, Robins JM. Estimating the causal eect of zidovudine on CD4 count with a marginal structural model for repeated measures. Statistics in Medicine 2002;21:1689–1709.

64. Hernan M, Robins J. Estimating causal effects from epidemiological data. Journal of Epidemiology and Community Health 2006;60(7):578-586.

65. Cole SR, Hernán MA, Robins JM, Anastos K, Chmiel J, Detels R, Ervin C, Feldman J, Greenblatt R, Kingsley L, others. Effect of highly active antiretroviral therapy on time to acquired immunodeficiency syndrome or death using marginal structural models. American journal of epidemiology 2003;158(7):687.

66. Cole SR, Hernan MA, Margolick JB, Cohen MH, Robins JM. Marginal structural models for estimating the effect of highly active antiretroviral therapy initiation on CD4 cell count. American journal of epidemiology 2005;162(5):471.

67. Brumback BA, Hernán MA, Haneuse S, Robins JM. Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures. Statistics in medicine 2004;23(5):749–767.

68. Lefebvre G, Delaney JAC, Platt RW. Impact of mis-specification of the treatment model on estimates from a marginal structural model. Statist. Med. 2008;27(18):3629-3642.

69. Epstein R. Six authors in search of a citation: Villains or victims of the Vancouver Convention? British Medical Journal 1993;306:765-767.

70. Bennett D, Taylor D. Unethical practices in Authorship of Scientific Papers. Emergency Medicine 2003;15:263-270.

71. Rennie D, Yank V, Emanuel L. When authorship fails: a proposal to make contributors accountable. Journal of the American Medical Association 1997;278:579-585.

72. Buehring G, Buehring J, Gerard P. Lost in citation: vanishing visibility of senior authors. Scientometrics. 2007;72(3):459-468.

73. Nosyk B, Sun H, Li X, Palepu A, Anis AH. Highly active antiretroviral therapy and hospital readmission: comparison of a matched cohort. BMC Infectious Diseases 2006;6(1):146.

74. Merito M, Pezzotti P, for the ICONA Study Group. Comparing costs and effectiveness of different starting points for highly active antiretroviral therapy in HIV-positive patients. Eur J Health Econ 2006;7(1):30-36.

75. Liu C, Weber K, Robison E, Hu Z, Jacobson LP, Gange SJ. Assessing the effect of HAART on change in quality of life among HIV-infected women. AIDS Research and Therapy 2006;3(1):6.

76. Liu C, Ostrow D, Detels R, Hu Z, Johnson L, Kingsley L, Jacobson LP. Impacts of HIV infection and HAART use on quality of life. Qual Life Res 2006;15(6):941-949.

77. Chu JH, Gange SJ, Anastos K, Minkoff H, Cejtin H, Bacon M, Levine A, Greenblatt RM. Hormonal contraceptive use and the effectiveness of highly active antiretroviral therapy. American journal of epidemiology 2005;161(9):881.

78. El-Bassel N, Gilbert L, Wu E. HIV and intimate partner violence among methadone-maintained women in New York City. Social Science and Medicine 2005;61:171-183.

79. Wenzel SL, Tucker JS, Elliott MN, Hambarsoomians K, Perlman J, Becker K, Kollross C, Golinelli D. Prevalence and co-occurrence of violence, substance use and disorder, and HIV risk behavior: A comparison of sheltered and low-income housed women in Los Angeles County. Preventive Medicine 2004;39(3):617–624.

80. Rotheram-Borus MJ, Song J, Gwadz M, Lee M, Rossem RV, Koopman C. Reductions in HIV risk among runaway youth. Prevention Science 2003;4(3):173–187.

81. Gangopadhyay DN, Chanda M, Sarkar K, Niyogi SK, Chakraborty S, Saha MK, Manna B, Jana S, Ray P, Bhattacharya SK, Detels R. Evaluation of Sexually Transmitted Diseases/Human Immunodeficiency Virus Intervention Programs for Sex Workers in Calcutta, India. Sexually Transmitted Diseases 2005;32(11):680-684.

82. Hogg RS, Bangsberg DR, Lima VD, Alexander C, Bonner S, Yip B, Wood E, Dong WW, Montaner JS, Harrigan PR. Emergence of drug resistance is associated with an increased risk of death among patients first starting HAART. PLoS Med 2006;3(9):e356.

83. De Luca A, Cozzi-Lepri A, Antinori A. Lopinavir/ritonavir or efavirenz plus two nucleoside analogues as first-line antiretroviral therapy: a non-randomized comparison. Antiviral Therapy 2006;11:609-618.

84. Sterne JA, Hernán MA, Ledergerber B, Tilling K, Weber R, Sendi P, Rickenbach M, Robins JM, Egger M. Long-term effectiveness of potent antiretroviral therapy in preventing AIDS and death: a prospective cohort study. The Lancet 2005;366(9483):378–384.

85. Casper C, Redman M, Huang ML, Pauk J, Lampinen TM, Hawes SE, Critchlow CW, Morrow RA, Corey L, Kiviat N, others. HIV infection and human herpesvirus-8 oral shedding among men who have sex with men. JAIDS Journal of Acquired Immune Deficiency Syndromes 2004;35(3):233.

86. Barron Y, Cole SR, Greenblatt RM, Cohen MH, Anastos K, DeHovitz JA, Delapenha R, Gange SJ. Effect of discontinuing antiretroviral therapy on survival of women initiated on highly active antiretroviral therapy. AIDS 2004;18(11):1579-1584.

87. Ko H, Hogan J, Mayer K. Estimating causal treatment effects from longitudinal HIV natural history studies using marginal structural models. Biometrics 2003;59:152-162.

88. Bachmann MO, Booysen FL. Economic causes and effects of AIDS in South African households. AIDS 2006;20(14):1861.

89. Brookhart A, van Der Laan M. A semiparametric model selection criterion with applications to the marginal structural model. Computational Statistics and Data Analysis 2006;50:475-498.

90. Wang C, Vlahov D, Galai N, Cole SR, Bareta J, Pollini R, Mehta SH, Nelson KE, Galea S. The effect of HIV infection on overdose mortality. AIDS 2005;19(9):935.

91. Eisenberg JNS, Wade TJ, Hubbard A, Abrams DI, Leiser RJ, Charles S, Vu M, Saha S, Wright CC, Levy DA, others. Associations between water-treatment methods and diarrhoea in HIV-positive individuals. Epidemiology and infection 2002;129(02):315–323.

92. Prado G, Feaster DJ, Schwartz SJ, Pratt IA, Smith L, Szapocznik J. Religious involvement, coping, social support, and psychological distress in HIV-seropositive African American mothers. AIDS and Behavior 2004;8(3):221–235.

93. Lim K. Managing HIV at the workplace: an empirical study of HIV and HR managers in Singapore. Journal of Occupational Health Psychology 2003;8(4):235-246.

94. Sengupta S, Strauss RP, DeVellis R, Quinn SC, DeVellis B, Ware WB. Factors affecting African-American participation in AIDS research. JAIDS Journal of Acquired Immune Deficiency Syndromes 2000;24(3):275.

Chapter 3

The Effect of Early Triple Therapy Among HIV-Infected Children: A Causal Inference Approach

3.1 ABSTRACT

Background
Though principally extremely rare in the United States today, the majority of human immunodeficiency virus (HIV) infections in children are attributed to mother-to-child-transmission (MTCT). Specifically, there are three avenues of infection via MTCT: in utero; at delivery; and through breastfeeding. It is estimated that the probability of infection in utero and at time of delivery is approximately 15-30%, though breastfeeding from 18-24 months the overall probability of vertical transmission increases to about 30-45%.(1) Worldwide, however, particularly in areas with no treatment availability or antenatal programs, approximately 1600 children are diagnosed with HIV every day,(2) and over 300,000 deaths among infected children occur annually worldwide.(3)

In 1996, the advent of highly active antiretroviral therapy (HAART) dramatically reduced the risk of mortality from HIV. However, the long-term effects of HAART, or triple therapy, are not yet fully understood. As a consequence, the costs and benefits of initiating therapy earlier rather than later are still at the forefront of pediatric HIV research and treatment guidance. Treatment recommendations vary between the Centers for Disease Control (CDC), World Health Organization (WHO), and Ministries of Health within individual European countries. Similar to the current CDC guidelines, the WHO recently changed the treatment guidelines to include all HIV infected children under 12 months regardless of immunologic status.(4)

In the present study, I use marginal structural models as estimated by G-Computation to estimate the causal effect of triple therapy (HAART) on time to C diagnosis, time to C diagnosis/death, and death alone among HIV-infected children.

Methods
The Pediatric Spectrum of Disease (PSD) is a multicenter active surveillance program specifically for children who have been exposed to HIV perinatally.(5) Through this program, I have identified and defined a population-based cohort of HIV positive northern Californian children who were vertically infected from 1988-2008.

To estimate the effect of therapies on outcomes of interest, traditional methods for controlling biases, like regression techniques, are often employed. However, they often fall victim to model misspecification, thus inherently biased. So-called causal inference methods are alternative techniques with causal effect interpretations.

I have looked at binary levels of A, $A \in \{0,1\}$ in triple therapy (HAART) initiated in the first 6 months versus no triple ARV therapy initiated in the first 6 months. Further, I have also looked at binary levels of A, $A \in \{0,1\}$ in triple therapy (HAART) initiated in the first 12 months versus no triple ARV therapy initiated in the first 12 months. Two subgroup analyses were performed by further restricting A to triple therapy initiated within the first 6 or 12 months of life among symptomatic children and triple therapy initiated within the first 6 or 12 months of life among asymptomatic children. I have defined a vector of baseline covariates, W, which includes immune status at treatment

initiation, length of pregnancy (full-term or less than full-term), child's race, sex, whether mothers received prenatal care, and birthweight (< 2500 grams or >= 2500 grams). Using a g-comp approach, where I define $\Psi_1 (p_0)(t_k) \equiv P(T_a > t_k)$ as all treated and $\Psi_0 (p_0)(t_k) \equiv P(T_a > t_k)$ as all untreated, I estimate the marginal additive difference and marginal log hazard ratios for each treatment scenario.

Results

The sample comprised of N=217 HIV infected children whose infection is assumed to occurred in utero or at delivery. The majority of the sample is female (56.2%) and non-White ethnicity (71.9%). Approximately half of the mothers of the children included received prenatal care. Over one-quarter of the children were born low birth weight (29.5%) and about forty-three percent were not full-term. Immune impairment at ARV treatment initiation was common as 40.1% were severely impaired and 32.3% were moderately impaired. Eight percent of the children received triple therapy in their first 6 months of life, while 45% received triple therapy within the first 12 months of life.

Though no results were statistically significant, there are some trends that should be highlighted. Among children who initiated triple therapy within 6 months of birth the causal effect of treatment in delaying a C diagnosis, $\Psi HZ(p_0)(t_k)$= -0.466 (95% CI -1.46-0.397), is seemingly stronger than children who initiated therapy within 12 months of birth ($\Psi HZ(p_0)(t_k)$= -0.321 (95% CI -0.588-0.212)). Additionally, the effect of triple therapy initiated within the first 6 or 12 months of life on time to C diagnosis is greater among *symptomatic* children (12 Months: $\Psi HZ_{symptomatic}$ (p0)(t_{36}): -0.587 (95% CI -1.217-0.480)) than among *asymptomatic* children (12 Months: $\Psi HZ_{asymptomatic}$ (p0)(t_{36}): -0.106 (95% CI -1.138-2.105)). In contrast, the effect of triple therapy initiated within the first 6 or 12 months of life on time to death is stronger among *asymptomatic* children (12 Months: $\Psi HZ_{asymptomatic}$ (p0)(t_{36}): -0.336 (95% CI -1.423-0.305)) than among *symptomatic* children (12 Months: $\Psi HZ_{symptomatic}$ (p0)(t_{36}): -0.165 (95% CI -15.297-0.621)).

Discussion

The WHO, in 2006, developed clinical and immunologic guidelines for treatment initiation in asymptomatic children in resource-limited settings based on HIV Pediatric Prognostic Markers Collaborative Study (HPPMCS) data.(4) In 2008, WHO amended their recommendations for treatment initiation for HIV-positive children as a result from an RCT in South Africa.(7) Though not statistically significant, the results from the present analysis may be interpreted as supportive of the current WHO treatment guidelines for initiating treatment among all HIV positive children, regardless of symptoms.

3.2 Introduction

The earliest known reports of Human Immunodeficiency Virus (HIV) and Acquired Immune Deficiency Syndrome (AIDS) were in 1983 among homosexual men in the United States.(8; 9) The epidemic quickly spread throughout all communities and continues today disproportionately affecting the most vulnerable populations. Early in the epidemic, childhood infections were often the result of blood transfusions and resulted in the loss of nearly the entire childhood hemophiliac community. However, since the early 1990s as researchers have learned more about the spread of the infection within children, the majority of HIV infections in children are attributed to mother-to-child-transmission (MTCT). Specifically, there are three avenues of infection via MTCT: in utero; at delivery; and through breastfeeding. It is estimated that the probability of infection in utero and at time of delivery is approximately 15-30%, though breastfeeding from 18-24 months the overall probability of vertical transmission increases to about 30-45%.(1) All three of these sources of infection have been neutralized in the United States through early prenatal diagnosis and treatment among mothers and recommendations for breastfeeding cessation in antenatal programs. In fact, in the United States, the risk for childhood infection has decreased such that it is very rare in all communities as the prophylaxis treatment for prevention of infection in children has improved and testing among mothers has become obligatory.(10) Worldwide, however, particularly in areas with no treatment availability or antenatal programs, approximately 1600 children are diagnosed with HIV every day.(2)

Infants born to HIV+ mothers are a particularly vulnerable population for myriad reasons. Pregnant women infected with HIV are more likely to give birth to a low birth weight (LBW) infant if they are not taking antiretroviral medications.(11) Without access to appropriate drugs and care, a reality in the developing world, these infants are likely to seroconvert, as well. In turn, a LBW child infected with HIV via MTCT has two significant threats to his immune function—LBW and HIV. Furthermore, while the burden of vertical transmissions is in Sub-Saharan Africa and practitioners know the ways to prevent MTCT, often in these developing countries food insecurity leaves HIV-positive mothers little choice but to continue exposing their children to HIV via breastfeeding.(12; 13)

Determining which children are more likely to convert than others may prove to be invaluable in perinatal HIV treatment and prevention. Embree et al found that in a cohort of HIV exposed and unexposed children enrolled at birth, postnatally infected children, those who seroconverted at some time after 3 months of age, were more likely to have lower, pre-seroconversion CD4+ counts than HIV unexposed children (p value < 0.007).(14) Research has also shown that HIV-infected newborns with fewer t-cells are more likely to progress quicker than HIV-infected newborns with more t-cells.(15) A study exploring seasonality and maternal factors among HIV exposed, but initially uninfected, children in Cameroon, birth weight was a significant predictor of

seroconversion in their first two months of life (p value < 0.01).(16) These results are particularly notable because all mothers enrolled in the study were treated with Nevirapine at the beginning of labor. In contrast, the use of antiretroviral therapy during pregnancy and labor is associated with an increase in birth weight in HIV exposed children.(11)

Not all children who are exposed to HIV perinatally seroconvert, independent of retroviral therapy. The more severe the infection in the mothers, the more likely she is to transmit the disease in utero and through breastfeeding, in the absence of prophylaxis. The standard treatment for prevention of MTCT is single dose nevirapine (sdNVP) perinatally, though about 10% of children exposed to sdNVP will still develop HIV even before breastfeeding.(17; 18)

Today, over 90% of the estimated 2.5 million HIV infected children worldwide live in sub-Saharan Africa.(3) Over 300,000 deaths among infected children occur annually worldwide.(3) The burden of pediatric HIV infections lies in the poorest regions of sub-Saharan Africa, where approximately only 10% of mothers have access to antenatal programs aimed at preventing MTCT.(19) As access to care in developing areas continues to affect the health of future mothers, so too does it affect their children's health. A recent study in South Africa found that 85% of HIV-infected, sdNVP exposed infants were moderately or severely immuno-compromised (CD4 % < 25) by 6 months post-partum,(17) suggesting these children were particularly vulnerable due to the severity of infection in their mothers.

Family planning services play a key role in reducing MTCT as both unwanted pregnancies and vertical transmission of HIV result from engaging in unprotected sex. Not only is an increase in contraception use more cost-effective at reducing MTCT than preventative drugs, namely Nevirapine,(20) but Sweat et al found that even a small reduction in unwanted pregnancies prevents an equal number of HIV infections as provision of Nevirapine.(21) In fact, an argument can be made that the risk of vertical transmission would be reduced if HIV positive women were able to space their births further apart because not only would their need for prevention of MTCT services be reduced, but her own health would improve, thus improving her virologic status.(22)

In 1996, the advent of highly active antiretroviral therapy (HAART) dramatically reduced the risk of mortality from HIV. Results from birth cohort studies of HIV+ children indicate that approximately 70-80% of children left untreated will survive to age five.(23-25) Previously, only three large, observational studies have evaluated the impact of HAART on mortality among HIV-infected children (regardless of symptoms),(26; 23) only one of which had a longer period of follow-up of HAART-exposed children. In turn, the long-term effects of HAART on mortality among children still need to be explored. Patel et al, with the use of marginal structural models, estimated the weighted, adjusted proportional hazard for mortality as 0.24 (95% CI 0.11–0.51) when comparing HAART treated children to untreated children.(27) Similarly, Gortmaker et al found a reduced hazard ratio for death (HR 0.33; 95% CI 0.19-0.58)(26) and de Martino et al found a reduced RH of death (RH 0.29; 95% CI 0.13-0.67) among triple therapy initiated

children compared to untreated children.(23) After the roll-out of HAART, survival expectedly has only improved.(23; 28) HAART is used as a first-line treatment now among HIV infected children in order to recuperate from HIV-associated illnesses and re-establish immuno-compentence.(29-33)

Researchers are still trying to establish the most ideal time to initiate antiretroviral therapy in vertically infected children. Weighing the benefits and risks of early initiation of HAART or triple therapy is a necessary component in making treatment guidelines recommendations. Treatment recommendations vary between the Centers for Disease Control (CDC), World Health Organization (WHO), and Ministries of Health within individual European countries. Nevertheless, the varying treatment guidelines for pediatric HIV do not seem to significantly affect the clinical outcomes.(34) For numerous reasons, not the least of which are the inherent ethical issues, randomized controlled trials (RCT) exploring the best time to initiate HAART in HIV positive children are very uncommon. In fact, the only published randomized trial estimating the effect of early HAART versus delayed HAART on mortality among HIV positive infants prematurely terminated in 2008 as a result of an unbalanced, disproportionate number of deaths in the delayed group.(35) One other RCT was conducted to explore the impact of delaying HAART initiation on clinical disease progression, however this study was a small feasibility study in preparation for another, larger RCT which will likely be completed in 2011. Moreover, the study population only included HIV positive children 1-12 years of age, excluding all positive infants.(36) Prior to the HAART era, the PENTA 1 study conducted a similar study of delayed versus early initiation of zidovudine monotherapy.(37) Their results suggest that early initiation of ART has no added benefit on clinical outcomes. Among non-RCTs, Newell et al from the European Collaborative Study, a prospective study of a birth cohort of 131 HIV infected children, conclude that initiating ART in the first 5 months of life and the use of HAART were both highly predictive of an improved CD4 z-score 6 months after treatment initiation.(38) In one of the only other identified observational studies evaluating the impact of delayed treatment initiation among HIV positive infants, Chiappini et al found children treated early with HAART had significantly lower viral load than deferred treatment children and they were also less likely to progress to a C diagnosis.(39)

Results from adult studies that have explored HAART and its effect on mortality are likely not reflective of its effect on mortality if applied to pediatric populations.(40) Typically CD4 counts and viral load are clinical parameters used to assess disease progression in adult populations, but these same parameters have a wide variability in children.(41-43) Additionally, in children HIV affects neuro-cognitive development, growth, and an immune system that is not yet fully mature.(44-46) Though adult observational studies have shown that patients who start HAART early (higher CD4 counts) have better clinical outcomes than adults who start at CD4 counts below 200 cells/μL, these results are likely not generalizeable to the pediatric HIV infected community.

As a result of new evidence from one prematurely terminated RCT estimating the impact of early HAART on mortality, the WHO recently changed the treatment guidelines to

include all children under 12 months regardless of immunologic status.(4) Previously, treatment for pediatric HIV infection was only recommended for children who presented symptomatically. The previous age-specific recommendations by the WHO are listed in Appendix A.2.(7) The implications for these treatment guidelines are particularly important for the infected infants who will now have to be on HAART for life. Once a child begins therapy, he must remain on treatment for life, or he risks developing drug resistance, which may in turn hasten his death.

The long-term exposure to HAART has known adverse health implications for HIV-infected adults. In fact, HAART has been associated with an increased risk of hyperlipidemia, lipodystrophy, and atherosclerosis.(47) This risk is not as well understood in infants, but there is evidence of an increased risk of cardiac abnormalities and mitochondrial damage in children, HIV positive and negative alike, whose mothers were treated with HAART.(48; 49)

The CDC's clinical categories of HIV disease among children helps determine the progress of the disease and establish immune suppression. Though the CDC's treatment guidelines include all HIV positive children under 12 months, previous algorithms were employed to determine treatment eligibility. Essentially, severe disease was determined by a combination of clinical presentations and immunologic measurements (CD4 count or preferably CD4%). These criteria are outlined in Appendix A.3. In contrast, the Pediatric European Network for the Treatment of AIDS (PENTA) group's treatment recommendations are less aggressive; essentially, treatment is recommended among infants if they have a C diagnosis or CD4% less than 20% (see Appendix A.4).(50)

Particularly a problem in HIV/AIDS literature, observational studies are often biased as traditional analysis methods are employed to estimate the effect of a treatment on an outcome of interest. Causal inference methods have been developed to overcome many of these biases and have been employed in the present study. Among the four previously published observational studies exploring the effect of HAART on mortality in a pediatric population, two applied causal inference methods, one of which was a follow-up study of the first study.(27; 51)

To overcome the inherent issue of correct model specification in time to event observational studies, I aimed to use g-computation, a marginal structural models (MSM) estimator, to estimate the causal effect of HAART (interchangeably referred to as triple therapy) on reducing AIDS/death among children who were infected in utero. Additionally, I performed a subanalysis of symptomatic and asymptomatic children.

3.3 Methods

3.3.1 Study Population

The Pediatric Spectrum of Disease (PSD) is a multicenter active surveillance program specifically for children who have been exposed to HIV perinatally.(5) Since 1988 this program has been located at Stanford University and has a surveillance catchment area of 12 counties in northern California with a total population of approximately 6 million. Through this program, I have identified and defined a population-based cohort of HIV positive northern Californian children.

Researchers working with the PSD database examine records from the California Children Services program, which provides case management services for HIV infected children, and medical records at hospital-based clinics. Study nurses visited pediatric HIV clinics biannually for data extraction from medical records and to identify new patients entering the PSD database. Medical records for all children under 18 years of age were followed until they were lost to follow up, died, or their status was definitely negative. Vertical transmission was determined by the CDC classification system for HIV in children younger than 13 years of age.(30; 52) An alphanumeric code combined with the birth date was used as a unique identifier to preserve confidentiality and avoid record duplication. For the ongoing surveillance for the PSD database, institutional review boards approval has been granted annually by the enrolling hospitals for the children and by Stanford University. For the present study, approval was obtained from the institutional review boards of Stanford University and University of California-Berkeley.

3.3.2 Statistical Methods

Data Structure
In the present analysis, I have done a time to event analysis to explore the effect of treatments (triple ARV therapy, or no triple ARV therapy), A, have on the amount of time until my event of interest occurs. That is to say, I have estimated the time, T, it takes for a child to experience an event (1: Category C; 2: Category C diagnosis or death; 3: death).

I have looked at binary levels of A, $A \in \{0,1\}$ in triple therapy (HAART) initiated in the first 6 months versus no triple ARV therapy initiated in the first 6 months. To allow for a less restrictive treatment assignment, I have also looked at binary levels of A, $A \in \{0,1\}$ in triple therapy (HAART) initiated in the first 12 months versus no triple ARV therapy initiated in the first 12 months. Two subgroup analyses were performed by further restricting A to triple therapy initiated within the first 6 or 12 months of life among symptomatic children and triple therapy initiated within the first 6 or 12 months of life among asymptomatic children. I have defined a vector of baseline covariates, W, which includes immune status at treatment initiation, length of pregnancy (full-term or less than

full-term), child's race, sex, whether mothers received prenatal care, and birthweight (< 2500 grams or >= 2500 grams). I have defined T as a discrete variable with values {1,….,K}, where K is last time point children are monitored. In contrast, I defined censoring, C, as the last time point children are observed. My data structure also defines whether an event has occurred as N_1 and a similar scenario for censoring with N_2. In turn, all time points until the event occurs are denoted by $dN_1(t) = 0$ and $dN_1(t) = 1$ at the time point the event occurs. All time points until a child is censored are denoted by $dN_2(t) = 0$ and $dN_2(t) = 1$ at the time of censoring. The long form of my observed data can be expressed as n iid observations of $O = (A,W,dN_1(t), dN_2(t): t=1, \ldots, K) \sim p_o$, where p_o is the density of the my observed data, O.

The likelihood of the observed data is described as:

$$L(0) = P(W)P(A|W) \prod_{t=1}^{k} P(dN_1(t) \mid dN_1(t-1) = 0, dN_2(t-1) =0,A,W) \quad (1)$$
$$P(dN_2(t) \mid dN_1(t) = 0, dN_2(t-1) = 0, A, W)$$

Where,

$Q_{10}(W) \equiv P(W)$ is the distribution of baseline covariates, W;
$Q_{20}(N_1(t),A,W) \equiv P(dN_1(t) \mid dN_1(t-1) = 0, dN_2(t-1) =0,A,W)$ is the conditional hazard of the event (C diagnosis and/or death) given the treatment (A) and baseline covariates, W;
$g_{10}(A,W) \equiv P(A \mid W)$ is the treatment mechanism;
$g_{20}(N_2(t), A,W) \equiv P(dN_2(t) \mid dN_1(t) = 0, dN_2(t-1) = 0, A, W)$ is the censoring mechanism—the conditional hazard of censoring given the subject did not yet experience an event, no previous censoring, and given the treatment, A, and baseline covariates.

In turn, the likelihood (equation 1) factorizes the distribution of W, baseline covariates, the missingness mechanism, g, and the conditional hazard of the outcome of interest. To estimate the survival, that is, the probability of surviving to time k given treatment, A, and baseline covariates, W, one would define $S_0(t_k \mid A, W) = P(T > t_k \mid A,W)$. Then one would take the cumulative product of 1 minus the conditional hazard of C diagnosis and/or death to estimate survival (see equation 2):

$$S_0(t_k \mid A, W) = \prod_{t=1}^{t_k}(1- Q_{20}(N_1(t),A,W)) \quad (2)$$

In time to event analyses, researchers will often a priori specify a parametric hazard model in traditional techniques exploring the effect of a treatment, A, on a time to event outcome, and test if A is different from zero. Covariates are selected for inclusion in these traditional approaches usually in a rudimentary way—selecting and deleting covariates based on their influence on A's effect on the outcome. One of the most common approaches for time to event scenarios is the Cox proportional hazard model, or

logistic regression in the case of discrete time outcomes. Furthermore, time will often be fit and a linear model will be employed to estimate the effect of A and the covariates. The parameters in conditional hazard models are estimated with a maximum likelihood approach. The central feature here is to evaluate whether the parameter representing the treatment, A, is significantly different from 0. Again, these estimates are heavily dependent on how well the model is specified. In turn, the parameters within hazard models may not be correctly specified, though this may go unnoticed unless the selected hazard model is contrasted with alternative models. It should be noted that the parameter estimating the effect of A on the outcome of interest is only relevant within that specific model. In the optimistic case of correctly specifying the hazard model, then this parameter represents the log odds ratios of the event occurring at each time point for A, and only in the context of that specific model. Essentially, this is the proportional hazards assumption—A's effect is identical at every time point.

Furthermore, the log-rank statistic in a Cox proportional hazards model can be described as the effect of A on time to event, but only if the model is correctly specified. Proportional hazards are similarly assumed in these models while fitting the baseline hazard. Either the conditional hazards or Cox proportional hazards models are often biased as a result of incorrectly specified models.

G-Computation Approach

Unfortunately, restrictive parametric models often seen in traditional methods are not usually representative of the data generating distribution of the outcome of interest, even if baseline covariates are included in the model.(53) Using an approach that uses parameters that are naturally selected based on the data as opposed to using parameters selected in the ways highlighted above allows for easier interpretation of the model and its parameters. Additionally, the models' parameters are only correct in the context in the specified model.

For example, let us define our parameter of interest as a function of our data generating distribution, $\Psi(p_0)$. If one is interested in the survival at time point, t_k, the specific survival curve may be expressed as $P(T_a > t_k)$. Once again referring to the counterfactual framework, if a subject's treatment level were to be set at a, then T_a would be the event time, T, one would have observed, regardless of whether that subject's true observed treatment is at level a. Had this subject been treated with a different level of a than what is expressed in $P(T_a > t_k)$, then this is a counterfactual description of his survival curve.

The causal assumptions for the present study are illustrated in Figures 3.1 and 3.2. In general, I could assign all my subjects to treatment a, triple therapy, and their censoring, $dN_2(t)$, to not censored at all t throughout the study (Figure 3.2). In turn, I will have my counterfactual outcome—treated with triple therapy and not censored. In order to quantify the causal effect of A, triple therapy, on mortality (or death/category C diagnosis), the effect was estimated using these counterfactuals. More specifically, I used a marginal additive difference $\Psi_{RD}(p_0)(t_k)$, in the probability of survival. That is,

$$\Psi_1(p_0)(t_k) \equiv P(T_a > t_k) \text{ [all treated]} - \Psi_0(p_0)(t_k) \equiv P(T_a > t_k) \text{ [all untreated]} \quad (3)$$

My parameters established in this fashion may also be influenced by baseline covariates, defined previously as W. As such, I can estimate the counterfactual survival at t_k setting other baseline variables at 0 or 1. Additionally, the marginal log hazard ratio is defined by:

$$\Psi_{RH}(p_0)(t_k) = \log((\log(\Psi_1(p_0)(t_k)))/(\log(\Psi_0(p_0)(t_k)))) \quad (4)$$

A directed acyclic graph (DAG) often helps researchers decide which variables are necessary for adjustment by identifying possible confounders. Elements within DAGs were described previously (Ch.2-Causal Inference Methods applied to HIV/AIDS Data). Figure 3.3 is a DAG describing the possible associations between initiation of HAART and HIV mortality among infected children. Similarly, Figure 3.4 describes the possible associations between initiation of HAART and a C Diagnosis. My analysis plan included exploring each of the measured baseline covariates within Figures 3.3 and 3.4 and define them as W.

In order to express the mean counterfactual outcomes, as described by equations 3 and 4, I have employed the G-computation estimator. Because it is important to estimate the distributions of my baseline covariates, W, the conditional hazard of the event given their treatment, A, and W, and the conditional survival of the outcome of interest, as related to the conditional hazard, I have employed super learner software (Deletion/Substitution/Addition). The empirical distribution of my baseline covariates in my data estimate non-parametrically the marginal distribution of W. By using this data-adaptive machine learning algorithm, and its cross-validation based on likelihood, I am avoiding the problems inherent with traditional approaches and model building.(60) All confidence intervals for G-computation estimates were calculated by bootstrap sampling.

Specifically, my data analysis plan includes using DSA on the entire data set in the repeated measures form for discrete time survival analysis. I have forced in A and indicators for time into the DSA, in turn, time was non-parametrically fit. Then, the standard g-comp was performed where everyone is be set equal to A=1 (treated) and the DSA fit was used to get conditional hazards from $1...t_k$. I then took the product integral of one minus the hazard. As a result, for each child I derived an estimated survival as though the child were treated and then I took the mean over those children to get the mean treatment specific survival at t_k. I repeated this process for children setting A=0.

In an intent-to-treat approach, I assumed the initial treatment, whether mono, dual, or triple therapy, was unmodified throughout the study. In order to establish t=0, that is the first day of follow-up for each child, I employed an algorithm to include only children who were assumed to be infected in utero or delivery. In turn, I excluded n=60 children who were likely infected postnatally via breastfeeding, which could have occurred any time throughout breastfeeding. This approach included children whose mothers were assumed to be knowledgeable about their infection as breastfeeding was likely discouraged among these women. To identify these children, an algorithm was applied

that identified children whose mothers showed HIV symptoms during their pregnancy, symptoms during their delivery, were known to have taken HIV medication, or were known to have received prenatal care, as mandatory HIV testing for pregnant women began in 1987. Two separate, primary analyses were explored regarding timing of initiation of HAART—a) starting triple therapy in the first six months of life, b) starting triple therapy in the first twelve months of life.

For my subgroup analyses of children who were asymptomatic and symptomatic at time of treatment initiation, I constructed an algorithm to identify their disease status. The definition of asymptomatic children was adapted from the CDC's definition of severe/moderate/mild immune suppression among children (see Appendix A.3). In short, somewhat similar to PENTA's previous treatment initiation guidelines (see Appendix A.4), asymptomatic children were described as not having a C diagnosis and not having a CD4% below 15%. In the absence of CD4% data, the CD4 count as it relates to the immune competence age-specific threshold was used. To ensure children who began ART were asymptomatic, another algorithm was applied to identify children who were diagnosed with a C diagnosis 4 weeks or more after the initial treatment. Previously, it has been shown that at least four weeks of ARV treatment are needed to have any clinical effectiveness.(54) Additionally, this algorithm identified the CD4% or CD4 counts within four weeks of ARV treatment initiation to ensure that the immunological data (CD4% and CD4 count) at (or near) treatment initiation were likely unaffected by ARV initiation.

The variables included in the primary analysis were length of pregnancy (full-term or not full-term), sex, birthweight (<2500 grams, >2499 grams), race (non-White or White ethnicity), prenatal care, and immune status at treatment initiation (severely, moderately, or mildly/not suppressed). The same variables were included in the subanalysis of asymptomatic/symptomatic children, except for immune status at treatment initiation as this information is captured in the definition of A in the subanalysis.

Analyses were limited to the first 3 years of life for C diagnosis and C diagnosis/death. Death was a somewhat rare event in the first 36 months; in turn, for death alone, survival analyses were limited to the first 60 months of life to capture more death cases.

In the present study, I have estimated the causal effect of triple therapy on mortality or C diagnosis in children enrolled in a population-based study using marginal structural models as estimated by G-computation methods. For comparison, assuming the model is correct, I estimated the coefficient in front of A in a Cox proportional hazards model, the log-rank statistic, for each treatment scenario comparison, which is the log-rank statistic.

3.4 Results

3.4.1 Demographics and Baseline Characteristics

After excluding children whose time of infection was difficult to ascertain or whose infection likely occurred via breastfeeding, n=60, the sample was comprised of N=217 HIV infected children whose infection was assumed to occur in utero or at delivery. This sample was the population used for the primary analysis—time to C diagnosis and/or death; time to C diagnosis; time to death. Patient characteristics are outlined in Table 3.1. The majority of the sample was female (56.2%) and non-White ethnicity (71.9%). Approximately half of the mothers of the children included received prenatal care. Over one-quarter of the children were born low birth weight (29.5%) and about forty-three percent were not full-term. Immune impairment at ARV treatment initiation was common as 40.1% were severely impaired and 32.3% were moderately impaired. Eight percent of the children received triple therapy in their first 6 months of life, while 45% received triple therapy within the first 12 months of life. About 55% of the sample either never received triple therapy or initiated therapy after 12 months of life.

The associations between the baseline covariates (W) and triple therapy initiation in the first 6 months and in the first 12 months are described in Tables 3.2 and 3.3. As the immune status deteriorated to moderate or severe immune suppression, children were more likely to initiate therapy within their first 6 months of life (cOR = 1.61; p value 0.12). Similarly, children born to mothers who received prenatal care were more likely to initiate triple therapy within 6 months of birth, though this relationship is not statistically significant (cOR = 1.47; p value 0.20). Though not significant, the data suggest that children of White ethnicity were half as likely to initiate triple therapy within their first 6 months of life (cOR = 0.49; p value 0.27). The associations between triple therapy initiation in the first 6 months of life and W were approximately the same at 12 months, though race and immune status at treatment initiation become significant. Namely, as the immune status worsens at treatment initiation, the odds of starting triple therapy in the first 12 months are increased 1.82 times (p value = 0.03). The odds of beginning triple therapy in the first 12 months among White children decrease to 0.28 when compared to non-White children (p value = 0.05).

A total of n=75 children were diagnosed with a C diagnosis within the first 36 months of life. Bivariate estimates of W and C diagnosis within the first 36 months of life are listed in Table 3.4. Both ethnicity and immune status at treatment initiation are strongly associated with C diagnosis in the first 36 months. Specifically, a child of White ethnicity is more than two times more likely to be diagnosed with a C diagnosis than a non-White child (cOR = 2.39; p value < 0.01); the worse the immune status at treatment initiation the more likely the child was to be diagnosed with a C diagnosis (cOR = 2.17; p value < 0.01). Using a traditional approach to identifying possible confounders of the triple therapy and C diagnosis relationship, ethnicity and immune status at treatment

initiation appear to be the only baseline covariates that should be considered confounders as they are both related to the treatment and the outcome.

A total of n=84 children were either diagnosed with a C diagnosis or died within 36 months of birth (see Table 3.5). Again, White children were twice as likely to either be diagnosed with a C diagnosis or die within the first 36 months of life (cOR = 2.01; p value < 0.01). A worsening of immune status at treatment initiation would increase a child's odds of being diagnosed with a C diagnosis or dying within the first 3 years of life by nearly two-fold (cOR = 1.90; p value < 0.01).

A total of n=58 children died by the end of 60 months of follow-up. Table 3.6 shows the bivariate estimates of death and W. The only baseline covariate with borderline statistical significance is ethnicity. That is to say, White children had about 1.7 times the odds of dying within the first 5 years as non-White children (cOR = 1.69; p value = 0.11).

For the sub-analyses, the sample populations were limited to children who were treated asymptomatically and children who were treated symptomatically at treatment initiation. N=10 symptomatic children were treated with triple therapy within their first 6 months of life, and n=8 asymptomatic children were similarly treated (see Table 3.7). With so few observations, expectedly there were no significant findings in the bivariate analyses.

Forty-eight children were treated with triple therapy symptomatically within their first 12 months of life. In contrast, n=50 children were asymptomatic at treatment initiation in their first 12 months of life (see Table 3.8). The bivariate estimates of triple therapy initiation and W are listed below. Just as the case with treatment initiated within 6 months among symptomatic and asymptomatic children, there are no significant associations between any baseline covariate and triple therapy initiation within 12 months among symptomatic and asymptomatic children.

Time to Event Analysis: A G-Computation Approach

The data were expanded such that time to event outcomes could be estimated. To estimate the survival, specifically the probability of surviving to time k given treatment, A, and baseline covariates, W, I defined $S_0(t_k \mid A, W) = P(T > t_k \mid A,W)$. Additionally, I estimated the cumulative product of 1 minus the conditional hazard of experiencing the event to estimate survival:

$$S_0(t_k \mid A, W) = \prod_{t=1}^{t_k}(1- P(dN_1(t) \mid dN_1(t-1) = 0, dN_2(t-1) =0,A,W))$$

The definitions of A vary based on the research question I am trying to answer. Treatment, A, is defined as: 1) Treatment with triple therapy within the first 6 months of life; 2) treatment with triple therapy within the first 12 months of life; 3) treatment with triple therapy within the first 6 months of life among asymptomatic children; 4) treatment

with triple therapy within the first 6 months of life among symptomatic children; 5) treatment with triple therapy within the first 12 months of life among asymptomatic children; 6) treatment with triple therapy within the first 12 months of life among symptomatic children.

The model selection was performed by an algorithm that selects the fit of the initial hazard data-adaptively. A super learner, Deletion/Substitution/Addition (D/S/A), was used to search through function forms using deletion, substitution, and addition actions. Sinisi and van der Laan have applied this algorithm to fit the initial hazard on pooled data over time.(55) The covariates selected by D/S/A from the candidate covariates in W for each model are identified in each subsection below.

The survival probabilities at time *k*, $\Psi_1(p_0)(t_k)$ and $\Psi_0(p_0)(t_k)$, are described in each subsection. Similarly, the marginal log hazard ratios at time *k*, denoted by $\Psi HZ(p_0)(t_k)$, and marginal additive differences at time *k*, $\Psi AD(p_0)(t_k)$, are given for each comparison. For comparison purposes, the log-rank statistic in a Cox proportional hazards model was estimated for each comparison, as well, while adjusting for the same baseline covariates selected by D/S/A in the marginal structural approach; this estimate is only correct is the model is correctly specified.

3.4.2 All Children—Regardless of Symptoms

Time to C Diagnosis (36 Months of follow-up)

A initiated in First 6 months of Life

In estimating the time to a C diagnosis, D/S/A selected sex, race, and pregnancy term as covariates within our treatment mechanism. The unadjusted binary estimates of W on time to a C diagnosis are listed in Table 3.9. Male children were less likely to be diagnosed with a C diagnosis than female children (OR = 0.52; 95% CI 0.31-0.84). Furthermore, children born at term were significantly less likely as children not born at term to be diagnosed with a C diagnosis (OR = 0.76; 95% CI 0.58-1.01).

A Initiated in First 12 months of Life

Similarly, the same baseline covariates were selected using D/S/A when estimating the effect of triple therapy initiated in the first 12 months of life on a C diagnosis.

In Figure 3.5a, the G-computation survival curves for the causal effect of triple therapy initiated in the first 6 months of life on time to C diagnosis among all children are illustrated, similarly for the estimates for therapy initiated in the first 12 months of life in Figure 3.5b.

The causal treatment specific parameters at time *k* related to triple therapy on C diagnosis had all children been treated or untreated within 6 or 12 months of birth are listed in Table 3.10. Using the counterfactual language, the estimates in the first column are the causal treatment specific parameters related to triple therapy on C diagnosis had all

children started therapy within 6 months of birth. The second column illustrates the causal treatment specific parameters related to triple therapy on C diagnosis had all children been untreated within 6 months of birth.

In Table 3.11, the marginal additive differences at time *k*, denoted by $\Psi AD(p_0)(t_k)$, are increasing over time for both treatment initiation scenarios; at 36 months the marginal additive difference is 0.120 (95% CI -0.127-0.270) among children who initiated therapy within 6 months. Among children who initiated within 12 months, the marginal additive differences at time *k* are also increasing over time; at 36 months the marginal additive difference is 0.087 (95% CI -0.065-0.151). This suggests that the causal effect of triple therapy initiated in the first 12 months of life on time to C diagnosis is not as strong as the causal effect of triple therapy initiated in the first 6 months of life as the marginal additive difference is less. However, neither of the marginal additive differences is statistically significant. The marginal log hazard ratios at time *k*, denoted by $\Psi HZ(p_0)(t_k)$, are described in Table 3.11 and stay somewhat proportional through time k-1 for both treatment initiation scenarios. At 36 months the log hazard ratio for children who initiated treatment within 6 months of birth is -0.466 (95% CI -1.46-0.397). The marginal log hazard at 36 months estimating the causal effect of triple therapy initiated in the first 6 months of life is greater than the same ratio for children who initiated treatment within the first year, $\Psi HZ(p_0)(t_k)$= -0.321 (95% CI -0.588-0.212). Though not statistically significant, these results suggest that among children who initiated triple therapy within 6 months of birth the causal effect of treatment in delaying a C diagnosis is stronger than children who initiated therapy within 12 months of birth.
It should be noted that $\Psi HZ(p_0)(t_k)$, averaged over t is equivalent to the Cox proportional hazards parameter, in turn the log rank test parameter. Comparing these results to a more traditional approach, the Cox proportional hazards parameter at 36 months comparing children treated within the first 6 months of life and children not treated within the first 6 months of life is HR=-0.476 (p value = 0.356). The Cox proportional hazards parameter at 36 months comparing children treated within the first 12 months of life and children not treated within the first 12 months of life is HR=-0.407 (p value = 0.089).

Time to C Diagnosis or Death (36 Months of Follow-up)

Initiated in First 6 months of Life

In estimating the time to a C diagnosis or death, often referred to as AIDS-free survival, among all children who started triple therapy in their first 6 months of life, D/S/A, selected no unforced terms for the treatment mechanism. Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, immune status at treatment initiation, and time indicator variables.

Initiated in First 12 months of Life

Similarly, no unforced terms were selected by D/S/A for the treatment mechanism in estimating the time to a C diagnosis or death among children who started triple therapy in the first 12 months of life.

The G-computation survival curves for the causal effect of triple therapy initiated in the first 6 months of life on time to C diagnosis or death among all children are illustrated in Figure 3.6a. Similarly, in Figure 3.6b are the survival curves for the estimates for therapy initiated in the first 12 months of life.

The causal treatment specific parameters at time *k* related to triple therapy on C diagnosis or death had all children been treated or untreated within 6 or 12 months of birth are listed in Table 3.12. Again, using the counterfactual language the estimates in the first column are the causal treatment specific parameters related to triple therapy on C diagnosis or death had all children started therapy within 6 months of birth. The second column illustrates the causal treatment specific parameters related to triple therapy on C diagnosis or death had all children been untreated within 6 months of birth. Similarly, the third and fourth columns relate to the counterfactuals related to initiating triple therapy within the first 12 months of life.

In Table 3.13, the marginal additive differences at time *k*, denoted by $\Psi AD(p_0)(t_k)$, are increasing over time for both treatment initiation scenarios; at 36 months the marginal additive difference is 0.108 (95% CI -0.110-0.311) among children who initiated therapy within 6 months. Among children who initiated within 12 months, the marginal additive differences at time *k* are also increasing over time; at 36 months the marginal additive difference is 0.055 (95% CI -0.150-0.158). The marginal additive difference of effect among children who initiated triple therapy within 12 months of life is approximately half the marginal additive difference among children who initiated within their first 6 months of life. The marginal log hazard ratios at time *k*, denoted by $\Psi HZ(p_0)(t_k)$, are described in Table 3.13 and stay somewhat proportional through time k-1 for both treatment initiation scenarios. At 36 months the log hazard ratio estimating time to a C diagnosis or death among children who initiated treatment within 6 months of birth is -0.369 (95% CI -1.588-0.300). In contrast, the marginal log hazard at 36 months estimating the causal effect of triple therapy initiated in the first 12 months of life is less pronounced ($\Psi HZ(p_0)(t_k) = -0.180$ (95% CI -0.599-0.445)). Though not significant, these results suggest that the causal effect of triple therapy initiated in the first 6 months on time to C diagnosis or death is twice the causal effect of triple therapy initiated in the first 12 months of life. The Cox proportional hazards parameter at 36 months comparing children treated within the first 6 months of life and children not treated within the first 6 months of life is HR=-0.346 (p value = 0.454). The Cox proportional hazards parameter at 36 months comparing children treated within the first 12 months of life and children not treated within the first 12 months of life is HR=-0.431 (p value = 0.058).

Time to Death (60 Months of Follow-up)

Initiated in First 6 months of Life

In estimating the time to death, D/S/A selected no unforced terms for the treatment mechanism. Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, immune status at treatment initiation, and time indicator variables.

Initiated in First 12 months of Life
No unforced terms were selected by D/S/A for the treatment mechanism in estimating the time to death among children who initiated triple therapy within 12 months of life.

The G-computation survival curves for the causal effect of triple therapy initiated in the first 6 months of life on time to death among all children are illustrated in Figure 3.7a. Similarly, in Figure 3.7b are the survival curves for the estimates for therapy initiated in the first 12 months of life.

The causal treatment specific parameters at time *k* related to triple therapy on death had all children been treated or untreated within 6 or 12 months of birth are listed in Table 3.14. The estimates in the first column are the causal treatment specific parameters related to triple therapy on death had all children started therapy within 6 months of birth. The second column illustrates the causal treatment specific parameters related to triple therapy on death had all children been untreated within 6 months of birth.

In Table 3.15, the marginal additive differences at time *k*, denoted by $\Psi AD(p_0)(t_k)$, are increasing over time for both treatment initiation scenarios. The marginal additive difference at 60 months for children who initiated therapy within 6 months of life is 0.046 (95% CI -0.229-0.200). When compared to the marginal additive difference among children who initiated therapy within 12 months of life, 0.048 (95% CI -0.249-0.130), there is no discernable difference in triple therapy's effect on time to death under the two treatment initiation scenarios. Similarly, the marginal log hazard ratios of triple therapy's effect on time to death at 60 months under each treatment plan are comparable ($\Psi HZ(p_0)(t_{60})$ = 6 months: -0.199 (95% CI -1.378-0.825); 12 Months: -0.205 (95% CI -0.526-0.821)). The Cox proportional hazards parameter at 60 months comparing children treated within the first 6 months of life and children not treated within the first 6 months of life is HR=-0.203 (p value = 0.695). The Cox proportional hazards parameter at 60 months comparing children treated within the first 12 months of life and children not treated within the first 12 months of life is HR=-0.432 (p value = 0.114).

3.4.3 All Children—Asymptomatic vs Symptomatic

Time to C Diagnosis (36 Months of follow-up)

Initiated in First 6 months of Life-Asymptomatic Children
In estimating the time to C diagnosis among *asymptomatically* treated children who started triple therapy within 6 months of birth, D/S/A selected no unforced terms for the treatment mechanism. Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, and time indicator variables.

Initiated in First 6 months of Life-Symptomatic Children

In estimating the time to a C diagnosis among *symptomatically* treated children who started triple therapy within 6 months of birth, D/S/A selected an interaction between prenatal care and a time indicator variable for 25-30 months, and selected an interaction between sex and pregnancy term as covariates within our treatment mechanism. Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, and time indicator variables.

The G-computation survival curves for the causal effect of triple therapy initiated in the first 6 months of life on time to C diagnosis among *asymptomatic* children are illustrated in Figure 3.8a. Similarly, in Figure 3.8b are the survival curves for the estimates for therapy initiated in the first 6 months of life among *symptomatic* children.

The causal treatment specific parameters at time *k* related to triple therapy on C diagnosis had all *asymptomatic* or *symptomatic* children been treated or untreated within 6 months of birth are listed in Table 3.16. The estimates in the first column are the causal treatment specific parameters related to triple therapy on C diagnosis had all *asymptomatic* children started therapy within 6 months of birth. The second column illustrates the causal treatment specific parameters related to triple therapy on C diagnosis had all *asymptomatic* children been untreated within 6 months of birth. Similarly, the third and fourth columns are the causal treatment specific parameters related to triple therapy on C diagnosis among *symptomatic* children.

In Table 3.17, the marginal additive differences at time *k*, denoted by $\Psi AD(p_0)(t_k)$, are increasing over time for both treatment initiation scenarios. The marginal additive difference at 36 months for *asymptomatic* children who initiated therapy within 6 months of life is 0.115 (95% CI -0.327-0.446). When compared to the marginal additive difference among *symptomatic* children who initiated therapy within 6 months of life, 0.146 (95% CI -0.214-0.355), there is a slight increase in the marginal additive difference in triple therapy's effect on time to C diagnosis among *symptomatically* treated children. Similarly, the marginal log hazard ratio of triple therapy's effect on time to C diagnosis at 36 months among *symptomatically* treated children is increased when compared to *asymptomatically* treated children ($\Psi HZ_{symptomatic}(p_0)(t_{36})$: -0.563 (95% CI -14.34-0.653); $\Psi HZ_{asymptomatic}(p_0)(t_{36})$: -0.429 (95% CI-16.23-0.948)). These results, though not statistically significant, suggest that the effect of triple therapy initiated within the first 6 months of life on time to C diagnosis is stronger among symptomatic children than among asymptomatic children. The Cox proportional hazards parameter at 36 months comparing children treated asymptomatically within the first 6 months of life and children not treated asymptomatically within the first 6 months of life is HR=-0.712 (p value = 0.057). The Cox proportional hazards parameter at 36 months comparing children treated symptomatically within the first 6 months of life and children not treated symptomatically within the first 6 months of life is HR=-0.531 (p value = 0.459).

Initiated in First 12 months of Life-Asymptomatic Children

In estimating the time to C diagnosis among *asymptomatic* children, D/S/A selected term of pregnancy for the treatment mechanism (see Table 3.18). Children who were full-term had a significantly reduced risk of a C diagnosis when compared to children who were not born full-term (cOR=0.71; 95% CI 0.54-0.95). Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, and time indicator variables.

Initiated in First 12 months of Life-Symptomatic Children

Among *symptomatically* treated children, D/S/A selected an interaction between prenatal care and a time indicator variable for 25-30 months, and selected an interaction between sex and pregnancy term as covariates within our treatment mechanism. Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, and time indicator variables.

The G-computation survival curves for the causal effect of triple therapy initiated in the first 12 months of life on time to C diagnosis among *asymptomatic* children are illustrated in Figure 3.9a. Similarly, in Figure 3.9b are the survival curves for the estimates for therapy initiated in the first 12 months of life among *symptomatic* children.

The causal treatment specific parameters at time *k* related to triple therapy on C diagnosis had all *asymptomatic* or *symptomatic* children been treated or untreated within 12 months of birth are listed in Table 3.19. The estimates in the first column are the causal treatment specific parameters related to triple therapy on C diagnosis had all *asymptomatic* children started therapy within 12 months of birth. The second column illustrates the causal treatment specific parameters related to triple therapy on C diagnosis had all *asymptomatic* children been untreated within 12 months of birth. Similarly, the third and fourth columns are the causal treatment specific parameters related to triple therapy within 12 months of birth on C diagnosis among *symptomatic* children.

In Table 3.20, the marginal additive differences at time *k*, denoted by $\Psi AD(p_0)(t_k)$, are increasing over time for both treatment initiation scenarios. The marginal additive difference at 36 months for *asymptomatic* children who initiated therapy within 12 months of life is 0.030 (95% CI-0.250-0.225). These results suggest that there is little causal effect of triple therapy started in the first 12 months of life among *asymptomatic* children on time to C diagnosis. This evidence is further supported by the marginal log hazard ratio -0.106 (95% CI -1.054-0.739). When compared to the marginal additive difference among *symptomatic* children who initiated therapy within 12 months of life, 0.152 (95% CI -0.153-0.240), with the marginal additive difference among *asymptomatic* children, there is a noticeable difference. In fact, there is approximately a five-fold difference in the marginal additive differences, though these differences are not statistically significant. Similarly, the marginal log hazard ratio of triple therapy's effect on time to C diagnosis at 36 months among *symptomatically* treated children ($\Psi HZ_{symptomatic}(p_0)(t_{36})$: -0.587 (95% CI -1.217-0.480)), is approximately five times larger than the marginal log hazard ratio among *asymptomatically* treated children

($\Psi HZ_{asymptomatic}$ $(p_0)(t_{36})$: -0.106 (95% CI -1.054-0.739)). These results, though not statistically significant, suggest that the effect of triple therapy initiated within the first 12 months of life on time to C diagnosis is stronger among symptomatic children than among asymptomatic children. The Cox proportional hazards parameter at 36 months comparing children asymptomatically treated within the first 12 months of life and children not treated asymptomatically within the first 12 months of life is HR=-0.516 (p value = 0.102). The Cox proportional hazards parameter at 36 months comparing children treated symptomatically within the first 12 months of life and children not treated symptomatically within the first 12 months of life is HR=-0.008 (p value = 0.997).

Time to C Diagnosis or Death (36 Months of follow-up)

Initiated in First 6 months of Life-Asymptomatic Children
Among *asymptomatically* treated children, in estimating the time to C diagnosis or death, D/S/A selected term of pregnancy for the treatment mechanism (see Table 3.21). Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, and time indicator variables.

Initiated in First 6 months of Life-Symptomatic Children
In estimating the time to C diagnosis or death among *symptomatically* treated children, D/S/A selected no unforced terms for the treatment mechanism. Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, and time indicator variables.

The G-computation survival curves for the causal effect of triple therapy initiated in the first 6 months of life on time to C diagnosis or death among *asymptomatic* children are illustrated in Figure 3.10a. Similarly, in Figure 3.10b are the survival curves for the estimates for therapy initiated in the first 6 months of life among *symptomatic* children.

The causal treatment specific parameters at time *k* related to triple therapy on C diagnosis or death had all *asymptomatic* or *symptomatic* children been treated or untreated within 6 months of birth are listed in Table 3.22. The estimates in the first column are the causal treatment specific parameters related to triple therapy on C diagnosis or death had all *asymptomatic* children started therapy within 6 months of birth. The second column illustrates the causal treatment specific parameters related to triple therapy on C diagnosis or death had all *asymptomatic* children been untreated within 6 months of birth. Similarly, the third and fourth columns are the causal treatment specific parameters related to triple therapy within 6 months of birth on C diagnosis or death among *symptomatic* children.

In Table 3.23, the marginal additive differences at time *k*, denoted by $\Psi AD(p_0)(t_k)$, are increasing over time for both treatment initiation scenarios. The marginal additive difference at 36 months for *asymptomatic* children who initiated therapy within 6 months of life is 0.140 (95% CI -0.380-0.490). The marginal additive difference for *symptomatic* children who initiated therapy within 6 months, 0.062 (95% CI -0.236-0.403), is

approximately half the difference among *asymptomatic* children. The marginal log hazard ratio estimating the causal effect of triple therapy on time to C diagnosis or death among *asymptomatically* treated children, -0.514 (95% CI -17.72-1.086), is more than twice the marginal log hazard ratio among *symptomatically* treated children (-0.204 (95% CI-1.475-0.683)). These results, though not statistically significant, suggest that the effect of triple therapy initiated within the first 6 months of life on time to C diagnosis or death is stronger among *asymptomatic* children than among *symptomatic* children. The Cox proportional hazards parameter at 36 months comparing children treated asymptomatically within the first 6 months of life and children not treated asymptomatically within the first 6 months of life is HR=-0.570 (p value = 0.091). The Cox proportional hazards parameter at 36 months comparing children treated symptomatically within the first 6 months of life and children not treated symptomatically within the first 6 months of life is HR=-0.291 (p value = 0.388).

Initiated in First 12 months of Life-Asymptomatic Children

Among *asymptomatically* treated children, in estimating the time to C diagnosis or death, D/S/A selected the pregnancy term as a covariate within our treatment mechanism. The unadjusted binary estimates of W on time to death are listed in Table 3.24.

Initiated in First 12 months of Life-Symptomatic Children

Among *symptomatically* treated children, D/S/A selected an interaction between prenatal care and a time indicator variable for 25-30 months, and selected an interaction between sex and pregnancy term as covariates within our treatment mechanism. Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, and time indicator variables.

The G-computation survival curves for the causal effect of triple therapy initiated in the first 12 months of life on time to C diagnosis or death among *asymptomatic* children are illustrated in Figure 3.11a. Similarly, in Figure 3.11b are the survival curves for the estimates for therapy initiated in the first 12 months of life among *symptomatic* children.

The causal treatment specific parameters at time *k* related to triple therapy on C diagnosis or death had all *asymptomatic* or *symptomatic* children been treated or untreated within 12 months of birth are listed in Table 3.25. The estimates in the first column are the causal treatment specific parameters related to triple therapy on C diagnosis or death had all *asymptomatic* children started therapy within 12 months of birth. The second column illustrates the causal treatment specific parameters related to triple therapy on C diagnosis or death had all *asymptomatic* children been untreated within 12 months of birth. Similarly, the third and fourth columns are the causal treatment specific parameters related to triple therapy within 12 months of birth on C diagnosis or death among *symptomatic* children.

In Table 3.26, the marginal additive differences at time *k*, denoted by $\Psi AD(p_0)(t_k)$, are increasing over time for both treatment initiation scenarios. The marginal additive difference at 36 months for *asymptomatic* children who initiated therapy within 12 months of life is 0.061 (95% CI -0.208-0.267). The marginal additive difference for

symptomatic children who initiated therapy within 12 months, 0.052 (95% CI -0.255-0.319), is only slightly less than the marginal additive difference among *asymptomatic* children. The marginal log hazard ratio estimating the causal effect of triple therapy on time to C diagnosis or death among *asymptomatically* treated children, -0.205 (95% CI -1.205-0.601), is greater than the marginal log hazard ratio among *symptomatically* treated children -0.168 (95% CI -1.526-0.709)). Though not statistically significant, the marginal log hazard ratios suggest that the effect of triple therapy initiated within 12 months on time to C diagnosis or death is stronger among *asymptomatic* children than among *symptomatic* children. The Cox proportional hazards parameter at 36 months comparing children treated asymptomatically within the first 12 months of life and children not treated asymptomatically within the first 12 months of life is HR=-0.204 (p value = 0.455). The Cox proportional hazards parameter at 36 months comparing children treated symptomatically within the first 12 months of life and children not treated symptomatically within the first 12 months of life is HR=-0.154 (p value = 0.571).

Time to Death (60 Months of follow-up)

Initiated in First 6 months of Life-Asymptomatic Children

In estimating the time to death among *asymptomatic* children treated within 6 months of birth, D/S/A selected the square of the term of pregnancy and the square of prenatal care as covariates within our treatment mechanism. The unadjusted binary estimates of W on time to death are listed in Table 3.27.

Initiated in First 6 months of Life-Symptomatic Children

Among *symptomatically* treated children, in estimating the time to death D/S/A selected pregnancy term as a covariate within our treatment mechanism. The unadjusted binary estimates of W on time to death are listed in Table 3.28.

The G-computation survival curves for the causal effect of triple therapy initiated in the first 6 months of life on time to death among *asymptomatic* children are illustrated in Figure 3.12a. Similarly, in Figure 3.12b are the survival curves for the estimates for therapy initiated in the first 6 months of life among *symptomatic* children.

The causal treatment specific parameters at time *k* related to triple therapy on death had all *asymptomatic* or *symptomatic* children been treated or untreated within 6 months of birth are listed in Table 3.29. The estimates in the first column are the causal treatment specific parameters related to triple therapy on death had all *asymptomatic* children started therapy within 6 months of birth. The second column illustrates the causal treatment specific parameters related to triple therapy on death had all *asymptomatic* children been untreated within 6 months of birth. Similarly, the third and fourth columns are the causal treatment specific parameters related to triple therapy within 6 months of birth on death among *symptomatic* children.

In Table 3.30, the marginal additive differences at time *k*, denoted by $\Psi AD(p_0)(t_k)$, are increasing over time for both treatment initiation scenarios. The marginal additive difference at 60 months for *asymptomatic* children who initiated therapy within 6 months

of life is 0.032 (95% CI -0.348-0.129). The marginal additive difference for *symptomatic* children who initiated therapy within 6 months, 0.057(95% CI -0.313-0.237), is nearly twice the difference among *asymptomatic* children. The marginal log hazard ratio estimating the causal effect of triple therapy on time to death among *asymptomatically* treated children, -0.160 (95% CI -0.774-1.008) is approximately half the marginal log hazard ratio among *symptomatically* treated children -0.301 (95% CI -15.297-0.820). Though these results are not statistically significant, they suggest that initiating treatment within 6 months of birth among *symptomatic* children has a greater effect on time to death than among *asymptomatic* children treated within 6 months of birth. The Cox proportional hazards parameter at 36 months comparing children treated asymptomatically within the first 6 months of life and children not treated asymptomatically within the first 6 months of life is HR=-0.745 (p value = 0.084). The Cox proportional hazards parameter at 36 months comparing children treated symptomatically within the first 6 months of life and children not treated symptomatically within the first 6 months of life is HR=-0.299 (p value = 0.459).

Initiated in First 12 months of Life-Asymptomatic Children

In estimating the time to death among *asymptomatic* children treated within the first 12 months of life, D/S/A selected the square terms length of pregnancy term and prenatal care as covariates within our treatment mechanism. The unadjusted binary estimates of W on time to death are listed in Table 3.31.

Initiated in First 12 months of Life-Symptomatic Children

Among *symptomatically* treated children, in estimating the time to death, D/S/A selected pregnancy term as a covariate within our treatment mechanism. The unadjusted binary estimates of W on time to death are listed in Table 3.32.

The G-computation survival curves for the causal effect of triple therapy initiated in the first 12 months of life on time to death among *asymptomatic* children are illustrated in Figure 3.13a. Similarly, in Figure 3.13b are the survival curves for the estimates for therapy initiated in the first 12 months of life among *symptomatic* children.

The causal treatment specific parameters at time *k* related to triple therapy on death had all *asymptomatic* or *symptomatic* children been treated or untreated within 12 months of birth are listed in Table 3.33. The estimates in the first column are the causal treatment specific parameters related to triple therapy on death had all *asymptomatic* children started therapy within 12 months of birth. The second column illustrates the causal treatment specific parameters related to triple therapy on death had all *asymptomatic* children been untreated within 12 months of birth.

In Table 3.34, the marginal additive differences at time *k*, denoted by $\Psi AD(p_0)(t_k)$, are increasing over time for both treatment initiation scenarios. The marginal additive difference at 60 months for *asymptomatic* children who initiated therapy within 12 months of life is 0.039 (95% CI -0.156-0.158). The marginal additive difference for *symptomatic* children who initiated therapy within 12 months, 0.075 (95% CI -0.084-0.205), is approximately twice the difference among *asymptomatic* children. The

marginal log hazard ratio estimating the causal effect of triple therapy on time to death among *asymptomatically* treated children, -0.165 (95% CI -15.297-0.621), is approximately half the marginal log hazard ratio among *symptomatically* treated children (-0.336 (95% CI -1.423-0.305)). Again, these results are not statistically significant. However, they suggest that initiating treatment within 12 months of birth among *symptomatic* children has a greater effect on time to death than among *asymptomatic* children treated within 12 months of birth. The Cox proportional hazards parameter at 36 months comparing children treated asymptomatically within the first 12 months of life and children not treated asymptomatically within the first 12 months of life is HR=-0.591 (p value = 0.103). The Cox proportional hazards parameter at 36 months comparing children treated symptomatically within the first 12 months of life and children not treated symptomatically within the first 12 months of life is HR=-0.098 (p value = 0.764).

Table 3.35 allows one to compare the MSM results as estimated by g-computation to results one would have found using traditional techniques. Specifically, the mean marginal log hazard ratio over tk is the standard analogue to the Cox proportional hazards parameter. If the Cox PH model is correct, its estimates should be similar to the estimates from the mean marginal log HR over tk. In most cases these estimates were similar, however there are a few estimates that are different. Of particular note, the standard Cox PH model estimated the effect of triple therapy initiated in the first 12 months on C diagnosis or death as -0.431(95% CI: -0.875 - 0.012) while the mean marginal log hazard ratio over tk was estimated as -0.179 (95% CI: -0.625 - 0.468). Similarly, the standard Cox PH model estimated the effect of triple therapy initiated in the first 12 months among asymptomatic children on C diagnosis as -0.516(95% CI: -1.134 - 0.102) while the mean marginal log hazard ratio over tk was estimated as -0.106 (95% CI: -1.138 – 2.105). Further, the standard Cox PH model estimated the effect of triple therapy initiated in the first 12 months among symptomatic children on C diagnosis as -0.068 (95% CI: -0.610 - 0.474) while the mean marginal log hazard ratio over tk was estimated as -0.594 (95% CI: -15.65 – 1.529).

3.5 Discussion

The optimal timing of initiation of HAART among HIV-infected children is an on-going debate and recommendations for treatment initiation vary.(56) Guidelines in the United States and in Europe in previous years were based on 2-5 year risk of disease progression estimates calculated from observational studies.(56) In contrast, more recent guidelines (2003) have been based on estimates of the 12-month risk of disease progression as reported by the HIV Pediatric Prognostic Markers Collaborative Study (HPPMCS) Group, a collection of studies conducted in the developed world or in high-resource settings.(57) The WHO, in 2006, developed clinical and immunologic guidelines for treatment initiation in asymptomatic children in resource-limited settings based on HPPMCS data.(4) In 2008, WHO amended their recommendations for treatment initiation for HIV-positive children as a result from an RCT in South Africa.(7) Though not statistically significant, the results from the present analysis may be interpreted as supportive of the current WHO treatment guidelines.

In an attempt to use the data available to fit the most efficient model, the present study estimated the causal effects of triple therapy initiated at different thresholds on time to C diagnosis, time to C diagnosis or death, and death alone. The sampled children were more likely to be non-White females. The unbalanced distribution of gender may not be important as studies in high-resource settings have previously found no difference in survival or disease progression between genders.(57) Additionally, studies have found increasing trends in risk of death among children of non-White ethnicities, though not statistically significant.(57) In contrast, the risk of disease progression is seemingly reduced among children of non-White ethnicities, though these too are not significant.(57) In the present study, over 70% of the children were either moderately or severely immunologically impaired at the time of triple therapy treatment initiation. Moreover, nearly 90% of the children remained untreated with HAART through the first 12 months of life. This could be both a reflection of the treatment guidelines at the time or the availability of some drugs at the time of disease progression.

Traditional approaches in estimating the effects of treatments on time to event data, to include Cox proportional hazards models, often rely heavily on correct model specification. Using a data generating distribution approach avoids the inherent problems of employing traditional Cox methods, even if baseline covariates are included in the Cox model.(53) In turn, as opposed to using parameters selected by stepwise inclusion techniques or similar approaches, using an approach that uses parameters that are naturally selected based on the data allows for easier interpretation of the model and its parameters. For example, using a more traditional approach, the data suggest that ethnicity is a baseline confounder because it is significantly associated with both treatment initiation and C diagnosis (and death). Furthermore, immune function at treatment initiation was significantly associated with treatment and our events of interest. However, D/S/A did not select either of these variables for inclusion in our analysis, therefore preventing any unnecessary loss in precision in my estimates.

The causal effect of triple therapy among all children in delaying the time to a C diagnosis and/or death, regardless of immune status at treatment initiation, appears to be stronger among children who initiated therapy within 6 months rather than within 12 months of birth. Though no study has explored optimal treatment initiation in a pediatric HIV population using causal inference methods, this study's results seem to be in tune with previous, traditional analyses in early therapy initiation. Chiappini et al found that children who were treated with HAART early, as defined by: treatment initiation within 6 months of birth; category N, A, or B disease before treatment initiation; in immunologic category CDC 1 or 2 before treatment initiation, had significantly lower risk of progression to category C disease than not-early treated children (p value < 0.0001).(39) Similarly, Newell et al found that HIV positive children who started ART before 5 months of age were significantly more likely to have an improved immunologic response (time to a 20% increase in CD4 z score), after adjusting for immunocompetence status at treatment initiation.(38) It should be noted, however, that the authors were unable to find any added benefit in early treatment on sustained CD4 cell count after 6 months. In contrast, some laboratory research suggests that the positive immune

response in children who have been treated early versus children who have been treated later in life may only be an artifact of younger age and not truly associated with early treatment.(57)

An RCT in South Africa recently concluded that children who were assigned to early treatment initiation, regardless of symptoms, had a significantly lower mortality risk and risk of progression to category C disease when compared to children whose treatments were deferred until they became symptomatic.(35) The results from this RCT were the catalyst in the WHO's revised treatment guidelines. Similarly, the results, though not statistically significant, from the present analysis suggest that the effect of triple therapy initiated within the first 12 months of life on time to death is stronger among *asymptomatic* children (12 Months: $\Psi HZ_{asymptomatic}$ $(p_0)(t_{36})$: -0.336 (95% CI -1.423-0.305)) than among *symptomatic* children (12 Months: $\Psi HZ_{symptomatic}$ $(p_0)(t_{36})$: -0.165 (95% CI -15.297-0.621)). This suggests that the mortality risk is, in fact, potentially reduced among HIV positive children who initiate HAART early rather than deferring treatment until symptoms arise. In contrast, the present study found that the effect of triple therapy initiated within the first 6 or 12 months of life on time to C diagnosis is greater among *symptomatic* children (12 Months: $\Psi HZ_{symptomatic}$ $(p_0)(t_{36})$: -0.587 (95% CI -1.217-0.480)) than among *asymptomatic* children (12 Months: $\Psi HZ_{asymptomatic}$ $(p_0)(t_{36})$: -0.106 (95% CI -1.054-0.739)). This suggests that the gained benefit in initiating HAART by reducing the risk of disease progression is perhaps more fully realized among children who are already severely immune-compromised.

In some cases, the g-comp estimates are quite different than estimates one would have found using traditional approaches. Using a traditional approach to covariate selection, one may have chosen race/ethnicity and length of pregnancy because of their significant effect on both treatment and outcomes of interest. Using a superlearner to help select covariates based on a cross-validation and likelihood framework has been shown to minimize the empirical risk of loss function over a subspace.(55) Even if the Cox model were correctly identified, the estimates are only correct within the context of that particular model. In the present study, I also used a standard Cox proportional hazards model to estimate the effect of different treatment scenarios on a C diagnosis. Despite adjusting for the same covariates DSA selected for g-comp, the standard Cox model's estimates were sometimes different. For example, the standard Cox PH model estimated the effect of triple therapy initiated in the first 12 months among symptomatic children on C diagnosis as -0.068 (95% CI: -0.610 - 0.474) while the mean marginal log hazard ratio over tk was estimated as -0.594 (95% CI: -15.65 – 1.529). This discrepancy may imply that the Cox model may simply be incorrect.

Despite the added efficiency in parameters as a result of using a causal inference approach, the present study has several limitations. Death was somewhat a rare outcome and the data do not provide enough power to show an effect of triple therapy. Though an attempt to increase the number of events seen during the follow up period was made by expanding the follow up to 60 months, there was no added significance in the findings. There were no data available regarding income, so the impact of socio-economic status could not be measured.

Though the data do not provide the opportunity to calculate the impact of socio-economic status, it is possible the children who were born to poor mothers were less likely to have access to the same health care as other children. Evaluating the use of prenatal care within the sample allows one to speculate about the proportion of children who had access to care. Furthermore, the proposed DAG showing the impacts of variables on triple therapy initiation and C diagnosis and/or death suggests that SES affects access to prenatal care and maternal ARVs at delivery and during pregnancy. Though I was not able to control for SES, I did have data on the child's race. Nearly three-quarters of this study sample was non-White ethnicity. Late initiation of HAART and low drug adherence have been seen more among non-White populations in the U.S. than among White populations, which in turn results in more advanced disease and increases in mortality risk.(58)

Though an attempt to isolate children who were likely infected in utero or at birth was made in the present analysis, there is still a risk that some children included in the study were infected postnatally via breastfeeding. Data on breastfeeding practices were not collected and it was assumed to have not occurred if women received prenatal care. However, even if low SES women were well-informed of the risk of transmission from breastfeeding, they may have felt they had little choice but to breastfeed for economic reasons. In Sub-Saharan Africa, the risk of MTCT is well understood by many women who consciously decide to breastfeeding despite the transmission risks because of economic hardships and fears of increased risk of illnesses prevented by breast milk's nutritional benefits.

These data were extracted from the children's medical records, which may not offer a complete picture of maternal ARV exposure. Unfortunately, it was difficult to accurately estimate the proportion of mothers who were exposed to ARV during pregnancy or at delivery. Moreover, maternal HIV disease severity was not assessed. The severity of disease in mothers has been linked not only to the transmission of HIV, but also the severity in disease in vertically infected children.(15)

Viral load was not a variable analyzed in the present study as most of the children did not have these data collected. While CD4% is the primary source of clinical treatment guidance, viral load would likely also be an acceptable proxy for immuno-competence. Though immune function (as measured by clinical diagnoses and CD4 count or percent) at treatment initiation was adjusted, there could be residual confounding from a child's overall immune function that is unaccounted for. Thus, it is possible that these results could be indicative of the unmeasured severity of immune suppression at treatment initiation rather than a true causal effect of triple therapy initiated at 6 or 12 months.

Losses to follow-up were a formidable threat in this study. Because the population was dynamic, children could have been any age at their first HIV clinic visit. I established t0, first day of study entry, as the birth date, assuming all children were infected in utero or at delivery. Unfortunately, some children may have first visited the clinic after the first 6 or 12 months, which would include them as untreated children in the study. At a

minimum n=68 children were considered untreated in this analysis because their first study data were collected after the first 6 months. Furthermore, n=44 children were untreated in the analyses with *A* defined by treatment starting in the first 12 months because their first study data were collected after the first 12 months. These children could have moved into the catchment area later and may have already been receiving triple therapy. Additionally, it is possible that the healthiest children stop going to clinic after a period of time. Similarly, it is possible that sicker children either begin to go to the clinic after symptoms arise or cease going to the clinic because they were too ill. A sensitivity analysis on the data showed that the children who were censored before the end of the follow up period in the C diagnosis analyses were slightly more likely to be severely immuno-compromised at treatment initiation than children who were uncensored (44% vs 38%). Similarly, the children who were censored before the end of the follow up period in the C diagnosis/death analyses were more likely to be severely immuno-compromised at treatment initiation than children who were uncensored (45% vs 38%). Not surprisingly, this suggests that the children who were censored, which included death, reached a C diagnosis, or lost to follow up, were sicker than the children who survived beyond the follow-up time. Children who were treated within 6 months were slightly more likely to be censored than children who were untreated within 6 months of birth (33% vs 28%). This was likely due to the fact the children who started treatment early were sicker than children who delayed treatment. Children who were treated within 12 months were slightly less likely to be censored than children who were untreated within 12 months of birth (24% vs 31%). This was likely due to the possibility that the children who started treatment early were able to regain their health quicker than children who delayed treatment.

A sample size calculation was performed to see what the sample size would have to be in order to see a statistically significant finding. Conditional on the proportion of children treated within the first 6 months remaining the same (approximately 10%), then with a power of 0.80 in order to see hazard ratio of 0.63 comparing all treated and untreated children estimating the time to a C diagnosis or death I would have needed to enroll approximately 405 untreated children with 41 treated children. See Appendix A.5 for details.

Time-dependent confounding, (e.g. immune status throughout the study in the present example), was not controlled for. A child's baseline immune status was adjusted in the analyses estimating the effect of triple therapy among all children, regardless of symptoms. A child who started monotherapy may have good immunocompentence at treatment initiation, but later his immune status may deteriorate thus influencing the treatment and clinical outcome of interest. In this study, n=165 children (76.0%) at some point in life after the first 6 months of life initiated triple therapy. Furthermore, n=92 children (42.4%) modified their treatment to triple therapy after the first 12 months of life.

Though the approach applied in this analysis was an intent-to-treat analysis, an analysis that adjusts for treatment modification would likely lead to a stronger effect of triple therapy on clinical outcomes. By not adjusting for treatment modification, the present

results are biased toward the null, suggesting that the effect of triple therapy on C diagnosis and/or death is likely stronger than the G-computation estimates.

Adherence and resistance are always a threat in studies of ARV effectiveness. The present study was not able to assess the rates of adherence among the HIV positive children; this is likely heavily dependent on the mother's own ARV adherence. As a result, it is possible that some children who were started on triple therapy treatment early did not continuously receive the therapy, which in turn could create a drug resistance. If this child later restarted triple therapy, he may have poorer clinical and immunologic outcomes than other children who were continuously treated.

Tables

Table 3.1. Patient Demographics and Baseline Characteristics (N=217)

Baseline Covariate (W)	N (%)
Male Sex	95 (43.8)
White Ethnicity	61 (28.1)
Mother had Prenatal Care	110 (50.7)
Not Low Birth weight	153 (70.5)
Full-Term Pregnancy	123 (56.7)
Immune Status at Treatment Initiation	
Untreated	14 (6.5)
No or Mild Impairment	46 (21.2)
Moderate Impairment	70 (32.3)
Severe Impairment	87 (40.1)
HAART Initiation	
First 6 Months of Life	18 (8.3)
First 12 Months of Life	98 (45.2)
After First 12 Months of Life or Never	119 (54.8)

Table 3.2: Sample Baseline Characteristics and Associations With Triple Therapy Initiation In First 6 Months (N=217)

Baseline Covariate (W)	Triple Therapy In First 6 Months (cOR)[1]	P Value
Male Sex	0.62	0.35
White Ethnicity	0.49	0.27
Mother had Prenatal Care	1.47	0.20
Not Low Birth weight	0.63	0.36
Full-Term Pregnancy	0.84	0.57
Immune Status at Treatment Initiation	1.61	0.12

1. cOR=crude odds ratio

Table 3.3: Sample Baseline Characteristics and Associations With Triple Therapy Initiation in First 12 Months (N=217)

	Triple Therapy In First 12 Months (cOR) [1]	P Value
Male Sex	0.87	0.73
White Ethnicity	0.28	0.05
Mother had Prenatal Care	1.41	0.17
Not Low Birth weight	0.81	0.64
Full-Term Pregnancy	0.86	0.56
Immune Status at Treatment Initiation	1.82	0.03

1. cOR=crude odds ratio

Table 3.4: Sample Characteristics and Associations With C Diagnosis With 36 Months

	C Diagnosis Within First 36 Months (N=75) (cOR) [1]	P Value
Male Sex	0.86	0.60
White Ethnicity	2.39	< 0.01
Mother had Prenatal Care	1.00	0.98
Not Low Birth weight	1.23	0.51
Full-Term Pregnancy	0.87	0.43
Immune Status at Treatment Initiation	2.17	< 0.01

1. cOR=crude odds ratio

Table 3.5: Sample Characteristics and Associations With C Diagnosis or Death

	C Diagnosis/Death Within First 36 Months (N=84) (cOR) [1]	P Value
Male Sex	0.87	0.62
White Ethnicity	2.01	0.02
Mother had Prenatal Care	0.96	0.82
Not Low Birth weight	0.98	0.95
Full-Term Pregnancy	0.83	0.31
Immune Status at Treatment Initiation	1.90	< 0.01

1. cOR=crude odds ratio

Table 3.6: Sample Characteristics and Associations With Death

	Death Within First 60 Months (N=58) (cOR) [1]	P Value
Male Sex	0.79	0.46
White Ethnicity	1.69	0.11
Mother had Prenatal Care	0.95	0.77
Not Low Birth weight	1.01	0.97
Full-Term Pregnancy	0.83	0.34
Immune Status at Treatment Initiation	1.10	0.56

1. cOR=crude odds ratio

Table 3.7: Sample Characteristics and Associations With Triple Therapy Initiation in First 6 Months By Symptomatic Status

	Triple Therapy In First 6 Months (N=10) (cOR) [1]		Triple Therapy In First 6 Months (N=8) (cOR) [1]	
	Symptomatic	P Value	Asymptomatic	P Value
Male Sex	0.54	0.38	0.76	0.72
White Ethnicity	1.10	0.89	--	n/a
Mother had Prenatal Care	1.59	0.26	1.30	0.54
Not Low Birth weight	0.61	0.46	0.69	0.61
Full-Term Pregnancy	0.65	0.26	1.25	0.65

1. cOR=crude odds ratio

Table 3.8: Sample Characteristics and Associations With Triple Therapy Initiation in First 12 Months By Symptomatic Status

	Triple Therapy In First 12 Months (N=48) (cOR) [1]		Triple Therapy In First 12 Months (N=50) (cOR) [1]	
	Symptomatic	P Value	Asymptomatic	P Value
Male Sex	0.85	0.76	0.91	0.88
White Ethnicity	0.62	0.47	--	n/a
Mother had Prenatal Care	1.36	0.34	1.40	0.36
Not Low Birth weight	1.16	0.80	0.57	0.35
Full-Term Pregnancy	0.67	0.22	1.25	0.58

1. cOR=crude odds ratio

Table 3.9 Binary Estimates of Selected Baseline Covariates on C Diagnosis Among Children Who Initiated A in First 6 Months of Life*

Baseline Covariate (W)	Unadjusted Odds Ratio
Sex	0.52 (0.31 – 0.84)
Term of Pregnancy	0.76 (0.58 – 1.01)
Race	1.19 (0.72 – 1.92)

*Covariates selected by D/S/A. Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birthweight, immune status at treatment initiation, and time indicator variables.

Table 3.10: Causal Treatment Specific Parameters Related to Triple Therapy on C Diagnosis: Therapy Initiation Under 6 Months or Under 12 Months (36 Months of Follow-up)

	Triple Therapy Initiated First 6 Months		**Triple Therapy Initiated First 12 Months**	
Time Interval	(all treated) $\Psi 1(p_0)(tk)$	(all untreated) $\Psi 0(p_0)(tk)$	(all treated) $\Psi 1(p_0)(tk)$	(all untreated) $\Psi 0(p_0)(tk)$
6 Months	0.92 (0.81-0.97)	0.90 (0.85-0.92)	0.91 (0.66-0.96)	0.88 (0.83-0.94)
12 Months	0.85 (0.66-0.94)	0.81 (0.73-0.85)	0.83 (0.58-0.93)	0.78 (0.70-0.84)
18 Months	0.79 (0.54-0.91)	0.74 (0.63-0.80)	0.76 (0.54-0.90)	0.69 (0.61-0.80)
24 Months	0.73(0.51-0.88)	0.67 (0.56-0.74)	0.73(0.52-0.87)	0.65 (0.56-0.75)
30 Months	0.72 (0.50-0.86)	0.66 (0.56-0.73)	0.73 (0.50-0.85)	0.64 (0.56-0.71)
36 Months	0.67 (0.41-0.84)	0.60 (0.55-0.69)	0.71 (0.49-0.84)	0.62 (0.53-0.67)

Table 3.11: Causal Effects of Triple Therapy on C Diagnosis: Therapy Initiation Under 6 Months or Under 12 Months (36 Months of Follow-up)

	Triple Therapy Initiated First 6 Months		**Triple Therapy Initiated First 12 Months**	
Time Interval	Marginal Additive Difference $\Psi AD(p_0)(tk)$	Marginal log Hazard Ratio $\Psi HZ(p_0)(tk)$	Marginal Additive Difference $\Psi AD(p_0)(tk)$	Marginal log Hazard Ratio $\Psi HZ(p_0)(tk)$
6 Months	0.043 (-0.079-0.101)	-0.475 (-1.231-0.616)	0.032 (-0.252-0.111)	-0.328 (-1.174-1.646)
12 Months	0.077 (-0.132-0.180)	-0.472 (-1.222-0.591)	0.057 (-0.260-0.195)	-0.325 (-1.162-1.136)
18 Months	0.103 (-0.166-0.241)	-0.468 (-1.213-0.574)	0.075 (-0.229-0.223)	-0.323 (-1.158-0.854)
24 Months	0.113 (-0.175-0.264)	-0.467 (-1.204-0.572)	0.082 (-0.187-0.251)	-0.322 (-1.153-0.632)
30 Months	0.115 (-0.176-0.270)	-0.467 (-1.203-0.571)	0.083 (-0.143-0.258)	-0.322 (-1.152-0.452)
36 Months	0.120 (-0.192-0.276)	-0.466 (-1.202-0.565)	0.087 (-0.099-0.264)	-0.321 (-1.151-0.300)

Table 3.12: Causal Treatment Specific Parameters Related to Triple Therapy on C Diagnosis or Death: Therapy Initiation Under 6 Months or Under 12 Months (36 Months of Follow-up)

	Triple Therapy Initiated First 6 Months		**Triple Therapy Initiated First 12 Months**	
Time Interval	(all treated) $\Psi 1(p_0)(tk)$	(all untreated) $\Psi 0(p_0)(tk)$	(all treated) $\Psi 1(p_0)(tk)$	(all untreated) $\Psi 0(p_0)(tk)$
6 Months	0.93 (0.80-0.95)	0.90 (0.81-0.93)	0.91 (0.70-0.95)	0.90 (0.81-0.93)
12 Months	0.86 (0.64-0.91)	0.80 (0.65-0.86)	0.83 (0.57-0.90)	0.80 (0.66-0.87)
18 Months	0.79 (0.52-0.86)	0.72 (0.60-0.81)	0.76 (0.51-0.86)	0.72 (0.59-0.81)
24 Months	0.74(0.48-0.83)	0.64 (0.54-0.76)	0.69 (0.46-0.81)	0.64 (0.53-0.76)
30 Months	0.73 (0.48-0.83)	0.64 (0.54-0.72)	0.69 (0.43-0.81)	0.64 (0.53-0.72)
36 Months	0.68 (0.44-0.81)	0.57 (0.54-0.68)	0.63 (0.40-0.77)	0.57 (0.49-0.68)

Table 3.13: Causal Effects of Triple Therapy on C Diagnosis or Death: Therapy Initiation Under 6 Months or Under 12 Months (36 Months of Follow-up)

	Triple Therapy Initiated First 6 Months		**Triple Therapy Initiated First 12 Months**	
Time Interval	Marginal Additive Difference $\Psi AD(p_0)(tk)$	Marginal log Hazard Ratio $\Psi HZ(p_0)(tk)$	Marginal Additive Difference $\Psi AD(p_0)(tk)$	Marginal log Hazard Ratio $\Psi HZ(p_0)(tk)$
6 Months	0.035 (-0.088-0.114)	-0.369 (-1.072-0.843)	0.016 (-0.205-0.080)	-0.179 (-0.803-1.296)
12 Months	0.057 (-0.143-0.197)	-0.369 (-1.064-0.807)	0.030 (-0.250-0.137)	-0.179 (-0.799-1.054)
18 Months	0.077 (-0.179-0.230)	-0.369 (-1.061-0.780)	0.040 (-0.242-0.158)	-0.179 (-0.795-0.853)
24 Months	0.094 (-0.203-0.244)	-0.369 (-1.060-0.759)	0.048 (-0.222-0.192)	-0.179 (-0.790-0.823)
30 Months	0.095 (-0.220-0.244)	-0.369 (-1.060-0.743)	0.049 (-0.240-0.195)	-0.180 (-0.790-0.806)
36 Months	0.108 (-0.231-0.261)	-0.369 (-1.058-0.730)	0.055 (-0.252-0.208)	-0.180 (-0.786-0.792)

Table 3.14: Causal Treatment Specific Parameters Related to Triple Therapy on Death: Therapy Initiation Under 6 or 12 Months (60 Months of Follow-up)

	Triple Therapy Initiated First 6 Months		**Triple Therapy Initiated First 12 Months**	
Time Interval	(all treated) $\Psi_1(p_0)(t_k)$	(all untreated) $\Psi_0(p_0)(t_k)$	(all treated) $\Psi_1(p_0)(t_k)$	(all untreated) $\Psi_0(p_0)(t_k)$
6 Months	0.97 (0.94-1.00)	0.97 (0.95-0.98)	0.97 (0.95-0.99)	0.96 (0.96-0.97)
12 Months	0.94 (0.89-1.00)	0.93 (0.90-0.95)	0.94 (0.90-0.97)	0.93 (0.90-0.95)
18 Months	0.92 (0.85-1.00)	0.90 (0.85-0.93)	0.92 (0.85-0.96)	0.90 (0.87-0.92)
24 Months	0.89 (0.73-0.96)	0.87 (0.77-0.90)	0.89 (0.61-0.95)	0.87 (0.81-0.87)
30 Months	0.86 (0.65-0.96)	0.84 (0.74-0.87)	0.86 (0.68-0.94)	0.84 (0.79-0.84)
36 Months	0.84 (0.57-0.95)	0.81 (0.72-0.85)	0.84 (0.55-0.88)	0.81 (0.76-0.82)
42 Months	0.82 (0.57-0.94)	0.78 (0.70-0.82)	0.82 (0.53-0.86)	0.78 (0.72-0.79)
48 Months	0.79 (0.57-0.93)	0.75 (0.68-0.80)	0.79 (0.51-0.85)	0.75 (0.69-0.77)
54 Months	0.77 (0.57-0.93)	0.73 (0.66-0.78)	0.77 (0.48-0.83)	0.72 (0.66-0.75)
60 Months	0.75 (0.53-0.92)	0.70 (0.63-0.76)	0.75 (0.46-0.92)	0.70 (0.63-0.72)

Table 3.15: Causal Effects of Triple Therapy on Death: Therapy Initiation Under 6 Months or Under 12 Months (60 Months of Follow-up)

	Triple Therapy Initiated First 6 Months		**Triple Therapy Initiated First 12 Months**	
Time Interval	Marginal Additive Difference $\Psi AD(p_0)(tk)$	Marginal log Hazard Ratio $\Psi HZ(p_0)(tk)$	Marginal Additive Difference $\Psi AD(p_0)(tk)$	Marginal log Hazard Ratio $\Psi HZ(p_0)(tk)$
6 Months	0.006 (-0.036-0.053)	-0.199 (-16.11-0.938)	0.007 (-0.019-0.030)	-0.205 (-1.966-0.464)
12 Months	0.012 (-0.065-0.104)	-0.199 (-16.11-0.907)	0.012 (-0.036-0.059)	-0.205 (-1.206-0.458)
18 Months	0.017 (-0.088-0.151)	-0.199 (-16.14-0.887)	0.018 (-0.051-0.086)	-0.205 (-1.216-0.451)
24 Months	0.023 (-0.145-0.102)	-0.199 (-1.368-0.851)	0.023 (-0.265-0.112)	-0.205 (-1.225-1.300)
30 Months	0.027 (-0.159-0.120)	-0.199 (-1.373-0.847)	0.028 (-0.264-0.136)	-0.205 (-1.234-1.167)
36 Months	0.032 (-0.217-0.138)	-0.199 (-1.376-0.846)	0.033 (-0.262-0.087)	-0.205 (-0.529-1.067)
42 Months	0.036 (-0.197-0.154)	-0.199 (-1.377-0.835)	0.037 (-0.260-0.092)	-0.205 (-0.527-0.987)
48 Months	0.040 (-0.206-0.170)	-0.199 (-1.378-0.832)	0.041 (-0.257-0.098)	-0.205 (-0.526-0.922)
54 Months	0.043 (-0.214-0.185)	-0.199 (-1.378-0.831)	0.045 (-0.253-0.109)	-0.205 (-0.525-0.867)
60 Months	0.046 (-0.229-0.200)	-0.199 (-1.378-0.825)	0.048 (-0.249-0.130)	-0.205 (-0.526-0.821)

Table 3.16: Causal Treatment Specific Parameters Related to Triple Therapy on C Diagnosis: Therapy Initiation Under 6 Months Among Asymptomatic or Symptomatic Children (36 Months of Follow-up)

	Triple Therapy Among Asymptomatic		**Triple Therapy Among Symptomatic**	
Time Interval	(all treated) $\Psi_1(p_0)(t_k)$	(all untreated) $\Psi_0(p_0)(t_k)$	(all treated) $\Psi_1(p_0)(t_k)$	(all untreated) $\Psi_0(p_0)(t_k)$
6 Months	0.94 (0.69-1.00)	0.91 (0.86-0.92)	0.94 (0.85-1.00)	0.90 (0.81-0.90)
12 Months	0.88 (0.50-1.00)	0.82 (0.74-0.84)	0.89 (0.73-1.00)	0.82 (0.66-0.82)
18 Months	0.83 (0.37-1.00)	0.75 (0.66-0.77)	0.84 (0.63-1.00)	0.74 (0.59-0.76)
24 Months	0.78 (0.34-1.00)	0.68 (0.62-0.71)	0.80(0.60-1.00)	0.67 (0.54-0.72)
30 Months	0.77 (0.33-1.00)	0.67 (0.62-0.71)	0.79 (0.59-1.00)	0.67 (0.54-0.71)
36 Months	0.72 (0.30-1.00)	0.61 (0.55-0.69)	0.75 (0.56-1.00)	0.61 (0.53-0.68)

Table 3.17: Causal Effects of Triple Therapy on C Diagnosis: Therapy Initiation Under 6 Months Among Asymptomatic or Symptomatic Children (36 Months of Follow-up)

	Triple Therapy Among Asymptomatic		**Triple Therapy Among Symptomatic**	
Time Interval	Marginal Additive Difference $\Psi_{AD}(p_0)(t_k)$	Marginal log Hazard Ratio $\Psi_{HZ}(p_0)(t_k)$	Marginal Additive Difference $\Psi_{AD}(p_0)(t_k)$	Marginal log Hazard Ratio $\Psi_{HZ}(p_0)(t_k)$
6 Months	0.031 (-0.184-0.115)	-0.429 (-16.26-1.022)	0.041 (-0.023-0.170)	-0.575 (-15.346-0.184)
12 Months	0.058 (-0.271-0.216)	-0.429 (-16.25-0.986)	0.075 (-0.039-0.304)	-0.572 (-15.317-0.181)
18 Months	0.080 (-0.312-0.303)	-0.429 (-16.24-0.960)	0.103 (-0.050-0.351)	-0.569 (-15.313-0.179)
24 Months	0.099 (-0.320-0.380)	-0.429 (-16.23-0.955)	0.126 (-0.052-0.397)	-0.566 (-15.311-0.179)
30 Months	0.100 (-0.322-0.380)	-0.429 (-16.23-0.953)	0.128 (-0.052-0.406)	-0.566 (-15.313-0.178)
36 Months	0.115 (-0.327-0.446)	-0.429 (-16.23-0.948)	0.146 (-0.055-0.415)	-0.563 (-15.314-0.178)

Table 3.18 Binary Estimates of Selected Baseline Covariates on C Diagnosis Among Children Who Initiated A Asymptomatically in the First 12 Months of Life*

Baseline Covariate (W)	Unadjusted Odds Ratio
Term of Pregnancy	0.71 (0.54 – 0.95)

*Covariates selected by D/S/A. Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, and time indicator variables.

Table 3.19: Causal Treatment Specific Parameters Related to Triple Therapy on C Diagnosis: Therapy Initiation Under 12 Months Among Asymptomatic or Symptomatic Children (36 Months of Follow-up)

	Triple Therapy Among Asymptomatic		**Triple Therapy Among Symptomatic**	
Time Interval	(all treated) $\Psi_1(p_0)(t_k)$	(all untreated) $\Psi_0(p_0)(t_k)$	(all treated) $\Psi_1(p_0)(t_k)$	(all untreated) $\Psi_0(p_0)(t_k)$
6 Months	0.91 (0.73-0.97)	0.90 (0.86-0.94)	0.95 (0.76-1.00)	0.90 (0.85-0.93)
12 Months	0.82 (0.58-0.94)	0.80 (0.74-0.88)	0.89 (0.70-1.00)	0.82 (0.73-0.85)
18 Months	0.74 (0.45-0.92)	0.72 (0.66-0.82)	0.85 (0.67-0.92)	0.74 (0.64-0.79)
24 Months	0.68 (0.42-0.89)	0.65 (0.63-0.77)	0.80 (0.61-0.90)	0.67 (0.62-0.74)
30 Months	0.67 (0.41-0.88)	0.64 (0.63-0.72)	0.80 (0.54-0.90)	0.66 (0.62-0.73)
36 Months	0.65 (0.38-0.85)	0.62 (0.56-0.68)	0.75 (0.49-0.87)	0.60 (0.58-0.68)

Table 3.20: Causal Effects of Triple Therapy on C Diagnosis: Therapy Initiation Under 12 Months Among Asymptomatic or Symptomatic Children (36 Months Follow-up)

	Triple Therapy Among Asymptomatic		**Triple Therapy Among Symptomatic**	
Time Interval	Marginal Additive Difference $\Psi_{AD}(p_0)(t_k)$	Marginal log Hazard Ratio $\Psi_{HZ}(p_0)(t_k)$	Marginal Additive Difference $\Psi_{AD}(p_0)(t_k)$	Marginal log Hazard Ratio $\Psi_{HZ}(p_0)(t_k)$
6 Months	0.010 (-0.204-0.064)	-0.107 (-1.207-1.608)	0.043 (-0.142-0.137)	-0.600 (-15.645-0.987)
12 Months	0.018 (-0.199-0.116)	-0.106 (-1.101-1.094)	0.078 (-0.119-0.254)	-0.597 (-15.643-0.584)
18 Months	0.024 (-0.234-0.151)	-0.106 (-1.092-0.749)	0.108 (-0.113-0.208)	-0.594 (-1.220-0.508)
24 Months	0.029 (-0.242-0.188)	-0.106 (-1.075-0.745)	0.132 (-0.131-0.224)	-0.591 (-1.219-0.497)
30 Months	0.029 (-0.244-0.196)	-0.106 (-1.069-0.743)	0.133 (-0.144-0.234)	-0.590 (-1.218-0.488)
36 Months	0.030 (-0.250-0.225)	-0.106 (-1.054-0.739)	0.152 (-0.153-0.240)	-0.587 (-1.217-0.480)

Table 3.21 Binary Estimates of Selected Baseline Covariates on C Diagnosis or Death Among Children Who Initiated A Asymptomatically in First 6 Months of Life*

Baseline Covariate (W)	Unadjusted Odds Ratio
Term of Pregnancy	0.72 (0.56 – 0.95)

*Covariates selected by D/S/A. Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, and time indicator variables.

Table 3.22: Causal Treatment Specific Parameters Related to Triple Therapy on C Diagnosis or Death: Therapy Initiation Under 6 Months Among Asymptomatic or Symptomatic Children (36 Months of Follow-up)

	Triple Therapy Among Asymptomatic		**Triple Therapy Among Symptomatic**	
Time Interval	(all treated) $\Psi_1(p_0)(t_k)$	(all untreated) $\Psi_0(p_0)(t_k)$	(all treated) $\Psi_1(p_0)(t_k)$	(all untreated) $\Psi_0(p_0)(t_k)$
6 Months	0.92 (0.62-1.00)	0.87 (0.85-0.91)	0.92 (0.81-0.96)	0.90 (0.81-0.90)
12 Months	0.85 (0.40-1.00)	0.76 (0.72-0.81)	0.84 (0.67-0.92)	0.80 (0.65-0.80)
18 Months	0.78 (0.31-1.00)	0.67 (0.62-0.74)	0.77 (0.55-0.89)	0.72 (0.58-0.72)
24 Months	0.75 (0.27-1.00)	0.62 (0.53-0.68)	0.70(0.49-0.88)	0.65 (0.53-0.68)
30 Months	0.75 (0.26-1.00)	0.62 (0.53-0.67)	0.70 (0.49-0.88)	0.64 (0.53-0.67)
36 Months	0.73 (0.22-1.00)	0.59 (0.51-0.65)	0.64 (0.42-0.88)	0.58 (0.49-0.65)

Table 3.23: Causal Effects of Triple Therapy on C Diagnosis or Death: Therapy Initiation Under 6 Months Among Asymptomatic or Symptomatic Children (36 Months Follow-up)

	Triple Therapy Among Asymptomatic		**Triple Therapy Among Symptomatic**	
Time Interval	Marginal Additive Difference $\Psi_{AD}(p_0)(t_k)$	Marginal log Hazard Ratio $\Psi_{HZ}(p_0)(t_k)$	Marginal Additive Difference $\Psi_{AD}(p_0)(t_k)$	Marginal log Hazard Ratio $\Psi_{HZ}(p_0)(t_k)$
6 Months	0.049 (-0.236-0.145)	-0.518 (-17.75-1.131)	0.018 (-0.063-0.135)	-0.204 (-1.494-0.452)
12 Months	0.088 (-0.339-0.269)	-0.516 (-17.74-1.108)	0.033 (-0.105-0.238)	-0.204 (-1.488-0.447)
18 Months	0.118 (-0.366-0.376)	-0.515 (-17.74-1.098)	0.045 (-0.131-0.285)	-0.204 (-1.486-0.443)
24 Months	0.131 (-0.374-0.466)	-0.514 (-17.73-1.094)	0.054 (-0.195-0.305)	-0.204 (-1.485-0.633)
30 Months	0.133 (-0.376-0.466)	-0.514 (-17.73-1.093)	0.055 (-0.188-0.308)	-0.204 (-1.485-0.605)
36 Months	0.140 (-0.380-0.490)	-0.514 (-17.72-1.086)	0.062 (-0.178-0.319)	-0.204 (-1.485-0.556)

Table 3.24 Binary Estimates of Selected Baseline Covariates on C Diagnosis or Death Among Children Who Initiated A Asymptomatically in the First 12 Months of Life*

Baseline Covariate (W)	Unadjusted Odds Ratio
Term of Pregnancy	0.72 (0.56 – 0.95)

*Covariates selected by D/S/A. Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, and time indicator variables.

Table 3.25: Causal Treatment Specific Parameters Related to Triple Therapy on C Diagnosis or Death: Therapy Initiation Under 12 Months Among Asymptomatic or Symptomatic Children (36 Months of Follow-up)

	Triple Therapy Among Asymptomatic		**Triple Therapy Among Symptomatic**	
Time Interval	(all treated) $\Psi_1(p_0)(t_k)$	(all untreated) $\Psi_0(p_0)(t_k)$	(all treated) $\Psi_1(p_0)(t_k)$	(all untreated) $\Psi_0(p_0)(t_k)$
6 Months	0.90 (0.58-0.97)	0.87 (0.84-0.91)	0.91 (0.81-0.96)	0.89 (0.81-0.90)
12 Months	0.80 (0.51-0.93)	0.76 (0.72-0.83)	0.83 (0.67-0.92)	0.80 (0.65-0.80)
18 Months	0.72 (0.47-0.92)	0.67 (0.63-0.76)	0.75 (0.55-0.89)	0.72 (0.58-0.72)
24 Months	0.68 (0.44-0.89)	0.62 (0.58-0.69)	0.69(0.49-0.88)	0.64 (0.53-0.68)
30 Months	0.67 (0.43-0.88)	0.62 (0.58-0.68)	0.68 (0.49-0.88)	0.64 (0.53-0.67)
36 Months	0.65 (0.38-0.85)	0.59 (0.51-0.66)	0.62 (0.42-0.88)	0.57 (0.49-0.65)

Table 3.26: Causal Effects of Triple Therapy on C Diagnosis or Death: Therapy Initiation Under 12 Months Among Asymptomatic or Symptomatic Children (36 Months Follow-up)

	Triple Therapy Among Asymptomatic		**Triple Therapy Among Symptomatic**	
Time Interval	Marginal Additive Difference $\Psi_{AD}(p_0)(t_k)$	Marginal log Hazard Ratio $\Psi_{HZ}(p_0)(t_k)$	Marginal Additive Difference $\Psi_{AD}(p_0)(t_k)$	Marginal log Hazard Ratio $\Psi_{HZ}(p_0)(t_k)$
6 Months	0.022 (-0.303-0.080)	-0.207 (-1.263-1.495)	0.016 (-0.063-0.135)	-0.173 (-1.494-0.452)
12 Months	0.039 (-0.276-0.148)	-0.206 (-1.247-1.027)	0.029 (-0.105-0.238)	-0.172 (-1.488-0.447)
18 Months	0.052 (-0.200-0.176)	-0.205 (-1.238-0.659)	0.038 (-0.131-0.285)	-0.171 (-1.486-0.443)
24 Months	0.057 (-0.200-0.221)	-0.205 (-1.222-0.607)	0.046 (-0.195-0.305)	-0.170 (-1.485-0.633)
30 Months	0.058 (-0.202-0.233)	-0.205 (-1.218-0.606)	0.046 (-0.188-0.308)	-0.169 (-1.485-0.605)
36 Months	0.061 (-0.208-0.267)	-0.205 (-1.205-0.601)	0.052 (-0.178-0.319)	-0.168 (-1.485-0.556)

Table 3.27. Binary Estimates of Selected Baseline Covariates on Death Among Children Who Initiated A Asymptomatically In the First 6 Months of Life*

Baseline Covariate (W)	Unadjusted Odds Ratio
Term of Pregnancy[2]	0.74 (0.54-1.02)
Prenatal Care[2]	0.79 (0.59 – 1.05)

*Covariates selected by D/S/A. Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, and time indicator variables.

Table 3.28 Binary Estimates of Selected Baseline Covariates on Death Among Children Who Initiated A Symptomatically in the First 6 Months of Life*

Baseline Covariate (W)	Unadjusted Odds Ratio
Term of Pregnancy	0.74 (0.54 – 1.02)

*Covariates selected by D/S/A. Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, and time indicator variables.

Table 3.29: Causal Treatment Specific Parameters Related to Triple Therapy on Death: Therapy Initiation Under 6 Months Among Asymptomatic or Symptomatic Children (60 Months of Follow-up)

	Triple Therapy Among Asymptomatic		**Triple Therapy Among Symptomatic**	
Time Interval	(all treated) $\Psi_1(p_0)(t_k)$	(all untreated) $\Psi_0(p_0)(t_k)$	(all treated) $\Psi_1(p_0)(t_k)$	(all untreated) $\Psi_0(p_0)(t_k)$
6 Months	0.97 (0.91-1.00)	0.96 (0.96-0.97)	0.97 (0.95-1.00)	0.96 (0.96-0.98)
12 Months	0.94 (0.84-1.00)	0.93 (0.90-0.95)	0.95 (0.91-1.00)	0.93 (0.91-0.95)
18 Months	0.91 (0.78-1.00)	0.89 (0.87-0.91)	0.92 (0.87-1.00)	0.89 (0.88-0.92)
24 Months	0.86 (0.71-1.00)	0.84 (0.81-0.89)	0.88 (0.45-1.00)	0.84 (0.82-0.88)
30 Months	0.85 (0.68-1.00)	0.82 (0.79-0.86)	0.86 (0.45-1.00)	0.82 (0.80-0.86)
36 Months	0.82 (0.44-0.90)	0.79 (0.76-0.83)	0.84 (0.45-1.00)	0.79 (0.77-0.83)
42 Months	0.81 (0.44-0.89)	0.79 (0.74-0.81)	0.84 (0.45-1.00)	0.79 (0.75-0.81)
48 Months	0.79 (0.45-0.87)	0.76 (0.71-0.79)	0.81 (0.45-1.00)	0.76 (0.72-0.78)
54 Months	0.76 (0.45-0.86)	0.73 (0.68-0.76)	0.79 (0.45-1.00)	0.73 (0.69-0.76)
60 Months	0.74 (0.45-0.84)	0.70 (0.65-0.74)	0.77 (0.45-1.00)	0.70 (0.65-0.74)

Table 3.30: Causal Effects of Triple Therapy on Death: Therapy Initiation Under 6 Months Among Asymptomatic or Symptomatic Children (60 Months Follow-up)

	Triple Therapy Among Asymptomatic		**Triple Therapy Among Symptomatic**	
Time Interval	Marginal Additive Difference $\Psi_{AD}(p_0)(t_k)$	Marginal log Hazard Ratio $\Psi_{HZ}(p_0)(t_k)$	Marginal Additive Difference $\Psi_{AD}(p_0)(t_k)$	Marginal log Hazard Ratio $\Psi_{HZ}(p_0)(t_k)$
6 Months	0.005 (-0.056-0.039)	-0.162 (-15.37-1.084)	0.010 (-0.020-0.039)	-0.304 (-15.298-0.577)
12 Months	0.001 (-0.099-0.076)	-0.161 (-15.37-1.045)	0.019 (-0.039-0.074)	-0.303 (-15.298-0.577)
18 Months	0.015 (-0.132-0.127)	-0.161 (-15.37-1.005)	0.027 (-0.056-0.109)	-0.303 (-15.298-0.576)
24 Months	0.022 (-0.157-0.146)	-0.161 (-15.37-0.964)	0.040 (-0.415-0.143)	-0.302 (-15.296-1.522)
30 Months	0.024 (-0.175-0.179)	-0.160 (-15.37-0.921)	0.043 (-0.389-0.175)	-0.302 (-15.297-1.522)
36 Months	0.028 (-0.397-0.103)	-0.160 (-0.774-1.519)	0.050 (-0.335-0.207)	-0.302 (-15.297-1.200)
42 Months	0.029 (-0.372-0.116)	-0.160 (-0.774-1.365)	0.051 (-0.260-0.092)	-0.302 (-15.297-1.088)
48 Months	0.032 (-0.348-0.129)	-0.160 (-0.774-1.231)	0.057 (-0.313-0.237)	-0.302 (-15.297-0.989)
54 Months	0.034 (-0.325-0.141)	-0.160 (-0.774-1.114)	0.063 (-0.292-0.293)	-0.301 (-15.297-0.901)
60 Months	0.038 (-0.302-0.153)	-0.160 (-0.774-1.008)	0.068 (-0.253-0.320)	-0.301 (-15.297-0.820)

Table 3.31. Binary Estimates of Selected Baseline Covariates on Death Among Children Who Initiated A Asymptomatically in the First 12 Months of Life*

Baseline Covariate (W)	Unadjusted Odds Ratio
Term of Pregnancy[2]	0.74 (0.54-1.02)
Prenatal Care[2]	0.79 (0.59 – 1.05)

*Covariates selected by D/S/A. Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, and time indicator variables.

Table 3.32. Binary Estimates of Selected Baseline Covariates on Death Among Children Who Initiated A Symptomatically in the First 12 Months of Life*

Baseline Covariate (W)	Unadjusted Odds Ratio
Term of Pregnancy	0.74 (0.54 – 1.02)

*Covariates selected by D/S/A. Variables from which the treatment mechanism model was selected included the following: pregnancy term, sex, race, prenatal care, birth weight, and time indicator variables.

Table 3.33: Causal Treatment Specific Parameters Related to Triple Therapy on Death: Therapy Initiation Under 12 Months Among Asymptomatic or Symptomatic Children (60 Months of Follow-up)

	Triple Therapy Among Asymptomatic		**Triple Therapy Among Symptomatic**	
Time Interval	(all treated) $\Psi_1(p_0)(t_k)$	(all untreated) $\Psi_0(p_0)(t_k)$	(all treated) $\Psi_1(p_0)(t_k)$	(all untreated) $\Psi_0(p_0)(t_k)$
6 Months	0.97 (0.84-1.00)	0.96 (0.95-0.97)	0.97 (0.96-0.99)	0.96 (0.96-0.98)
12 Months	0.94 (0.80-1.00)	0.93 (0.91-0.95)	0.95 (0.92-0.98)	0.93 (0.91-0.94)
18 Months	0.91 (0.77-1.00)	0.89 (0.87-0.92)	0.93 (0.87-0.97)	0.90 (0.87-0.92)
24 Months	0.86 (0.74-1.00)	0.84 (0.77-0.90)	0.88 (0.71-0.96)	0.84 (0.80-0.88)
30 Months	0.85 (0.72-1.00)	0.82 (0.75-0.88)	0.86 (0.70-0.96)	0.81 (0.79-0.87)
36 Months	0.82 (0.53-0.91)	0.79 (0.73-0.85)	0.84 (0.69-0.95)	0.78 (0.76-0.83)
42 Months	0.81 (0.53-0.90)	0.79 (0.70-0.83)	0.84 (0.68-0.94)	0.78 (0.75-0.81)
48 Months	0.79 (0.53-0.88)	0.76 (0.68-0.81)	0.81 (0.67-0.94)	0.75 (0.72-0.79)
54 Months	0.76 (0.53-0.87)	0.73 (0.66-0.79)	0.79 (0.64-0.93)	0.72 (0.70-0.77)
60 Months	0.74 (0.53-0.86)	0.70 (0.63-0.77)	0.77 (0.61-0.92)	0.70 (0.67-0.75)

Table 3.34: Causal Effects of Triple Therapy on Death: Therapy Initiation Under 12 Months Among Asymptomatic or Symptomatic Children (60 Months Follow-up)

	Triple Therapy Among Asymptomatic		**Triple Therapy Among Symptomatic**	
Time Interval	Marginal Additive Difference $\Psi_{AD}(p_0)(t_k)$	Marginal log Hazard Ratio $\Psi_{HZ}(p_0)(t_k)$	Marginal Additive Difference $\Psi_{AD}(p_0)(t_k)$	Marginal log Hazard Ratio $\Psi_{HZ}(p_0)(t_k)$
6 Months	0.007 (-0.119-0.046)	-0.167 (-15.137-1.497)	0.010 (-0.010-0.028)	-0.340 (-1.445-0.306)
12 Months	0.012 (-0.126-0.089)	-0.167 (-15.137-1.073)	0.020 (-0.020-0.065)	-0.340 (-1.368-0.306)
18 Months	0.016 (-0.126-0.131)	-0.166 (-15.137-0.846)	0.028 (-0.028-0.085)	-0.339 (-1.387-0.306)
24 Months	0.027 (-0.119-0.170)	-0.166 (-15.137-0.684)	0.043 (-0.112-0.104)	-0.338 (-1.398-0.562)
30 Months	0.025 (-0.109-0.208)	-0.166 (-15.137-0.554)	0.050 (-0.095-0.123)	-0.338 (-1.405-0.445)
36 Months	0.029 (-0.261-0.105)	-0.166 (-15.297-1.009)	0.057 (-0.079-0.140)	-0.338 (-1.411-0.348)
42 Months	0.030 (-0.232-0.119)	-0.166 (-15.297-0.871)	0.058 (-0.070-0.157)	-0.338 (-1.415-0.305)
48 Months	0.033 (-0.203-0.133)	-0.165 (-15.297-0.775)	0.064 (-0.075-0.174)	-0.337 (-1.418-0.305)
54 Months	0.036 (-0.176-0.145)	-0.165 (-15.297-0.693)	0.070 (-0.080-0.190)	-0.337 (-1.421-0.305)
60 Months	0.039 (-0.156-0.158)	-0.165 (-15.297-0.621)	0.075 (-0.084-0.205)	-0.336 (-1.423-0.305)

Table 3.35: Comparison of Estimates from G-Comp and Traditional Techniques Therapy Initiation Under 6 Months or Under 12 Months (36 Months of Follow-up)

	Triple Therapy Initiated First 6 Months				**Triple Therapy Initiated First 12 Months**			
Outcome	Marginal Additive Difference $\Psi AD(p_0)(tk)$	Marginal log Hazard Ratio $\Psi HZ(p_0)(tk)$	Mean Marginal Log HR over tk	Cox PH Model	Marginal Additive Difference $\Psi AD(p_0)(tk)$	Marginal log Hazard Ratio $\Psi HZ(p_0)(tk)$	Mean Marginal Log HR over tk	Cox PH Model
	Overall				Overall			
C Diagnosis	0.120 (-0.127, 0.270)[1]	-0.466 (-1.457, 0.397)[1]	-0.470 (-1.468, 1.353)[1]	-0.476 (-1.486, 0.534)[1]	0.087 (-0.065, 0.151)[1]	-0.321 (-0.588, 0.212)[1]	-0.324 (-0.617, 0.217)[1]	-0.407 (-0.876, 0.062)[1]
C Diagnosis or Death	0.108 (-0.110, 0.311)	-0.369 (-1.588, 0.300)	-0.369 (-1.634, 0.822)	-0.346 (-1.249, 0.558)	0.055 (-0.150, 0.158)	-0.180 (-0.599, 0.445)	-0.179 (-0.625, 0.468)	-0.431 (-0.875, 0.012)
	Asymptomatic				Asymptomatic			
C Diagnosis	0.115 (-0.327- 0.446)	-0.429 (-16.23, 0.948)	-0.429 (-16.27, 1.438)	-0.712 (-1.445, 0.022)	0.030 (-0.250, 0.225)[2]	-0.106 (-1.054, 0.739)[2]	-0.106 (-1.138, 2.105)[2]	-0.516 (-1.134, 0.102)[2]
C Diagnosis or Death	0.140 (-0.380- 0.490)[2]	-0.514 (-17.72, 1.086)[2]	-0.515 (-17.76, 1.316)[2]	-0.570 (-1.231, 0.091)[2]	0.061 (-0.208, 0.267)[2]	-0.205 (-1.205, 0.601)[2]	-0.206 (-1.279, 1.906)[2]	-0.471 (-0.523, 0.004)[2]
	Symptomatic				Symptomatic			
C Diagnosis	0.146 (-0.214- 0.355)[3]	-0.563 (-14.34- 0.653)[3]	-0.570 (-14.42, 0.755)[3]	-0.634 (-2.041, 0.774)[3]	0.152 (-0.153- 0.240)[3]	-0.587 (-1.217- 0.480)[3]	-0.594 (-15.65, 1.529)[3]	-0.068 (-0.610, 0.474)[3]
C Diagnosis or Death	0.062 (-0.236- 0.403)	-0.204 (-14.75- 0.683)	-0.204 (-14.73, 0.766)	-0.291 (-0.951, 0.369)	0.052 (-0.255- 0.319)[3]	-0.168 (-1.526- 0.709)[3]	-0.171 (-15.36, 2.807)[3]	-0.214 (-0.749, 0.320)[3]

1 Adjusted for sex, race, and pregnancy term

2 Adjusted for pregnancy term

3 Adjusted for an interaction between prenatal care and a time indicator variable for 25-30 months and an interaction between sex and pregnancy term

Figures

Figure 3.1: Causal Assumptions Within This Study of the Effect of Triple Therapy Initiated in First 6 Months of Life on Time to C Diagnosis and/or Death

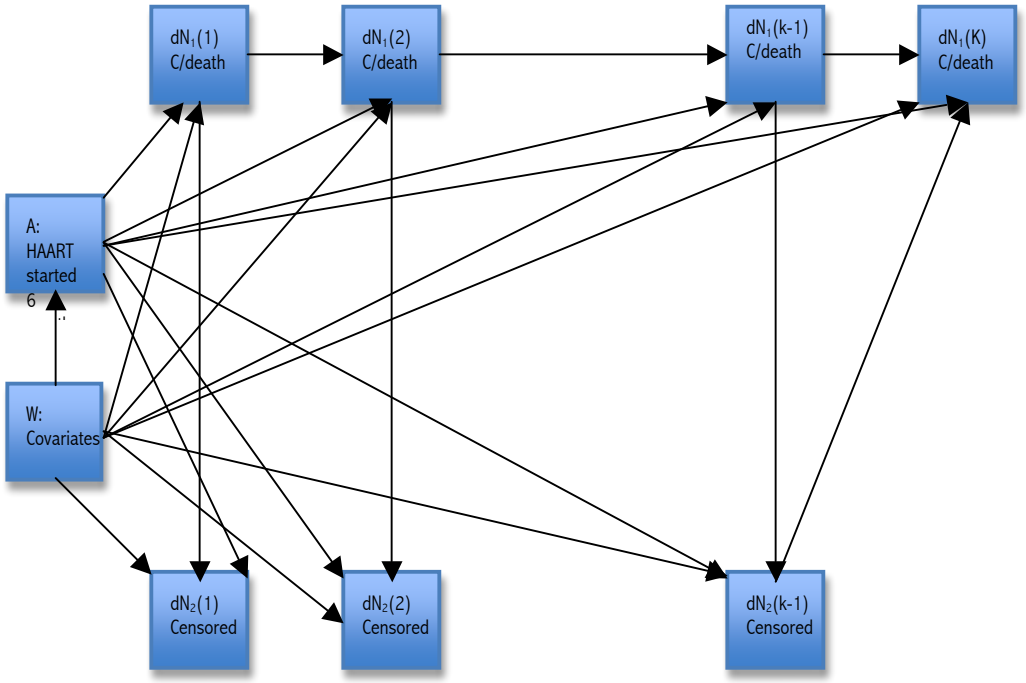


Figure 3.2: Causal Assumptions Within This Study of the Effect of Triple Therapy Initiated in First 6 Months of Life on Time to C Diagnosis and/or Death After Adjustment

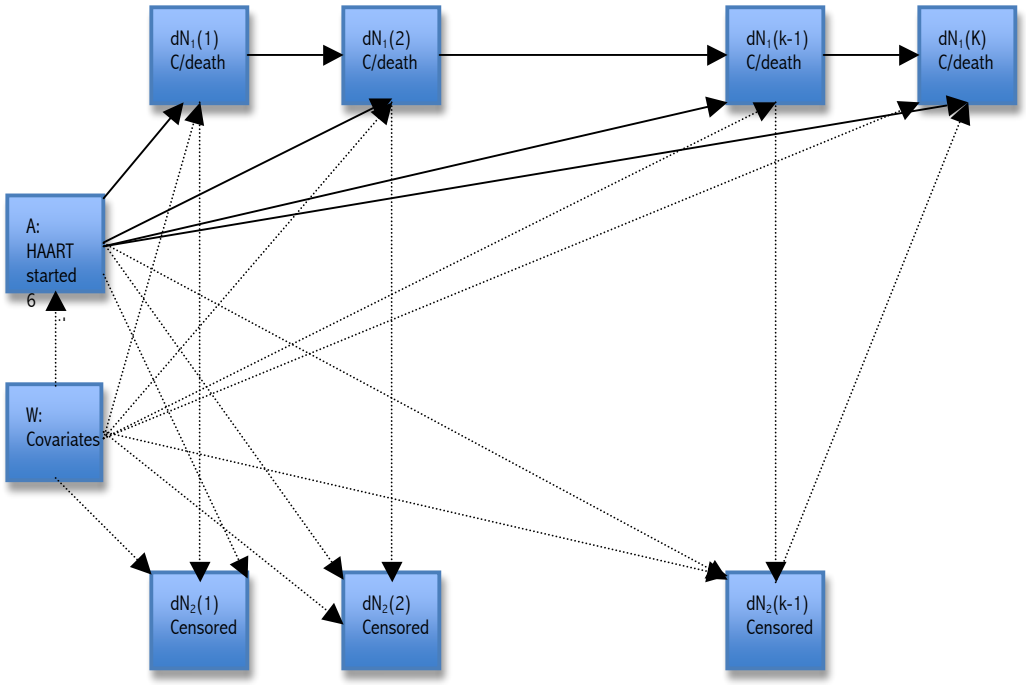


Figure 3.3: Directed Acyclic Graph (DAG) describing possible association between initiation of HAART and mortality from HIV/AIDS

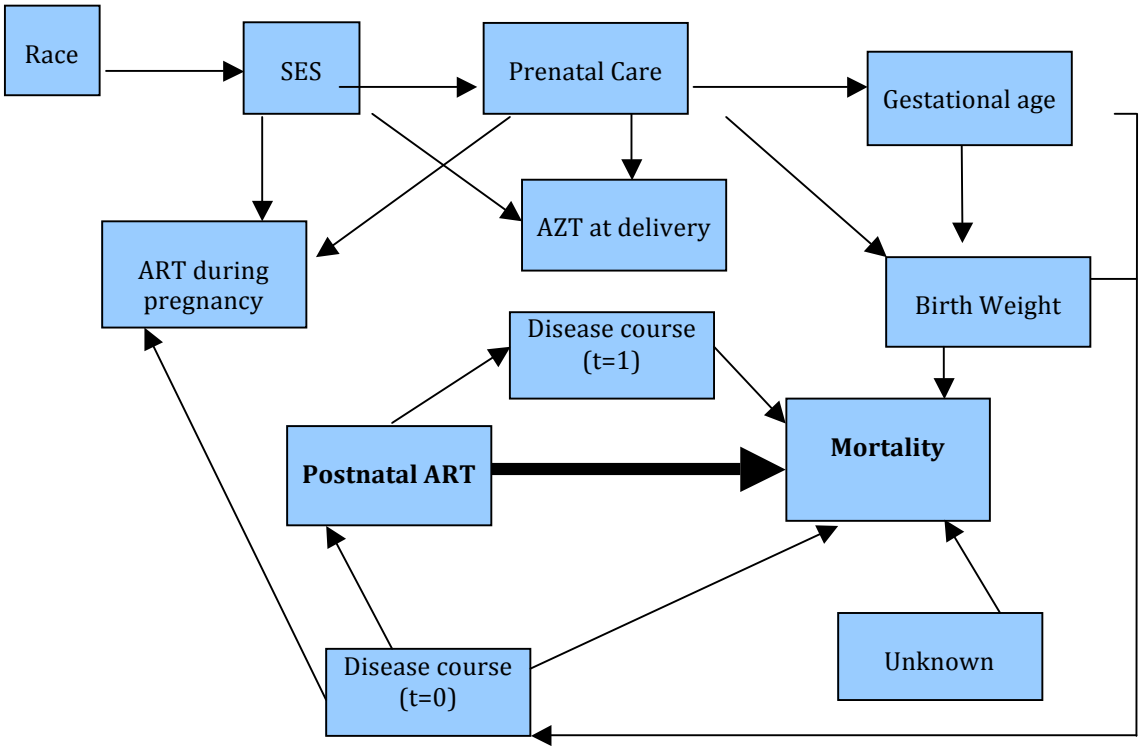


Figure 3.4: Directed Acyclic Graph (DAG) describing possible association between initiation of HAART and C Diagnosis

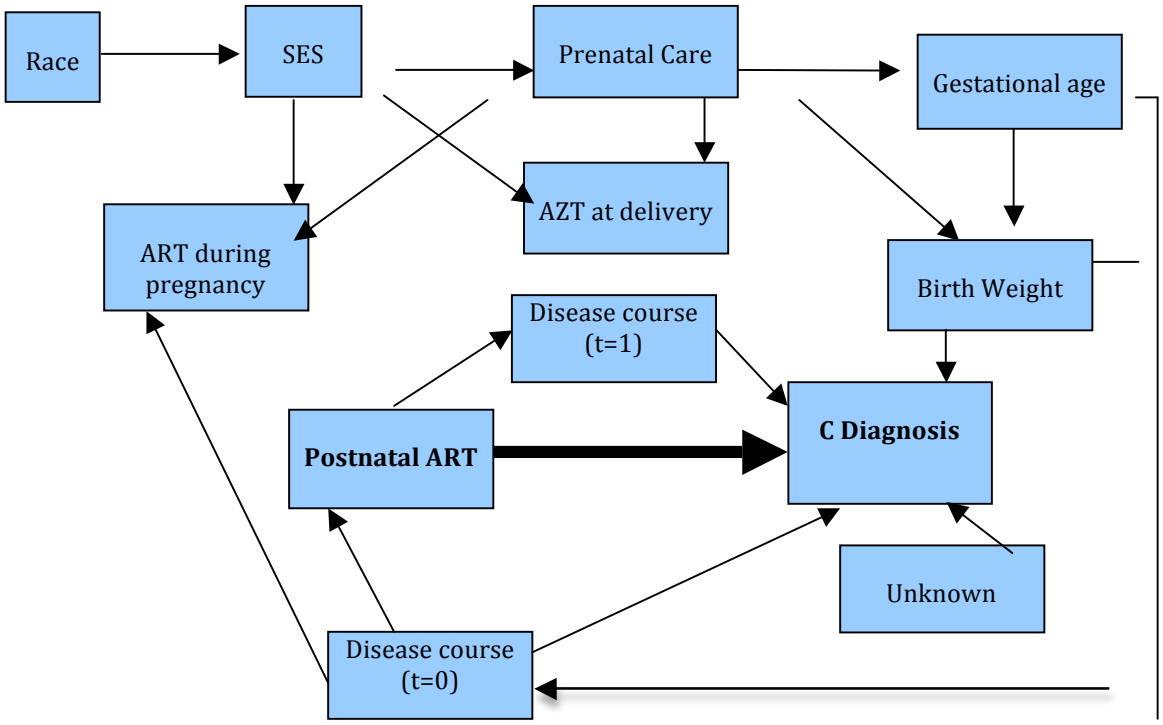


Figure 3.5a:

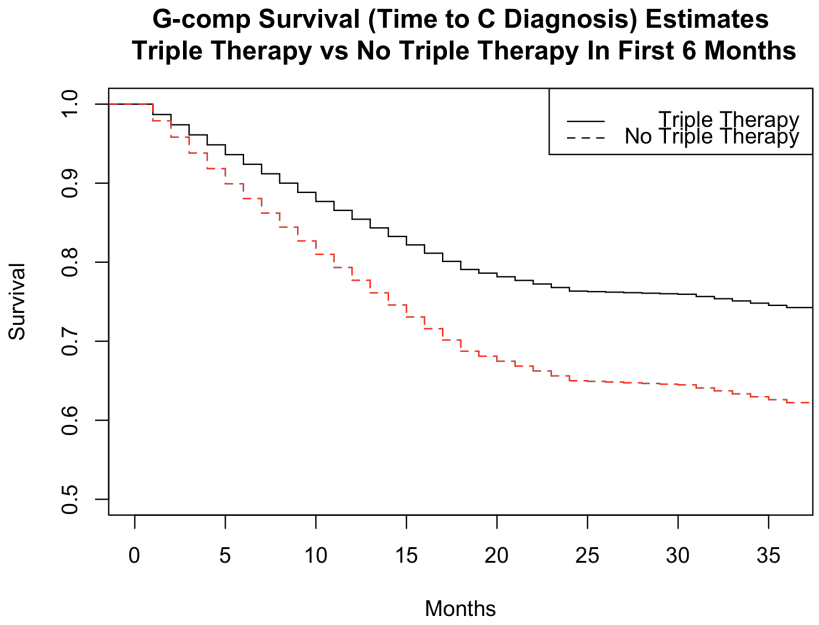


Figure 3.5b:

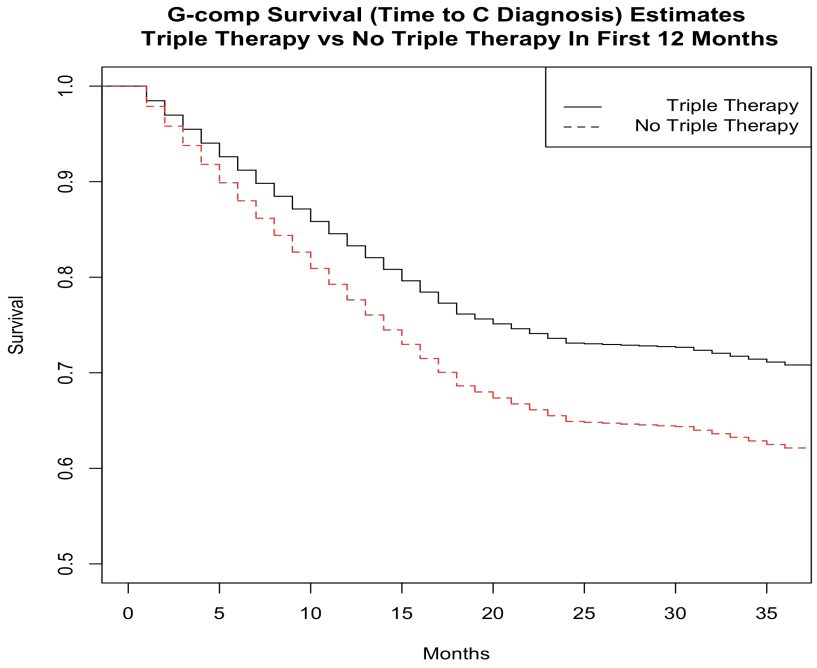


Figure 3.6a:

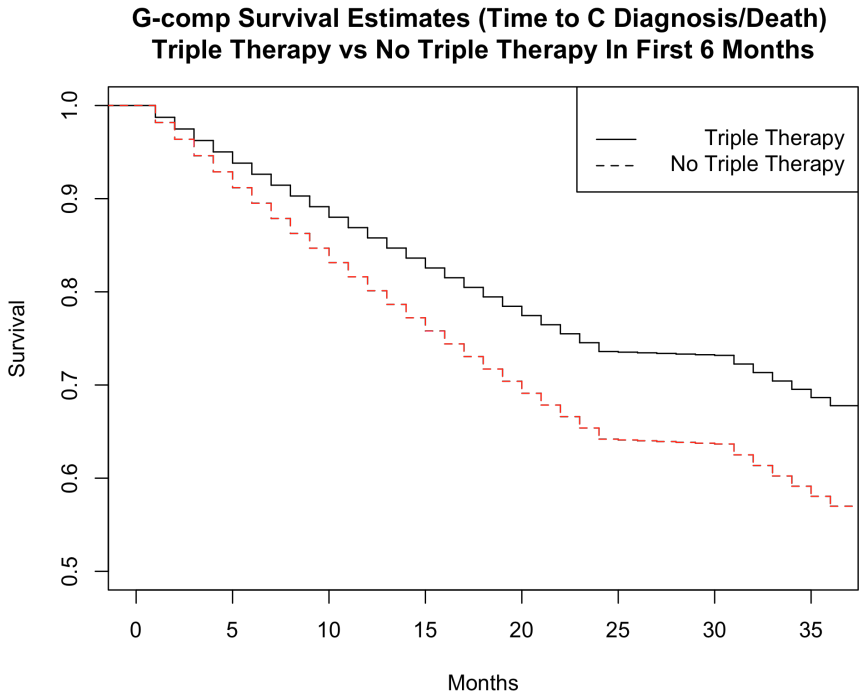


Figure 3.6b:

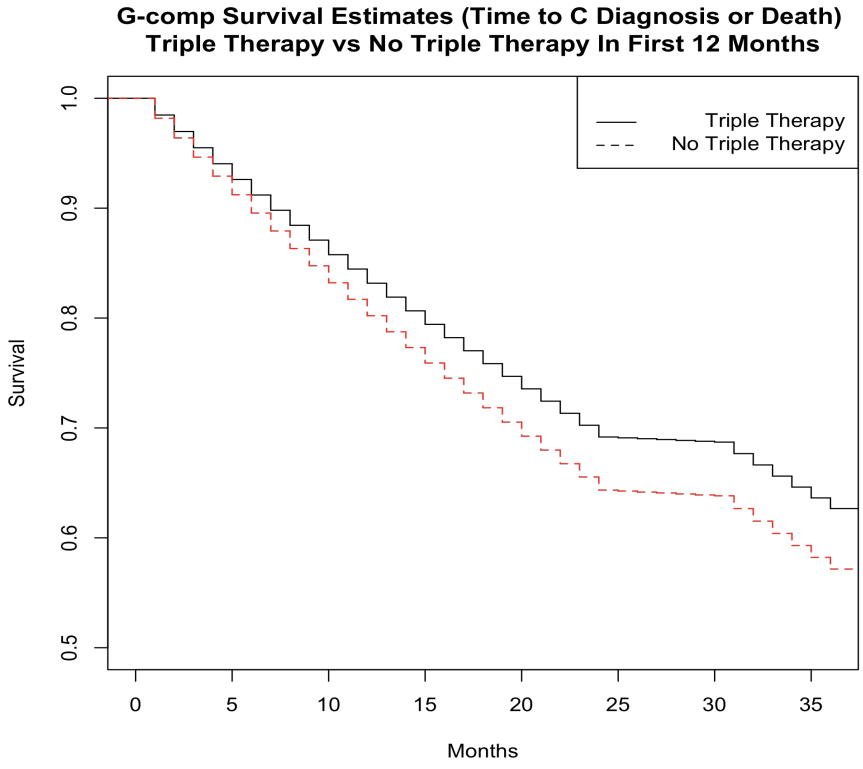


Figure 3.7a:

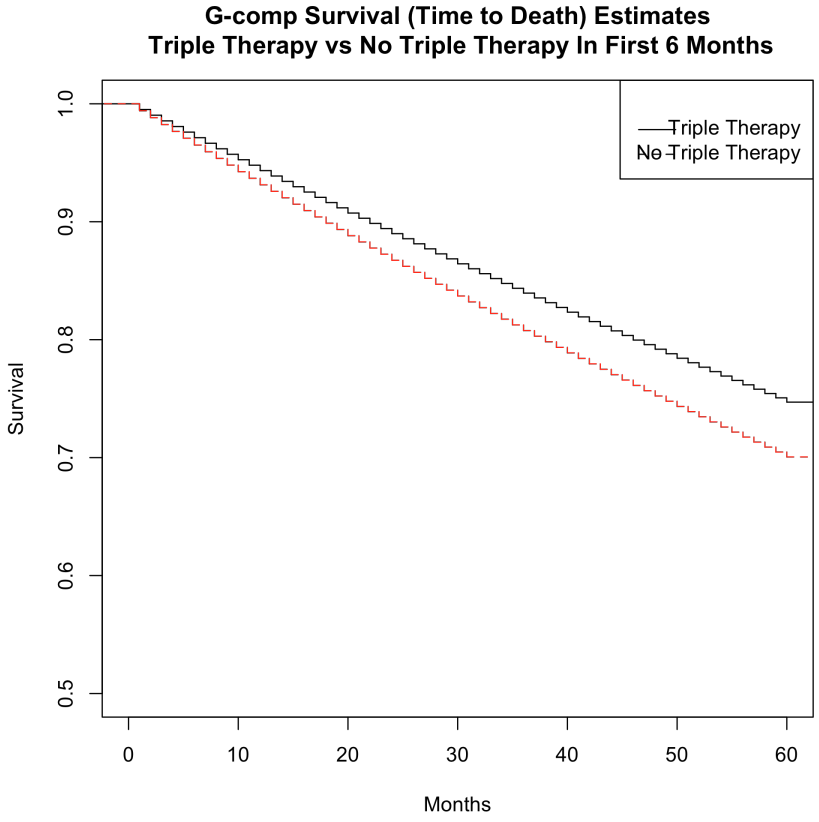


Figure 3.7b:

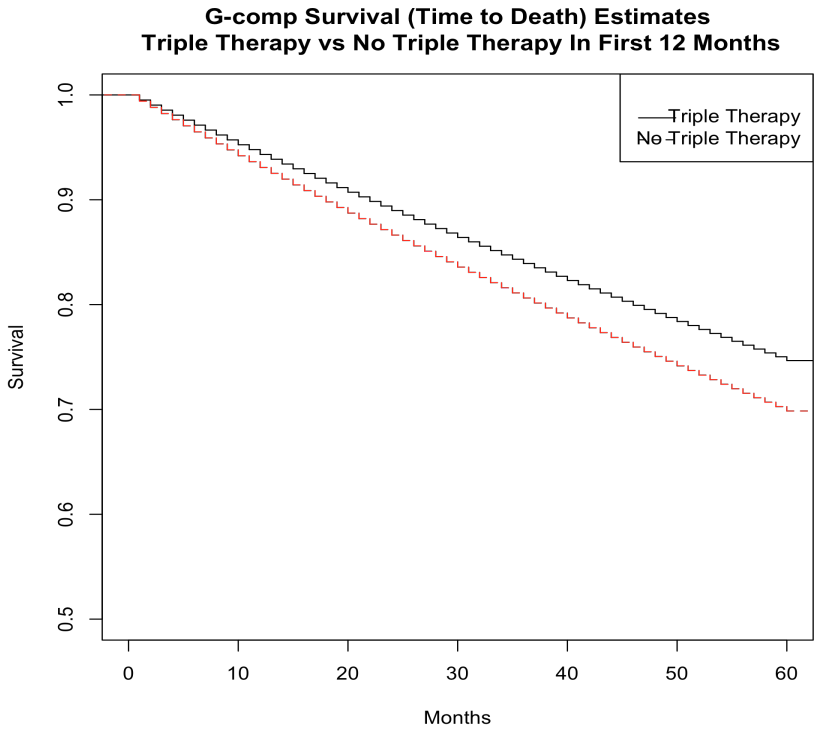


Figure 3.8a:

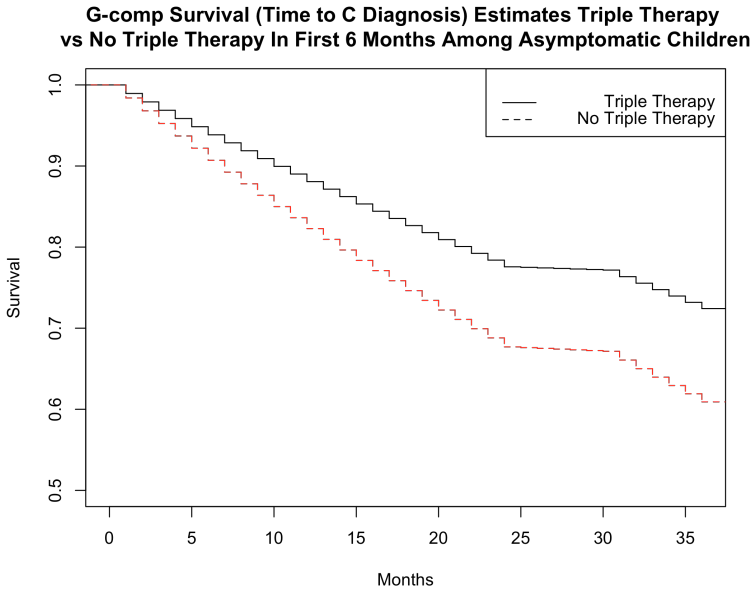


Figure 3.8b:

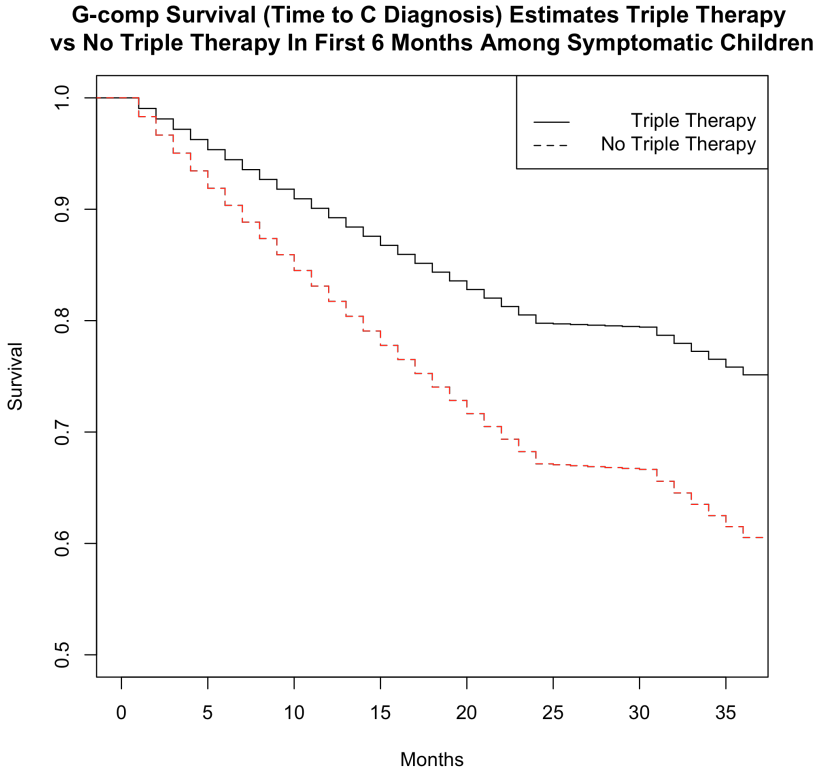


Figure 3.9a:

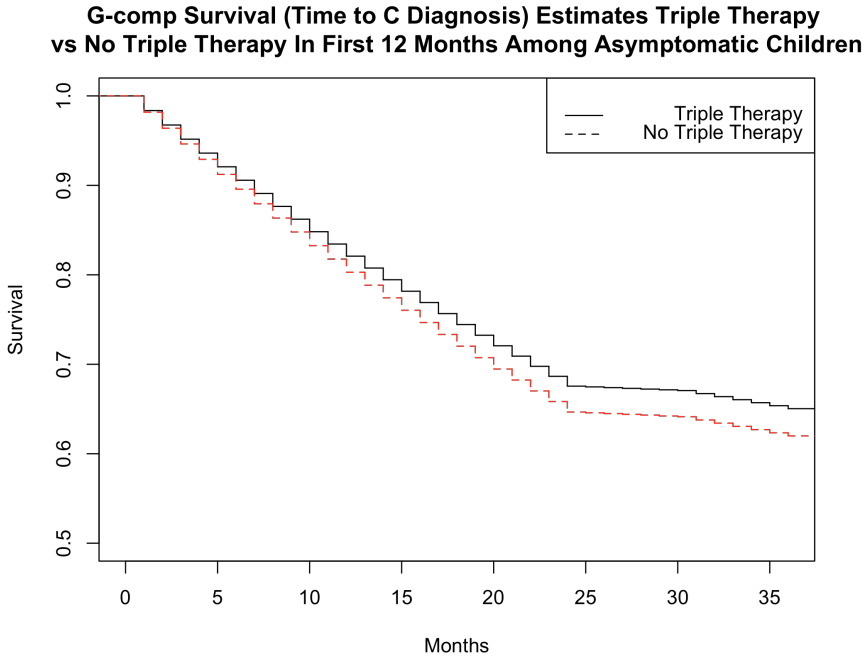


Figure 3.9b:

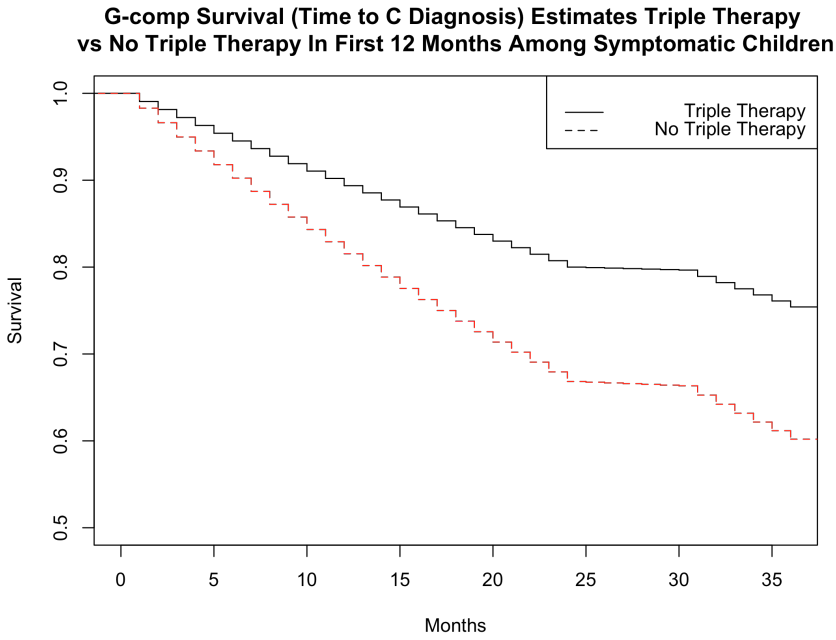


Figure 3.10a:

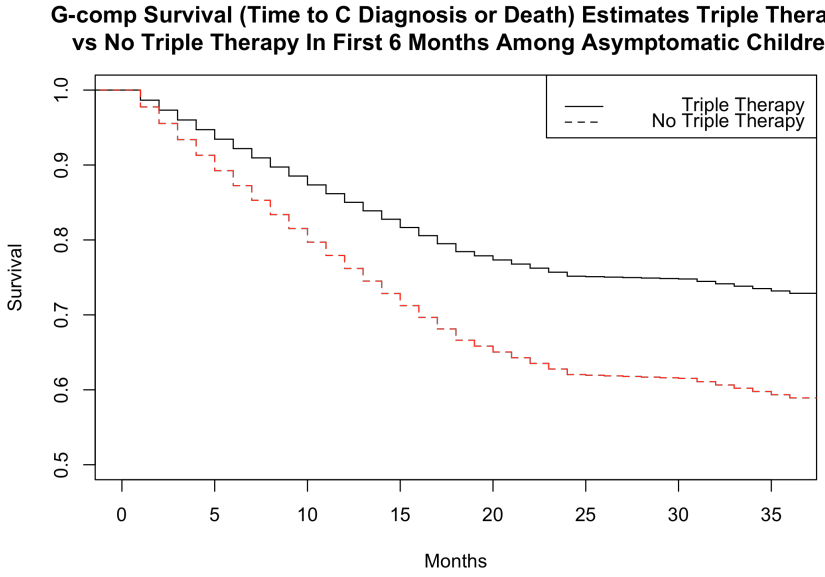


Figure 3.10b:

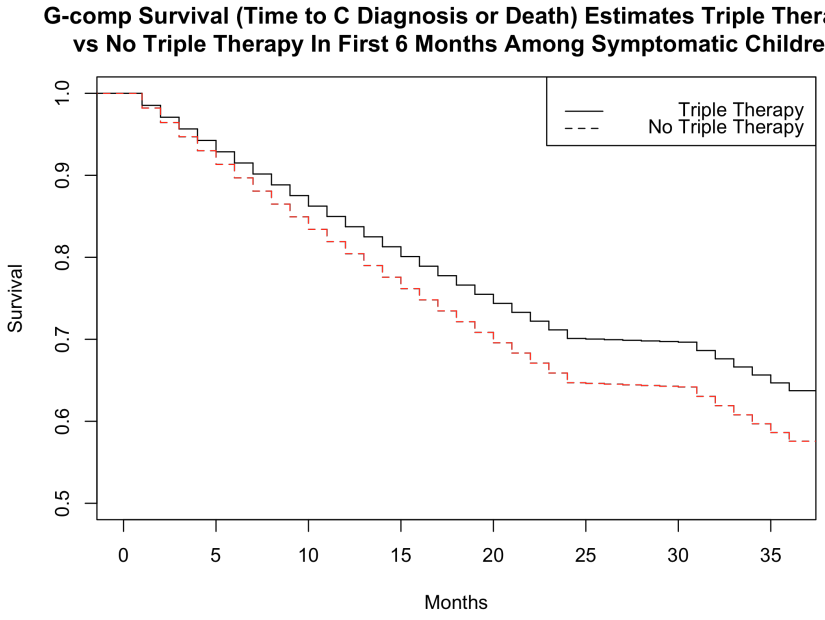


Figure 3.11a:

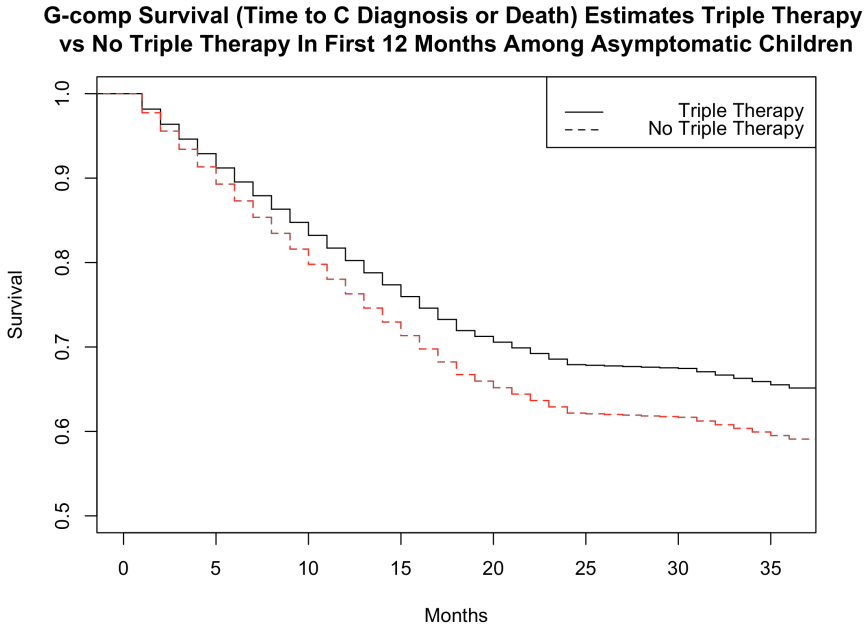


Figure 3.11b:

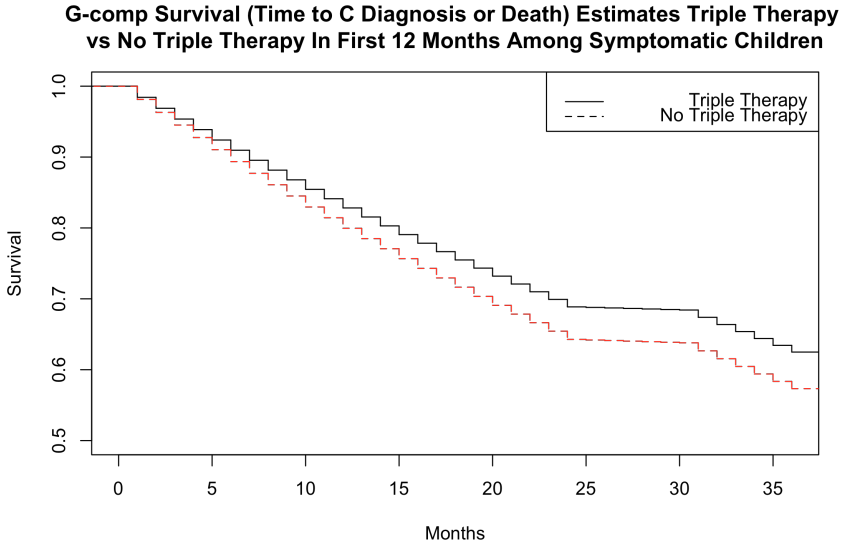


Figure 3.12a:

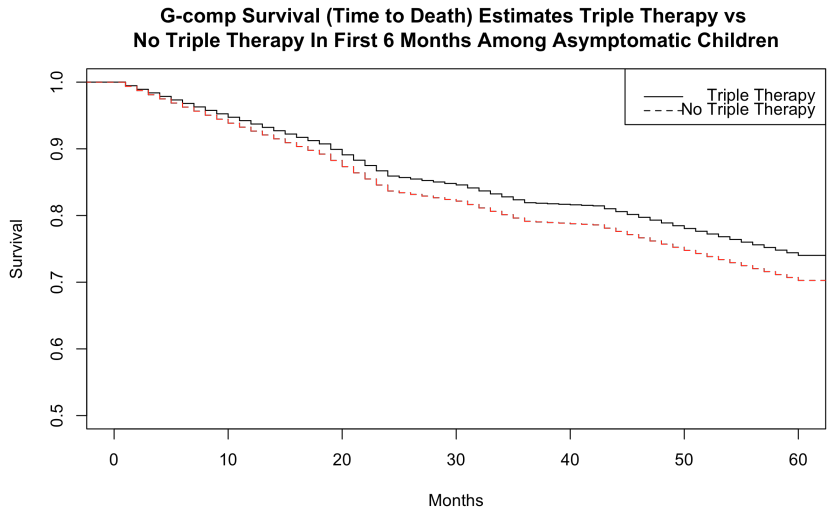


Figure 3.12b:

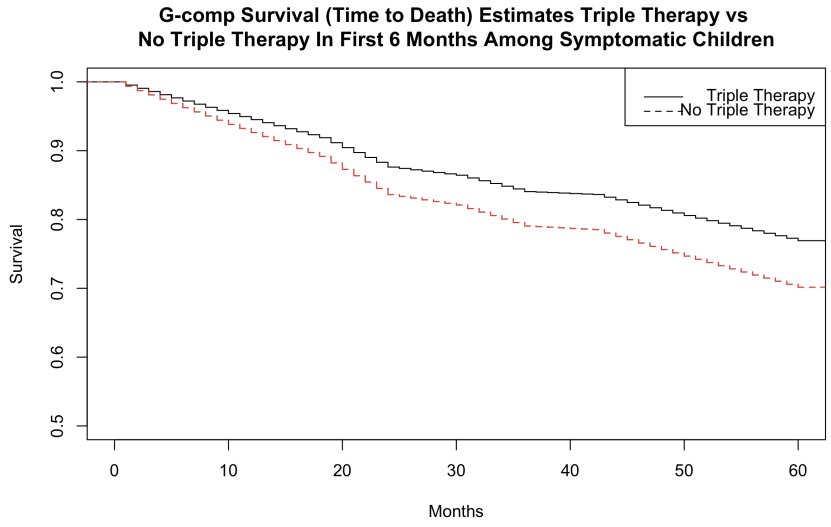


Figure 3.13a:

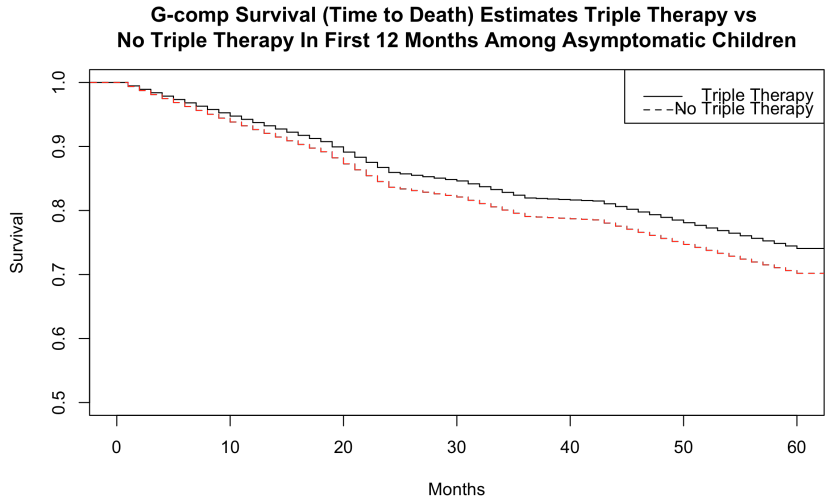
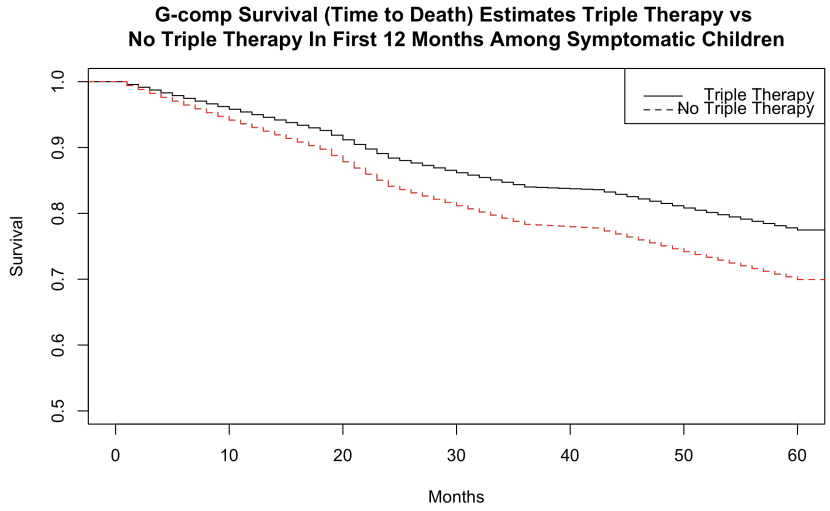


Figure 3.13b:



Bibliography

1. De Cock KM, Fowler MG, Mercier E, de Vincenzi I, Saba J, Hoff E, Alnwick DJ, Rogers M, Shaffer N. Prevention of mother-to-child HIV transmission in resource-poor countries: translating research into policy and practice. Jama 2000;283(9):1175.

2. HIV/AIDS Update [Internet]. [date unknown];Available from: http://www.unaids.org/

3. UNAIDS/World Health Organization. HIV/AIDS Update [Internet]. [date unknown].Available from: http://data.unaid-s.org/pub/EPISlides/2007/2007_epiupdate_en.pdf

4. World Health Organization. Report of the WHO Technical Reference Group, Paediatric HIV/ART Care Guideline Group Meeting. Geneva, Switzerland: 2008.

5. Maldonado Y, Wang N, Caldwell B. Factors associated with early clinical recognition of children with perinatal human immunodeficiency virus infection. Northern California Pediatric HIV Consortium. J Infect Dis 1995;171:689-692.

6. Shah B, Laupacis A, Hux J, Austin P. Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review. Journal of Clinical Epidemiology 2005;58(6):550-559.

7. World Health Organization. Guidelines for Anitretroviral Therapy for HIV Infection in Adults and Adolescents: Recommendations for a public health approach. 2006.

8. BARRE-SINOUSSI F, Chermann C, Rey F, Nugeyre M. Isolation of a T-Lymphotropic Retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science 1983;220(4599):868-871.

9. Gallo R, Sarin P, Gelmann E, Robert-Guroff M. Isolation of Human T-Cell Leukemia Virus in Acquired Immune Deficiency Syndrome (AIDS). Science 1983;220(4599):865-867.

10. Rutherford GW, Oliva GE, Grossman M, Green JR, Wara DW, Shaw NS, Echenberg DF, Wofsy CB, Weinstein DH, Stroud F, others. Guidelines for the control of perinatally transmitted human immunodeficiency virus infection and care of infected mothers, infants and children. Western Journal of Medicine 1987;147(1):104.

11. Schulte J, Dominguez K, Sukalac T, et al. Declines in low birth weight and preterm birth among infants who were born to HIV-infected women during an era of increased use of maternal antiretroviral drugs: pediatric spectrum of HIV disease, 1989-2004. Pediatrics 2007;119:e900-e906.

12. Taha TE, Kumwenda NI, Hoover DR, Kafulafula G, Fiscus SA, Nkhoma C, Chen S, Broadhead RL. The impact of breastfeeding on the health of HIV-positive mothers and their children in sub-Saharan Africa. Bulletin of the World Health Organization 2006;84:546–554.

13. Adair T. UNMET NEED FOR CONTRACEPTION AMONG HIV-POSITIVE WOMEN IN LESOTHO AND IMPLICATIONS FOR MOTHER-TO-CHILD TRANSMISSION. J. Biosoc. Sci 2008;41(02):269.

14. Embree J, Bwayo J, Nagelkerke N, Njenga S, Nyange P, Ndinya-Achola J, Pamba H, Plummer F. Lymphocyte subsets in human immunodeficiency virus type 1-infected and uninfected children in Nairobi. The Pediatric infectious disease journal 2001;20(4):397.

15. Petru A. [date unknown];

16. AYOUBA A, NERRIENET E, MENU E, LOBE MM, THONNON J, LEKE RJ, BARRE-SINOUSSI F, MARTIN P, CUNIN P. Mother-to-child transmission of human immunodeficiency virus type 1 in relation to the season in Yaounde, Cameroon. The American journal of tropical medicine and hygiene 2003;69(4):447.

17. Mphatswe W, Blanckenberg N, Tudor-Williams G, Prendergast A, Thobakgale C, Mkhwanazi N, McCarthy N, Walker BD, Kiepiela P, Goulder P. High frequency of rapid immunological progression in African infants infected in the era of perinatal HIV prophylaxis. Aids 2007;21(10):1253.

18. Guay LA, Musoke P, Fleming T, Bagenda D, Allen M, Nakabiito C, Sherman J, Bakaki P, Ducar C, Deseyve M, others. Intrapartum and neonatal single-dose nevirapine compared with zidovudine for prevention of mother-to-child transmission of HIV-1 in Kampala, Uganda: HIVNET 012 randomised trial. The Lancet 1999;354(9181):795–802.

19. Global Partners Forum. Achieving Universal Access to Comprehensive PMTCT Services. [Internet]. [date unknown];Available from: http://www.procaare.org/archive/ procaare/200711/msg00013.php

20. Reynolds HW, Janowitz B, Homan R, Johnson L. The Value of Contraception to Prevent Perinatal HIV Transmission. Sexually Transmitted Diseases 2006;33(6):350-356.

21. Sweat MD, O'Reilly KR, Schmid GP, Denison J, de Zoysa I. Cost-effectiveness of nevirapine to prevent mother-to-child HIV transmission in eight African countries. AIDS 2004;18(12):1661-1671.

22. Rutenberg N, Baek C. Field experiences integrating family planning into programs to prevent mother-to-child transmission of HIV. Studies in Family Planning 2005;36(3):235–245.

23. De Martino M, Tovo PA, Balducci M, Galli L, Gabiano C, Rezza G, Pezzotti P, others. Reduction in mortality with availability of antiretroviral therapy for children with perinatal HIV-1 infection. Jama 2000;284(2):190.

24. The European Collaborative Study. Fluctuations in Symptoms in Human Immunodeficiency Virus-Infected Children: The First 10 Years of Life. PEDIATRICS 2001;108(1):116-122.

25. Blanche S, Newell M, Mayoux, M, et al. Morbidity and Mortality in European Children Vertically Infected by HIV-1. Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology. 1997; 14(5): 442-450.

26. Gortmaker SL, Hughes M, Cervia J, Brady M, Johnson GM, Seage III GR, Song LY, Dankner WM, Oleske JM, others. Effect of combination therapy including protease inhibitors on mortality among children and adolescents infected with HIV-1. The New England Journal of Medicine 2001;345(21):1522.

27. Patel K, Hernan MA, Williams PL, Seeger JD, McIntosh K, Dyke RBV, Seage III GR, Pediatric AIDS Clinical Trials Group 219/219C Study Team. Long□Term Effectiveness of Highly Active Antiretroviral Therapy on the Survival of Children and Adolescents with HIV Infection: A 10□Year Follow□Up Study. CLIN INFECT DIS 2008;46(4):507-515.

28. Gibb DM, Duong T, Tookey PA, Sharland M, Tudor-Williams G, Novelli V, Butler K, Riordan A, Farrelly L, Masters J, others. Decline in mortality, AIDS, and hospital admissions in perinatally HIV-1 infected children in the United Kingdom and Ireland. Bmj 2003;327(7422):1019.

29. Borleffs J, Boer K, Boucher C. Antiretroviral Treatment in the Netherlands [Internet]. [date unknown].Available from: www.cbo.nl

30. Centers for Disease Control and Prevention. Guidelines for the use of antiretroviral agents in pediatric HIV infection. [Internet]. [date unknown];Available from: www.aidsinfo.nih.gov

31. De Martino M, Tovo P, Giaquinto C, de Rossi A, Galli L, Gabiano C. Italian Guidelines for antiretroviral therapy in children with human immunodeficiency virus-type 1 infection. Acta Paediatrica 1999;88:228-32.

32. Sharland M, Blanche S, Castelli G, Ramos J, Gibb DM, others. PENTA guidelines for the use of antiretroviral therapy, 2004. HIV medicine 2004;5:61.

33. Draenert R, Goebel F. Recommendations and prospects of antiretroviral therapy. Dtsch Med Wochenschr 2001;126(18):539-43.

34. Paediatrica A. Management of vertically HIV-infected children in Europe. Acta Paediatrica [date unknown];92(2):246–250.

35. Violari A, Cotton MF, Gibb DM, Babiker AG, Steyn J, Madhi SA, Jean-Philippe P, McIntyre JA. Early antiretroviral therapy and mortality among HIV-infected infants. New England Journal of Medicine 2008;359(21):2233.

36. Ananworanich J, Kosalaraksa P, Siangphoe U, Engchanil C, Pancharoen C, Lumbiganon P, Intasan J, Apateerapong W, Chuenyam T, Ubolyam S, others. A feasibility study of immediate versus deferred antiretroviral therapy in children with HIV infection. AIDS Research and Therapy 2008;5(1):24.

37. Five year follow up of vertically HIV infected children in a randomised double blind controlled trial of immediate versus deferred zidovudine: the PENTA 1 trial. Archives of Disease in Childhood 2001;84(3):230-236.

38. Membership I, In VC, Falsey AR, Walsh EE, Capellan J, Gravenstein S, Zambon M, Yau E, Gorse GJ, Edelman R, others. CD4 cell response to antiretroviral therapy in children with vertically acquired HIV infection: is it associated with age at initiation? The Journal of Infectious Diseases 2006;193(7):954–962.

39. Chiappini E, Galli L, Tovo P. Five year follow up of children with perinatal HIV-1 infection receiving early highly active antiretroviral therapy. BMC Infectious Diseases 2009;9:140.

40. Keating MJ, Flinn I, Jain V, Binet JL, Hillmen P, Byrd J, Albitar M, Brettman L, Santabarbara P, Wacker B, others. Therapeutic role of alemtuzumab (Campath-1H) in patients who have failed fludarabine: results of a large international study. Blood 2002;99(10):3554.

41. McIntosh K, Shevitz A, Zaknun D, et al. Age and Time related changes in extracellular viral load in children vertically infected by human immunodeficiency virus. The Pediatric Infectious Disease Journal. 1996; 15(12): 1087-1091.

42. Palumbo PE, Kwok S, Waters S, Wesley Y, Lewis D, McKinney N, Bardeguez A, Connor EM, Oleske JM. Viral measurement by polymerase chain reaction-based assays in human immunodeficiency virus-infected infants. The Journal of pediatrics 1995;126(4):592–595.

43. Shearer WT, Quinn TC, LaRussa P, Lew JF, Mofenson L, Almy S, Rich K, Handelsman E, Diaz C, Pagano M, others. Viral load and disease progression in infants infected with human immunodeficiency virus type 1. The New England journal of medicine 1997;336(19):1337.

44. Burns D, Mofenson L. Paediatric HIV-1 Infection. The Lancet 1999;354((Supplement II)):1.

45. de Martino M, Tovo PA, Galli L, Gabiano C, Chiarelli F, Zappa M, Gattinara GC, Bassetti D, Giacomet V, Chiappini E, others. Puberty in perinatal HIV-1 infection: a multicentre longitudinal study of 212 children. Aids 2001;15(12):1527.

46. Macmillan C, Magder L, Browers P, et al. Head growth and neurodevelopment of infants born to HIV-1-infected drug-using women. Neurology 2001;57:1402-1411.

47. Dubé MP, Sprecher D, Henry WK, Aberg JA, Torriani FJ, Hodis HN, Schouten J, Levin J, Myers G, Zackin R, others. Preliminary guidelines for the evaluation and management of dyslipidemia in adults infected with human immunodeficiency virus and receiving antiretroviral therapy: recommendations of the Adult AIDS Clinical Trial Group Cardiovascular Disease Focus Group. Clinical Infectious Diseases 2000;31(5):1216–1224.

48. Alimenti A, Burdge DR, Ogilvie GS, Money DM, Forbes JC. Lactic acidemia in human immunodeficiency virus-uninfected infants exposed to perinatal antiretroviral therapy. The Pediatric infectious disease journal 2003;22(9):782.

49. Poirier MC, Divi RL, Al-Harthi L, Olivero OA, Nguyen V, Walker B, Landay AL, Walker VE, Charurat M, Blattner WA. Long-term mitochondrial toxicity in HIV-uninfected infants born to HIV-infected mothers. JAIDS Journal of Acquired Immune Deficiency Syndromes 2003;33(2):175.

50. Sharland M, Gattinara di Zub GC, Tomas Ramos J, Blanche S, Gibb DM. PENTA guidelines for the use of antiretroviral therapy in paediatric HIV infection. HIV medicine 2002;3(3):215–226.

51. Patel K, Hernán M, Williams P, Seeger J, McIntosh K, Dyke R, Seage III G, Pediatric AIDS Clinical Trials Group 219/219C Study Team. Long□Term Effects of Highly Active Antiretroviral Therapy on CD4 Cell Evolution among Children and Adolescents Infected with HIV: 5 Years and Counting. CLIN INFECT DIS 2008;46(11):1751-1760.

52. Virus HI, Addenda OA, Codes I. ICD-9-CM. [date unknown];

53. Stitleman O. G-Computation Methods. (In Press) [date unknown];

54. The Working Group on Antiretroviral Therapy and Medical Management of HIV Infected Children. Guidelines for the use of antiretroviral agents in pediatric HIV infection. [Internet]. US Department of Health and Human Services; [date unknown].Available from: http://aidsinfo.nih.gov/ContentFiles/PediatricGuidelines.pdf

55. Sinisi S, van der Laan MJ. Loss-based cross-validated deletion/substitution/addition algorithms in estimation. UC Berkeley Division of Biostatistics Working Paper Series 2004;

56. Verweel G, Saavedra-Lozano J, van Rossum AMC, Ramilo O, de Groot R. Initiating Highly Active Antiretroviral Therapy in Human Immunodeficiency Virus Type 1-Infected Children in Europe and the United States. The Pediatric Infectious Disease Journal 2006;25(11):987-994.

57. HIV Paediatric Prognostic Markers Collaborative Study Group. Short-term risk of disease progression in HIV-1-infected children receiving no antiretroviral therapy or zidovudine monotherapy: a meta-analysis. Lancet 2003;362:1605-11.

58. Campbell SE, Stone WL, Whaley SG, Qui M, Krishnan K. Gamma tocopherol upregulates peroxisome proliferator activated receptor (PPAR) gamma expression in SW 480 human colon cancer cell lines. BMC Cancer 2003;3(1):25.

Chapter 4

The Effect of Highly Active Antiretroviral Therapy Use Among HIV Positive Children on AIDS or Death Using Calendar Year as an Instrumental Variable

4.1 Abstract

Researchers view randomized controlled trials (RCTs) as the gold standard when estimating the effect of highly active antiretroviral therapy (HAART) on human immunodeficiency virus (HIV) disease progression and/or death. In the absence of RCT data, epidemiologists are tasked with asserting causal inference, assuming no unmeasured confounding, from observational data. Previously calendar periods have been used as a proxy for HAART use as HAART was introduced in mid-1996 in the U.S.(1-12) This approach, referred to as an instrumental variable analysis, can be biased because of misclassification of HAART use.(1)

Methods

In the present study I perform an adapted instrumental variable analysis of 267 HIV-positive children living in Northern California from 1988 to 2009 in order to estimate the causal effect of HAART on the hazard of AIDS events or death. In this adapted instrumental variable analysis, previously proposed by Cain et al, researchers can adjust for noncompliance corrections similar to adjustments performed in RCTs.(1) The two instruments used for the primary analyses were defined as: 1) before 1997/1997 and after; 2) before 1998/1998 and after. Further, I performed separate analyses estimating the causal effect of HAART on the hazard of AIDS events alone under each instrumental variable scenario.

Results

During 61,847 person-days, 113 HIV-positive children received an AIDS diagnosis or died. The naïve rate ratio comparing the early-defined instrumental variable (1997 cut-off) non-HAART era with the HAART era was estimated at 2.17 (95% CI 1.34-3.52). As a result of HAART use misclassification by calendar era, an instrumental variable estimator was used, yielding an instrumental variable rate ratio of 3.91 (95% CI 2.41, 6.34), 80% higher than the naïve result. For the latter-defined instrumental variable (cut-off 1998), the naïve rate ratio comparing the non-HAART era with the HAART era was estimated at 2.27 (95% CI 1.34, 3.85). As a result of HAART use misclassification by calendar era, an instrumental variable estimator was used, yielding an instrumental variable rate ratio of 4.87 (95% CI 2.87, 8.26), more than 2 times higher than the naïve result. To adjust for variables associated with both calendar era and the outcome, weighting by the inverse probability of calendar era given selected covariates was performed. Weighted estimates were not noticeably different than unweighted estimates.

Discussion

Noncompliance adjustments in instrumental variable analyses may help bridge the gap between RCT and observational study evidence. In this analysis, I assumed that calendar period is an appropriate instrument for HAART use as defined by Greenland et al.(13) Further, I assumed exchangeability between calendar eras. This assumption is likely satisfied as Detels et al have previously shown that alternative explanations of the effect of HAART, including HIV-related, non-HAART therapies or use of health care, are not supported because neither of these factors vary significantly across calendar periods.(7)

4.2 Introduction

At a time when the call for action for better treatment for patients with human immunodeficiency virus (HIV) was a prevailing theme in the scientific community, researchers introduced in 1996, a new breed of antiretroviral therapy, highly active antiretroviral therapy (HAART). Unlike its predecessors, which included antiretroviral therapy (ART) regimens in mono and dual therapy form, HAART consisted of a drug regimen of 2 nucleoside reverse transcriptase inhibitors *plus* a protease inhibitor.(14) In fact, HAART has proved to be a significant factor in delaying time to acquired immune deficiency syndrome (AIDS) and death, while previous ART regimens did not fare as well in either randomized controlled trials (RCTs)(15; 16) or in observational studies(3; 5; 7; 17).

The gold standard in research, results from randomized controlled trials are often the first steps in estimating treatment effects of specific therapies. Research from RCTs has established HAART's protective effect on time to AIDS or death before researchers had the ability to estimate HAART's population-level effect. In 1997, using AIDS Clinical Trials Group 320 data, Hammer et al found a reduced risk of AIDS or death among patients who were assigned to receive indinavir, zidovudine, and lamivudine combination therapy when compared to patients assigned to receive zidovudine and lamivudine therapy alone (HR = 0.11; 95% 0.33-0.76).(15) The risk of mortality alone was estimated as HR = 0.43; 95% CI 0.19-0.99. In early clinical trials of the protease inhibitor, ritonavir, the risk of AIDS or death among patients assigned to receive ritonavir therapy decreased when compared a placebo patient group (HR = 0.53; 95% CI 0.42-0.66).(16) Ritonavir's effect on viral load and CD4 count was also noted by Markowitz et al in 1995.(18) Similarly, Gulick et al found that the effect of a three drug combination therapy, consisting of indinavir, zidovudine, and lamivudine, reduced the viral load over 24 weeks significantly more than treatment groups not receiving the triple therapy (p value < 0.001).(19)

The protective effect of HAART on delaying time to an AIDS-defining illness and/or death is well-established in observational study findings in several study settings and countries and support the findings from clinical trial research.(3; 5; 7; 17) Furthermore, results from observational studies have allowed researchers to estimate HAART's effect on a population level.(7) In a Swiss cohort, Egger et al found among the later time periods 1991-92 [RR=0.82 (95% CI 0.73-0.93)], 1993-94 [RR=0.77 (95% CI 0.65-0.91)], and 1995-96 [RR=0.27 (95% CI 0.18-0.39)], reduced relative risks of progression to AIDS diagnoses when compared to earlier calendar periods (1988-90).(17) Similarly, mortality was reduced by 19%, 26%, and 62% over respective time periods when compared to the mortality risks in 1988-90. When considering the role of antiretroviral therapies has played on AIDS risk, Egger et al found the risk of AIDS diagnoses after CD4 counts drop below 200 cells decreased by 16% with monotherapy, by 24% dual therapy, and by 42% with triple therapy when compared to no antiretroviral treatment.(17) Similarly, mortality was reduced by 23%, 31%, and 65% over respective

time periods when compared to the mortality risks with no antiretroviral therapy.(17) The EuroSIDA Study group found that among European patients the incidence of AIDS-defining illnesses (ADIs) declined from 30.7 per 100 patient years (1994) to 2.5 per 100 patient years (1998) (p value < 0.001 test for trend).(3) Furthermore, after stratification by CD4 count (<=50, 51-200, and > 200 cell/mL), patients not on HAART had a higher rate of ADIs when compared to patients taking HAART. In a US-based cohort, Detels et al found the relative hazard of AIDS was similar between the time periods of 1990-July 1993 and 1993-July 1995 (RH 1.04; 95% CI 0.73-1.48).(7) However, the relative hazard decreased in July 1995-July 1997 when compared to the earliest time period (RH 0.35; 95% CI 0.20-0.61). Similarly, the relative hazard of death decreased in July 1995-July 1997 when compared to the earliest time period (RH 0.62; 95% CI 0.38-1.01). Lastly, the CASCADE collaboration studied HIV patient populations in Europe, Australia, and Canada. Researchers compared the risk of death between the pre-HAART era, defined by the researchers as before 1997, and post-ART. The hazard ratio for death was estimated as 0.47 (95% CI 0.39-0.56) in 1997 and 0.16 (95% CI 0.12–0.22) in 2001 when compared to the pre-HAART era.(5) The hazard ratio for AIDS progression also decreased and was estimated as 0.46 (95% CI 0.38-0.55) in 1997 and 0.13 (95% CI 0.09-0.21) by 2001 when compared to the pre-HAART era.(5)

Data from RCTs are most often analyzed using the intention to treat (ITT) principle. Under this rule, once person A is randomized to a treatment X, he should be included in any future analyses comparing treatment assignment arms as if he actually received treatment X despite actual receipt of treatment X. In this respect, ITT is actually determining the effect of assigning person A to treatment X and/or everything else downstreatm (all pathways) of treatment assignment. In a well-run RCT in which few participants are censored or change treatment assignments, the causal effect of the treatment and ITT effects are similar. Alternatively, if randomization failed or treatment assignments were rarely followed, non-compliance adjustments are needed to ensure the results are a true reflection of the causal effect of treatment X.(20-22) In this respect, the estimates produced using an ITT approach will likely be biased toward the null hypothesis if 100% compliance is not realized.

Though results seen in RCTs, with strict random allocation principles, are often supported by observational data, confounding biases are a persistent, inherent problem found in observational studies. For example, while an RCT randomizes study participants to receive treatments, the physicians or the participants themselves select receipt of treatments in an observational study. Therefore, in the latter scenario, an argument for causality is difficult because the effect seen could either be a result of the treatment or it could be a result arising from the reason for selecting the treatment.(23)

Though RCTs provide the best epidemiologic scenario for estimating the causal effect of therapy on disease progression, often researchers are somewhat restricted to analyze observational data because of ethical concerns that are inherent with implementing RCTs. For example, when randomizing children to either receive a far superior drug (e.g. HAART) or receive mono or dual therapy drug regimens, it would be hard to justify the excess illnesses and deaths for this trial if there were already a strong effect suspected.

One of the methods that were developed to deal with the difficulties with asserting causality in observational data, assuming no unmeasured confounding, was instrumental variables (IV). The use of IVs can be dated back over a half-century when they have been found in econometric theory.(24) Unlike the case with many epidemiologic studies, data found in economics are often sparse and lack randomization. To account for these disadvantages in their data economists developed IVs. The earliest application of IVs analyzing HIV/AIDS data, with methods explicitly identifying an instrumental variable, was in 2001.(9) In a study of population effectiveness of antiretroviral therapy in reducing AIDS diagnoses in an HIV positive population, Tarwater et al used calendar year as an external time-dependent variable.(9) This allowed the researchers to account for different infection durations. As discussed in Ch. 2, of all causal inference techniques developed to better account for biases commonly found in observational data, instrumental variables are one of the most under-utilized.

Like the principle of random allocation to treatments in RCTs, traditionally instrumental variables are variables that only affect the outcome through their effect on the treatment or exposure alone.(23) Instrumental variables methods attempt to estimate mean difference (rates) in counterfactual distributions as if everyone complied with the intended treatment. Moreover, though it is often difficult to assert a variable can serve as an IV, randomization assures that the treatment assignment is completely exogenous, and thus independent of any future counterfactual outcomes. Unless noncompliance is rampant, the IV will be highly correlated with the treatment assignment. In the current context, noncompliance does not refer to clinical adherence, but rather use of HAART during the HAART era. In the present analysis, I assume that calendar period is an appropriate instrument for HAART use. This assumption is based on three key characteristics of the IV: 1) is independent of variables that affect both HAART and outcome; 2) is associated with HAART; 3) is independent of the outcome given HAART and covariates that affect both HAART and the outcome.(13) The third assumption is particularly important as if it does not hold true and the instrument is directly related to the outcome, the results will be biased.(25)

The use of directed acyclic graphs (DAGs) helps justify the use of a specific IV. In Figures 4.1 calendar year is only related to the outcome (death) through the exposure (ART). In this manner, calendar year presumably makes for a plausible instrumental variable. As a result, the researcher will be able to estimate how much the variation in ART that is explained by the calendar year affects death. Detels et al has previously shown that alternative explanations of the effect of HAART, including HIV-related, non-HAART therapies or use of health care, are not supported as neither of these factors vary significantly across calendar periods.(7)

While using the era of HAART (post HAART introduction in 1996) as a proxy for actual HAART use in an instrumental variable approach is not a panacea for inherent problems found in HIV/AIDS observational data, there are some unique benefits. Researchers have shown that any confounding by indication is presumably removed.(26) That is to say, the

confounding bias seen in observational studies where there are differences in underlying health conditions between the treated and untreated populations is minimized. In situations where there is no misclassification of HAART exposure, calendar year performs well as a proxy for actual HAART use. However, if the calendar period, e.g. pre-1997 and 1997 onwards in the present case, is not 100% representative of actual HAART exposure, information bias can be introduced.(1) Furthermore, some covariates may have associations with calendar period *and* the outcome of interest, violating a principle concept of instrumental variables—their independence of the outcome given treatment and covariates that affect both treatment and the outcome.

To account for possible information bias introduced by calendar year, noncompliance correction in RCTs has been previously adapted for use in observational HIV/AIDS data.(1) In clinical trials, this correction method is useful in situations where randomization fails.(22) Unfortunately, often treatment contamination (the use of the intervention among controls) remains a possibility in RCTs, especially in prevention trials, or is an inherent attribute of the trial design.(1) As a result, baseline risks for compliers and non-compliers may be different, negating the benefits of randomization. As opposed to an ITT analysis, which ignores compliance and simply estimates the mean difference between treatment assignment arms, and to adjust for covariates that may be associated with calendar period and AIDS or death, I have employed methods previously adapted and modified by Cain et al(1) to estimate the effect of HAART on AIDS or death in a population of HIV infected children in Northern California.

4.3 Methods

4.3.1 Study Population

The Pediatric Spectrum of Disease (PSD) is a multicenter active surveillance program specifically for children who have been exposed to HIV perinatally.(27) Since 1988 this program has been located at Stanford University and has a surveillance catchment area of 12 counties in northern California with a total population of approximately 6 million. Through this program, I have identified and defined a population-based cohort of HIV positive northern Californian children.

Researchers working with the PSD database examine records from the California Children Services program, which provides case management services for HIV infected children, and medical records at hospital-based clinics. Study nurses visited pediatric HIV clinics biannually for data extraction from medical records and to identify new patients entering the PSD database. Medical records for all children under 18 years of age were followed until they were lost to follow up, died, or their status was definitely negative. Vertical transmission was determined by the CDC classification system for HIV in children younger than 13 years of age.(28; 29) An alphanumeric code combined with the birth date was used as a unique identifier to preserve confidentiality and avoid

record duplication. For the ongoing surveillance for the PSD database, institutional review board approval has been granted annually by the enrolling hospitals for the children and by Stanford University. For the present study, approval was obtained from the institutional review boards of Stanford University and University of California-Berkeley. In the present study, all N=113 children reached one of two possible endpoints—a C diagnosis (AIDS) or death, and were all assumed to be infected in utero.

4.3.2 Exposure Assessment

Antiretroviral therapy use is identified by information from medical record extracts. At each study visit, I defined the treatment received as either HAART or non-HAART and defined as a drug regimen of 2 nucleoside reverse transcriptase inhibitors *plus* a protease inhibitor or nonnucleotide reverse transcriptase inhibitor.

To estimate each child's person-time within calendar periods, I separated the calendar periods by pre-HAART and HAART eras. To explore the possibility of allowing enough time for complete HAART availability, two definitions of calendar year partitioning were explored: 1) pre-HAART (before 1997) and HAART (1997 and beyond); 2) pre-HAART (before 1998) and HAART (1998 and beyond). To act as a proxy for HAART use, an indicator variable for HAART calendar eras was created.

To help illustrate the appropriateness and applicability of this instrumental variable approach, Figure 4.2 shows the association of calendar period, Z, therapy use, X, the outcome of interest Y, measured covariates V, and unmeasured covariates U.
The properties of instrumental variables, as defined by Greenland,(13) are assumed to be satisfied. Again, calendar period: 1) is independent of variables that affect both HAART and outcome; 2) is associated with HAART; 3) is independent of the outcome given HAART and covariates that affect both HAART and the outcome. In Figure 4.2, the absence of a link between Z and U illustrates that Z cannot be affected by indications for treatment with HAART, satisfying condition 1. Additionally, condition 2 is supported by the link between Z and X resulting from antiretroviral therapies being introduced over time; calendar period is associated with HAART. Lastly, from previous published work by Detels et al,(7) condition 3 is satisfied as Z has been shown to be independent of AIDS given HAART treatment initiation indications and actual HAART use. This condition is represented by the absence of an arrow between Z and Y.

There remains the possibility that condition 3 may be too restrictive as there may be some covariates, V, that are related with both calendar period, Z, and the outcome, Y (see Figure 4.2). Potential examples of these variables are length of infection, age at seroconversion, or race/ethnicity. Therefore, I removed the arrow from V to Z in Figure 4.2 (e.g. removed the association between V and Z) by creating a weighted pseudo-population by using inverse probability weighting as described in Robins et al.(30) As a result, the new observations in this pseudo-population are now weighted by the inverse of the probability of calendar period given V.

4.3.3 Endpoint Assessment

The outcomes of interest were time from assumed HIV seroconversion to a category C diagnosis (AIDS) and/or death. While children vertically infected with HIV can seroconvert in utero, at delivery, or post-partum, I assumed all children were infected at birth. The presence of an AIDS-defining illness was determined by a clinician and noted within the child's medical records. Using the Centers for Disease Prevention and Control's guidelines for disease classification, children with one of the identified illnesses (e.g. C diagnosis) were determined as progressing to AIDS. See Appendix A.3 for details. Censoring occurred at three potential time points—time of AIDS diagnosis, time of death, or never attaining either outcome by the end of the study period in January 2009.

4.3.4 Statistical Methods

Using similar script found in Cain et al,(1) subscript i indexes the 1 to N=113 children, j indexes the 1 to J_i visits for each child i. The maximum number of visits was 58. HAART use is indexed by subscript x, where 1 is HAART use and 0 non-HAART use. When a child experiences the outcome, a diagnosis of an AIDS defining illness or death, the script $D_{ijxz} = 1$ is used to indicate that child i experienced the outcome between visits j and $j + 1$ during calendar period z while using therapy x. $D_{ijxz} = 0$ indicates that child did not experience the outcome. The number of person-days that each child i contributed between visits $j - 1$ and j while using therapy x during calendar period z is indicated by T_{ijxz}. Identified covariates for each child i at visit j, variables include both time-varying and time-fixed alike, are included in vector V_{ij}. Among the possible covariates included in V_{ij} are race/ethnicity, age at randomization, age (i.e. time since randomization).

Furthermore, $T_{xz}=\sum_{i=1}^{113}\sum_{j=1}^{Ji} T_{ijxz}$ is the total number of person-days contributed

while using therapy x during calendar period z summed over all children I and visits.

$D_{xz}=\sum_{i=1}^{113}\sum_{j=1}^{Ji} D_{ijxz}$ is the total number of events experienced while using therapy x during calendar period z summed over all children I and visits J. As described in Cain et al, let α_{xz} be the conditional probability of using therapy x given calendar period z, as estimated by the proportion of person-days while treated with therapy x during calendar period z. That is to say, $\alpha_{xz} = P(X = x| Z = z) = T_{xz}/T_{+z}$ where

$T_{+z} = \sum_{x=0}^{1} T_{xz}$.

In a traditional approach, a researcher could analyze the observational data with a standard ITT analysis—compare rates between calendar periods (before HAART and

HAART eras), regardless of actual HAART use. Specifically, the estimator for ITT of the average causal effect is:

$$\beta_{ITT} = \frac{\alpha_{10} \text{ X } (D_{10}/T_{10}) + \alpha_{00} \text{ X } (D_{00}/T_{00})}{\alpha_{11} \text{ X } (D_{11}/T_{11}) + \alpha_{01} \text{ X } (D_{01}/T_{01})} = \frac{(D_{+0}/T_{+0})}{(D_{+1}/T_{+1})}.$$

Akin to the noncompliance corrections proposed by Curzick et al for RCT data,(22) I will compare rates between calendar periods among those who would have used non-HAART in the era prior to HAART and those who would have used HAART in the era of HAART—therapy "compliers". Assuming that calendar periods are exchangeable and that calendar period is a valid instrument, the estimator is:

$$\beta_{IV} = \frac{\alpha_{00} \text{ X } (D_{00}/T_{00}) - \alpha_{01} \text{ X } (D_{01}/T_{01})}{\alpha_{11} \text{ X } (D_{11}/T_{11}) - \alpha_{10} \text{ X } (D_{10}/T_{10})} =$$

$$\frac{\beta_{ITT}}{\dfrac{[\alpha_{00} \text{ X } (D_{00}/T_{00}) - \alpha_{01} \text{ X } (D_{01}/T_{01})] \text{ X } [\alpha_{00} \text{ X } (D_{00}/T_{00}) + \alpha_{01} \text{ X } (D_{01}/T_{01})]}{[\alpha_{11} \text{ X } (D_{11}/T_{11}) - \alpha_{10} \text{ X } (D_{10}/T_{10})] \text{ X } [\alpha_{10} \text{ X } (D_{10}/T_{10}) + \alpha_{01} \text{ X } (D_{01}/T_{01})]}}$$

The ITT estimator divided by the estimator of the association between the exposure and the IV, as illustrated in the lower ratio, depicts a traditional β_{IV} analysis. In the absence of contamination or non-compliers, which would occur if no one contributes person-time to the non-HAART calendar period while using HAART and if no one contributes person-time to the HAART calendar period while not using HAART, then $\beta_{IV} = \beta_{ITT}$ and $\alpha_{10} = \alpha_{01} = 0$.

Although I duplicate precisely the analysis of Cain et al, I note that a different approach would be used if V_{ij} were on the causal pathway of calendar year. Inverse probability of calendar period weights, W_{ij+Z} , were estimated to adjust for measured confounders, V_{ij}. Specifically, we used stabilized standardized weights(1) $W_{ij+z} = P(Z=z)/P(Z=z|V_{ij} = v)$ for $i = 1$ to 113, $j = 1$ to J_i, where $\max(J_i) = 58$ and $z = 0$ or 1. By reintroducing the observed distribution of Z into the weights, maximum efficiency and stabilization of weights are realized.(31) This stabilization is accomplished by the numerator in the weights—an estimate of the probability of being in the same calendar period as what is observed. In contrast, the denominator represents the probability of being in the same calendar period as what is observed, given the covariates.

To select the covariates V_{ij} from a set of potentially influential variables associated with both AIDS/death and calendar period, I have employed super learner software -- Deletion/Substitution/Addition (D/S/A). By using this data-adaptive machine learning algorithm, and its cross-validation based on likelihood, I am avoiding the problems inherent with traditional approaches and model building. Specifically, D/S/A was used to

search through function forms using deletion, substitution, and addition actions. Sinisi and van der Laan have applied this algorithm to fit the initial hazard on pooled data over time.(32) The covariates selected by D/S/A from the candidate covariates in V for each instrumental variable analysis are identified within each analysis subsection. Similar to the covariates used by Cain et al, the pool of covariates D/S/A selected from included age (e.g. time since seroconversion), age at randomization, and race/ethnicity (defined as White or non-White).

The new weighted instrumental variable estimator of the causal rate ratio among compliers can be written as:

$$ {}_w\beta_{IV} = \frac{{}_w\alpha_{00} \text{ X } ({}_wD_{00}/{}_wT_{00}) - {}_w\alpha_{01} \text{ X } ({}_wD_{01}/{}_wT_{01})}{{}_w\alpha_{11} \text{ X } ({}_wD_{11}/{}_wT_{11}) - {}_w\alpha_{10} \text{ X } ({}_wD_{10}/{}_wT_{10})} $$

where ${}_w\alpha_{xz} = {}_wT_{xz}/{}_wT_{+z}$, ${}_wD_{xz}=\sum_{i=1}^{113}\sum_{j=1}^{Ji} D_{ijxz} \text{ X } W_{ijxz}$, and ${}_wT_{xz}=\sum_{i=1}^{113}\sum_{j=1}^{Ji} T_{ijxz} \text{ X } W_{ijxz}$.

Confidence intervals for unweighted ITT and IV estimates were calculated by formulas described by Rothman et al.(33) The 95% confidence intervals for the weighted ITT and IV estimates were estimated by bootstrap.

4.4 Results

To better understand the patient population, basic demographics and baseline characteristics are outlined in Table 4.1. The estimates are proportions from the main analysis which included children who either progressed to AIDS or died. The sample was mostly females (55.8%) of non-White ethnicity (69.0%) and at least 42 percent of mothers received prenatal care.

4.4.1 Instrumental Variable: Calendar Periods Pre-1997 and 1997-Beyond

AIDS or Death

The distribution of AIDS events or death, person-days, and rates by calendar period and HAART use are described in Table 4.2. Overall, 113 AIDS events or deaths occurred over 61,847 person-days. In the pre-HAART era there were no misclassified events, though a small proportion of person-time was misclassified as 2,501 of 42,195 (5.9%) person-days were observed while the participant was using HAART. During the HAART era, 12 out of 20 events (60.0%) and 10,437 out 19,652 person-days (53.1%)

were misclassified as the participants were not observed using HAART during this period. The rate of AIDS progression or death was estimated at 2.09 events per 1,000 person-days for the children in the non-HAART therapy group. For children in the HAART therapy group, the rate of AIDS progression or death was estimated at 0.68 events per 1,000 person-days. Overall, the rate of AIDS progression or death was 1.83 events per 1,000 person-days.

The distribution of events of AIDS or death and person-days by calendar period for the weighted and unweighted data is described in Table 4.3. When considering the unweighted data, the intent-to-treat rate ratio was estimated at $uRR_{ITT} = 2.17$ (95% CI 1.34-3.52) when comparing the pre-HAART era with the HAART era. The intent-to-treat rate difference was estimated at $uRD_{ITT} = 1.19$ (95% CI 0.56-1.82). The instrumental variable rate ratio using the unweighted data was estimated at $uRR_{IV} = 3.91$ (95% CI 2.41-6.34) when comparing the pre-HAART era with the HAART era.

When considering the weighted data, the super-learner, D/S/A, only selected one covariate to include in the most efficient vector of variables for V_{ij}--race/ethnicity. The weighted intent-to-treat rate ratio was estimated at $wRR_{ITT} = 2.14$ (95% CI 0.86–3.38) when comparing the pre-HAART era with the HAART era. The intent-to-treat rate difference was estimated at $wRD_{ITT} = 1.16$ (95% CI -0.04-2.21). The instrumental variable rate ratio using the weighted data was estimated at $wRR_{IV} = 3.84$ (95% CI 2.45-12.13) when comparing the pre-HAART era with the HAART era.

Adapted from Cain et al, Figure 4.3 helps illustrate how the calculation of the unweighted estimators in Table 4.3 was derived for β_{ITT} and β_{IV}. There are 3 divisions of person-days and events: 1) the first division shows the total number of events and total amount of person-days in the study sample; 2) the second division illustrates the how the events and person-time by calendar period were divided; 3) lastly, the third row shows within calendar period eras the use of HAART division. With the assumption that calendar period as defined is an appropriate instrument for HAART use, the last row depicts how events and person-days would have been classified in the cases of correctly classified children. Under any given calendar period $z = 0,1$ the possible therapy use is shown as x^z.

The person-days of last row of the tree diagram are calculated before the number of events. The conditional probability of HAART use (α_{xz}) of one calendar period is used to divide up the person-time in the other calendar period. For example, in the HAART era and HAART use groups, the 9,215 person-days are divided based on the conditional probabilities of non-HAART use (α_{00}) and HAART use (α_{01}) in the pre-HAART era group: 8,668.8 ~ (1-0.06) X 9,215 and 546.2 ~ 0.06 X 9,215, respectively. (Note: rounding error of α_{xz} does not provide for exact answers in this example.) Next, the events for one calendar period are subdivided such that the rate of those who always use HAART or the rate of those who never use HAART is equitable to the rate of non-compliance in the other calendar period. So, in my present example, the number of AIDS events or deaths in the group of users of HAART from the HAART period who would have used HAART had they been in the non-HAART era is selected such that their rate is

equitable to the rate in the group of users of HAART in the non-HAART era: 0 = (0/2,501) X 546.2. And lastly, I distributed the remaining events 8.0 = (8.0 - 0) as users of HAART from the HAART calendar era who would not have used HAART had they been in the non-HAART era.

To compare the unweighted data ITT estimates comparing pre-HAART calendar era with the HAART era from Figure 4.3 to Table 4.3, one would use the rates from the second level of Figure 4.3. In the present example, the estimates from Figure 4.3 and Table 4.3 are identical: 2.17=(93/42,195)/(20/19,652). Similarly, the unweighted estimate of the instrumental variable, an estimate of the complier-average causal effect, is calculated by using the rates from the bottom row of Figure 3 and is comparable to the instrumental variable estimate in Table 4.3. The complier-average causal effect can be interpreted as the rate ratio for the children in the pre-HAART era who were non-users of HAART but would have used HAART if they were in the HAART era, compared to the HAART-using children in the HAART era who would not have used HAART had they been members of the pre-HAART era: 3.68 = (71.6/21,081.1)/(8.0/8,668.8). In the present example, the uRR_{IV} from Table 4.3 is 3.84—a slightly higher estimate than one would have estimated from Figure 4.3. A likely artifact of rounding error, had the weighted number of HAART-using children who would not have used HAART had they been member of the pre-HAART era been estimated at 7.5 rather than 8.0, the RR_{IV} from both Table 4.3 and Figure 4.3 would have been identical.

AIDS Alone

The distribution of AIDS events, person-days, and rates by calendar period and HAART use are described in Table 4. Overall, 100 AIDS events occurred over 61,860 person-days. In the pre-HAART era there were no misclassified events, though a small proportion of person-time was misclassified as 2,501 of 42,208 (5.9%) person-days were observed while the participant was using HAART. During the HAART era, 8 out of 16 events (50.0%) and 10,437 out 19,652 person-days (53.1%) were misclassified as the participants were not observed using HAART during this period. The rate of AIDS progression was estimated at 1.83 events per 1,000 person-days for the children in the non-HAART therapy group. For children in the HAART therapy group, the rate of AIDS progression was estimated at 0.68 events per 1,000 person-days. Overall, the rate of AIDS progression was 1.62 events per 1,000 person-days.

The distribution of events of AIDS and person-days by calendar period for the weighted and unweighted data is described in Table 5. When considering the unweighted data, the intent-to-treat rate ratio was estimated at uRR_{ITT} = 2.44 (95% CI 1.44-4.18) when comparing the pre-HAART era with the HAART era. The intent-to-treat rate difference was estimated at uRD_{ITT} = 1.18 (95% CI 0.75-1.61). The instrumental variable rate ratio using the unweighted data was estimated at uRR_{IV} = 3.89 (95% CI 2.28-6.64) when comparing the pre-HAART era with the HAART era.

When considering the weighted data, the super-learner, D/S/A, only selected one covariate to include in the most efficient vector of variables for V_{ij}--race/ethnicity. The weighted intent-to-treat rate ratio was estimated at wRR_{ITT} = 2.39 (95% CI 0.91-5.40)

when comparing the pre-HAART era with the HAART era. The intent-to-treat rate difference was estimated at $wRR_{ITT} = 1.14$ (95% CI -0.17-2.19). The instrumental variable rate ratio using the weighted data was estimated at $uRR_{IV} = 3.79$ (95% CI 1.97-13.77) when comparing the pre-HAART era with the HAART era.

Figure 4.4 helps illustrate how the calculation of the unweighted estimators in Table 4.5 was derived for β_{ITT} and β_{IV}. The person-days of last row of the tree diagram are calculated before the number of events. The conditional probability of HAART use (α_{xz}) of one calendar period is used to divide up the person-time in the other calendar period. For example, in the HAART era and HAART use groups, the 9,215 person-days are divided based on the conditional probabilities of non-HAART use (α_{00}) and HAART use (α_{01}) in the pre-HAART era group: 8,669 ~ (1-0.06) X 9,215 and 546 ~ 0.06 X 9,215, respectively. (Note: rounding error of α_{xz} does not provide for exact answers in this example.) Next, the events for one calendar period are subdivided such that the rate of those who always use HAART or the rate of those who never use HAART is equitable to the rate of non-compliance in the other calendar period. So, in my present example, the number of AIDS events in the group of users of HAART from the HAART period who would have used HAART had they been in the non-HAART era is selected such that their rate is equitable to the rate in the group of users of HAART in the non-HAART era: 0 = (0/2,501) X 546.0. And lastly, I distributed the remaining events 8.0 = (8.0 - 0) as users of HAART from the HAART calendar era who would not have used HAART had they been in the non-HAART era.

To compare the unweighted data ITT estimates comparing pre-HAART calendar era with the HAART era from Figure 4.4 to Table 4.5, one would use the rates from the second level of Figure 4.4. In the present example, the estimates from Figure 4.4 and Table 4.6 are identical: 2.44=(84/42,208)/(16/19,652). Similarly, the unweighted estimate of the instrumental variable, an estimate of the complier-average causal effect, is calculated by using the rates from the bottom row of Figure 4.4 and is comparable to the instrumental variable estimate in Table 4.5. The complier-average causal effect can be interpreted as the rate ratio for the children in the pre-HAART era who were non-users of HAART but would have used HAART if they were in the HAART era, compared to the HAART-using children in the HAART era who would not have used HAART had they been members of the pre-HAART era: 3.58 = (69.73/21,088)/(8.0/8,669). In the present example, the RR_{IV} from Table 4.5 is 3.79—a slightly higher estimate than one would have estimated from Figure 4.4. A likely artifact of rounding error, had the weighted number of HAART-using children who would not have used HAART had they been member of the pre-HAART era been estimated at 7.5 rather than 8.0, the RR_{IV} from both Table 4.5 and Figure 4.4 would have been identical.

4.4.2 Instrumental Variable: Calendar Periods Pre-1998 and 1998-Beyond

AIDS or Death

The distribution of AIDS events or death, person-days, and rates by calendar period and HAART use are described in Table 4.6. Overall, 113 AIDS events or deaths occurred over 61,847 person-days. In the pre-HAART era there were only 2 misclassified events out of 97 (2.1%), while a small proportion of person-time was misclassified as 3,099 out of 45,000 (6.9%) person-days were observed while the participant was using HAART. During the HAART era, 10 out of 16 events (62.5%) and 8,230 out of 16,847 person-days (48.9%) were misclassified as the participants were not observed using HAART during this period. The rate of AIDS progression or death was estimated at 2.09 events per 1,000 person-days for the children in the non-HAART therapy group. For children in the HAART therapy group, the rate of AIDS progression or death was estimated at 0.68 events per 1,000 person-days. Overall, the rate of AIDS progression or death was 1.83 events per 1,000 person-days.

The distribution of events of AIDS or death and person-days by calendar period for the weighted and unweighted data is described in Table 7. When considering the unweighted data, the intent-to-treat rate ratio was estimated at $uRR_{ITT} = 2.27$ (95% CI 1.34-3.85) when comparing the pre-HAART era with the HAART era. The intent-to-treat rate difference was estimated at $uRD_{ITT} = 1.21$ (95% CI 0.58-1.84). The instrumental variable rate ratio using the unweighted data was estimated at $uRR_{IV} = 4.87$ (95% CI 2.87-8.26) when comparing the pre-HAART era with the HAART era.

When considering the weighted data, the super-learner, D/S/A, only selected one covariate to include in the most efficient vector of variables for V_{ij}--race/ethnicity. The weighted intent-to-treat rate ratio was estimated at $wRR_{ITT} = 2.25$ (95% CI 0.90-4.19) when comparing the pre-HAART era with the HAART era. The intent-to-treat rate difference was estimated at $wRD_{ITT} = 1.19$ (95% CI -0.27-1.95). The instrumental variable rate ratio using the weighted data was estimated at $wRR_{IV} = 4.83$ (95% CI 2.62-15.29) when comparing the pre-HAART era with the HAART era.

Figure 4.5 helps illustrate how the calculation of the unweighted estimators in Table 4.7 was derived for β_{ITT} and β_{IV}. The person-days of last row of the tree diagram are calculated before the number of events. The conditional probability of HAART use (α_{xz}) of one calendar period is used to divide up the person-time in the other calendar period. For example, in the HAART era and HAART use groups, the 8,617 person-days are divided based on the conditional probabilities of non-HAART use (α_{00}) and HAART use (α_{01}) in the pre-HAART era group: 8,014 ~ (1-0.07) X 8,617 and 603 ~ 0.07 X 8,617, respectively. (Note: rounding error of α_{xz} does not provide for exact answers in this example.) Next, the events for one calendar period are subdivided such that the rate of those who always use HAART or the rate of those who never use HAART is equitable to the rate of non-compliance in the other calendar period. So, in my present example, the

number of AIDS events or deaths in the group of users of HAART from the HAART period who would have used HAART had they been in the non-HAART era is selected such that their rate is equitable to the rate in the group of users of HAART in the non-HAART era: 0.38 = (2/3,099) X 593.4. And lastly, I distributed the remaining events 5.62 = (6 – 0.38) as users of HAART from the HAART calendar era who would not have used HAART had they been in the non-HAART era.

To compare the unweighted data ITT estimates comparing pre-HAART calendar era with the HAART era from Figure 4.5 to Table 4.7, one would use the rates from the second level of Figure 4.5. In the present example, the estimates from Figure 4.5 and Table 4.7 are identical: 2.27=(97/45,000)/(16/16,847). Similarly, the unweighted estimate of the instrumental variable, an estimate of the complier-average causal effect, is calculated by using the rates from the bottom row of Figure 4.5 and is comparable to the instrumental variable estimate in Table 4.7. The complier-average causal effect can be interpreted as the rate ratio for the children in the pre-HAART era who were non-users of HAART but would have used HAART if they were in the HAART era, compared to the HAART-using children in the HAART era who would not have used HAART had they been members of the pre-HAART era: 4.81 = (69.0/20,469.2)/(5.62/8,023.6). In the present example, the RR_{IV} from Table 4.7 is 4.87—a slightly higher estimate than one would have estimated from Figure 4.5—and is a likely artifact of rounding error.

AIDS Alone

The distribution of AIDS events, person-days, and rates by calendar period and HAART use are described in Table 4.8. Overall, 100 AIDS events occurred over 61,860 person-days. In the pre-HAART era there were 2 misclassified events out of 88 (2.3%), while a small proportion of person-time was misclassified as 3,099 of 45,013 (6.9%) person-days were observed while the participant was using HAART. During the HAART era, 6 out of 12 events (50.0%) and 8,230 out 16,847 person-days (48.9%) were misclassified as the participants were not observed using HAART during this period. The rate of AIDS progression was estimated at 1.83 events per 1,000 person-days for the children in the non-HAART therapy group. For children in the HAART therapy group, the rate of AIDS progression was estimated at 0.68 events per 1,000 person-days. Overall, the rate of AIDS progression was 1.62 events per 1,000 person-days.

The distribution of events of AIDS and person-days by calendar period for the weighted and unweighted data is described in Table 4.9. When considering the unweighted data, the intent-to-treat rate ratio was estimated at uRR_{ITT} = 2.74 (95% CI 1.50 – 5.01) when comparing the pre-HAART era with the HAART era. The intent-to-treat rate difference was estimated at uRD_{ITT} = 1.24 (95% CI 0.67-1.81). The instrumental variable rate ratio using the unweighted data was estimated at uRR_{IV} = 4.99 (95% CI 2.73 – 9.12) when comparing the pre-HAART era with the HAART era.

When considering the weighted data, the super-learner, D/S/A, only selected one covariate to include in the most efficient vector of variables for V_{ij}--race/ethnicity. The weighted intent-to-treat rate ratio was estimated at wRR_{ITT} = 2.71 (95% CI 1.11-6.45) when comparing the pre-HAART era with the HAART era. The intent-to-treat rate

difference was estimated at wRR_{ITT} = 1.22 (95% CI 0.16-2.13). The instrumental variable rate ratio using the weighted data was estimated at wRR_{IV} = 4.94 (95% CI 2.37-16.85) when comparing the pre-HAART era with the HAART era.

Figure 4.6 helps illustrate how the calculation of the unweighted estimators in Table 4.9 was derived for β_{ITT} and β_{IV}. The person-days of last row of the tree diagram are calculated before the number of events. The conditional probability of HAART use (α_{xz}) of one calendar period is used to divide up the person-time in the other calendar period. For example, in the HAART era and HAART use groups, the 8,617 person-days are divided based on the conditional probabilities of non-HAART use (α_{00}) and HAART use (α_{01}) in the pre-HAART era group: 8,014 ~ (1-0.07) X 8,617 and 603 ~ 0.07 X 8,617, respectively. (Note: rounding error of α_{xz} does not provide for exact answers in this example.) Next, the events for one calendar period are subdivided such that the rate of those who always use HAART or the rate of those who never use HAART is equitable to the rate of non-compliance in the other calendar period. So, in my present example, the number of AIDS events in the group of users of HAART from the HAART period who would have used HAART had they been in the non-HAART era is selected such that their rate is equitable to the rate in the group of users of HAART in the non-HAART era: 0.38 = (2/3,099) X 593.3. And lastly, I distributed the remaining events 5.62 = (6 – 0.38) as users of HAART from the HAART calendar era who would not have used HAART had they been in the non-HAART era.

To compare the unweighted data ITT estimates comparing pre-HAART calendar era with the HAART era from Figure 4.6 to Table 4.9, one would use the rates from the second level of Figure 4.6. In the present example, the estimates from Figure 4.6 and Table 4.9 are identical: 2.74=(88/45,013)/(12/16,847). Similarly, the unweighted estimate of the instrumental variable, an estimate of the complier-average causal effect, is calculated by using the rates from the bottom row of Figure 4.6 and is comparable to the instrumental variable estimate in Table 4.9. The complier-average causal effect can be interpreted as the rate ratio for the children in the pre-HAART era who were non-users of HAART but would have used HAART if they were in the HAART era, compared to the HAART-using children in the HAART era who would not have used HAART had they been members of the pre-HAART era: 4.91 = (70.37/20,475.6)/(5.62/8,023.7). In the present example, the RR_{IV} from Table 9 is 4.99—a slightly higher estimate than one would have estimated from Figure 4.6—and is a likely artifact of rounding error.

Table 4.10 summarizes the causal estimates derived from the two different instruments used, weighted and unweighted. Regardless of the definition of the HAART era and weighted or unweighted estimates, the effect of HAART on prevention of AIDS events in this population is just as strong as the effect of HAART on prevention AIDS events or deaths. By redefining the instrument as before 1998/1998 and after rather than before 1997/1997 and after, the unweighted and weighted estimates of the effect of HAART on prevention of AIDS events or death increases (uRR_{IV}=4.87 vs uRR_{IV}=3.91; wRR_{IV}=4.83 v wRR_{IV}=3.84, respectively). Similarly, the estimate of the effect of HAART on prevention of AIDS events alone increases when the instrument is redefined to the latter period (uRR_{IV}=4.99 vs uRR_{IV}=3.89; wRR_{IV}=4.94 v wRR_{IV}=3.79, respectively). When

the instrumental variable is defined as before 1997/1997 and after, the weighted RR_{IV} decreases slightly, suggesting that V_{ij} only mildly confounded the unweighted RR_{IV}. Similarly, when the instrumental variable is defined as before 1998/1998 and after, the weighted RR_{IV} decreases, again suggesting that V_{ij} mildly confounded the unweighted RR_{IV}. The complier average causal effect of HAART on AIDS events or deaths increases in the latter defined calendar period when compared to the earlier defined calendar period (complier causal $effect_{avg}$= 4.83 vs complier causal $effect_{avg}$= 3.70). This effect is similar in the estimate of the complier average causal effect of HAART on AIDS events alone (complier causal $effect_{avg}$= 4.91 vs complier causal $effect_{avg}$= 3.60).

4.5 Discussion

In the early-defined instrument—before 1997/1997 and after--analyses exploring the effect of HAART on AIDS events and deaths suggest that the ITT rate ratios are biased toward the null, likely a result from the substantial event and person-time exposure misclassifications (60.0% and 53.1%, respectively). Exposure to non-HAART increased the hazard of an AIDS event or death 3.91 times when compared to HAART exposure, using an instrumental variable estimator. This IV estimate is 80% (3.91/2.17) higher than the result one would see from a traditional ITT approach using calendar period. I weighted the number of events and amount of person-days by inverse probability of calendar period given race/ethnicity. In turn, the weighted rate ratios were estimated, which were expanded from basic instrumental variable methods,(22) after adjusting for measured covariates.

In the latter-defined instrument—before 1998/1998 and after--analyses exploring the effect of HAART on AIDS events and deaths suggest that the rate ratios are similarly biased toward the null, likely a result from the substantial event and person-time exposure misclassifications (62.5% and 48.9%, respectively). Exposure to non-HAART increased the hazard of an AIDS event or death 4.87 times when compared to HAART exposure, using an instrumental variable estimator. This IV estimate is more than twice (4.87/2.27) higher than the result one would see from a traditional ITT approach using calendar period.

Analyses exploring the effect of HAART on AIDS events alone using the early-defined instrument—before 1997/1997 and after--suggest that the rate ratios are biased toward the null, likely a result from the substantial event and person-time exposure misclassifications (50.0% and 53.1%, respectively). Exposure to non-HAART increased the hazard of an AIDS event or death 2.10 times when compared to HAART exposure, using an instrumental variable estimator. This IV estimate is 15% (2.10/1.82) higher than the result one would see from a traditional ITT approach using calendar period.

Analyses exploring the effect of HAART on AIDS events alone using the latter-defined instrument—before 1998/1998 and after--suggest that the rate ratios are biased toward the null, likely a result from the substantial event and person-time exposure misclassifications (62.5% and 48.9%, respectively). Exposure to non-HAART increased

the hazard of an AIDS event or death 4.44 times when compared to HAART exposure, using an instrumental variable estimator. This IV estimate is 68% (4.44/2.65) higher than the result one would see from a traditional ITT approach using calendar period.

In the present analysis, I assume that calendar period is an appropriate instrument for HAART use. As previously described, the three assumptions regarding instrumental variables are assumed satisfied in the present analysis. Specifically, conditional on controlled confounders (e.g. race/ethnicity), the instrumental variable should have the following characteristics: 1) is independent of variables that affect both HAART and outcome; 2) is associated with HAART; 3) is independent of the outcome given HAART and covariates that affect both HAART and the outcome.(13) The second principle has been repeatedly shown to be true.(2-12) Unfortunately, in our data the first and third principles are not testable. In order to relax the third principle such that the calendar period is presumably independent of AIDS events or death conditional on HAART exposure and adjusted confounders, the use of inverse probability of calendar period weights were used. This has been used previously to address this third principle.(1)

Furthermore, similar to Cuzick et al's assumptions,(22) this instrumental variable estimator analysis assumes exchangeability between calendar eras. For example, among the children who used HAART during the HAART era, my analysis assumes that, had these children been observed during the non-HAART era, the same proportion of their person-time would have been on HAART as for the children who in fact were observed in the pre-HAART era. Similarly, for children who were non-HAART users in the non-HAART era, had they been observed during the HAART era, the same proportion of their person-time would have been on non-HAART therapy as for the children who were in fact observed in the HAART era. This assumption is expanded such that the rate among HAART using children during the HAART era who would have used HAART had they been observed in the non-HAART era is equal to the rate among the HAART using children during the non-HAART era. Similarly, the rate among non-HAART using children during the non-HAART era who would have used non-HAART therapies had they been observed in the HAART era is equal to the rate among the non-HAART using children during the HAART era. It has been previously suggested that one of the few times this assumption could be violated is if a new, non-HAART HIV therapy that could decrease the risk of AIDS or death was introduced during the HAART era.(1; 7) In this unlikely scenario, time trends would affect calendar era comparability.

My results are contingent on the assumption that the model for weights has within it all possible determinants of calendar era and AIDS events or death. The algorithm D/S/A selected only race/ethnicity for the final model, ignoring age at seroconversion and time since seroconversion.

The impact of the choice of calendar period on the estimates is not negligible. In fact, when using the latter calendar period cut-off for the analysis of risk of AIDS events or death, the unweighted instrumental variable estimates are 25% higher than the unweighted instrumental variable estimates from the earlier calendar period (uRR_{IV}=4.87 and uRR_{IV}=3.91, respectively); and weighted instrumental estimates are 26% higher than

the estimates from the earlier calendar period cut-off. In contrast, the unweighted ITT estimates are 5% higher when using the 1998 cut-off rather than the 1997 cut-off; similarly, the weighted ITT estimates are 5% higher when using the 1998 cut-off rather than the 1997 cut-off. Similar differences were also seen with the analysis of risk of AIDS events alone. I chose 1997 as a cut-off for the first calendar period in the one analysis as HAART was only first introduced in mid-1996. To allow for the possibility that the use of HAART was not widely available until later, I explored 1998 as a cut-off in a second analysis. Cain et al performed a similar analysis using 1996 and 1998 as separate instrumental variable cut-offs.(1) They found the unweighted and weighted ITT estimates 3% and 8% higher when using the 1998 calendar year cut-off than using the 1996 calendar year cut-off. The researchers also found that the 1998 unweighted and weighted instrumental variables estimates were 5% and 4% higher than the 1996 instrumental variable estimates.

The current analysis only considered HAART or non-HAART therapy exposure and only 2 calendar periods. Future research using calendar periods as instrumental variables may benefit from using separate calendar periods representing different treatment eras: no therapy; monotherapy; dual therapy; and HAART. In fact, Detels et al performed a similar analysis in which therapy exposures were classified as monotherapy, combination therapy, and potent antiretroviral therapy groups. The calendar period in which protease inhibitors (a component of the most potent antiretroviral therapy groups) were introduced had the lowest relative hazard (RH=0.35; 95% CI 0.20-0.61) versus the comparison calendar period (1990-1983).(7) Perhaps expanding this instrumental variable estimator method to allow for multiple calendar years and multiple therapies is the next step.

Despite efforts to estimate the effect of HAART using an adapted instrumental variable approach, the present study still has limitations. Children in this population may have already been treated with HAART before entering the cohort if they moved from another state or region of California. If this occurred during the pre-HAART era, misclassification of therapy exposure would have increased. Additionally, there were only 8 events (i.e. AIDS events or death) among children treated with HAART, which could have increased the variability in the weighted estimates. The small number of events may have also influenced the complier average causal effect of HAART. In turn, the complier causal effects, which are estimated from the data in the figures representing each subsection, are not exactly equal to the effects that would have been obtained from the information from the tables alone. It is important to note, however, that rounding error is the likely explanation for the differences between the unweighted instrumental variable estimates and the complier causal effects. For example, α_{01} in Figure 3 is indicated as 0.06 when in fact the exact estimate is 0.055. In turn, rounded estimates influence the complier causal effects. In the early-defined calendar era, the estimates for the effect of HAART on the hazard of AIDS events or death are similar but not exact (e.g. uRR_{IV}=3.91 vs complier causal effect=3.70). In the latter-defined calendar era, the estimates for the effect of HAART on the hazard of AIDS events or death are much more similar (e.g. uRR_{IV}=4.87 vs complier causal effect=4.83).

If the instrument used is before 1998/1998 and after, the unweighted effect of HAART on prevention of AIDS events or death appears to be positively confounded by race/ethnicity. This is seen as the unweighted estimate of effect for HAART in this scenario is uRR_{IV}=4.87 (2.87, 8.26) and the weighted estimate of effect for HAART is wRR_{IV}=4.83 (2.62, 15.29) suggesting a slight overestimation of the true strength of association in the unweighted estimates.

This mild confounding is also noticed if the instrument used is before 1997/1997 and after--the unweighted effect of HAART on prevention of AIDS events or death appears to be positively confounded by race/ethnicity. This is seen as the unweighted estimate of effect for HAART in this scenario is uRR_{IV}=3.91 (2.41- 6.34) and the weighted estimate of effect for HAART is wRR_{IV}=3.84 (2.45-12.13) suggesting an underestimation of the true strength of association in the unweighted estimates.

Tables

Table 4.1. Patient Demographics and Baseline Characteristics for Main Analysis Including Children Who Progressed to AIDS or Died (N=113)

Baseline Covariate	N (%)
Male Sex	50 (44.2)
White Ethnicity	35 (31.0)
Mother had Prenatal Care	
Yes	48 (42.5)
No	14 (12.4)
Unknown	51 (45.1)

Table 4.2: Distribution of Events (AIDS events or deaths), Person-Days, and Rates by Calendar Period (Before 1997/1997 and After) and HAART Use

Calendar Period	No. of AIDS Events or Death	No. of Person Days	Rate
	Non-HAART Therapy		
Pre-HAART	93	39,694	2.34
HAART	12	10,437	1.15
Total	105	50,131	2.09
	HAART Therapy		
Pre-HAART	0	2,501	0.00
HAART	8	9,215	0.87
Total	8	11,716	0.68
	Total		
Pre-HAART	93	42,195	2.20
HAART	20	19,652	1.02
Total	113	61,847	1.83

Table 4.3: Distribution of Events (AIDS events or deaths) and Person-Days by Calendar Period (Before 1997/1997 and After)

Calendar Period	No. of AIDS Events or Death	No. of Person-Days	Rate	Intent To Treat				Instrumental Variable	
				Rate Difference	95% CI	Rate Ratio	95% CI	Rate Ratio	95% CI
				Unweighted					
Pre-HAART	93	42,195	2.20	1.19	(0.56, 1.82)	2.17	(1.34, 3.52)	3.91	(2.41, 6.34)
HAART	20	19,652	1.02	0		1		1	
Total	113	61,847	1.83						
				Weighted					
Pre-HAART	90.79	41,774	2.17	1.16	(0.04, 2.21)	2.14	(0.86, 3.38)	3.84	(2.45, 12.1)
HAART	15.44	15,173	1.02	0		1		1	
Total	106.2	56,947	1.87						

Table 4.4: Distribution of Events (AIDS events alone), Person-Days, and Rates by Calendar Period (Before 1997/1997 and After) and HAART Use

Calendar Period	No. of AIDS Events	No. of Person Days	Rate
	Non-HAART Therapy		
Pre-HAART	84	39,707	2.12
HAART	8	10,437	0.77
Total	92	50,144	1.83
	HAART Therapy		
Pre-HAART	0	2,501	0.00
HAART	8	9,215	0.87
Total	8	11,716	0.68
	Total		
Pre-HAART	84	42,208	1.99
HAART	16	19,652	0.81
Total	100	61,860	1.62

Table 4.5: Distribution of Events (AIDS events alone) and Person-Days by Calendar Period (Before 1997/1997 and After)

Calendar Period	No. of AIDS Events	No. of Person-Days	Rate	Intent To Treat				Instrumental Variable	
				Rate Difference	95% CI	Rate Ratio	95% CI	Rate Ratio	95% CI
				Unweighted					
Pre-HAART	84	42,208	1.99	1.18	(0.75, 1.61)	2.44	(1.44, 4.18)	3.89	(2.28, 6.64)
HAART	16	19,652	0.81	0		1		1	
Total	100	61,860	1.62						
				Weighted					
Pre-HAART	81.45	41,784	1.95	1.14	(-0.2, 2.2)	2.39	(0.91, 5.40)	3.79	(1.97, 13.8)
HAART	12.36	15,178	0.81	0		1		1	
Total	93.81	56,963	1.65						

Table 4.6: Distribution of Events (AIDS events or deaths), Person-Days, and Rates by Calendar Period (Before 1998/1998 and After) and HAART Use

Calendar Period	No. of AIDS Events or Death	No. of Person Days	Rate
	Non-HAART Therapy		
Pre-HAART	95	41,901	2.27
HAART	10	8,230	1.22
Total	105	50,131	2.09
	HAART Therapy		
Pre-HAART	2	3,099	0.65
HAART	6	8,617	0.70
Total	8	11,716	0.68
	Total		
Pre-HAART	97	45,000	2.16
HAART	16	16,847	0.95
Total	113	61,847	1.83

Table 4.7: Distribution of Events (AIDS events or deaths) and Person-Days by Calendar Period (Before 1998/1998 and After)

Calendar Period	No. of AIDS Events or Death	No. of Person-Days	Rate	Intent To Treat				Instrumental Variable	
				Rate Difference	95% CI	Rate Ratio	95% CI	Rate Ratio	95% CI
Unweighted									
Pre-HAART	97	45,000	2.16	1.21	(0.58, 1.84)	2.27	(1.34, 3.85)	4.87	(2.87, 8.26)
HAART	16	16,847	0.95	0		1		1	
Total	113	61,847	1.83						
Weighted									
Pre-HAART	95.71	44,746	2.14	1.19	(-0.3, 2.0)	2.25	(0.90, 4.19)	4.83	(2.62, 15.3)
HAART	12.35	13,007	0.95	0		1		1	
Total	108.1	57,753	1.87						

Table 4.8: Distribution of Events (AIDS events alone), Person-Days, and Rates by Calendar Period (Before 1998/1998 and After) and HAART Use

Calendar Period	No. of AIDS Events	No. of Person Days	Rate
Non-HAART Therapy			
Pre-HAART	86	41,914	2.05
HAART	6	8,230	0.73
Total	92	50,144	1.83
HAART Therapy			
Pre-HAART	2	3,099	0.64
HAART	6	8,617	0.70
Total	8	11,716	0.68
Total			
Pre-HAART	88	45,013	1.95
HAART	12	16,847	0.71
Total	100	61,860	1.62

Table 4.9: Distribution of Events (AIDS events alone) and Person-Days by Calendar Period (Before 1998/1998 and After)

Calendar Period	No. of AIDS Events	No. of Person-Days	Rate	Intent To Treat				Instrumental Variable	
				Rate Difference	95% CI	Rate Ratio	95% CI	Rate Ratio	95% CI
				Unweighted					
Pre-HAART	88	45,013	1.95	1.24	(0.67, 1.81)	2.74	(1.50, 5.01)	4.99	(2.73, 9.12)
HAART	12	16,847	0.71	0		1		1	
Total	100	61,860	1.62						
				Weighted					
Pre-HAART	86.55	44,758	1.93	1.22	(0.16, 2.13)	2.71	(1.11, 6.45)	4.94	(2.37, 16.9)
HAART	9.27	13,012	0.71	0		1		1	
Total	95.82	57,770	1.66						

Table 4.10: Comparison of Causal Estimates Derived from Different Instrumental Variables

AIDS Events or Deaths			
Before 1997/1997 and After		Before 1998/1998 and After	
Unweighted RR_{IV}	Weighted RR_{IV}	Unweighted RR_{IV}	Weighted RR_{IV}
3.91 (2.41- 6.34)	3.84 (2.45-12.13)	4.87 (2.87-8.26)	4.83 (2.62-15.29)
Complier Average Causal Effect of HAART[1]			
3.70		4.83	
AIDS Events Alone			
Before 1997/1997 and After		Before 1998/1998 and After	
Unweighted RR_{IV}	Weighted RR_{IV}	Unweighted RR_{IV}	Weighted RR_{IV}
3.89 (2.28- 6.64)	3.79 (1.97-13.77)	4.99 (2.73-9.12)	4.94 (2.37-16.85)
Complier Average Causal Effect of HAART[1]			
3.60		4.91	

1. Complier Average Causal Effect is defined as the rate ratio for children in the non-HAART era who did not use HAART but would have used HAART had they been in the HAART era, compared with the children in the HAART era who used HAART therapy but would not have used HAART had they been in the non-HAART era.

Figures

Figure 4.1. Traditional instrumental variable approach

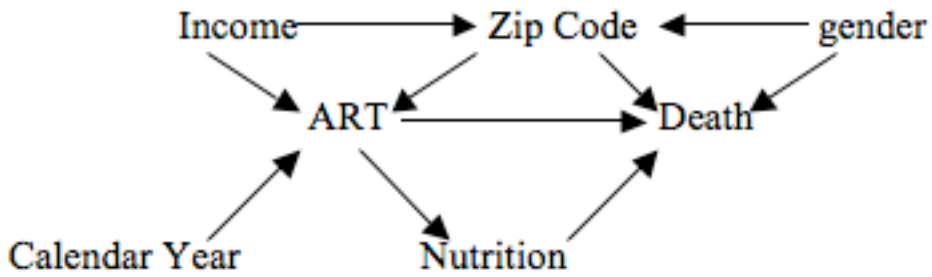


Figure 4.2: Adapted Instrumental Variable Approach

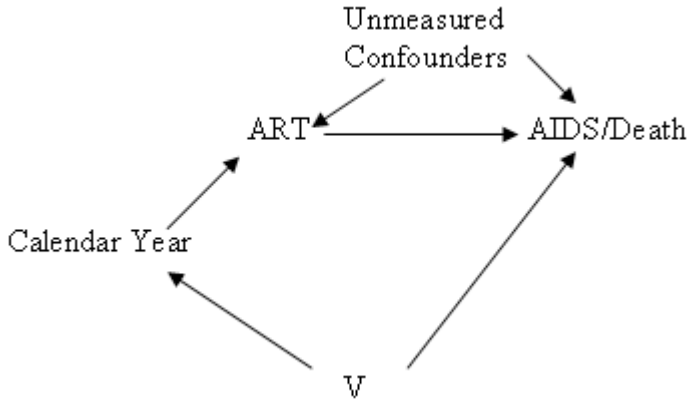


Figure 4.3: Tree Diagram Illustrating the Subdivision of AIDS Events or Deaths and Person-Days By: Calendar Period, z (Before 1997/1997 and After); HAART use, x; and Potential Therapy Use, x^z among PSD HIV-Infected Children.

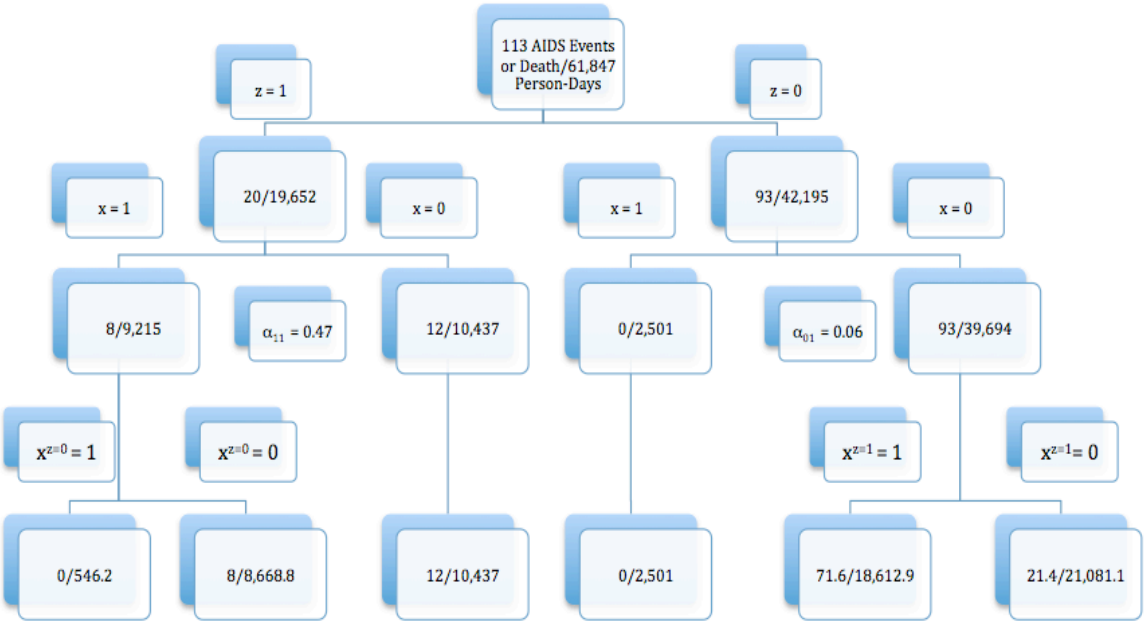


Figure 4.4: Tree Diagram Illustrating the Subdivision of AIDS Events and Person-Days By: Calendar Period, z (Before 1997/1997 and After); HAART use, x; and Potential Therapy Use, x^z among PSD HIV-Infected Children.

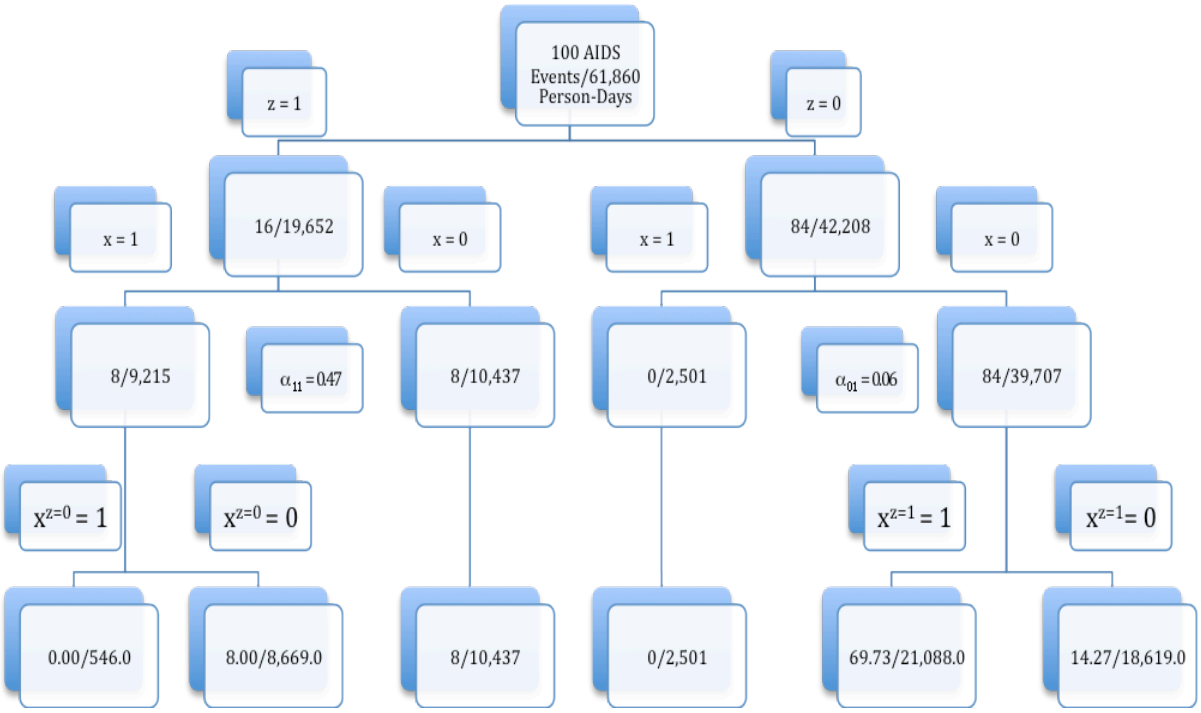


Figure 4.5: Tree Diagram Illustrating the Subdivision of AIDS Events or Deaths and Person-Days By: Calendar Period, z (Before 1998/1998 and After); HAART use, x; and Potential Therapy Use, x^z among PSD HIV-Infected Children.

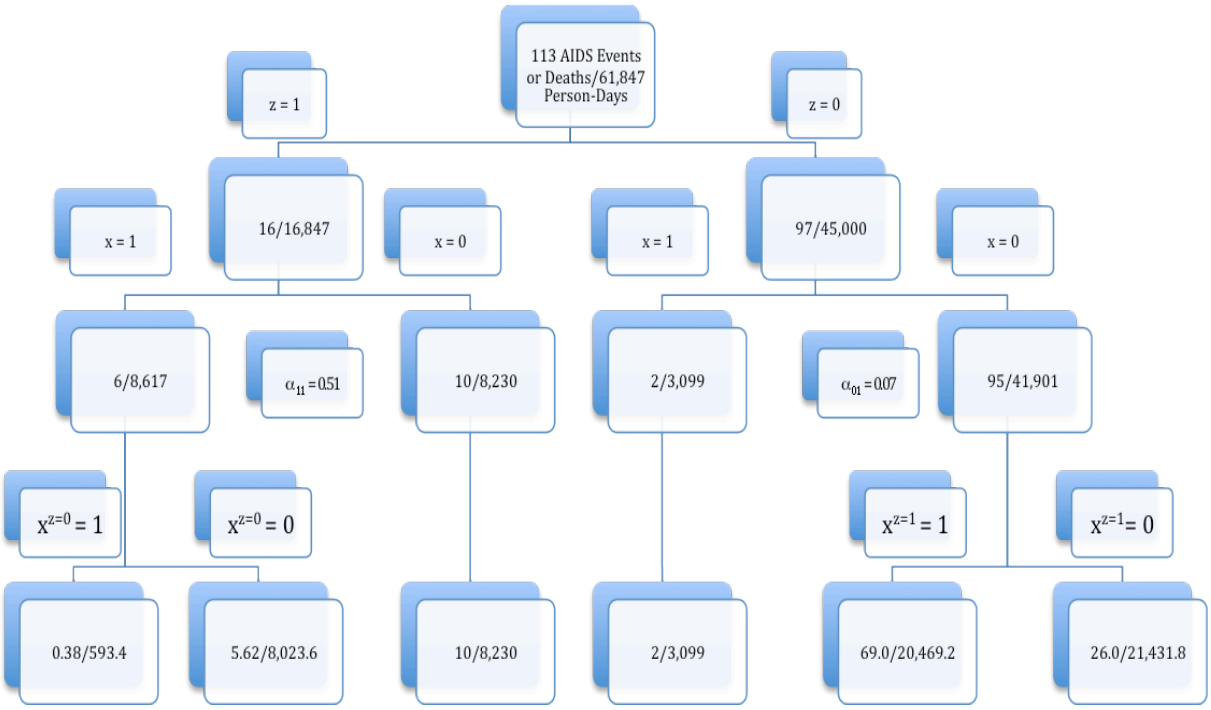
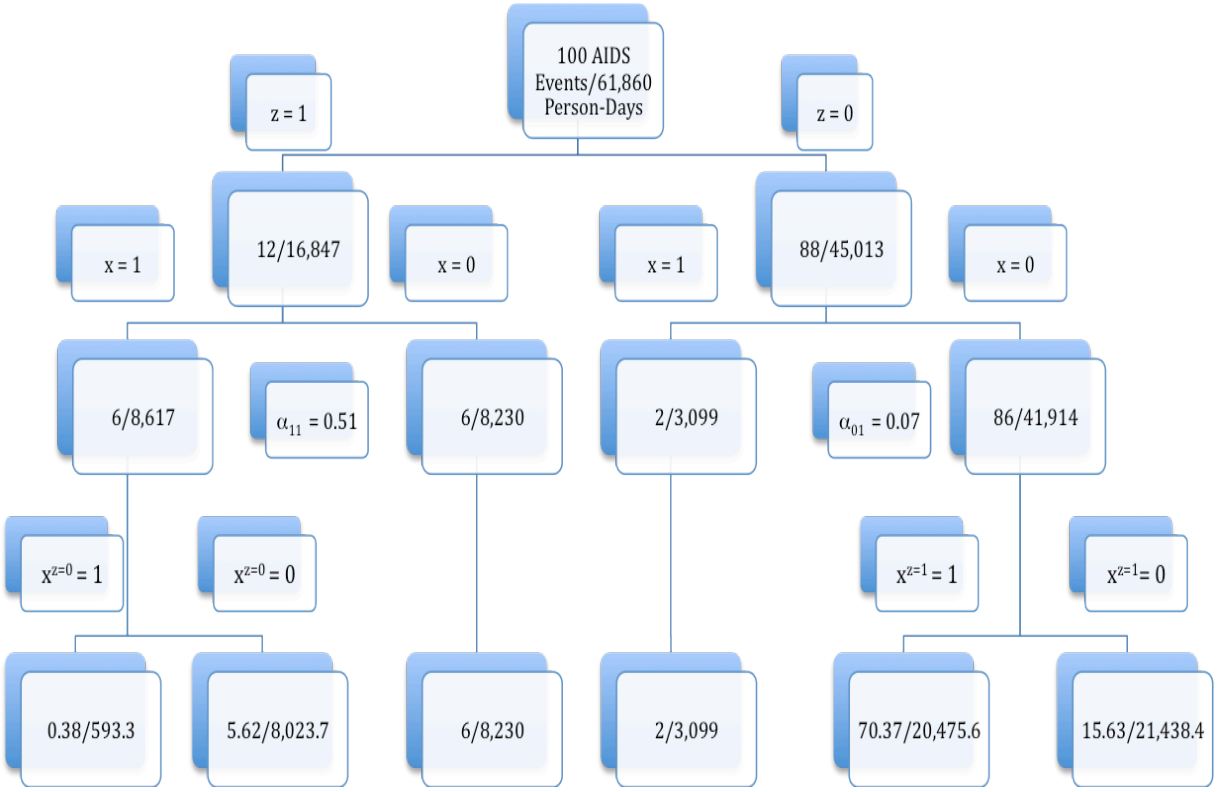


Figure 4.6: Tree Diagram Illustrating the Subdivision of AIDS Events and Person-Days By: Calendar Period, z (Before 1998/1998 and After); HAART use, x; and Potential Therapy Use, x^z among PSD HIV-Infected Children.



Bibliography

1. Cain LE, Cole SR, Greenland S, Brown TT, Chmiel JS, Kingsley L, Detels R. Effect of Highly Active Antiretroviral Therapy on Incident AIDS Using Calendar Period as an Instrumental Variable. American Journal of Epidemiology 2009;

2. Cain LE, Cole SR, Chmiel JS, Margolick JB, Rinaldo Jr CR, Detels R. Effect of highly active antiretroviral therapy on multiple AIDS-defining illnesses among male HIV seroconverters. American journal of epidemiology 2006;163(4):310.

3. Mocroft A, Ledergerber B, Katlama C, Kirk O, Reiss P, Monforte AA, Knysz B, Dietrich M, Phillips AN, Lundgren JD. Decline in the AIDS and death rates in the EuroSIDA study: an observational study. The Lancet 2003;362(9377):22–29.

4. Palella FJ, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, Aschman DJ, Holmberg SD. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. New England Journal of Medicine 1998;338(13):853.

5. Porter K, Babiker A, Bhaskaran K. Determinants of survival following HIV-1 seroconversion after the introduction of HAART. Lancet 2003;362(9392):1267-1274.

6. Gange SJ, Barron Y, Greenblatt RM, Anastos K, Minkoff H, Young M, Kovacs A, Cohen M, Meyer III WA, Munoz A. Effectiveness of highly active antiretroviral therapy among HIV-1 infected women. British Medical Journal 2002;56(2):153.

7. Detels R, Munoz A, McFarlane G, Kingsley LA, Margolick JB, Giorgi J, Schrager LK, Phair JP, others. Effectiveness of potent antiretroviral therapy on time to AIDS and death in men with known HIV infection duration. Jama 1998;280(17):1497.

8. Ahdieh-Grant L, Li R, Levine AM, Massad LS, Strickler HD, Minkoff H, Moxley M, Palefsky J, Sacks H, Burk RD, others. Highly active antiretroviral therapy and cervical squamous intraepithelial lesions in human immunodeficiency virus-positive women. JNCI Journal of the National Cancer Institute 2004;96(14):1070.

9. Tarwater PM, Mellors J, Gore ME, Margolick JB, Phair J, Detels R, Munoz A. Methods to assess population effectiveness of therapies in human immunodeficiency virus incident and prevalent cohorts. American journal of epidemiology 2001;154(7):675.

10. Moore RD, Chaisson RE. Natural history of HIV infection in the_era of combination antiretroviral therapy. Aids 1999;13(14):1933.

11. Schneider MF, Gange SJ, Williams CM, Anastos K, Greenblatt RM, Kingsley L, Detels R, Mu\ noz A. Patterns of the hazard of death after AIDS through the evolution of antiretroviral therapy: 1984-2004. Aids 2005;19(17):2009.

12. Tassie JM, Grabar S, Lancar R, Deloumeaux J, Bentata M, Costagliola D, others. Time to AIDS from 1992 to 1999 in HIV-1-infected subjects with known date of infection. JAIDS Journal of Acquired Immune Deficiency Syndromes 2002;30(1):81.

13. Greenland S. An introduction to instrumental variables for epidemiologists. International Journal of Epidemiology 2000;29(4):722-729.

14. Cole SR, Hernán MA, Robins JM, Anastos K, Chmiel J, Detels R, Ervin C, Feldman J, Greenblatt R, Kingsley L, others. Effect of highly active antiretroviral therapy on time to acquired immunodeficiency syndrome or death using marginal structural models. American journal of epidemiology 2003;158(7):687.

15. Hammer SM, Squires KE, Hughes MD, Grimes JM, Demeter LM, Currier JS, Eron JJ, Feinberg JE, Balfour HH, Deyton LR, others. A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. New England Journal of Medicine 1997;337(11):725.

16. Cameron DW, Heath-Chiozzi M, Danner S, Cohen C, Kravcik S, Maurath C, Sun E, Henry D, Rode R, Potthoff A, others. Randomised placebo-controlled trial of ritonavir in advanced HIV-1 disease. The Lancet 1998;351(9102):543–549.

17. Egger M, Hirschel B, Francioli P, Sudre P, Wirz M, Flepp M, Rickenbach M, Malinverni R, Vernazza P, Battegay M. Impact of new antiretroviral combination therapies in HIV infected patients in Switzerland: prospective multicentre study. Bmj 1997;315(7117):1194.

18. Markowitz M, Saag M, Powderly WG, Hurley AM, Hsu A, Valdes JM, Henry D, Sattler F, Marca AL, Leonard JM, others. A preliminary study of ritonavir, an inhibitor of HIV-1 protease, to treat HIV-1 infection. New England Journal of Medicine 1995;333(23):1534.

19. Gulick RM, Mellors JW, Havlir D, Eron JJ, Gonzalez C, McMahon D, Richman DD, Valentine FT, Jonas L, Meibohm A, others. Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. The New England journal of medicine 1997;337(11):734.

20. Nagelkerke N, Fidler V, Bernsen R. Estimating treatment effects in randomized clinical trials in the presence of non-compliance. STATISTICS IN MEDICINE 2000;19(14):1849-1864.

21. Cole SR, Chu H. Effect of acyclovir on herpetic ocular recurrence using a structural nested model. Contemporary clinical trials 2005;26(3):300–310.

22. Cuzick J, Edwards R, Segnan N. ADJUSTING FOR NON-COMPLIANCE AND CONTAMINATION IN RANDOMIZED CLINICAL TRIALS JACK CUZICK*, ROBERT EDWARDS AND NEREO SEGNAN 2. STATISTICS IN MEDICINE 1997;16:1017–1029.

23. Bond SJ, White IR, Sarah Walker A. Instrumental variables and interactions in the causal analysis of a complex clinical trial. Statist. Med. 2007;26(7):1473-1496.

24. Riersol E. Confluence analysis by means of lag moments and other methods of confluence analysis. Econometrica 1941;9:1-24.

25. Newhouse JP, McClellan M. Econometrics in outcomes research: the use of instrumental variables. Annual Review of Public Health 1998;19(1):17–34.

26. Miettinen O. The need for randomization in the study of intended effects. Statistics in Medicine 1983;2(2):267-271.

27. Maldonado Y, Wang N, Caldwell B. Factors associated with early clinical recognition of children with perinatal human immunodeficiency virus infection. Northern California Pediatric HIV Consortium. J Infect Dis 1995;171:689-692.

28. Centers for Disease Control and Prevention. Guidelines for the use of antiretroviral agents in pediatric HIV infection. [Internet]. [date unknown];Available from: www.aidsinfo.nih.gov

29. Virus HI, Addenda OA, Codes I. ICD-9-CM. [date unknown];

30. Robins JM, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000;11(5):550–560.

31. Hernan M, Brumback B, Robins J. Marginal Structural Models to estimate the joint causal effect of nonrandomized treatments. Journal of the American Statistical Association 2001;96(454):440-448.

32. Sinisi S, van der Laan MJ. Loss-based cross-validated deletion/substitution/addition algorithms in estimation. UC Berkeley Division of Biostatistics Working Paper Series 2004;

33. Rothman K, Greenland S. Modern Epidemiology. 2nd ed. Philadelphia, PA: Lippincott-Raven; 1998.

Chapter 5

Integrated Discussion

5.1 Synopsis of Results

Below I summarize the findings from Chapter 2, 3, and 4 and each chapter's main principles. Descriptive summaries and overarching conclusions about causal inference methods as they relate to HIV/AIDS data and their application to pediatric observational data are summarized from each chapter and in accompanying tables.

Chapter 2

The use of causal inference techniques has become increasingly popular over the last ten years. Observational HIV/AIDS data are a significant contribution to the understanding of treatment effects and can have public health implications. In Chapter 2 I briefly introduce four different causal inference techniques--propensity scores; instrumental variables, marginal structural models, and structural equation models—and describe the temporal trends in publications. Additionally, I performed a quality assessment and constructed network diagrams of the authors and institutions of the described techniques within HIV/AIDS settings. As the application of these techniques increases, the likelihood of the studies publishing comparative results with traditional techniques, or other details about study quality and interpretability is not increasing. However, there is a geographic shift of institutions publishing results as the earliest papers were published more at East Coast institutions and more recently it seems more are published at West Coast institutions. A summary of my goals and key conclusions from Chapter 2 is described in Table 5.1.

Table 5.1 Summary of Goals, Key Conclusions, and Page References from Chapter 2

Goals	Conclusions	Page References
Quantify the temporal trends in publications using causal inference techniques on HIV/AIDS data	Trends are increasing for all methods, but this is especially true for studies using MSMs.	30, 33, 40, 55
Identify the proportion of studies comparing causal inference results with traditional results	The majority of all studies compared results with traditional methods, though studies using SEMs made no comparisons.	32, 33, 60
Describe the proportion of studies that identify causal inference-specific assumptions	Only IV and MSM studies were likely to discuss assumptions, while studies using propensity scores or SEMs were less likely.	32, 33, 34, 60

(Continued on next page)

Table 5.1 Summary of Goals, Key Conclusions, and Page References from Chapter 2 (Continued)

Identify the proportion of studies that discuss the details of the instrument or treatment model selection	A large majority of all studies described the model or instrument selection. MSM publications were the least likely to publish details about model or instrument selection as only about half described the selection process.	32, 33, 34, 60
Identify the most commonly published authors using these techniques on HIV/AIDS data	Hernan MA has the most affiliated publications. Petersen M has the most first authorships out of all authors.	31, 32, 56, 57, 58
Identify the journals with the most publications using causal inference techniques.	AIDS, American Journal of Epidemiology, Journal of Acquired Immune Deficiency Syndromes, and Statistics in Medicine were the most common journals.	31, 32, 59
Describe how MSM studies derived the inference for estimates.	The majority of MSM studies described how they derived inference. The most common technique was bootstrapping.	35, 51
Identify the networks of institutions of affiliated authors for published studies.	The institutions with the most publications shifted more recently to other institutions in recent years for most causal inference methods.	35, 36, 61-66

Chapter 3

In Chapter 3, among a population-based cohort of HIV positive children the effect of triple therapy (e.g. HAART) on the time to AIDS or death was estimated using the MSM estimator, g-comp. Over all children, regardless of symptoms at treatment initiation, the effect of treatment in the first 6 or 12 months of life was estimated. Similarly, the effect was estimated for two subgroups—asymptomatically and symptomatically treated children. The goal of this analysis was to determine if the recently adopted treatment initiation guidelines for HIV positive children could be supported with observational study data. To mitigate the effects of unknown confounding and to avoid inherent problems with traditional model building techniques, I used a causal inference approach with a data-adaptive model selection procedure. A summary of my findings and goals is

described in Table 5.2. In general, though not statistically significant, the effect of HAART in the first 6 months of life was stronger than the effect of HAART in the first 12 months of life—suggesting that the current treatment guidelines, which dictate treatment as soon as the child's status is known, are supported with these data. For comparative purposes, traditional Cox proportional hazards models were used to estimate similar effects; generally the Cox estimates were only slightly more null-biased. When considering time to a C diagnosis, children symptomatically treated appear to benefit the most from early HAART as the estimated effect is stronger than among children asymptomatically treated. However, delaying treatment until symptoms may have a detrimental effect on a child's risk for death as asymptomatically treated children appear to have a lower mortality risk than symptomatically treated.

Table 5.2 Summary of Goals, Key Conclusions, and Page References from Chapter 3

Goals	Conclusions	Page References
Determine the effect of HAART in first 6 months of life on time to AIDS or death using g-comp	HAART in the first 6 months of life seemingly has a protective effect on time to AIDS or death.	90, 91, 92, 110, 111, 112
Determine the effect of HAART in first 12 months of life on time to AIDS or death using g-comp	HAART in the first 12 months of life has a protective effect on time to AIDS or death, but the effect is not as strong as the effect for children treated earlier in life.	90, 91, 92, 110, 111, 112
Determine if g-comp estimates are qualitatively different than traditional approaches	For estimates of the effect of therapy within the first 6 months of life, estimates from both approaches yield similar results.	100, 118
Describe the effects of HAART in first 6 or 12 months of life within subgroups of asymptomatic or symptomatic children on C diagnosis	Children symptomatically treated appear to benefit the most from early HAART as the estimated effect is stronger than among children asymptomatically treated.	93, 94, 95, 96, 112-115
Describe the effects of HAART in first 6 or 12 months of life within subgroups of asymptomatic or symptomatic children on C diagnosis or death	In contrast, children asymptomatically treated appear to benefit the most from early HAART as the estimated effect on C diagnosis or death is stronger than among children symptomatically treated.	96, 97, 98, 112-115

(Continued on next page)

Table 5.2 Summary of Goals, Key Conclusions, and Page References from Chapter 3 (continued)

Determine if treatment guidelines are supported by results	The recently changed treatment guidelines to initiate HAART early for all children regardless of symptoms appear to be supported by these results.	101

Chapter 4

In Chapter 4, I perform an adapted instrumental variable analysis of 267 HIV-positive children living in Northern California from 1988 to 2009 in order to estimate the causal effect of HAART on the hazard of AIDS events or death. I summaryize my goals and key conclusions from Chapter 4 in Table 5.3. I adjusted for noncompliance and used inverse probability weighting to remove any possible confounding from variables associated with both calendar year and the outcome. As a result of HAART use misclassification by calendar era, the instrumental variable estimator yielded a rate ratio 80% higher than the naïve result when using an earlier calendar year cut-off. The complier average causal effect of HAART, the rate ratio for children in the non-HAART era who did not use HAART but would have used HAART had they been in the HAART era compared with the children in the HAART era who used HAART but would not have used HAART had they been in the non-HAART era, was estimated at 3 to 4 depending on the selection of calendar year cut-off. Weighted estimates were not noticeably different than unweighted estimates. The decision of which calendar year cut-off, however, has an impact on both the unweighted and weighted estimates.

Table 5.3 Summary of Goals, Key Conclusions, and Page References from Chapter 4

Goals	Conclusions	Page References
Use traditional ITT approach on pediatric HIV observational data to assess population-level impact of HAART	Children in the pre-HAART era have a hazard of AIDS or death more than twice that of children in the HAART era.	145-151, 157-159
Determine if adjusting for noncompliance affects estimated effects	Without noncompliance adjustments, the effects are biased toward the null.	145-151, 157-159

(Continued on next page)

Table 5.3 Summary of Goals, Key Conclusions, and Page References from Chapter 4 (Continued)

Determine if covariates associated with calendar year and outcome have impact on IV estimates	The selected covariates had little impact on the IV estimator.	146, 147, 149-151, 157-159
Describe the impact of choice of calendar year cut-off.	The selection of a latter calendar period for a cut-off slightly increased the estimate sizes.	151-152, 157-160

5.2 Conclusions

Causal Inference Applications Using HIV/AIDS Observational Data

In this dissertation I have described the history of use of so-called causal inference methods applied to HIV/AIDS data and applied two techniques to describe the effect of therapy among HIV-positive children. Though temporal trends suggest that these techniques are being used more often than in previous years (see Table 2.3 and Figure 2.7), there is still a gap in their application with pediatric HIV/AIDS data as only 3 previous studies have employed any causal inference technique using pediatric data.

Ideally, HIV/AIDS researchers will attempt to minimize as much bias as possible in their study design and analysis. Unfortunately, as described in sections 2.4.6-2.4.9, the year of a study's publication does not seem to influence the likelihood that the authors were transparent about all their analysis methods. As causal inference methods appear in the HIV/AIDS epidemiological literature more frequently, the readership should be able to understand the basis for the technique used so that in turn the technique can be replicated in their own research.

Contributions to Pediatric HIV Research

The use of these techniques on pediatric data is of particular importance now because the long-term effects of HAART are not fully understood in a context of pediatric HIV populations. In turn, as HIV and health governing bodies consider mandating changes in treatment guidelines, they will need guidance and support for the proposed recommendations. Previously, treatment guidelines have been modified, opening the door for considerable criticism from evidence-based medicine advocates for a lack of evidence. In 2009, the National Institutes of Health updated their pediatric HIV

treatment initiation guidelines, but cited only 1 randomized controlled trial as the backbone of their supporting evidence.(1) Some health governing bodies, such as the WHO, have mandated new approaches for evidence-based medicine (EBM) to include using approaches like GRADE.(2) While this tool is useful in evaluating health care information,(3) critics have called into question its usefulness regarding rare disease studies(4) or observational data.(5) I have shown that the proposed guidelines to treat all HIV positive children as soon as their status is known is supported with these data using marginal structural models. In particular, I have shown that if the child waits to be treated until later in life (some time in the first 12 months as opposed to the first 6 months) or until he is symptomatic, the effect of HAART on time to a C diagnosis or death is somewhat nullified when compared earlier treatment (see Tables 3.13, 3.23, 3.26). Unfortunately, despite the use of causal inference techniques, most evidence grading approaches would likely not consider the evidence very strong at all as it is derived from observational study data.

Estimating the population level effects of HAART among an HIV positive pediatric population can yield important supporting evidence to implementing a more conservative national treatment initiation recommendation. My results in Chapter 4 give credence to the new "treat all infected infants" approach recommended by the CDC, WHO, and PENTA.(6) Specifically, children have a much lower risk for AIDS or death during the HAART era than children had during the pre-HAART era (see Tables 4.3 and 4.7). It is unlikely that this effect is explained by other HIV-related, non-HAART therapies or use of health care.(7)

Future Directions

The data used for Chapters 3 and 4 could also be used to explore the impact of treatment modification among HIV positive children. To date, there are no studies using causal inference methods that have considered treatment modification. Furthermore, an instrumental variable approach that considered sub-categories of treatment, such as monotherapy or dual therapy, may be a reasonable extension of the adapted instrumental variable approach used in Chapter 4.

Though not possible with the PSD dataset, research is currently underway to explore the possibility of interrupting ARV treatment in children after several years.(8) The theory is that since the immune systems are already reconstituted after several years of therapy, they will likely live a healthy life even without continued therapy. Though this treatment termination approach was explored among HIV positive adults with disastrous outcomes,(9) researchers believe that the developing immune system in children will provide better protection for the children.(8)

Final Remarks

As non-randomized studies will likely always outnumber clinical trial evidence, the experience and tools regarding causal inference techniques within this dissertation aim to provide direction for HIV researchers and investigators. Furthermore, the information contained within this project will likely inform researchers such that they can better

interpret the results of studies using causal inference techniques within their own scientific fields. Similarly, using the present examples as a backdrop, treatment guideline working groups may have a better understanding of the observational evidence supporting treatment recommendations about when and why to initiate HAART among HIV infected infants.

Bibliography

1. Working Group on Antiretrotrival Therapy and Medical Management of HIV Infected Children. Guidelines for the use of antiretroviral agents in pediatric HIV infection. National Institutes of Health 2009 Feb;:1-139.

2. Grade Working Group. Grading quality of evidence and strength of recommendations. British Medical Journal 2004;328:1490.

3. Haynes R. What kind of evidence is it that evidence based medicine advocates want health care providers and consumers to pay attention to? BMC Health Service Research 2002;2(1):3.

4. Dearlove O, Sharples A, O'Brien K. Evidence based medicine. Many questions cannot be answered by evidence based medicine. Bmj 1995;311(6999):257-8.

5. Cohen A, Stavri P, Hersh W. A categorization and analysis of the criticisms of evidence based medicine. International Journal of Medical Informatics 2004;73(1):35-43.

6. PENTA Steering Committee. PENTA 2009 guidelines for the use of antiretroviral therapy in paediatric HIV-1 infection. HIV Medicine 2009;10(10):591-613.

7. Detels R, Munoz A, McFarlane G, Kingsley LA, Margolick JB, Giorgi J, Schrager LK, Phair JP, others. Effectiveness of potent antiretroviral therapy on time to AIDS and death in men with known HIV infection duration. Jama 1998;280(17):1497.

8. Borkowsky W, Yogev R, Muresan P. Structured treatment interruptions in HIV-1 infected pediatric populations increases interferon gamma production and reduces viremia. Vaccine 2008;26(24):3086-3089.

9. El-Sadr WM, Lundgren JD, Neaton JD, Gordin F, Abrams D, Arduino RC, Babiker A, Burman W, Clumeck N, Cohen CJ, others. CD4+ count-guided interruption of antiretroviral treatment. New England Journal of Medicine 2006;355(22):2283–2296.

Appendix A

Additional Reference Materials

A.1 List of All Institutions and Their Abbreviations

Harvard; Massachusetts General Hospital (Mass.Gen); University of California Los Angeles (UCLA); Stanford; RAND; Johns Hopkins; University of California San Francisco (UCSF); Columbia; Calcutta National Medical College (Calcutta); Cook County Hospital (Cook County); Georgetown; Sant'Anna School of Advanced Studies, Pisa, Italy (Sant.Anna); Centro Operativo AIDS, Istituto Superiore di Sanita, Rome, Italy (Rome); St Paul's Hospital, Vancouver, BC, Canada (Saint. Pauls); University of British Columbia, Vancouver, BC, Canada (UBC); Centers for Disease Control and Prevention (CDC); Duke; Yale; Institut National de la Sante et de la Recherche Medical (INSERM); Vanderbilt; Thailand Ministry of Public Health (TMOPH); Phuket PH Office (Phuket); Nigerian Institute of Social and Economic Research (Nigeria); Bamrasnaradura Institute, Nonthaburi, Thailand (Bamrasnaradura); RTI, North Carolina (RTI); Northwestern; Brown; Cambridge; Chinese Academy of Sciences; University of Amsterdam; Ghent University (Ghent); CONRAD, Arlington, Virginia (CONRAD); University of Hong Kong; Queen Elizabeth Hospital; National Institute of Public Health, Oslo, Norway (NIPH Norway); University of Connecticut; Rhode Island College; Medical College of Wisconsin; University of North Carolina Chapel Hill (UNC Chapel Hill); National University of Singapore; Chinese University; University of Miami; Wayne State University; Uppsala University; Cung-Ang University; University of Pittsburgh; University of Colorado Denver; University of California Berkeley (UC Berkeley); Samsung; University of Washington; NY Academy of Medicine; University of East

Anglia; University of the Free State, SA; Escuela Valenciana de Estudios en Salud, Valencia, Spain (EVES, Spain); Instituto Nacional de Ciencias Medicas y Nutricion, Salvador Zubiran, Mexico (INCMN, Mexico); Lincoln Medical, NY; Kaiser; Universite de Lyon; Institut de Recherche pour le Developpement, Montpellier, France (IRD, France); University of Cape Town; Boston University (Boston); Ministry of Health, Mexico (MOH, Mexico); Beth Israel Hospital; Brandeis University ; National Development and Research Institutes, New York (National Development and Research Institutes); Utrecht University; Drug Policy Research Center; Montefiore Medical Center; University of Southern California (USC); National Institute of Cholera and Enteric Diseases Calcutta; STD/HIV Intervention Project Calcutta; Foundation for Innovative New Diagnostics Geneva; Veterans' Affairs, Connecticut (VA Connecticut); George Washington University; Baylor; Veterans' Affairs Houston, TX (VA Houston); Veterans' Affairs Palo Alto, California (VA Palo Alto); Comprehensive Care Center Nashville; Office of Disease Prevention and Control Thailand; Chiang Rai Provincial PH Office; Research Institute of Tuberculosis Tokyo; Bangkok Metropolitan Health Administrtation; National Institute of Allergy and Infectious Diseases; Medical Research Council Clinical Trials Unit London; Imperial College London; University of Arizona; University of Western Ontario; State University of New York Buffalo (SUNY Buffalo); University of Puerto Rico; University of Illinois Chicago; California School of Profession Psychology (CA School of Professional Psychology); Hunter College; City University of NY; Children's Hospital of Michigan; Innsbruck Medical University Austria; Karolinska University Hospital Sweden; University of Manchester; Denver PH Dept; Veterans' Affairs San Francisco (VA SF); State University of New York, Brooklyn (SUNY Brooklyn); Kenneth Norris Cancer Hospital LA; Howard University; Fred Hutchinson Cancer Research Center Seattle; The Polyclinic Seattle; Universitat d'Alacant Spain; Universidad Miguel Hernandez Spain; Makerere University Uganda; Chiang Mai University Thailand; University of Zimbabwe; Case Western Reserve University; Family Health International; Fann University Teaching Hospital Senegal; Military Hospital Senegal; Hopital Bichat-Claude Bernard France; National AIDS Program Senegal; University of Toronto; Childrens Hospital LA; Lusaka District Health Management Team Zambia; University of Zambia; Tulane University; University of Pennsylvania; University of Zurich; University of Bristol; Basel University; University of Berne; Swiss HIV Cohort Study; Cornell University; i3 Drug Safety; Naval Medical Center; Walter Reed Army Medical Center, Washington, DC (Walter Reed); Wilford Hall United States Air Force Medical Center, San Antonio, TX (Wilford Hall); National Naval Medical Center; Uniformed Services University of the Health Sciences, Bethesda, Maryland (Uniformed Services University); Institute of Clinical Infectious Diseases, Catholic University of the Sacred Heart, Rome, Italy (University of the Sacred Heart); Royal Free and University College Medical School, UCL, London, UK (University College Medical School UK); National Institute for Infectious Diseases 'Lazzaro Spallanzani', Rome, Italy; Institute of Infectious and Tropical Diseases, University of Milan, Italy (University of Milan); SS Annunziata Hospital, Taranto, Italy (SS Annunziata Hospital); Ospedali Riuniti, Foggia, Italy (Ospedali Riuniti); University La Sapienza, Rome, Italy (University La Sapienza).

A.2 WHO's Treatment Initiation Guidelines Among HIV+ Children (pre-2008)

Immunological marker	Age-specific recommendation to initiate ART			
	<= 11 months	12 months to 35 months	36 to 59 months	>= 5 years
%CD4+	<25%	<20%	<15%	<15%
CD4 Count	<1500 cells/mm^3	<750 cells/mm^3	<350 cells/mm^3	<200 cells/mm^3

A.3 Centers for Disease Prevention and Control Definition of Clinical HIV Disease Progression Among Children

	Clinical Categories			
Immunologic Categories	**N: No Signs/ Symptoms**	**A: Mild Signs/ Symptoms**	**B: Moderate Signs/ Symptoms**	**C: Severe Signs/ Symptoms**
1. No evidence of suppression <12 mo, ≥1,500 cells/µl ≥25% 1–5 y, ≥1,000 cells/µl ≥25% 6–12 y, ≥500 cells/µl ≥25%	N1	A1	B1	C1
2. Evidence of moderate suppression <12 mo, 750–1,499 cells/µl 15–24% 1–5 y, 500–999 cells/µl 15–24% 6–12 y, 200–499 cells/µl 15–24%	N2	A2	B2	C2
3. Severe suppression <12 mo, <750 cells/µl <15% 1–5 y, <500 cells/µl <15% 6–12 y, <200 cells/µl <15%	N3	A3	B3	C3

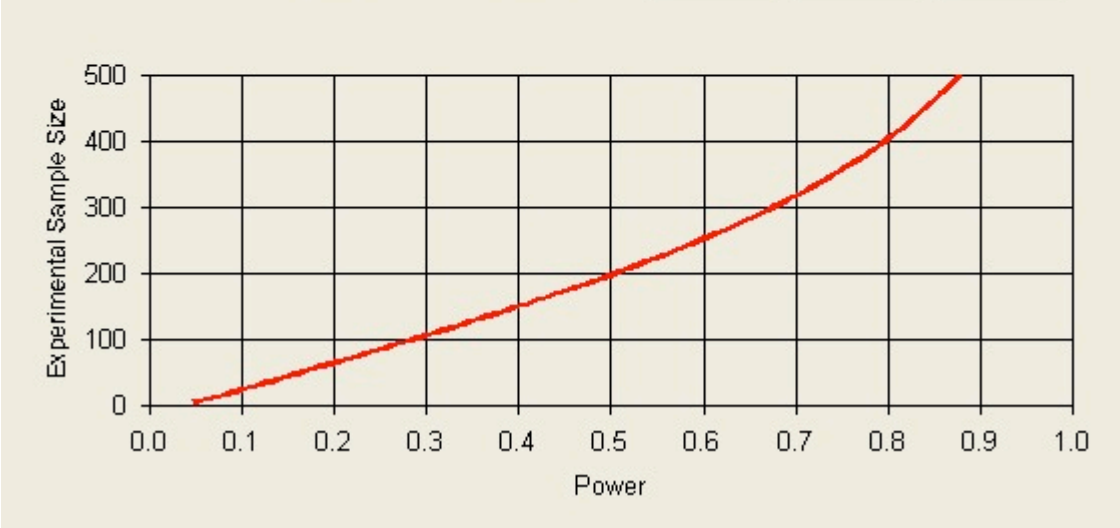
* If infection status is unknown but infant is exposed, use E (eg, EN2)

Modified from Centers for Disease Control and Prevention. 1994 revised classification system for human immunodeficiency virus infection in children less than 13 years of age. MMWR Morb Mortal Wkly Rep. 1994;43(No. RR-12):1–19.

A.4 PENTA's Treatment Initiation Guidelines For HIV+ Children (2002)

Infants	Children over 12 months of age
Always start if any: • Clinical stage C or • CD4 < 20% • Rapidly falling CD4% (irrespective of value), and for viral load persistently > 104 copes/ml	Always start ART if: • Clinical stage C or • CD4 < 15%
Consider ART if: Irrespective of clinical or immunological stage	Consider ART if: • Clinical stage B or • CD4 < 20% or • Viral load > 5 log
	Defer ART if: • Stage N or A disease and • CD4 > 20% and • Low viral load < 5 log

Appendix A.5: Power Calculation For Chapter 3 Analyses



Appendix B

Manuscripts Submitted for Publication

B.1 Review of Causal Inference Methods Used in HIV/AIDS Epidemiologic Studies

This review will be submitted for publication in the American Journal of Epidemiology in May 2010.

Review of Causal Inference Methods Used in HIV/AIDS Epidemiologic Studies

Word Count exclusive of abstract, tables, and figures: 3,955

Word Count of abstract: 199

Abstract:
In this review of causal inference methods used with HIV data--propensity scores, instrumental variables, marginal structural models, and structural equation models, we explored the temporal trends of their appearance in the literature, the frequency of publications by authors and journals, assessed the transparency and quality of the methods employed, and explored the networks of affiliated institutions. We included 70 studies that satisfied the eligibility criteria. Approximately 43% of all included studies were published in 2007 and 2008. Hernan has the most affiliated publications with 12 included studies; Petersen had the most first authorships with 5 publications. Studies using MSMs were more likely to relate causal inference results to those generated from more traditional methods than studies using any other causal inference method (OR=4.87; 95% CI 1.77-14.71). Studies using MSMs were also more likely to discuss specific model assumptions than other studies (OR=5.87; 95% CI 1.98-20.23). Among MSM studies published before 2007, the largest network was associated with Harvard University with 9 affiliated publications; the largest networks shifted to Johns Hopkins and University of California-San Francisco for MSM studies published in 2007/2008. Regardless of year of publication, all HIV studies are deficient by varying degrees in all assessed areas.

Key terms: Acquired Immunodeficiency Syndrome* HIV* propensity score* structural models*

To adequately prevent disease spread or treat patients with Human Immunodeficiency Virus (HIV) and Acquired Immune Deficiency Syndrome (AIDS), epidemiologists and biostatisticians are charged with finding the true impact of interventions or treatments. Unfortunately, the true effects are often obscured by nuisance factors and special analysis techniques are needed to address these factors. The analytic techniques that estimate a true causal relationship are referred to as causal inference methods. In this review of causal inference methods used with HIV/AIDS data, we will explore the temporal trends of their appearance in the literature, the frequency of publications by authors and journals, assess the transparency and quality of the methods employed, and explore the networks of affiliated institutions. We will review 4 causal inference methods that have relevance to addressing common data and statistical problems found in HIV/AIDS studies: propensity scores, instrumental variables, marginal structural models, and structural equation models.

The accepted gold standard for testing clinical effectiveness of treatments is the randomized controlled trial (RCT). However, often an RCT cannot be performed because of ethical or plausibility reasons and researchers are left with data from observational studies. Randomized controlled trials ensure in study design that person A's treatment allocation is not dependent on information connected to outcome Y, also known as the randomization assumption.(1) Essentially, this assumption ensures a lack of confounding.

Confounding, however, is not limited to baseline covariate or treatment distribution between two exposure groups; it can vary over time and may predict future treatment exposure. For example, antiretroviral therapy is often started in patients with worse baseline HIV disease (as measured by lower CD4+ counts); ART will have an impact on later HIV disease progression, which will later affect the probability of treatment.

One of the traditional methods for controlling the biases introduced by measured confounders is by using multivariable regression techniques. However, regression techniques are only as good as the measured confounder data.(2) Moreover, model building can become cumbersome even with comprehensive confounder data collected; often investigators will settle for model interpretability over adequate adjustment for bias.(2)

Time-dependent confounding is often addressed by using methods such as extended Cox regression models in which one would add an interaction term made up of the time-dependent variable of interest and some function of time.(3) Under two specific conditions, these conventional methods fail: C1) when there is a time-dependent risk factor of an event that also predicts subsequent treatment; and C2) when previous treatment history predicts a future risk factor.(4) These conditions are reintroduced later in the introduction of marginal structural models.

Propensity scores

Propensity scores (PS) were developed in 1983 by Rosenbaum and Rubin to control for known confounding bias and yet preserve model interpretability.(5) Sixteen years later,

McLaughlin et al published what is assumed to be the first HIV/AIDS study using propensity scores.(6) The propensity score is defined as a subject's conditional probability of treatment or exposure as opposed to another treatment or exposure, given the observed potential confounders and are described more thoroughly in Rosenbaum and Rubin.(2; 6)

Instrumental Variables

The use of IVs can be dated back over a half-century when they have been found in econometric theory.(7) Unlike the case with many epidemiologic studies, data found in economics is often sparse and lack randomization. To account for these disadvantages in their data economists developed IVs. The earliest application of IVs analyzing HIV/AIDS data and meeting our inclusion criteria was in 2001.(8) Like the principle of random allocation to treatments in RCTs, instrumental variables (IV) are variables that only affect the outcome through their effect on the treatment or exposure alone.(9)

Marginal Structural Models

A method proposed by Robins as early as 1997 and later emerging as a significant step forward in causal inference methods is the marginal structural model (MSM).(10) This method is unique in that despite the presence of time-dependent covariates, often simultaneously intermediate variables and confounders, it can estimate the causal effect of a time-dependent treatment.(11) By using inverse probability of treatment weighted (IPTW) estimators, the MSM parameters can be estimated, though G-computation and double robust are other appropriate methods, as well. Selecting the correct treatment model is imperative to ensure unbiased estimates. There are numerous ways of selecting models, including Monte Carlo cross-validation,(12) the use of Akaike Information Criterion (AIC), stepwise regression techniques, and learning programs like the algorithm Deletion/Substitution/Addition (DSA). Some options for producing confidence intervals for the MSM parameters include robust or sandwich estimators. Bootstrapping also helps to make a case for a probability-based inference about an effect based on an estimated effect using a population-based sample.(10; 13)

Structural Equation Models

Often used in psychology research, structural equation modeling (SEM) involves a network of independence assumptions and equations.(14) In this network of equations, each variable may only appear as a dependent variable once, but may appear in any equation as a causal variable. As such, the network of equations allows the researchers to see how each dependent variable changes as its causal variables change.

Materials and Methods

Literature Search and Study eligibility

We performed a systematic review of studies that used causal inference methods in the context of HIV/AIDS research. The initial search strategy collected all publications through December 2008 using Medline. Articles containing the textwords "propensity", or "instrumental AND variable", or "marginal structural model", or "structural equation model" and indexed to include "HIV" or "AIDS" textwords were selected with MeSH subject headings "Acquired Immunodeficiency Syndrome", "AIDS-Associated

Nephropathy", "AIDS Dementia Complex", "AIDS Serodiagnosis", "AIDS-Related Opportunistic Infections", "AIDS-Related Complex", "AIDS Vaccines", "HIV Seropositivity", "HIV Long-Term Survivors", "HIV", or "HIV Infections". Once the eligible articles were identified, a cross-reference search using Web of Science was performed. All studies citing the included Medline publications were included for the initial review. Following identification of eligible cross-referenced publications from Web of Science, a final search was performed on the bibliographies of methods-based researchers who are either first or senior authors of more than 2 eligible publications. These authors include: Cole S, Hernán M, van der Laan M, Petersen M, and Robins J. The eligibility criteria are summarized in Table 1. All analyses were performed using R.(15)

Qualitative Study Assessment

The interpretability of the study design to measure the causal effect of antiretroviral therapy and an HIV-related outcome was evaluated for each study based on the following 4 fields: 1) traditional interpretability; 2) discussion of statistical analysis-specific assumptions; 3) discussion of confounding measures; and 4) model or instrument selection. The rubric used to assess these fields is illustrated in Table 2.

Traditional Interpretability

To assess the interpretability of results from HIV/AIDS studies that employ causal inference methods, epidemiology and biostatistics literacy criteria were established. Our aim was to explore the frequency HIV/AIDS researchers compare their causal inference results to results they would have achieved had they used traditional methods. With studies using propensity scores, we were also exploring whether the researchers demonstrated the benefit of using propensity adjustment by showing the distribution of covariates before and after adjustment.

Statistical Analysis Assumptions

All statistical methods have important assumptions that should either be tested or at least discussed in papers using them, particularly if the methods are more contemporary or highly specific. Often studies will not list the methods-specific assumptions, but the researchers will rather acknowledge the assumptions through their discussion of methods or limitations.

Confounding Measures

As estimating the causal effect of x on y while minimizing bias is the main objective of causal inference methods, it is important for studies to identify specifically which covariates may bias this causal effect. A general discussion of confounder adjustment may not help future studies trying to expound on the study's results.

Model or Instrument Selection

It is not enough to simply employ these methods to control for confounding and reduce bias. The technique by which the researcher selects his model or instrument is just as important as the researcher recognizing *a priori* the necessity to perform the causal inference. In studies using propensity scores, a case for the specific variables used to

model the propensity of treatment is helpful to understand their influence on the estimated effect and control of bias. Similarly, the variables included in the treatment model for marginal structural models should not only be identified but also justified. Some of the model selection procedures may be as basic as an acknowledged prior knowledge or research, but may also include stepwise addition or deletion techniques, Akaike information criterion, or super learning applications like Deletion/Substitution/ Addition (DSA) algorithm. A discussion about the ways the authors selected the included variables (e.g. based on prior studies) is enough to satisfy this criterion. A justification for using a specific instrument in studies using instrumental variables is necessary to understand its influence, or lack thereof, on the estimated causal effect.

Results

Literature Search

The initial search on Medline, Web of Science, and selected bibliographies yielded 1535 potential studies, of which 932 were later removed based on publication date, duplication, title, or abstract details. The remaining 603 papers were reviewed for eligibility and 70 satisfied the eligibility criteria and were included in this review. The selection flow is detailed in Figure 1.

Temporal Trends

Though some of the included causal inference methods were developed by the onset of the HIV epidemic, the highest concentration of these methods in the HIV/AIDS literature was not apparent for about two decades. Propensity scores were introduced near the early stages of the HIV epidemic, however the first appearance of these methods in the HIV/AIDS literature was in 1999 (McLaughlin et al).(6) The trend in propensity score publications is markedly increasing as about 42% of all HIV studies using these methods were published in 2007 and 2008.(16-23) The appearance of IVs in HIV/AIDS literature occurred in 2001 (Tarwater et al).(8) Though only seven IV studies are known to have been published before 2009,(8; 9; 24-28) four of the seven were published in the last four years.(9; 24; 27; 28) In 2000 the first MSM study appeared in the literature (Hernan et al),(29) and every subsequent year saw an increase in publications using these methods on HIV/AIDS data. Approximately half (47%) of the studies using MSMs were published in 2007 and 2008.(24; 30-45) The use of SEMs on HIV/AIDS data has its origin in 1991 (Van der Velde et al).(46) Only one other publication using this method would appear in the 1990s,(47) while the majority (54.6%) was published since the end of 2005.(48-53) Approximately forty-three percent of all HIV/AIDS studies using the listed methods were published in 2007 and 2008. Temporal trends in the appearance of causal inference methods in HIV/AIDS publications are described in Table 3 and illustrated in Figure 2.

Causal Inference Publication Characteristics

Figure 3 is a dot-chart showing the frequency of causal inference publications analyzing HIV/AIDS data among all affiliated researchers with at least two included papers. Hernan has the most affiliated publications with 12 included studies; Cole had 10 and Robins had 9 associated publications. Petersen had the most first authorships with 5 publications, followed by Hernan with 4 and Cole with 3 publications (data not

displayed). About one in six (17.1%) of all included studies was principally authored by one of these three researchers. It should be noted, however, that all of the publications that were principally authored by these researchers employed marginal structural models. Robins had the most senior authorships with 7 publications, followed by van der Laan with 3 publications (data not displayed).

Over one-third (37.1%) of all the studies were published in *AIDS, American Journal of Epidemiology*, *Journal of Acquired Immune Deficiency Syndromes*, or *Statistics in Medicine* (see Figure 4). The most common exposure or outcome was antiretroviral therapy or HIV disease progression (64.3%).

Epidemiology of Studies

All studies using propensity scores had unique first authors and all but two (*AIDS*) were published in unique journals. Antiretroviral therapy or HIV disease progression were the most common exposure or outcome studied (47.4%). High-risk behaviors were the second most common exposure or outcome studied (21.1%) in publications using propensity scores (data not displayed).

All studies using IVs had unique first authors and all but two (*American Journal of Epidemiology*) were published in unique journals (data not displayed). Additionally, all but one of these studies explored the impact of ART on disease progression or high-risk behavior.

Among publications using MSMs, several researchers were the primary authors on more than one publication (data not displayed). Peterson was the first author of five HIV/AIDS studies using MSMs,(33-36; 43) Hernan was the primary author on four,(4; 29; 54; 55) and Cole(31; 56; 57) was the primary author for three studies. Lopez-Gatell,(32; 40) Brumback,(58; 59) and Patel(41; 42) were the primary authors for two studies each. Over a third (36.1%) of all HIV/AIDS studies using MSMs were published in *AIDS* or *American Journal of Epidemiology*. Antiretroviral therapy or HIV disease progression was the most common exposure or outcome studied (80.6%).

Studies using SEMs were authored by unique authors and generally published in different journals, though *AIDS and Behavior* had two publications (data not displayed). All publications using SEMs explored either ART adherence or high-risk behaviors.

Summary of Study Quality Assessment Results

Application of the quality assessment tool for these studies revealed that the most common weaknesses were traditional interpretability, a discussion of model or instrument selection, and a discussion of assumptions, though these results are highly method-dependent. The results are described in detail in Tables 4-7. Figure 5 illustrates the proportion of causal inference studies which satisfied the specific study assessment criteria.

Nearly two-thirds (63.0%) of all studies made comparisons between causal inference results and results using traditional methods; two-thirds (68.5%) of all HIV/AIDS studies

referenced any causal inference method-specific assumptions. Nearly all studies (95.9%) had a discussion about the type of confounding being controlled for using the causal inference method and listed the specific potential confounders. Over two-thirds (67.1%) of all studies discussed in detail the instrument or (treatment) model selection.

Over a third of all HIV/AIDS studies using propensity scores (36.8%) did not relate the causal inference results with results using traditional methods or failed to show the benefit of applying these techniques to the study data. Furthermore, propensity score studies were not likely to discuss any causal inference method-specific assumptions (42.1%).

Studies using IVs were likely to have causal effects compared to traditional methods (71.4%). Additionally, studies using IVs were just as likely to discuss the method-specific assumptions made as specific instrument selection (85.7%).

No studies using SEMs made comparisons of their results with results from traditional methods. Nearly half (45.5%) of these studies discussed their inherent assumptions. All studies using SEMs discussed their specific mediating variables and model selection.

Studies using MSMs were likely to have causal effects compared to traditional methods (80.6%). Additionally, studies using MSMs were likely to discuss the method-specific assumptions made (86.1%). Nearly all MSM studies discussed in detail the confounding variables they aimed to control (97.2%). Just over half of MSM studies (52.8%) stated the method by which the treatment model was selected.

Relationships between year of publication and temporal trends, traditional interpretability discussion of assumptions, discussion of confounding, and a discussion of treatment model/instrument selection were explored and no significant associations were noted.

Comparison of Studies Using MSMs Vs Any Other Causal Inference Method
Studies using MSMs are more likely to relate causal inference results to those generated from more traditional methods than studies using any other causal inference method (OR=4.87; 95% CI 1.77-14.71). Studies using MSMs are also more likely to discuss specific model assumptions than studies using any other causal inference method (OR=5.87; 95% CI 1.98-20.23). However, studies using MSMs are no more likely to discuss specific confounding or to discuss model selection than studies using any other causal inference method.

MSM-Specific Study Assessment
Among studies using MSMs, whether the authors stated specifically how they estimated their confidence intervals or standard errors for the causal effect of interest was investigated. Approximately eighty-six percent of all MSM studies reported their source of inference (data not dispayed). Specifically, approximately 22% used the bootstrapping method to estimate their standard errors or confidence intervals, 14% used the "sandwich" method, eight percent used generalized estimating equations, and 42% used a non-specific robust method.

There are three common methods of estimating the parameters in MSMs: G-comp, double robust, or IPTW. Though all three of these methods control for confounding, albeit in different ways, only two studies used any method other than IPTW to estimate the MSM parameters.(35; 36) Both of these studies used all three methods to compare results.

Network of Publishing Institutions
The network of institutions represented by all the listed authors is illustrated in Figures 6a-7b. Institution abbreviations are detailed in Appendix A.

Among MSM studies published before 2007 (Figure 6a), the largest network is associated with Harvard University with 9 affiliated publications. Among MSM studies published in 2007 or 2008 (Figure 6b), the largest networks were among Johns Hopkins and University of California-San Francisco each with 7 publications, followed by University of California-Berkeley with 6 publications. Among propensity scores studies published before 2007 (Figure 7a), the largest network is associated with University of California-Los Angeles with 4 publications, followed by Johns Hopkins University with 3 affiliated publications. The networks for propensity score papers published in 2007 or 2008 are shown in Figure 7b. The Centers for Disease Control and Prevention has the largest network with 3 affiliated publications among these publications. No networks of studies using IVs and SEMs had a network larger than 2 (data not shown).

Discussion
We have performed a systematic review of the HIV/AIDS literature for studies using causal inference methods including propensity scores, IVs, MSMs, and SEMs. We have found an increasing trend in appearance of most of these methods as over forty percent of studies using one of the listed methods were published in 2007 or 2008. Compared to all other methods, publications using MSMs had the highest proportion published in 2007 or 2008 (47.2%), followed by propensity score studies (42.1%). HIV/AIDS studies using IVs and SEMs have not seen the resurgence that other studies using other causal inference methods have.

The journals in which the studies were published may have some impact on the method's future use. Likely due to method-specific technical issues and readership, some methods are more often found in statistical or economics journals (e.g. IVs). Moreover, IVs have their origins in economics and have yet to be adopted as a common technique in epidemiology.(riersol)[38] Figure 4 illustrates the frequency of appearance of causal inference publications in specific journals. The highest frequencies are found in the journals *AIDS*, *American Journal of Epidemiology*, *Journal of Acquired Immune Deficiency Syndrome*, and *Statistics in Medicine*. As some of the most technical methods are published more often in journals such as *AIDS* and *American Journal of Epidemiology*, the readership may begin to employ the methods within their own research.

Some authors and affiliated institutions have contributed greatly to the dissemination of causal inference methods used for HIV data. As described in Figure 3, Hernan, Robins, and Cole have authored more HIV/AIDS studies using causal inference methods than any other researcher. As a result, their respective affiliated institutions, Johns Hopkins and Harvard University have some of the largest networks. It should be noted, however, that there was a geographical and institutional shift in network size for MSM studies from pre-2007 to 2007-2008 as Johns Hopkins, University of California-San Francisco, University of California-Berkeley, and University of California-Los Angeles had the largest networks most recently, while Harvard University was the most prolific producer of MSM studies prior to 2007.

The study assessments found that, regardless of year of publication, all HIV studies are deficient by varying degrees in traditional interpretability, assumption discussion, covariate and confounding discussion, and model and instrument selection discussion. Over all studies, traditional interpretability was the most common deficiency, but this is likely due to no study using structural equation models making any comparisons to traditional methods. Though not seen in our review, as a method has been used long enough, a traditional results comparison may be less important in the eyes of the researcher as limited space may be predicated on other results. Studies using IVs were most deficient in having a traditional interpretability component. While this is likely a reflection of the difficulty in making such comparisons with these methods, it remains an important aspect of promoting the use of a method.

Often some assumptions are not testable, but acknowledging them in the analysis should be policy to ensure study validity. Some assumptions, like the experimental treatment assignment (ETA) for MSMs, are often testable(60) and should be adequately described. Publications using propensity scores were most deficient in discussing inherent methods-specific assumptions.

The most deficient area for studies using MSMs was a discussion of their treatment model selection. Not only is this key for other researchers to understand fully how to implement MSMs, but it is also particularly important to ensure unbiased results. If incorrect covariates are used in building the treatment model, biased estimators are possible.(61) In particular, the inclusion of variables which predict only treatment, i.e. not confounders, can affect the estimator's performance. Regarding confounders and other covariates, most studies, regardless of causal inference method, described or listed the variables the method hoped to control for. In turn, future studies of similar research questions may be able to control for similar confounding.

Most studies using MSMs reported robust or sandwich confidence intervals. By explicitly stating the source of their confidence intervals, researchers once again have the opportunity for not only making a case for a true causal effect, but allow other researchers to learn from their methods.

Studies using MSMs are more likely to discuss their results as they relate to conventional methods than studies using any other causal inference method. Moreover, MSM-based

HIV publications are more likely to discuss the (treatment) model selection than the treatment or instrument model in any other causal inference method. Though in comparison to the most technically difficult methods MSMs are not terribly difficult to understand, perhaps the fact that most of the MSM studies make comparisons to traditional methods partly describes their increase in publication frequency.

There are several limitations concerning this review that should be considered. Firstly, though every attempt was made to capture all relevant studies, some may have been missed. A problem encountered when performing this review was that often studies will employ a method that is technically an included causal inference method, but the researchers fail to identify the method as specifically one of the included methods. We did not include these studies, as it would not have been reasonable to capture all of these studies using unidentified causal inference methods. In turn, we could have ended up with a skewed sample of only the studies with an identified causal inference method found on Medline and studies with an unidentified causal inference method found by our cross-referencing search. Secondly, as inherent in any review, we have to consider publication and author bias as a potential issue. If a prominent researcher is one of the authors on a causal inference paper, it may be more likely to be published. As the techniques and methods increase in complexity, many technical experts are required to collaborate. As a result, multiple institutions will likely have representation as either the problem necessitates or as the method becomes more recognized.(62) Thirdly, non-health related journals, like economics journals, might have more studies with HIV data. These results may not be generalizeable to studies of other diseases. In fact, the trends found in HIV studies may not be indicative of the trends among *all* epidemiologic studies. Some of the study assessments may not be as high as they could be because authors would reference previous work on the data in an attempt to avoid discussing technical details. Journal limitations may prevent authors from discussing further.

Tables

Table 1 Causal Inference Study Eligibility Criteria

1) Prospective or retrospective study published in peer-reviewed journal before calendar year 2009
2) At least 1 of the following endpoints was ascertained: i. Incident AIDS ii. Incident HIV iii. Stage of HIV disease iv. Response of HIV/AIDS to therapy v. HIV/AIDS disease progression vi. Death Or, behavioral or clinical risks were assessed within an exclusively HIV positive population. 3) One of the causal inference techniques described above explicitly stated and used in analysis (propensity scores, instrumental variables, MSM, structural equation model) 4) Results from application of method to dataset (not a subset of data for illustrative purposes) are published and have not been published previously elsewhere

Table 2: Quality Assessment Rubric Tool Applied To All Studies

Traditional Interpretability	**Is the method described as appropriate when compared to other methods? Or, are results from both causal inference and traditional methods given? (Alternatively, are before- and after- propensity score analyses done?)**
Discussion of Statistical Analysis-Specific Assumptions	**Are assumptions discussed generally and as they apply to the study data?**
Discussion of Confounding Measures	**Are the specific confounders the causal inference method aims to control or the type of (potential) confounding adjusted for discussed? This may also include mediating factors found in SEMs.**
Model and Instrument Selection	**Is the model or instrument selection technique discussed?**

Table 3 Trends of Causal Inference Methods Used in Studies of HIV and HIV Risk

Statistical method	Total	No. of articles	No. of articles	No. of articles	No. of articles	No. of articles
		Pre-00	**2001-02**	**2003-04**	**2005-06**	**2007-08**
Propensity scoring[±,1]	19 (26.0%)	1 (5.3%)	0 (0.0%)	3 (15.8%)	7 (36.8%)	8 (42.1%)
Instrumental Variables[±,2]	7 (9.6%)	0 (0.0%)	1 (14.3%)	2 (28.6%)	2 (28.6%)	2 (28.6%)
Marginal Structural Models[±]	36 (49.3%)	1 (2.8%)	3 (8.3%)	7 (19.4%)	8 (22.2%)	17 (47.2%)
Structural Equation Models[±]	11 (15.1%)	3 (27.3%)	0 (0.0%)	2 (18.2%)	2 (18.2%)	4 (36.4%)
Total[3]	73	5 (6.8%)	4 (5.5%)	14 (19.2%)	19 (26.0%)	31 (42.5%)

[±]Proportions reported are among studies using that statistical method

1. One of these publications is also included in Marginal Structural Models
2. Two of these publications are also included in Marginal Structural Models
3. Sum of total is more than sum of included studies due to papers using multiple methods

Table 4: Results of Study Quality Assessment of Propensity Scores Publications

Lead Author	Data Country of Origin	Traditional Interpretability	Assumptions Discussed	Confounding Or Mediating Variables Discussed	Model/ Instrument Selection Discussed
Zule(48) (2008)	US	Yes	No	Yes	Yes
Sanguan-wongse(41) (2008)	Thailand	Yes	Yes	Yes	Yes
Mahal(47) (2008)	Nigeria	No	Yes	Yes	Yes
Anuwatnon-thakate(46) (2008)	Thailand	Yes	Yes	Yes	Yes
Tai(45) (2007)	US	Yes	No	Yes	Yes
Potard(42) (2007)	France	No	Yes	Yes	Yes
Braithwaite(43) (2007)	US	Yes	No	Yes	No
Albalak(44) (2007)	US	No	No	No	No
Nosyk(73) (2006)	Canada	Yes	No	Yes	No
Merito(74) (2006)	Italy	Yes	No	Yes	Yes
Liu(76) (2006)	US	No	No	Yes	Yes
Liu(75) (2006)	US	Yes	Yes	Yes	Yes
Gangopadhyay(81) (2005)	India	No	No	Yes	No
El-Bassel(78) (2005)	US	No	Yes	Yes	Yes
Chu(77) (2005)	US	Yes	Yes	Yes	Yes
Wenzel(79) (2004)	US	Yes	No	Yes	Yes
Rotheram-Borus(80) (2003)	US	Yes	No	No	No
Brumback(40) (2003)	US	Yes	Yes	Yes	Yes
McLaughlin(32) (1999)	US	No	No	Yes	No

Table 5: Results of Study Quality Assessment of Instrumental Variable Publications

Lead Author	Data Country of Origin	Traditional Interpretability	Assumptions Discussed	Confounding Or Mediating Variables Discussed	Model/ Instrument Selection Discussed
Instrumental Variables					
Shiels(24) (2008)	US	No	Yes	Yes	Yes
Bond(6) (2007)	Europe	Yes	Yes	Yes	Yes
Lakdawalla(51) (2006)	US	Yes	Yes	Yes	Yes
Cain(52) (2006)	US	No	Yes	Yes	Yes
Hogan(50) (2004)	US	Yes	Yes	Yes	Yes
Bhattacharya(49) (2003)	US	Yes	Yes	Yes	Yes
Tarwater(34) (2001)	US	No	No	Yes	No

Table 6: Results of Study Quality Assessment of Marginal Structural Model Publications

Lead Author	Data Country of Origin	Traditional Interpretability	Assumptions Discussed	Confounding Or Mediating Variables Discussed	Model/ Instrument Selection Discussed
Marginal Structural Models					
Shiels(24) (2008)	US	No	No	Yes	No
Peterson(23) (2008)	US	No	Yes	Yes	Yes
Patel.2(21) (2008)	US	Yes	Yes	Yes	No
Patel.1(22) (2008)	US	Yes	Yes	Yes	No
Lopez-Gatell(20) (2008)	US	Yes	Yes	Yes	No
Fox(25) (2008)	Zambia	Yes	Yes	No	Yes
Fairall(19) (2008)	South Africa	Yes	Yes	Yes	No
Dolev(18) (2008)	US	No	No	Yes	No
De Beaudrap(17) (2008)	Senegal	No	Yes	Yes	Yes

Table 6: Results of Study Quality Assessment of Marginal Structural Model Publications (continued)

Lead Author	Data Country of Origin	Traditional Interpretability	Assumptions Discussed	Confounding Or Mediating Variables Discussed	Model/ Instrument Selection Discussed
Peterson.1(13) (2007)	US	Yes	Yes	Yes	Yes
Peterson.2(14) (2007)	US	No	Yes	Yes	Yes
Peterson.3(16) (2007)	US	Yes	Yes	Yes	Yes
Peterson.4(15) (2007)	US	Yes	Yes	Yes	Yes
Lopez-Gatell(12) (2007)	US	Yes	Yes	Yes	No
Cole(11) (2007)	US	Yes	Yes	Yes	No
Brown(10) (2007)	Zim-babwe/ Uganda	Yes	Yes	Yes	Yes
Perez(53) (2007)	Spain	Yes	Yes	Yes	No
Hogg(82) (2006)	Canada	Yes	Yes	Yes	Yes
Hernan(64) (2006)	France	Yes	Yes	Yes	Yes
De Luca(83) (2006)	Italy	Yes	No	Yes	No
Brookhart(89) (2006)	US	Yes	Yes	Yes	Yes
Bachmann(88) (2006)	South Africa	Yes	No	Yes	No
Wang(90) (2005)	US	Yes	Yes	Yes	Yes
Sterne(84) (2005)	Switzer-land	Yes	Yes	Yes	No
Cole(66) (2005)	US	Yes	Yes	Yes	Yes
Hogan(50) (2004)	US	Yes	Yes	Yes	Yes
Casper(85) (2004)	US	No	No	Yes	Yes
Brumback(67) (2004)	US	No	Yes	Yes	No
Barron(86) (2004)	US	Yes	Yes	Yes	No

Table 6: Results of Study Quality Assessment of Marginal Structural Model Publications (continued)

Lead Author	Data Country of Origin	Traditional Interpretability	Assumptions Discussed	Confounding Or Mediating Variables Discussed	Model/ Instrument Selection Discussed
Ko(87) (2003)	US	Yes	Yes	Yes	Yes
Cole(65) (2003)	US	Yes	Yes	Yes	Yes
Brumback(40) (2003)	US	Yes	Yes	Yes	Yes
Hernan(63) (2002)	US	Yes	Yes	Yes	No
Eisenberg(91) (2002)	US	Yes	Yes	Yes	Yes
Hernan(62) (2001)	US	Yes	Yes	Yes	No
Hernan(37) (2000)	US	Yes	Yes	Yes	Yes

Table 7: Results of Study Quality Assessment of Structural Equation Models Publications

Lead Author	Data Country of Origin	Traditional Interpretability	Assumptions Discussed	Confounding Or Mediating Variables Discussed	Model/ Instrument Selection Discussed
Structural Equation Models					
Rice(61) (2008)	US	No	No	Yes	Yes
Cha(60) (2008)	US	No	No	Yes	Yes
Bull(59) (2008)	US	No	Yes	Yes	Yes
Sodergard(58) (2007)	Sweden	No	No	Yes	Yes
Naar-King(57) (2006)	US	No	No	Yes	Yes
Llabre(56) (2006)	US	No	Yes	Yes	Yes
Prado(92) (2004)	US	No	No	Yes	Yes

Table 7: Results of Study Quality Assessment of Structural Equation Models Publications (continued)

Lead Author	Data Country of Origin	Traditional Interpretability	Assumptions Discussed	Confounding Or Mediating Variables Discussed	Model/ Instrument Selection Discussed
Structural Equation Models					
Lim(93) (2003)	Singapore	No	No	Yes	Yes
Sengupta(94) (2000)	US	No	Yes	Yes	Yes
Kraft(55) (1995)	Norway	No	Yes	Yes	Yes
Van der Velde(54) (1991)	Nether-lands	No	Yes	Yes	Yes

Figures

Figure 1. ***Study Selection Flow***

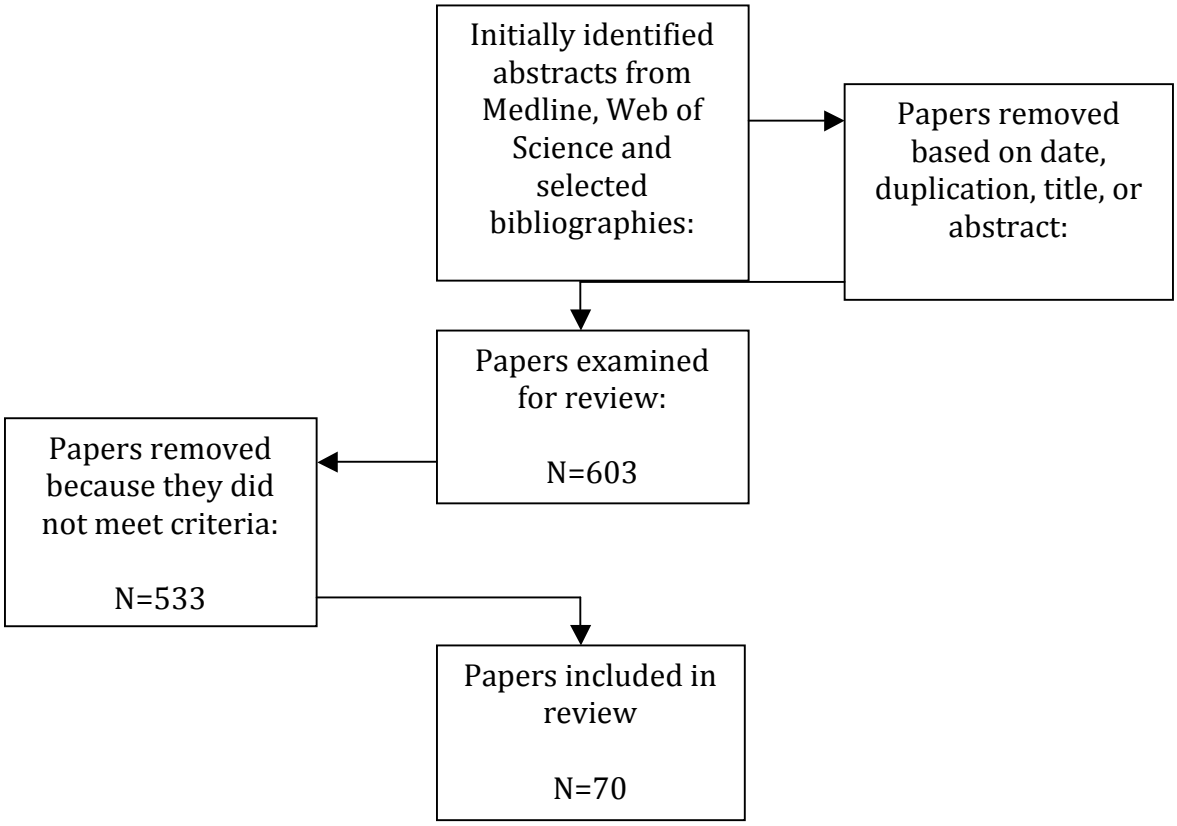


Figure 2. Histograms Showing Temporal Trends in Appearance of Causal Inference Methods in HIV/AIDS Publications

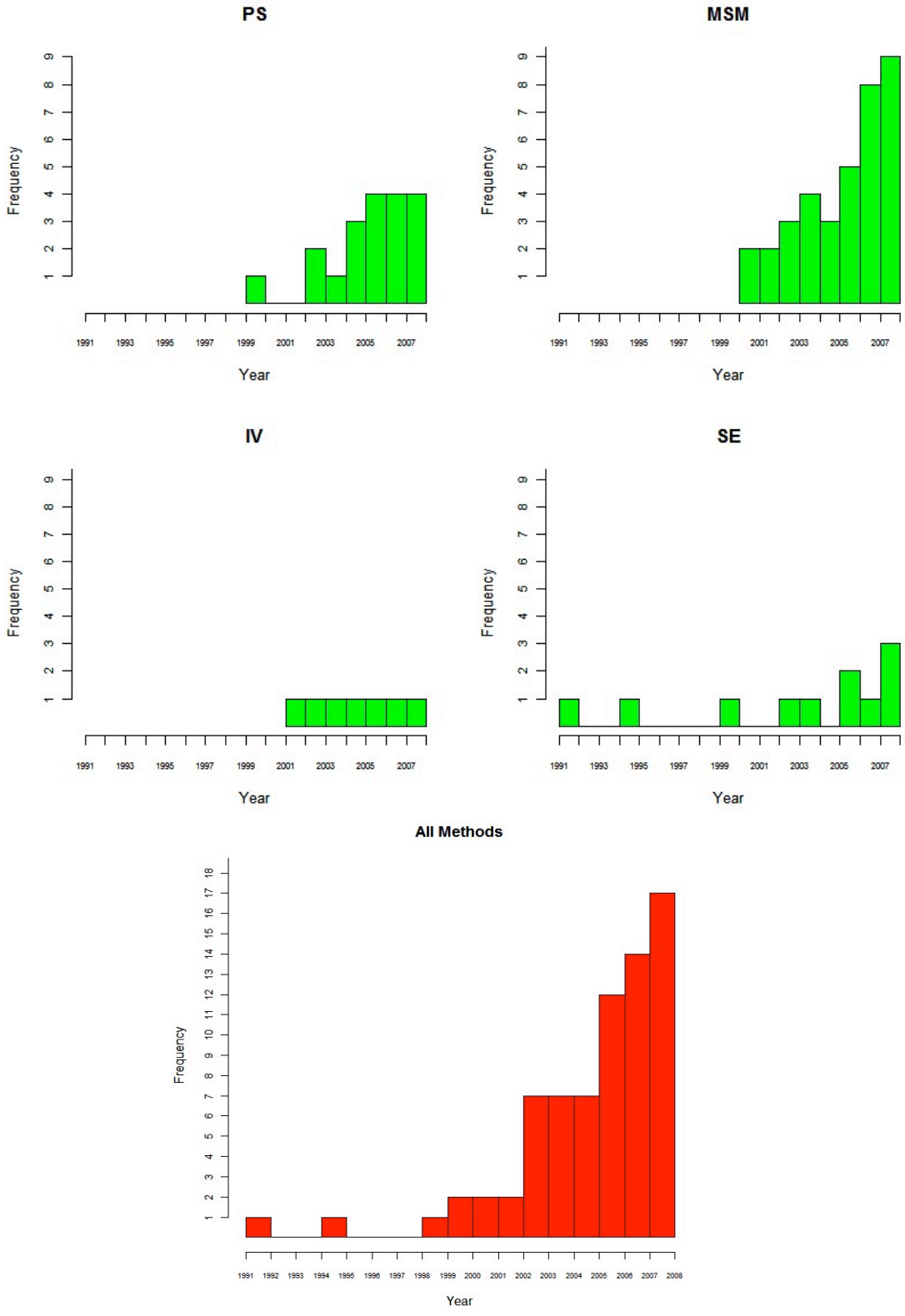


Figure 3 Dot Chart of Frequency of Publications (with a minimum of 2) By All Associated Authors Using Causal Inference Methods With HIV/AIDS

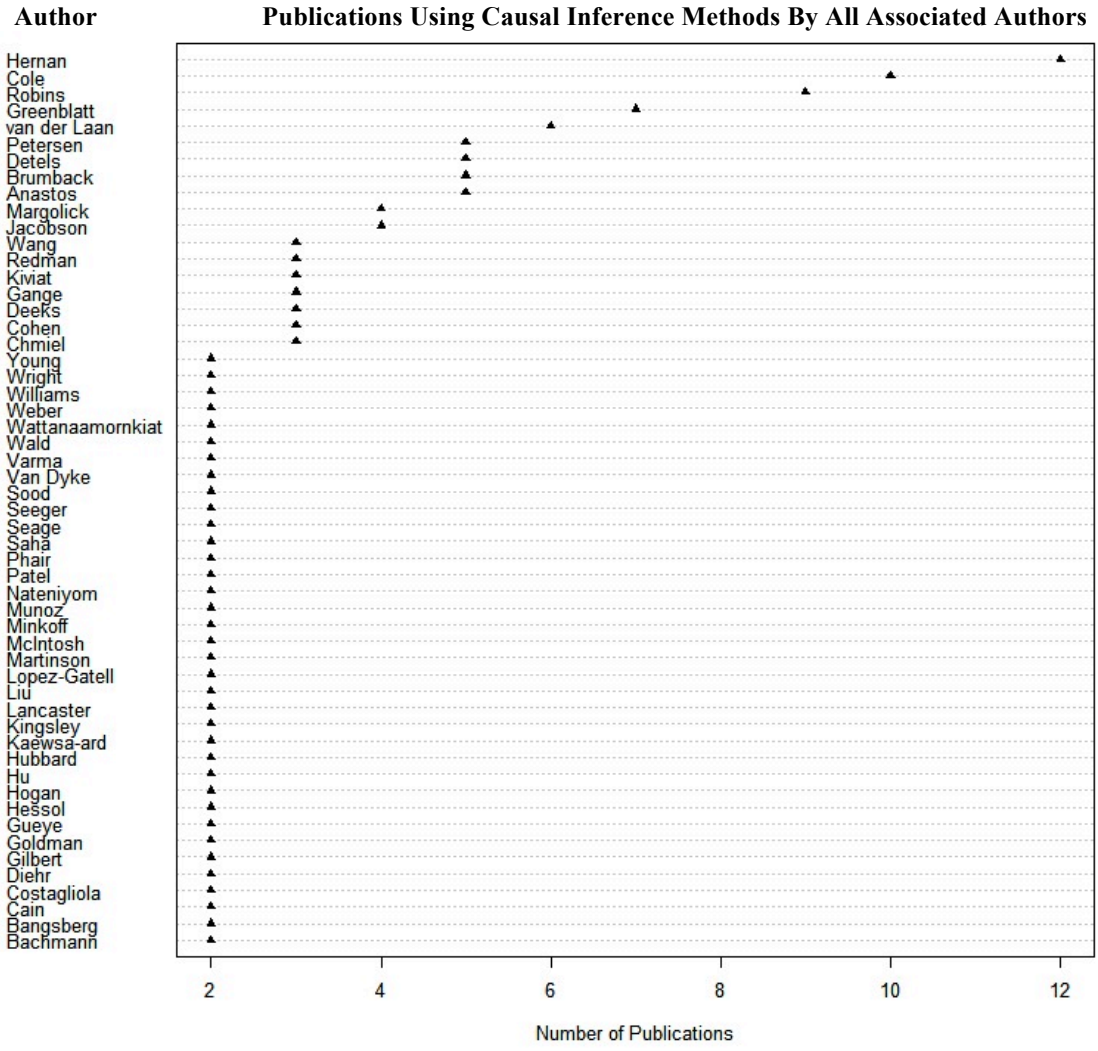


Figure 4. Dot Chart of Frequency of Appearance of Publications Using Causal Inference Methods With HIV/AIDS Data By Journal

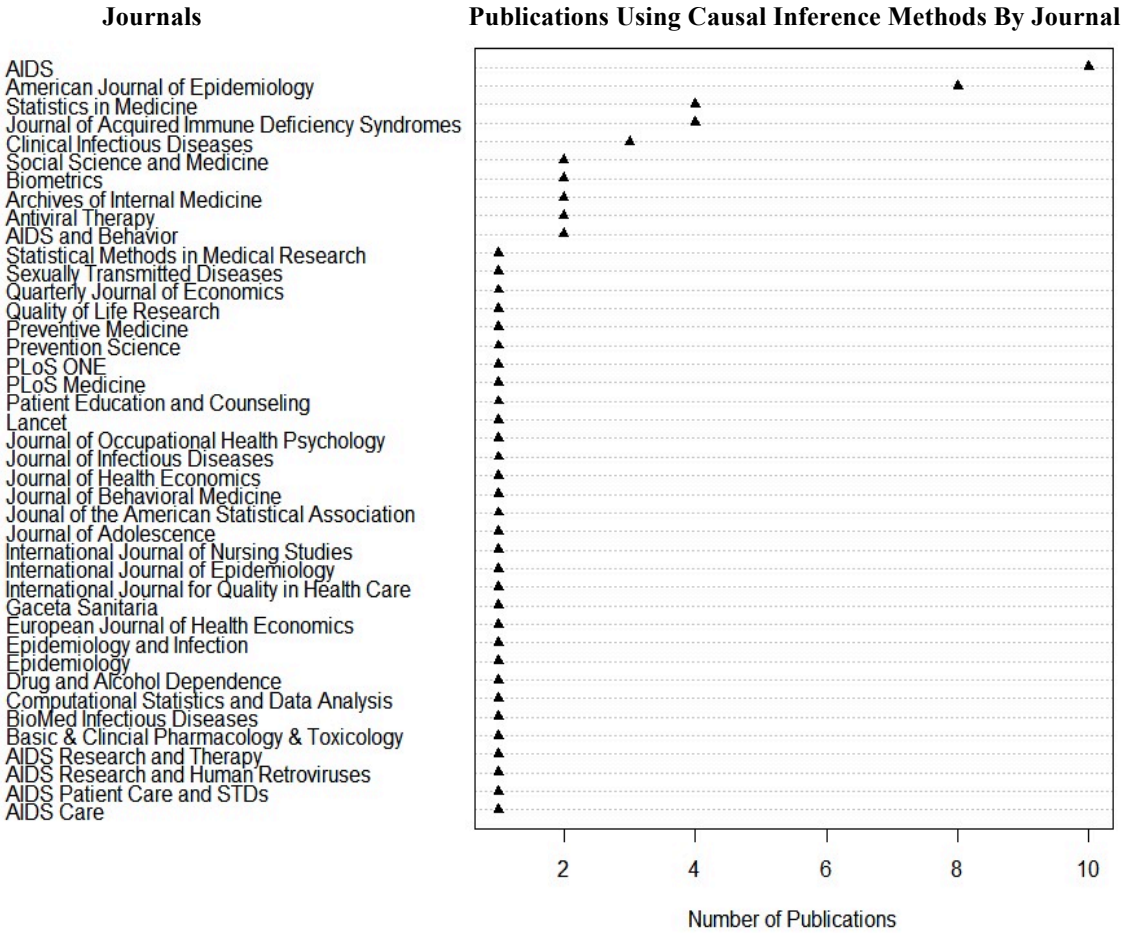


Figure 5. Bar-plots of Results of Study Quality Assessments By Method

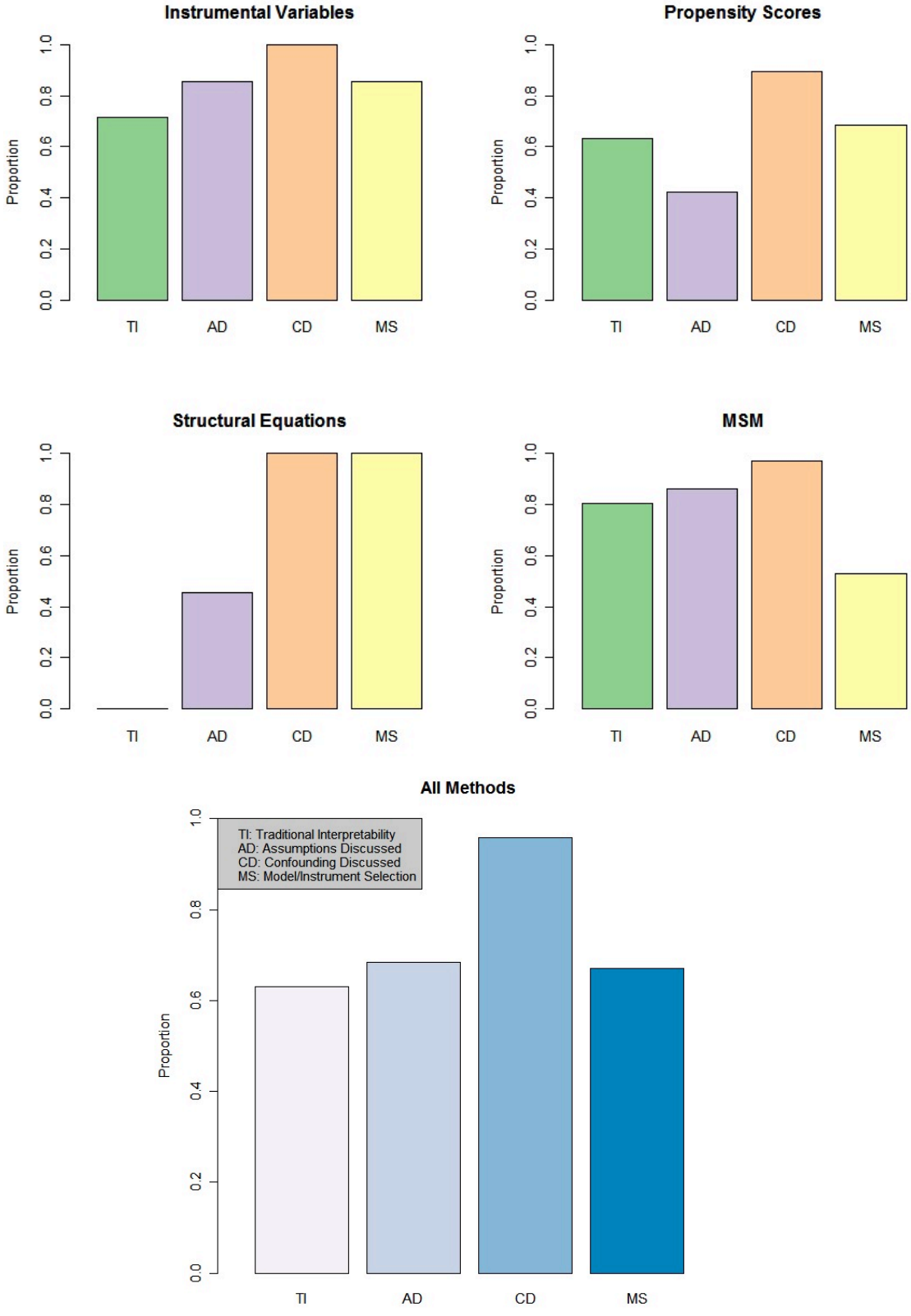


Figure 6a. Network of Institutions and Affiliated Authors of HIV-Related Studies Published Before 2007 Using Marginal Structural Models

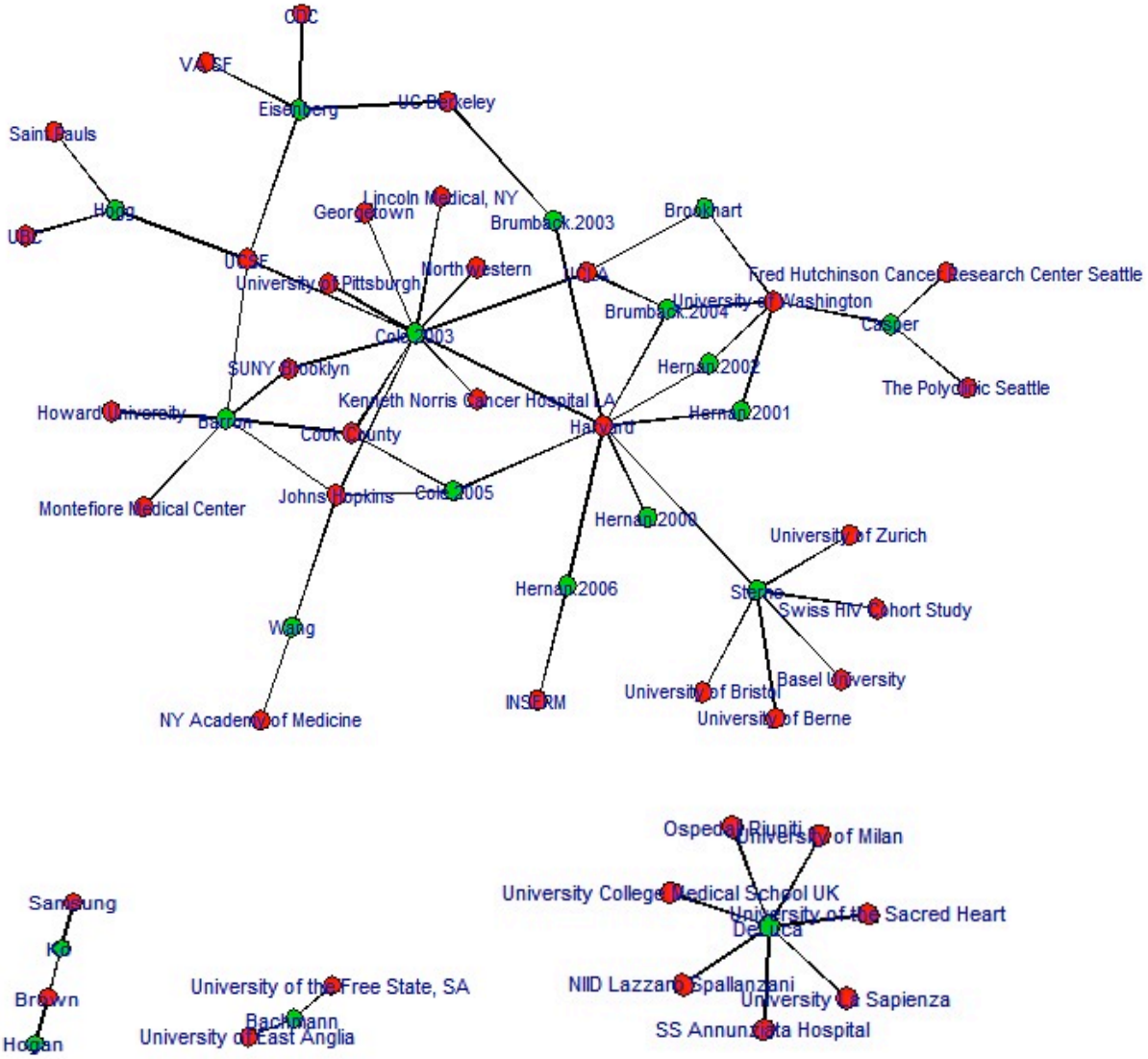


Figure 6b. Network of Institutions and Affiliated Authors of HIV-Related Studies Published In 2007-08 Using Marginal Structural Models

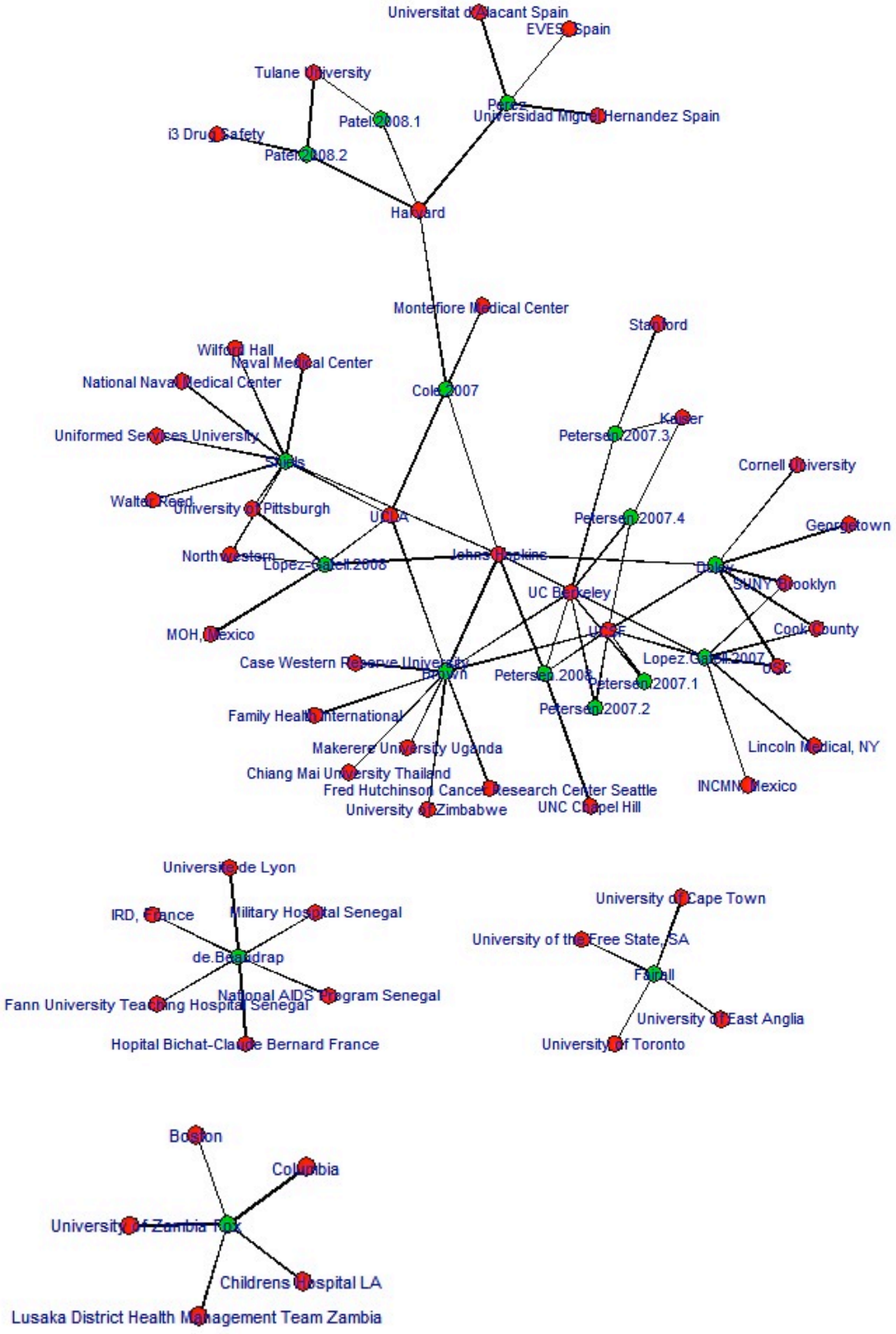


Figure 7a. Network of Institutions and Affiliated Authors of HIV-Related Studies Published Before 2007 Using Propensity Scores

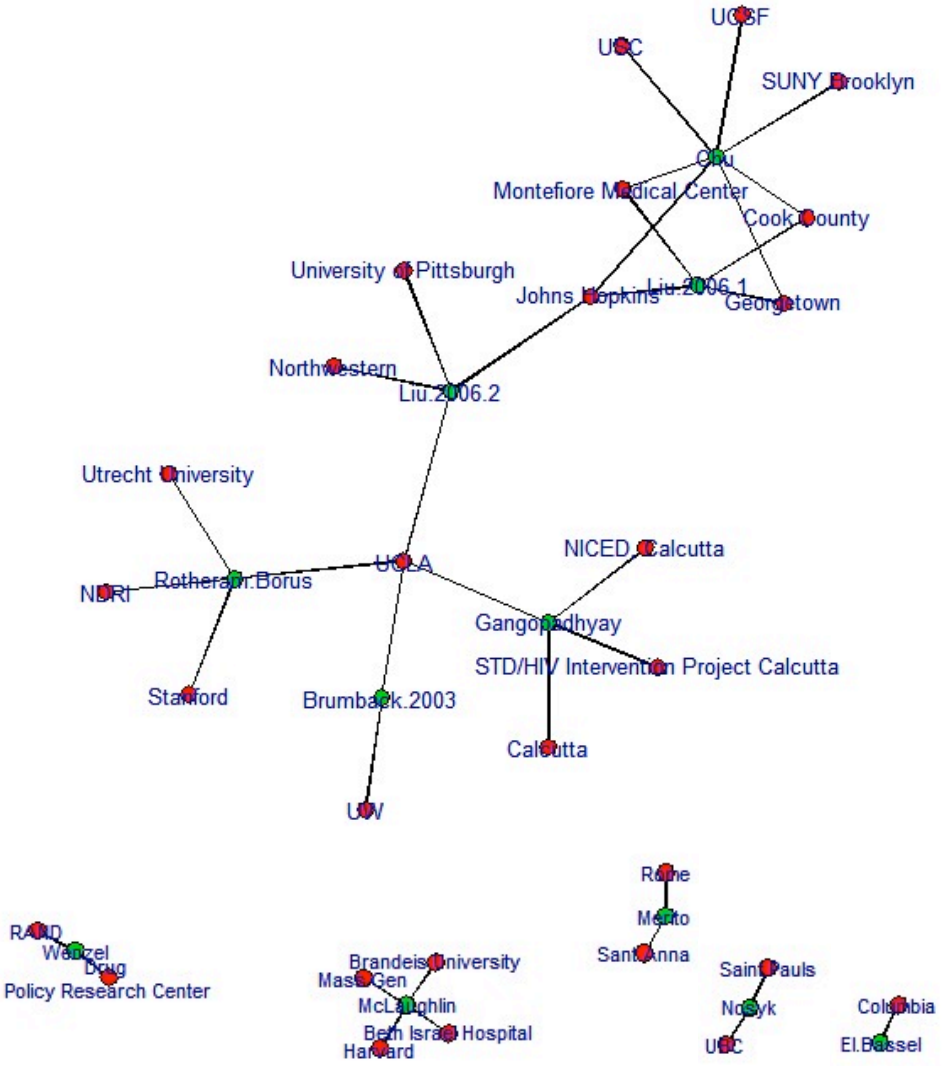
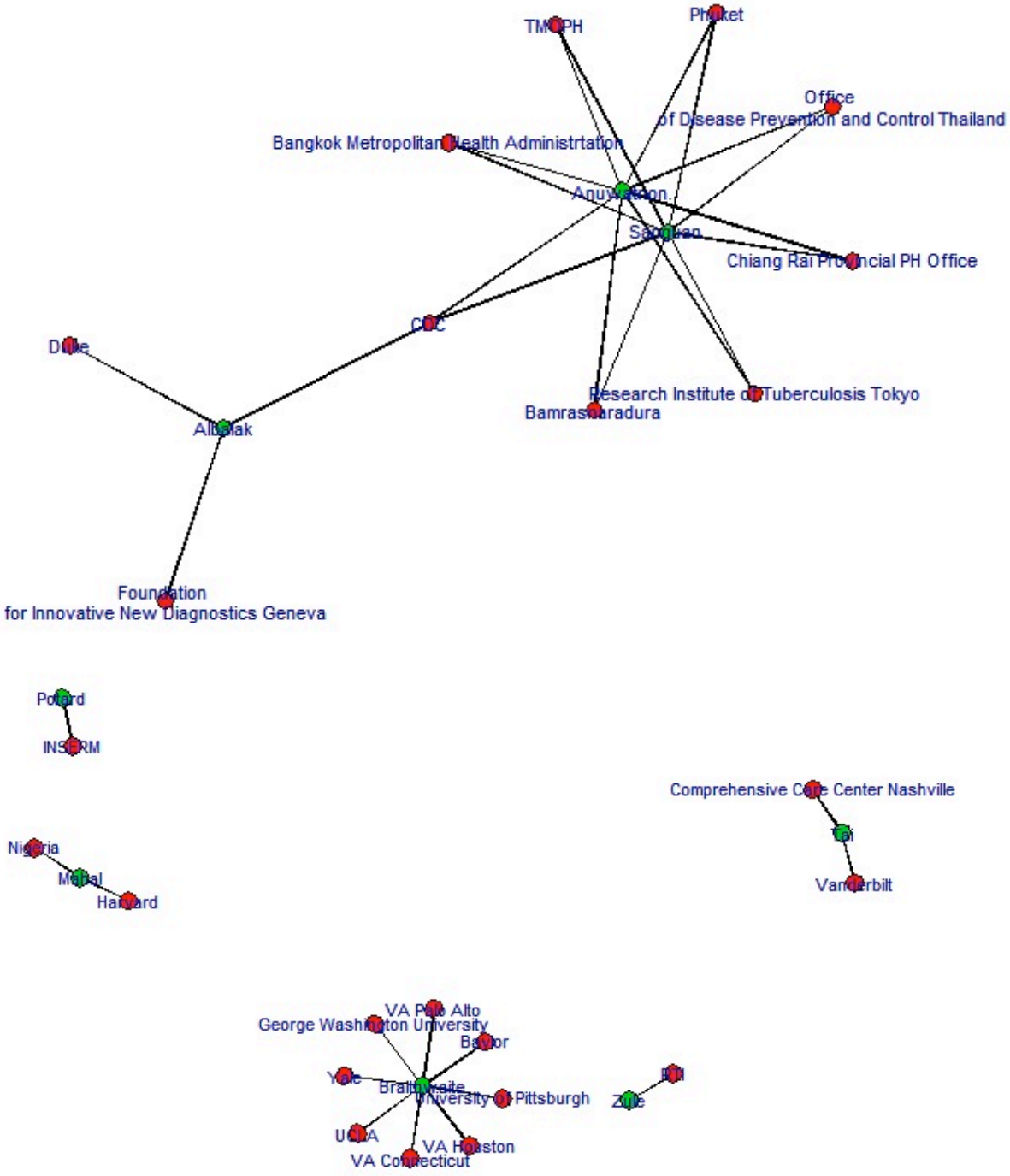


Figure 7b. Network of Institutions and Affiliated Authors of HIV-Related Studies Published In 2007 Or 2008 Using Propensity Scores



References

1. Jewell N. Statistics for Epidemiology. Chapman and Hall; 2004.

2. Shah B, Laupacis A, Hux J, Austin P. Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review. Journal of Clinical Epidemiology 2005;58(6):550-559.

3. Kleinbam D, Klein M. Survival Analysis. Second Edition. Springer-Verlag; 2005.

4. Hernan MA, Brumback B, Robins J. Marginal structural models to estimate joint causal effect of nonrandomized treatments. Journal of the American Statistical Association 2001;96(454):440-448.

5. Rosenbaum P, Rubin D. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70:41-55.

6. McLaughlin TJ, Soumerai SB, Weinrib D, Aupont O, Cotton D. The association between primary source of ambulatory care and access to a outcomes of treatment among AIDS patients. International Journal for Quality in Health Care 1999;11(4):293.

7. Riersol E. Confluence analysis by means of lag moments and other methods of confluence analysis. Econometrica 1941;9:1-24.

8. Tarwater PM, Mellors J, Gore ME, Margolick JB, Phair J, Detels R, Munoz A. Methods to assess population effectiveness of therapies in human immunodeficiency virus incident and prevalent cohorts. American journal of epidemiology 2001;154(7):675.

9. Bond SJ, White IR, Sarah Walker A. Instrumental variables and interactions in the causal analysis of a complex clinical trial. Statist. Med. 2007;26(7):1473-1496.

10. Robins JM. Marginal structural models. 1998;

11. Robins JM, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000;11(5):550–560.

12. van Der Laan M, Dudoit S. Unified cross-validation methodology for selection among estimators and a general cross-validated adaptive epsilon-net estimator: finite sample oracle inequalities and examples. UC Berkeley Division of Biostatistics Working Paper Series 2003;

13. Robins J. Marginal structural models versus structural nested models as tools for causal inference. In: Statistical Models in Epidemiology: The Environment and Clinical Trials. New York, NY: Springer-Verlag; 1999 p. 95-134.

14. Greenland S, Brumback B. An overview of relations among causal modelling methods. International journal of epidemiology 2002;31(5):1030.

15. R-Project. www.r-project.org: [date unknown].

16. Sanguanwongse N, Cain KP, Suriya P, Nateniyom S, Yamada N, Wattanaamornkiat W, Sumnapan S, Sattayawuthipong W, Kaewsa-ard S, Ingkaseth S, others. Antiretroviral therapy for HIV-infected tuberculosis patients saves lives but needs to be used more frequently in Thailand. JAIDS Journal of Acquired Immune Deficiency Syndromes 2008;48(2):181.

17. Potard V, Rey D, Mokhtari S. First-line highly active antiretroviral regimens in 2001-2002 in the French Hospital Database on HIV: combination prescribed and biological outcomes. Antiviral Therapy 2007;12:317-324.

18. Braithwaite RS, Kozal MJ, Chang CC, Roberts MS, Fultz SL, Goetz MB, Gibert C, Rodriguez-Barradas M, Mole L, Justice AC. Adherence, virological and immunological outcomes for HIV-infected veterans starting combination antiretroviral therapies. Aids 2007;21(12):1579.

19. Albalak R, O'Brien R, Kammerer S. Trends in tuberculosis/human immunodeficiency virus comorbidity, United States, 1993-2004. Archives of Internal Medicine 2007;167(22):2443-2452.

20. Tai J, Udoji M, Barkanic G, Byrne D, Rebeiro P, Byram B, Kheshti A, Carter J, Graves C, Raffanti S, Sterling T. Pregnancy and HIV Disease Progression during the Era of Highly Active Antiretroviral Therapy. J INFECT DIS 2007;196(7):1044-1052.

21. Anuwatnonthakate A, Limsomboon P, Nateniyom S, Wattanaamornkiat W, Komsakorn S, Moolphate S, Chiengsorn N, Kaewsa-ard S, Sombat P, Siangphoe U, Mock PA, Varma JK. Directly Observed Therapy and Improved Tuberculosis Treatment Outcomes in Thailand. PLoS ONE 2008;3(8):e3089.

22. Mahal A, Canning D, Odumosu K, Okonkwo P. Assessing the economic impact of HIV/AIDS on Nigerian households: a propensity score matching approach. AIDS 2008;22:S95.

23. Zule WA, Bobashev G. High dead-space syringes and the risk of HIV and HCV infection among injecting drug users. Drug and Alcohol Dependence 2009;100(3):204–213.

24. Shiels MS, Cole SR, Wegner S, Armenian H, Chmiel JS, Ganesan A, Marconi VC, Martinez-Maza O, Martinson J, Weintrob A, others. Effect of HAART on incident cancer and noncancer AIDS events among male HIV seroconverters. JAIDS Journal of Acquired Immune Deficiency Syndromes 2008;48(4):485.

25. Hogan JW, Lancaster T. Instrumental variables and inverse probability weighting for causal inference from longitudinal observational studies. Statistical Methods in Medical Research 2004;13(1):17-48.

26. Bhattacharya J, Goldman D, Sood N. The link between public and private insurance and HIV-related mortality. Journal of Health Economics 2003;22(6):1105–1122.

27. Lakdawalla D, Sood N, Goldman D. HIV Breakthroughs and Risky Sexual Behavior. The Quarterly Journal of Economics 2006;121(3):1063–1102.

28. Cain LE, Cole SR, Chmiel JS, Margolick JB, Rinaldo Jr CR, Detels R. Effect of highly active antiretroviral therapy on multiple AIDS-defining illnesses among male HIV seroconverters. American journal of epidemiology 2006;163(4):310.

29. Hernan M, Brumback B, Robins J. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. Epidemiology 2000;11(5):561–570.

30. Brown JM, Wald A, Hubbard A, Rungruengthanakit K, Chipato T, Rugpao S, Mmiro F, Celentano DD, Salata RS, Morrison CS, others. Incident and prevalent herpes simplex virus type 2 infection increases risk of HIV acquisition among women in Uganda and Zimbabwe. Aids 2007;21(12):1515.

31. Cole SR, Hernán MA, Robins JM, Anastos K, Chmiel J, Detels R, Ervin C, Feldman J, Greenblatt R, Kingsley L, others. Effect of highly active antiretroviral therapy on time to acquired immunodeficiency syndrome or death using marginal structural models. American journal of epidemiology 2003;158(7):687.

32. Lopez-Gatell H, Cole SR, Hessol NA, French AL, Greenblatt RM, Landesman S, Preston-Martin S, Anastos K. Effect of tuberculosis on the survival of women infected with human immunodeficiency virus. American journal of epidemiology 2007;165(10):1134.

33. Petersen ML, Deeks SG, Martin JN, van der Laan MJ. History-adjusted marginal structural models for estimating time-varying effect modification. American journal of epidemiology 2007;166(9):985.

34. Petersen ML, Deeks SG, van der Laan MJ. Individualized treatment rules: Generating candidate clinical trials. Statist. Med. 2007;26(25):4578-4601.

35. Petersen ML, Wang Y, van der Laan MJ, Guzman D, Riley E, Bangsberg DR. Pillbox Organizers Are Associated with Improved Adherence to HIV Antiretroviral Therapy and Viral Suppression: A Marginal Structural Model Analysis. CLIN INFECT DIS 2007;45(7):908-915.

36. Petersen ML, Wang Y, van der Laan MJ, Rhee SY, Shafer RW, Fessel WJ. Virologic efficacy of boosted double versus boosted single protease inhibitor therapy. AIDS 2007;21(12):1547–54.

37. de Beaudrap P, Etard J, Guèye FN, Guèye M, Landman R, Girard P, Sow PS, Ndoye I, Delaporte E, for the ANRS 1215/1290 Study Group. Long-Term Efficacy and Tolerance of Efavirenz- and Nevirapine-Containing Regimens in Adult HIV Type 1 Senegalese Patients. AIDS Research and Human Retroviruses 2008;24(6):753-760.

38. Dolev JC, Maurer T, Springer G, Glesby MJ, Minkoff H, Connell C, Young M, Schowalter K, Cox C, Hessol NA. Incidence and risk factors for verrucae in women. EPIDEMIOLOGY AND SOCIAL AIDS. 22 (10): 1213-1219, June 19, 2008. [date unknown];

39. Fairall LR, Bachmann MO, Louwagie G, van Vuuren C, Chikobvu P, Steyn D, Staniland GH, Timmerman V, Msimanga M, Seebregts CJ, others. Effectiveness of antiretroviral treatment in a South African program: a cohort study. Archives of Internal Medicine 2008;168(1):86.

40. López-Gatell H, Cole SR, Margolick JB, Witt MD, Martinson J, Phair JP, Jacobson LP. Effect of tuberculosis on the survival of HIV-infected men in a country with low tuberculosis incidence. AIDS 2008;22(14):1869.

41. Patel K, Hernán M, Williams P, Seeger J, McIntosh K, Dyke R, Seage III G, Pediatric AIDS Clinical Trials Group 219/219C Study Team. Long-term effect of highly active antiretroviral therapy on CD4 cell evoloution among children and adolescents infected with HIV: 5 years and counting. CLIN INFECT DIS 2008;46(11):1751-1760.

42. Patel K, Hernan MA, Williams PL, Seeger JD, McIntosh K, Dyke RBV, Seage III GR, Pediatric AIDS Clinical Trials Group 219/219C Study Team. Long□Term Effectiveness of Highly Active Antiretroviral Therapy on the Survival of Children and Adolescents with HIV Infection: A 10□Year Follow□Up Study. CLIN INFECT DIS 2008;46(4):507-515.

43. Petersen ML, van der Laan MJ, Napravnik S, Eron JJ, Moore RD, Deeks SG. Long-term consequences of the delay between virologic failure of highly active antiretroviral therapy and regimen modification. AIDS 2008;22(16):2097.

44. Fox MP, Brooks DR, Kuhn L, Aldrovandi G, Sinkala M, Kankasa C, Horsburgh R, Thea DM. Role of breastfeeding cessation in mediating the relationship between maternal HIV disease stage and increased child mortality among HIV-exposed uninfected children. International Journal of Epidemiology 2009;38(2):569.

45. METODOLÓGICA N. Aplicación de modelos estructurales marginales para estimar los efectos de la terapia antirretroviral en 5 cohortes de seroconvertores al virus de la inmunodeficiencia humana Marginal structural models application to estimate the effects of antiretroviral therapy in 5 cohorts of HIV seroconverters. Gac Sanit 2007;21(1):76–83.

46. Velde FW, Pligt J. AIDS-related health behavior: Coping, protection motivation, and previous behavior. Journal of behavioral medicine 1991;14(5):429–451.

47. Kraft P, Rise J. Prediction of attitudes towards restrictive AIDS policies: a structural equation modeling approach. Social Science and Medicine 1995;40(5):711.

48. Llabre MM, Weaver KE, Durán RE, Antoni MH, McPherson-Baker S, Schneiderman N. A measurement model of medication adherence to highly active antiretroviral therapy and its relation to viral load in HIV-positive adults. AIDS Patient Care & STDs 2006;20(10):701–711.

49. Naar-King S, Wright K, Parsons J, Frey M, Templin T, Ondersma S. Transtheoretical Model and substance use in HIV-positive youth. AIDS Care 2006;18(7):839-845.

50. S\öderg\aard B, H\öfer S, Halvarsson M, S\önnerborg A, Tully MP, Lindblad \K. A structural equation modeling approach to the concepts of adherence and readiness in antiretroviral treatment. Patient Education and Counseling 2007;67(1-2):108–116.

51. Bull S, Pratte K, Whitesell N, Rietmeijer C, McFarlane M. Effects of an Internet-Based Intervention for HIV Prevention: The Youthnet Trials. AIDS Behav 2008;13(3):474-487.

52. Cha ES, Erlen JA, Kim KH, Sereika SM, Caruthers D. Mediating roles of medication–taking self-efficacy and depressive symptoms on self-reported medication adherence in persons with HIV: A questionnaire survey. International journal of nursing studies 2008;45(8):1175.

53. Rice E, Stein J, Milburn N. Countervailing social network influences on problem behaviors among homeless youth. Journal of Adolescence 2008;31:625-639.

54. Hernan MA, Brumback BA, Robins JM. Estimating the causal eect of zidovudine on CD4 count with a marginal structural model for repeated measures. Statistics in Medicine 2002;21:1689–1709.

55. Hernan M, Robins J. Estimating causal effects from epidemiological data. Journal of Epidemiology and Community Health 2006;60(7):578-586.

56. Cole SR, Hernan MA, Margolick JB, Cohen MH, Robins JM. Marginal structural models for estimating the effect of highly active antiretroviral therapy initiation on CD4 cell count. American journal of epidemiology 2005;162(5):471.

57. Cole SR, Hernan MA, Anastos K, Jamieson BD, Robins JM. Determining the effect of highly active antiretroviral therapy on changes in human immunodeficiency virus type 1 RNA viral load using a marginal structural left-censored mean model. American journal of epidemiology 2007;

58. Brumback BA, Hernán MA, Haneuse S, Robins JM. Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures. Statistics in medicine 2004;23(5):749–767.

59. Brumback B, Greenland S, Redman M, Kiviat N, Diehr P. The intensity-score approach to adjusting for confounding. Biometrics 2003;59(2):274–285.

60. Tager IB, Haight T, Sternfeld B, Yu Z, van Der Laan M. Effects of Physical Activity and Body Composition on Functional Limitation in the Elderly. Epidemiology 2004;15(4):479-493.

61. Lefebvre G, Delaney JAC, Platt RW. Impact of mis□specification of the treatment model on estimates from a marginal structural model. Statist. Med. 2008;27(18):3629-3642.

62. Epstein R. Six authors in search of a citation: Villains or victims of the Vancouver Convention? British Medical Journal 1993;306:765-767.

63. Nosyk B, Sun H, Li X, Palepu A, Anis AH. Highly active antiretroviral therapy and hospital readmission: comparison of a matched cohort. BMC Infectious Diseases 2006;6(1):146.

64. Merito M, Pezzotti P, for the ICONA Study Group. Comparing costs and effectiveness of different starting points for highly active antiretroviral therapy in HIV-positive patients. Eur J Health Econ 2006;7(1):30-36.

65. Liu C, Ostrow D, Detels R, Hu Z, Johnson L, Kingsley L, Jacobson LP. Impacts of HIV infection and HAART use on quality of life. Qual Life Res 2006;15(6):941-949.

66. Liu C, Weber K, Robison E, Hu Z, Jacobson LP, Gange SJ. Assessing the effect of HAART on change in quality of life among HIV-infected women. AIDS Research and Therapy 2006;3(1):6.

67. Gangopadhyay DN, Chanda M, Sarkar K, Niyogi SK, Chakraborty S, Saha MK, Manna B, Jana S, Ray P, Bhattacharya SK, Detels R. Evaluation of Sexually Transmitted Diseases/Human Immunodeficiency Virus Intervention Programs for Sex Workers in Calcutta, India. Sexually Transmitted Diseases 2005;32(11):680-684.

68. El-Bassel N, Gilbert L, Wu E. HIV and intimate partner violence among methadone-maintained women in New York City. Social Science and Medicine 2005;61:171-183.

69. Chu JH, Gange SJ, Anastos K, Minkoff H, Cejtin H, Bacon M, Levine A, Greenblatt RM. Hormonal contraceptive use and the effectiveness of highly active antiretroviral therapy. American journal of epidemiology 2005;161(9):881.

70. Wenzel SL, Tucker JS, Elliott MN, Hambarsoomians K, Perlman J, Becker K, Kollross C, Golinelli D. Prevalence and co-occurrence of violence, substance use and disorder, and HIV risk behavior: A comparison of sheltered and low-income housed women in Los Angeles County. Preventive Medicine 2004;39(3):617–624.

71. Rotheram-Borus MJ, Song J, Gwadz M, Lee M, Rossem RV, Koopman C. Reductions in HIV risk among runaway youth. Prevention Science 2003;4(3):173–187.

72. Hogg RS, Bangsberg DR, Lima VD, Alexander C, Bonner S, Yip B, Wood E, Dong WW, Montaner JS, Harrigan PR. Emergence of drug resistance is associated with an increased risk of death among patients first starting HAART. PLoS Med 2006;3(9):e356.

73. Hernan M, Lanoy E, Costagliola D. Comparison of dynamic treatment regimes via inverse probability weighting. Basic and Clinical Pharmacology and Toxicology 2006;98:237-242.

74. De Luca A, Cozzi-Lepri A, Antinori A. Lopinavir/ritonavir or efavirenz plus two nucleoside analogues as first-line antiretroviral therapy: a non-randomized comparison. Antiviral Therapy 2006;11:609-618.

75. Brookhart A, van Der Laan M. A semiparametric model selection criterion with applications to the marginal structural model. Computational Statistics and Data Analysis 2006;50:475-498.

76. Bachmann MO, Booysen FL. Economic causes and effects of AIDS in South African households. AIDS 2006;20(14):1861.

77. Wang C, Vlahov D, Galai N, Cole SR, Bareta J, Pollini R, Mehta SH, Nelson KE, Galea S. The effect of HIV infection on overdose mortality. AIDS 2005;19(9):935.

78. Sterne JA, Hernán MA, Ledergerber B, Tilling K, Weber R, Sendi P, Rickenbach M, Robins JM, Egger M. Long-term effectiveness of potent antiretroviral therapy in preventing AIDS and death: a prospective cohort study. The Lancet 2005;366(9483):378–384.

79. Casper C, Redman M, Huang ML, Pauk J, Lampinen TM, Hawes SE, Critchlow CW, Morrow RA, Corey L, Kiviat N, others. HIV infection and human herpesvirus-8 oral shedding among men who have sex with men. JAIDS Journal of Acquired Immune Deficiency Syndromes 2004;35(3):233.

80. Barron Y, Cole SR, Greenblatt RM, Cohen MH, Anastos K, DeHovitz JA, Delapenha R, Gange SJ. Effect of discontinuing antiretroviral therapy on survival of women initiated on highly active antiretroviral therapy. AIDS 2004;18(11):1579-1584.

81. Ko H, Hogan J, Mayer K. Estimating causal treatment effects from longitudinal HIV natural history studies using marginal structural models. Biometrics 2003;59:152-162.

82. Eisenberg JNS, Wade TJ, Hubbard A, Abrams DI, Leiser RJ, Charles S, Vu M, Saha S, Wright CC, Levy DA, others. Associations between water-treatment methods and diarrhoea in HIV-positive individuals. Epidemiology and infection 2002;129(02):315–323.

83. Prado G, Feaster DJ, Schwartz SJ, Pratt IA, Smith L, Szapocznik J. Religious involvement, coping, social support, and psychological distress in HIV-seropositive African American mothers. AIDS and Behavior 2004;8(3):221–235.

84. Lim K. Managing HIV at the workplace: an empirical study of HIV and HR managers in Singapore. Journal of Occupational Health Psychology 2003;8(4):235-246.

85. Sengupta S, Strauss RP, DeVellis R, Quinn SC, DeVellis B, Ware WB. Factors affecting African-American participation in AIDS research. JAIDS Journal of Acquired Immune Deficiency Syndromes 2000;24(3):275.

B.2 The Effect of Early Triple Therapy Among HIV-Infected Children: A Causal Inference Approach

This study will be submitted for publication in the journal AIDS in Spring 2010.

The Effect of Early Triple Therapy in HIV-Infected Children:
A Causal Inference Approach

Word Count exclusive of abstract, tables and figures: 5,600

Word Count of abstract: 500

ABSTRACT

Background

Worldwide, particularly in areas with no treatment availability or antenatal programs, approximately 1600 children are diagnosed with HIV every day,(2) and over 300,000 deaths among infected children occur annually worldwide.(3) In 1996, the advent of highly active antiretroviral therapy (HAART) dramatically reduced the risk of mortality from HIV. However, the long-term effects of HAART, or triple therapy, are not yet fully understood. In the present study, I use marginal structural models as estimated by G-Computation to estimate the causal effect of triple therapy (HAART) on time to C diagnosis, and time to C diagnosis/death among HIV-infected children.

Methods

The Pediatric Spectrum of Disease (PSD) is a multicenter active surveillance program specifically for children who have been exposed to HIV perinatally.(5) Through this program, I have identified and defined a population-based cohort of HIV positive northern Californian children who were vertically infected from 1988-2008. Causal inference methods are alternative techniques with causal effect interpretations. I defined binary levels of A, $A \in \{0,1\}$ in triple therapy (HAART) initiated in the first 6 months versus no triple ARV therapy initiated in the first 6 months and defined binary levels of A, $A \in \{0,1\}$ in triple therapy (HAART) initiated in the first 12 months versus no triple ARV therapy initiated in the first 12 months. Two subgroup analyses will be performed by further restricting A to triple therapy initiated within the first 6 or 12 months of life among symptomatic children and triple therapy initiated within the first 6 or 12 months of life among asymptomatic children. I have defined a vector of baseline covariates, W, which includes immune status at treatment initiation, length of pregnancy (full-term or less than full-term), child's race, sex, whether mothers received prenatal care, and birthweight (< 2500 grams or >= 2500 grams).

Results

The sample comprised of N=217 HIV infected children whose infection is assumed to occurred in utero or at delivery. The majority of the sample is female (56.2%) and non-White ethnicity (71.9%). Immune impairment at ARV treatment initiation was common as 40.1% were severely impaired and 32.3% were moderately impaired. Eight percent of the children received triple therapy in their first 6 months of life, while 45% received triple therapy within the first 12 months of life.

Though no results were statistically significant, there are some trends that should be highlighted. Among children who initiated triple therapy within 6 months of birth the

causal effect of treatment in delaying a C diagnosis, $\Psi HZ(p_0)(t_k)$= -0.466 (95% CI -1.20-0.565), is seemingly stronger than children who initiated therapy within 12 months of birth ($\Psi HZ(p_0)(t_k)$= -0.321 (95% CI -1.151-0.300)). Additionally, the effect of triple therapy initiated within the first 6 or 12 months of life on time to C diagnosis is greater among *symptomatic* children (12 Months: $\Psi HZ_{symptomatic}$ (p0)(t_{36}): -0.587 (95% CI -1.217-0.480)) than among *asymptomatic* children (12 Months: $\Psi HZ_{asymptomatic}$ (p0)(t_{36}): -0.106 (95% CI -1.054-0.739)).

Discussion

The WHO, in 2006, developed clinical and immunologic guidelines for treatment initiation in asymptomatic children in resource-limited settings based on HIV Pediatric Prognostic Markers Collaborative Study (HPPMCS) data.(4) In 2008, WHO amended their recommendations for treatment initiation for HIV-positive children as a result from an RCT in South Africa.(7) The results from the present analysis may be interpreted as supportive of the current WHO treatment guidelines for initiating treatment among all HIV positive children, regardless of symptoms.

Introduction

Today, over 90% of the estimated 2.5 million HIV infected children worldwide live in sub-Saharan Africa.(3) Over 300,000 deaths among infected children occur annually worldwide.(3) The burden of pediatric HIV infections lies in the poorest regions of sub-Saharan Africa, where approximately only 10% of mothers have access to antenatal programs aimed at preventing mother-to-child-transmission (MTCT).(19) As access to care in developing areas continues to affect the health of future mothers, so too does it affect their children's health. A recent study in South Africa found that 85% of HIV-infected, infants who were exposed to single dose nevirapine (sdNVP) perinatally were moderately or severely immuno-compromised (CD4 % < 25) by 6 months post-partum,(17) suggesting these children were particularly vulnerable due to the severity of infection in their mothers. Though it is the standard treatment for prevention of MTCT, about 10% of children exposed to sdNVP will still develop HIV even before breastfeeding.(17; 18)

In 1996, the advent of highly active antiretroviral therapy (HAART) dramatically reduced the risk of mortality from HIV. Results from birth cohort studies of HIV+ children indicate that approximately 70-80% of children left untreated will survive to age five.(23-25) Patel et al, with the use of marginal structural models, estimated the weighted, adjusted proportional hazard for mortality as 0.24 (95% CI 0.11–0.51) when comparing HAART treated children to untreated children.(27) Similarly, Gortmaker et al found a reduced hazard ratio for death (HR 0.33; 95% CI 0.19-0.58)(26) and de Martino et al found a reduced RH of death (RH 0.29; 95% CI 0.13-0.67) among triple therapy initiated children compared to untreated children.(23) HAART is used as a first-line treatment now among HIV infected children in order to recuperate from HIV-associated illnesses and re-establish immuno-compentence.(29-33)

Researchers are still trying to establish the most ideal time to initiate antiretroviral therapy in vertically infected children. Weighing the benefits and risks of early initiation of HAART or triple therapy is a necessary component in making treatment guidelines recommendations. Treatment recommendations vary between the Centers for Disease Control (CDC), World Health Organization (WHO), and Ministries of Health within individual European countries.(4) The WHO recently changed the treatment guidelines to include all children under 12 months regardless of immunologic status. Previously, treatment for pediatric HIV infection was only recommended for children who presented symptomatically. The implications for these treatment guidelines are particularly important for the infected infants who will now have to be on HAART for life. Once a child begins therapy, he must remain on treatment for life, or he risks developing drug resistance, or other health concerns.(48; 49)

For numerous reasons, not the least of which are the inherent ethical issues, randomized controlled trials (RCT) exploring the best time to initiate HAART in HIV positive children are very uncommon. In fact, the only published randomized trial estimating the effect of early HAART versus delayed HAART on mortality among HIV positive infants prematurely terminated in 2008 as a result of an unbalanced, disproportionate number of deaths in the delayed group.(35) One other RCT was conducted to explore the impact of

delaying HAART initiation on clinical disease progression, however this study was a small feasibility study in preparation for another, larger RCT which will likely be completed in 2011. Moreover, the study population only included HIV positive children 1-12 years of age, excluding all positive infants.(36) Prior to the HAART era, the PENTA 1 study conducted a similar study of delayed versus early initiation of zidovudine monotherapy.(37) Their results suggest that early initiation of ART has no added benefit on clinical outcomes. Among non-RCTs, Newell et al from the European Collaborative Study, a prospective study of a birth cohort of 131 HIV infected children, conclude that initiating ART in the first 5 months of life and the use of HAART were both highly predictive of an improved CD4 z-score 6 months after treatment initiation.(38) In one of the only other identified observational studies evaluating the impact of delayed treatment initiation among HIV positive infants, Chiappini et al found children treated early with HAART had significantly lower viral load than deferred treatment children and they were also less likely to progress to a C diagnosis.(39)

Particularly a problem in HIV/AIDS literature, observational studies are often biased as traditional analysis methods are employed to estimate the effect of a treatment on an outcome of interest. Causal inference methods have been developed to overcome many of these biases and will be employed in the present study.

To overcome the inherent issue of correct model specification in time to event observational studies, I intend to use g-computation, a marginal structural models (MSM) estimator, to estimate the causal effect of HAART (interchangeably referred to as triple therapy) on reducing AIDS/death among children who were infected in utero. Additionally, I will perform a subanalysis of symptomatic and asymptomatic children.

Methods

The Pediatric Spectrum of Disease (PSD) is a multicenter active surveillance program specifically for children who have been exposed to HIV perinatally.(5) Maldonado et al have previously described this population.

Statistical Methods

Data Structure

In the present analysis, I have done a time to event analysis to explore the effect of treatments (triple ARV therapy, or no triple ARV therapy), A, have on the amount of time, T, it takes for a child to experience an event (1: Category C; 2: Category C diagnosis or death).

I have looked at binary levels of A, $A \in \{0,1\}$ in triple therapy (HAART) initiated in the first 6 months versus no triple ARV therapy initiated in the first 6 months. To allow for a less restrictive treatment assignment, I have also looked at binary levels of A, $A \in \{0,1\}$ in triple therapy (HAART) initiated in the first 12 months versus no triple ARV therapy initiated in the first 12 months. Two subgroup analyses were performed by further restricting A to triple therapy initiated within the first 6 or 12 months of life among symptomatic children and triple therapy initiated within the first 6 or 12 months of life among asymptomatic children. I have defined a vector of baseline covariates, W, which

includes immune status at treatment initiation (not included in subgroup analyses), length of pregnancy (full-term or less than full-term), child's race, sex, whether mothers received prenatal care, and birthweight (< 2500 grams or >= 2500 grams). I have defined T as a discrete variable with values {1,….,K}, where K is last time point children are monitored. In contrast, I defined censoring, C, as the last time point children are observed. My data structure also defines whether an event has occurred as N_1 and a similar scenario for censoring with N_2. In turn, all time points until the event occurs are denoted by $dN_1(t) = 0$ and $dN_1(t) = 1$ at the time point the event occurs. All time points until a child is censored are denoted by $dN_2(t) = 0$ and $dN_2(t) = 1$ at the time of censoring. The long form of my observed data can be expressed as n iid observations of $O = (A,W,dN_1(t), dN_2(t): t=1, \ldots, K) \sim p_o$, where p_o is the density of the my observed data, O.

The likelihood of the observed data is described as:

$$L(0) = P(W)P(A|W) \prod_{t=1}^{k} P(dN_1(t) \mid dN_1(t-1) = 0, dN_2(t-1) =0,A,W) \quad (1)$$
$$P(dN_2(t) \mid dN_1(t) = 0, dN_2(t-1) = 0, A, W)$$

Where,

$Q_{10}(W) \equiv P(W)$ is the distribution of baseline covariates, W;
$Q_{20}(N_1(t),A,W) \equiv P(dN_1(t) \mid dN_1(t-1) = 0, dN_2(t-1) =0,A,W)$ is the conditional hazard of the event (C diagnosis and/or death) given the treatment (A) and baseline covariates, W;
$g_{10}(A,W) \equiv P(A \mid W)$ is the treatment mechanism;
$g_{20}(N_2(t), A,W) \equiv P(dN_2(t) \mid dN_1(t) = 0, dN_2(t-1) = 0, A, W)$ is the censoring mechanism—the conditional hazard of censoring given the subject did not yet experience an event, no previous censoring, and given the treatment, A, and baseline covariates.

In turn, the likelihood (equation 1) factorizes the distribution of W, baseline covariates, the missingness mechanism, g, and the conditional hazard of the outcome of interest. To estimate the survival, that is, the probability of surviving to time k given treatment, A, and baseline covariates, W, one would define $S_0(t_k \mid A, W) = P(T > t_k \mid A,W)$. Then one would take the cumulative product of 1 minus the conditional hazard of C diagnosis and/or death to estimate survival (see equation 2):

$$S_0(t_k \mid A, W) = \prod_{t=1}^{t_k} (1- Q_{20}(N_1(t),A,W)) \quad (2)$$

Traditional Methods for Confounding Adjustment
One of the traditional methods for controlling the biases introduced by measured confounders is by using multivariable regression techniques. However, regression techniques are only as good as the measured confounder data.(6) Moreover, model building can become cumbersome even with comprehensive confounder data collected;

often investigators will settle for model interpretability over adequate adjustment for bias.(6) Researchers will often a priori specify a parametric hazard model in traditional techniques exploring the effect of a treatment, A, on a time to event outcome, and test if A is different from zero. Covariates are selected for inclusion in these traditional approaches usually in a rudimentary way—selecting and deleting covariates based on their influence on A's effect on the outcome. One of the most common approaches for time to event scenarios is the Cox proportional hazard model, or logistic regression in the case of discrete time outcomes. Furthermore, time will often be fit and a linear model will be employed to estimate the effect of A and the covariates. The parameters in conditional hazard models are estimated with a maximum likelihood approach. The central feature here is to evaluate whether the parameter representing the treatment, A, is significantly different from 0. Again, these estimates are heavily dependent on how well the model is specified. In turn, the parameters within hazard models may not be correctly specified, though this may go unnoticed unless the selected hazard model is contrasted with alternative models. It should be noted that the parameter estimating the effect of A on the outcome of interest is only relevant within that specific model. In the optimistic case of correctly specifying the hazard model, then this parameter represents the log odds ratios of the event occurring at each time point for A, and only in the context of that specific model.

G-Computation Approach

For example, let us define our parameter of interest as a function of our data generating distribution, $\Psi(p_0)$. If one is interested in the survival at time point, t_k, the specific survival curve may be expressed as $P(T_a > t_k)$. If a subject's treatment level were to be set at a, then T_a would be the event time, T, one would have observed, regardless of whether that subject's true observed treatment is at level a, often referred to as the counterfactual. Had this subject been treated with a different level of a than what is expressed in $P(T_a > t_k)$, then this is a counterfactual description of his survival curve.

In order to quantify the causal effect of A, triple therapy, on death/category C diagnosis, the effect will be estimated using these counterfactuals. More specifically, I will be using a marginal additive difference $\Psi_{RD}(p_0)(t_k)$, in the probability of survival. That is,

$$\Psi_1(p_0)(t_k) \equiv P(T_a > t_k) \text{ [all treated]} - \Psi_0(p_0)(t_k) \equiv P(T_a > t_k) \text{ [all untreated]} \quad (5)$$

My parameters established in this fashion may also be influenced by baseline covariates, defined previously as W. As such, I can estimate the counterfactual survival at t_k setting other baseline variables at 0 or 1. Additionally, the marginal log hazard ratio is defined by:

$$\Psi_{RH}(p_0)(t_k) = \log((\log(\Psi_1(p_0)(t_k)))/(\log(\Psi_0(p_0)(t_k)))) \quad (6)$$

In order to express the mean counterfactual outcomes, as described by equations 5 and 6, I will employ the G-computation estimator. Because it is important to estimate the distributions of my baseline covariates, W, the conditional hazard of the event given their treatment, A, and W, and the conditional survival of the outcome of interest, as related to

the conditional hazard, I will employ super learner software (D/S/A-Deletion/Substitution/Addition). Sinisi and van der Laan have applied this algorithm to fit the initial hazard on pooled data over time.(62) The empirical distribution of my baseline covariates in my data will estimate non-parametrically the marginal distribution of W. By using this data-adaptive machine learning algorithm, and its cross-validation based on likelihood, I am avoiding the problems inherent with traditional approaches and model building. All confidence intervals for G-computation estimates were calculated by bootstrap sampling.

In the present study, I will estimate the causal effect of triple therapy on mortality or C diagnosis in the first 36 months of life in children enrolled in a population-based study using marginal structural models as estimated by G-computation methods. For comparison, the log-rank statistic in a Cox proportional hazards model was estimated for each comparison, as well, while adjusting for the same baseline covariates selected by D/S/A in the marginal structural approach.

Results

Demographics and Baseline Characteristics

The sample was comprised of N=217 HIV infected children whose infection was assumed to occur in utero or at delivery. Patient characteristics are outlined in Table 1. The majority of the sample was female (56.2%) and non-White ethnicity (71.9%). Approximately half of the mothers of the children included received prenatal care. Over one-quarter of the children were born low birth weight (29.5%) and about forty-three percent were not full-term. Immune impairment at ARV treatment initiation was common as 40.1% were severely impaired and 32.3% were moderately impaired. Eight percent of the children received triple therapy in their first 6 months of life, while 45% received triple therapy within the first 12 months of life. About 55% of the sample either never received triple therapy or initiated therapy after 12 months of life.

The associations between the baseline covariates (W) and triple therapy initiation in the first 6 months and in the first 12 months are described in Table 2. As the immune status deteriorated to moderate or severe immune suppression, children were more likely to initiate therapy within their first 6 months of life (cOR = 1.61; p value 0.12). as the immune status worsens at treatment initiation, the odds of starting triple therapy in the first 12 months are increased 1.82 times (p value = 0.03). The odds of beginning triple therapy in the first 12 months among White children decrease to 0.28 when compared to non-White children (p value = 0.05).

A total of n=75 children were diagnosed with a C diagnosis within the first 36 months of life. Bivariate estimates of W and C diagnosis within the first 36 months of life are listed in Table 3. Both ethnicity and immune status at treatment initiation are strongly associated with C diagnosis in the first 36 months. Specifically, a child of White ethnicity is more than two times more likely to be diagnosed with a C diagnosis than a non-White child (cOR = 2.39; p value < 0.01); the worse the immune status at treatment initiation the more likely the child was to be diagnosed with a C diagnosis (cOR = 2.17; p value < 0.01). A total of n=84 children were either diagnosed with a C diagnosis or died

within 36 months of birth. Again, White children were twice as likely to either be diagnosed with a C diagnosis or die within the first 36 months of life (cOR = 2.01; p value < 0.01). A worsening of immune status at treatment initiation would increase a child's odds of being diagnosed with a C diagnosis or dying within the first 3 years of life by nearly two-fold (cOR = 1.90; p value < 0.01).

For the sub-analyses, the sample populations were limited to children who were treated asymptomatically and children who were treated symptomatically at treatment initiation. N=10 symptomatic children were treated with triple therapy within their first 6 months of life, and n=8 asymptomatic children were similarly treated (see Table 7). Forty-eight children were treated with triple therapy symptomatically within their first 12 months of life. In contrast, n=50 children were asymptomatic at treatment initiation in their first 12 months of life (see Table 8). there are no significant associations between any baseline covariate and triple therapy initiation within 6 or 12 months among symptomatic and asymptomatic children.

Time to Event Analysis: A G-Computation Approach

The data were expanded such that time to event outcomes could be estimated. To estimate the survival, specifically the probability of surviving to time k given treatment, A, and baseline covariates, W, I defined $S_0(t_k \mid A, W) = P(T > t_k \mid A,W)$. Additionally, I estimated the cumulative product of 1 minus the conditional hazard of experiencing the event to estimate survival:

$$S_0(t_k \mid A, W) = \prod_{t=1}^{t_k}(1- P(dN_1(t) \mid dN_1(t-1) = 0, dN_2(t-1) =0,A,W))$$

The marginal log hazard ratios at time *k=36*, denoted by $\Psi HZ(p_0)(t_k)$, and marginal additive differences at time *k=36*, $\Psi AD(p_0)(t_k)$, are given for each comparison in Table 4. For illustrative purposes, the G-computation survival curves for the causal effect of triple therapy initiated in the first 6 months of life on time to C diagnosis among all children are displayed in Figure 1.

Over all children, regardless of symptoms, the marginal additive difference of effect of triple therapy on time to C diagnosis among children who initiated therapy within 6 months was estimated at 0.120 (95% CI -0.192-0.276), while the marginal additive difference of effect among children treated within 12 months was 0.087 (95% CI -0.099-0.264). The marginal log hazard ratio for children who initiated treatment within 6 months of birth is -0.466 (95% CI -1.20-0.565) and -0.321 (95% CI -1.151-0.300) among children who were treated within 12 months of birth. The Cox proportional hazards parameter at 36 months comparing children treated within the first 6 months of life and children not treated within the first 6 months of life is HR=-0.476 (p value = 0.356) and HR=-0.407 (p value = 0.089) comparing children treated within the first 12 months of life and children not treated within 12 months.

The marginal additive difference of effect of triple therapy on time to C diagnosis or death among children who initiated therapy within 6 months was estimated at 0.108 (95% CI -0.231-0.261) and 0.055 (95% CI -0.252-0.208) for children treated within 12 months. The marginal log hazard ratio for children who initiated treatment within 6 months of birth is -0.369 (95% CI -1.058-0.730) and -0.180 (95% CI -0.786-0.792) for children treated within 12 months of birth. For comparison purposes, the Cox proportional hazards parameter at 36 months comparing children treated within the first 6 months of life is HR=-0.346 (p value = 0.454) and HR=-0.431 (p value = 0.058) for children treated within the first 12 months of life.

For subgroup analyses, I explored the effect of triple therapy among asymptomatically and symptomatically treated children. The marginal additive difference of effect of triple therapy on time to C diagnosis among *asymptomatic* children who initiated therapy within 6 months was estimated at 0.115 (95% CI -0.327-0.446) and 0.146 (95% CI -0.055-0.415) among *symptomatic* children. The marginal log hazard ratio for children treated *asymptomatically* within the first 6 months of life is -0.429 (95% CI-16.225-0.948) and -0.563 (95% CI -15.314-0.178) for children treated *symptomatically*. The Cox proportional hazards parameter at 36 months comparing children *asymptomatically* treated within the first 6 months of life is HR=-0.712 (p value = 0.057) and HR=-0.531 (p value = 0.459) for *symptomatically* treated children.

Looking at an expanded definition of treatment, the marginal additive difference of effect of triple therapy on time to C diagnosis among *asymptomatic* children who initiated therapy within 12 months was estimated at 0.030 (95% CI-0.250-0.225) and 0.152 (95% CI -0.153-0.240) among *symptomatic* children. The marginal log hazard ratio for children treated *asymptomatically* within the first 12 months of life is -0.106 (95% CI -1.054-0.739) and -0.587 (95% CI -1.217-0.480) for *symptomatically* treated children. The Cox proportional hazards parameter at 36 months comparing children *asymptomatically* treated within the first 12 months of life is HR=-0.516 (p value = 0.102) and HR=-0.008 (p value = 0.997) for *symptomatically* treated children.

The marginal additive difference of effect of triple therapy on time to C diagnosis or death among *asymptomatic* children who initiated therapy within 6 months was estimated at 0.140 (95% CI -0.380-0.490) and 0.062 (95% CI -0.178-0.319) among *symptomatic* children. The marginal log hazard ratio estimating the causal effect of triple therapy on time to C diagnosis or death among *asymptomatically* treated children was estimated as -0.514 (95% CI -17.719-1.086) and -0.204 (95% CI-1.485-0.556) among *symptomatically* treated children. The Cox proportional hazards parameter at 36 months comparing children *asymptomatically* treated within the first 6 months of life is HR=-0.570 (p value = 0.091) and HR=-0.291 (p value = 0.388) for *symptomatically* treated children.

The marginal additive difference of effect of triple therapy on time to C diagnosis or death among *asymptomatic* children who initiated therapy within 12 months was estimated at 0.061 (95% CI -0.208-0.267) and 0.052 (95% CI -0.178-0.319) for *symptomatically* treated children. The marginal log hazard ratio estimating the causal

effect of triple therapy on time to C diagnosis or death among *asymptomatically* treated children, -0.205 (95% CI -1.205-0.601), is greater than the marginal log hazard ratio among *symptomatically* treated children -0.168 (95% CI -1.485-0.556). The Cox proportional hazards parameter at 36 months comparing children *asymptomatically* treated within the first 12 months of life is HR=-0.204 (p value = 0.455) and HR=-0.154 (p value = 0.571) for children *symptomatically* treated.

Discussion

The optimal timing of initiation of HAART among HIV-infected children is an on-going debate and recommendations for treatment initiation vary.(63) Guidelines in the United States and in Europe in previous years were based on 2-5 year risk of disease progression estimates calculated from observational studies.(63) In contrast, more recent guidelines (2003) have been based on estimates of the 12-month risk of disease progression as reported by the HIV Pediatric Prognostic Markers Collaborative Study (HPPMCS) Group, a collection of studies conducted in the developed world or in high-resource settings.(64) The WHO, in 2006, developed clinical and immunologic guidelines for treatment initiation in asymptomatic children in resource-limited settings based on HPPMCS data.(4) In 2008, WHO amended their recommendations for treatment initiation for HIV-positive children partially as a result from an RCT in South Africa.(7) Though not statistically significant, the results from the present analysis may be interpreted as supportive of the current WHO treatment guidelines.

In an attempt to use the data available to fit the most efficient model, the present study estimated the causal effects of triple therapy initiated at different thresholds on time to C diagnosis, and time to C diagnosis or death. The sampled children were more likely to be non-White females. The unbalanced distribution of gender may not be important as studies in high-resource settings have previously found no difference in survival or disease progression between genders.(64) Additionally, studies have found increasing trends in risk of death among children of non-White ethnicities, though not statistically significant.(64) In the present study, over 70% of the children were either moderately or severely immunologically impaired at the time of triple therapy treatment initiation. Moreover, nearly 90% of the children remained untreated with HAART through the first 12 months of life. This could be both a reflection of the treatment guidelines at the time or the availability of some drugs at the time of disease progression.

To illustrate the causal assumptions for the present study, I could assign all my subjects to treatment a, triple therapy, and their censoring, $dN_2(t)$, to not censored at all t throughout the study. In turn, I will have my counterfactual outcome—treated with triple therapy and not censored. In an intent-to-treat approach, I assumed the initial treatment, whether mono, dual, or triple therapy, was unmodified throughout the study. For a detailed description of the algorithm employed to identify the children who were assumed to be infected in utero, see Appendix X. Additionally, in Appendix Y I have highlighted the algorithm that identified which children were asymptomatic and symptomatic at time of treatment initiation. For reference to treatment guidelines, see Appendix X2.

Traditional approaches in estimating the effects of treatments on time to event data, to include Cox proportional hazards models, often rely heavily on correct model specification. Using a data generating distribution approach avoids the inherent problems of employing traditional Cox methods, even if baseline covariates are included in the Cox model.(60) In turn, as opposed to using parameters selected by stepwise inclusion techniques or similar approaches, using an approach that uses parameters that are naturally selected based on the data allows for easier interpretation of the model and its parameters. For example, using a more traditional approach, the data suggest that ethnicity is a baseline confounder because it is significantly associated with both treatment initiation and C diagnosis (and death). Furthermore, immune function at treatment initiation was significantly associated with treatment and our events of interest. However, D/S/A did not select either of these variables for inclusion in our analysis, therefore preventing any unnecessary loss in precision in my estimates.

The causal effect of triple therapy among all children in delaying the time to a C diagnosis and/or death, regardless of immune status at treatment initiation, appears to be stronger among children who initiated therapy within 6 months rather than within 12 months of birth. Though no study has explored optimal treatment initiation in a pediatric HIV population using causal inference methods, this study's results seem to be in tune with previous, traditional analyses in early therapy initiation. Chiappini et al found that children who were treated with HAART early, as defined by: treatment initiation within 6 months of birth; category N, A, or B disease before treatment initiation; in immunologic category CDC 1 or 2 before treatment initiation, had significantly lower risk of progression to category C disease than not-early treated children (p value < 0.0001).(39) Similarly, Newell et al found that HIV positive children who started ART before 5 months of age were significantly more likely to have an improved immunologic response (time to a 20% increase in CD4 z score), after adjusting for immunocompetence status at treatment initiation.(38) It should be noted, however, that the authors were unable to find any added benefit in early treatment on sustained CD4 cell count after 6 months. In contrast, some laboratory research suggests that the positive immune response in children who have been treated early versus children who have been treated later in life may only be an artifact of younger age and not truly associated with early treatment.(64)

An RCT in South Africa recently concluded that children who were assigned to early treatment initiation, regardless of symptoms, had a significantly lower mortality risk and risk of progression to category C disease when compared to children whose treatments were deferred until they became symptomatic.(35) The results from this RCT were the catalyst in the WHO's revised treatment guidelines. Similarly, the results, though not statistically significant, from the present analysis suggest that the effect of triple therapy initiated within the first 12 months of life on time to death is stronger among *asymptomatic* children (12 Months: $\Psi HZ_{asymptomatic}$ $(p_0)(t_{36})$: -0.336 (95% CI -1.423-0.305)) than among *symptomatic* children (12 Months: $\Psi HZ_{symptomatic}$ $(p_0)(t_{36})$: -0.165 (95% CI -15.297-0.621)). This suggests that the mortality risk is, in fact, potentially reduced among HIV positive children who initiate HAART early rather than deferring treatment until symptoms arise. In contrast, the present study found that the effect of

triple therapy initiated within the first 6 or 12 months of life on time to C diagnosis is greater among *symptomatic* children (12 Months: $\Psi HZ_{symptomatic}$ $(p_0)(t_{36})$: -0.587 (95% CI -1.217-0.480)) than among *asymptomatic* children (12 Months: $\Psi HZ_{asymptomatic}$ $(p_0)(t_{36})$: -0.106 (95% CI -1.054-0.739)). This suggests that the gained benefit in initiating HAART by reducing the risk of disease progression is perhaps more fully realized among children who are already severely immune-compromised.

Despite the added efficiency in parameters as a result of using a causal inference approach, the present study has several limitations. Death was somewhat a rare outcome and the data do not provide enough power to show an effect of triple therapy. Though an attempt to increase the number of events seen during the follow up period was made by expanding the follow up to 60 months, there was no added significance in the findings. There were no data available regarding income, so the impact of socio-economic status could not be measured.

Though the data do not provide the opportunity to calculate the impact of socio-economic status, it is possible the children who were born to poor mothers were less likely to have access to the same health care as other children. Evaluating the use of prenatal care within the sample allows one to speculate about the proportion of children who had access to care. Furthermore, the proposed DAG showing the impacts of variables on triple therapy initiation and C diagnosis and/or death suggests that SES affects access to prenatal care and maternal ARVs at delivery and during pregnancy. Though I was not able to control for SES, I did have data on the child's race. Nearly three-quarters of this study sample was non-White ethnicity. Late initiation of HAART and low drug adherence have been seen more among non-White populations in the U.S. than among White populations, which in turn results in more advanced disease and increases in mortality risk.(65)

Though an attempt to isolate children who were likely infected in utero or at birth was made in the present analysis, there is still a risk that some children included in the study were infected postnatally via breastfeeding. Data on breastfeeding practices were not collected and it was assumed to have not occurred if women received prenatal care.

These data were extracted from the children's medical records, which may not offer a complete picture of maternal ARV exposure. Unfortunately, it was difficult to accurately estimate the proportion of mothers who were exposed to ARV during pregnancy or at delivery. Moreover, maternal HIV disease severity was not assessed. The severity of disease in mothers has been linked not only to the transmission of HIV, but also the severity in disease in vertically infected children.(15)

Viral load was not a variable analyzed in the present study as most of the children did not have these data collected. While CD4% is the primary source of clinical treatment guidance, viral load would likely also be an acceptable proxy for immuno-competence. Though immune function (as measured by clinical diagnoses and CD4 count or percent) at treatment initiation was adjusted, there could be residual confounding from a child's overall immune function that is unaccounted for. Thus, it is possible that these results

could be indicative of the unmeasured severity of immune suppression at treatment initiation rather than a true causal effect of triple therapy initiated at 6 or 12 months.

Losses to follow-up were a formidable threat in this study. Because the population was dynamic, children could have been any age at their first HIV clinic visit. I established t0, first day of study entry, as the birth date, assuming all children were infected in utero or at delivery. Unfortunately, some children may have first visited the clinic after the first 6 or 12 months, which would include them as untreated children in the study. At a minimum n=68 children were considered untreated in this analysis because their first study data were collected after the first 6 months. Furthermore, n=44 children were untreated in the analyses with *A* defined by treatment starting in the first 12 months because their first study data were collected after the first 12 months. These children could have moved into the catchment area later and may have already been receiving triple therapy. Additionally, it is possible that the healthiest children stop going to clinic after a period of time. Similarly, it is possible that sicker children either begin to go to the clinic after symptoms arise or cease going to the clinic because they were too ill. A sensitivity analysis on the data showed that the children who were censored before the end of the follow up period in the C diagnosis analyses were slightly more likely to be severely immuno-compromised at treatment initiation than children who were uncensored (44% vs 38%). Similarly, the children who were censored before the end of the follow up period in the C diagnosis/death analyses were more likely to be severely immuno-compromised at treatment initiation than children who were uncensored (45% vs 38%). Not surprisingly, this suggests that the children who were censored, which included death, reached a C diagnosis, or lost to follow up, were sicker than the children who survived beyond the follow-up time. Children who were treated within 6 months were slightly more likely to be censored than children who were untreated within 6 months of birth (33% vs 28%). This was likely due to the fact the children who started treatment early were sicker than children who delayed treatment. Children who were treated within 12 months were slightly less likely to be censored than children who were untreated within 12 months of birth (24% vs 31%). This was likely due to the possibility that the children who started treatment early were able to regain their health quicker than children who delayed treatment.

Adherence and resistance are always a threat in studies of ARV effectiveness. The present study was not able to assess the rates of adherence among the HIV positive children; this is likely heavily dependent on the mother's own ARV adherence. As a result, it is possible that some children who were started on triple therapy treatment early did not continuously receive the therapy, which in turn could create a drug resistance. If this child later restarted triple therapy, he may have poorer clinical and immunologic outcomes than other children who were continuously treated.

Tables

Table 1. Patient Demographics and Baseline Characteristics (N=217)

Baseline Covariate (W)	N (%)
Male Sex	95 (43.8)
White Ethnicity	61 (28.1)
Mother had Prenatal Care	110 (50.7)
Not Low Birth weight	153 (70.5)
Full-Term Pregnancy	123 (56.7)
Immune Status at Treatment Initiation	
Untreated	14 (6.5)
No or Mild Impairment	46 (21.2)
Moderate Impairment	70 (32.3)
Severe Impairment	87 (40.1)
HAART Initiation	
First 6 Months of Life	18 (8.3)
First 12 Months of Life	98 (45.2)
After First 12 Months of Life or Never	119 (54.8)

Table 2: Sample Baseline Characteristics and Associations With Triple Therapy Initiation (N=217)

Baseline Covariate (W)	Triple Therapy In First 6 Months (cOR)	P Value	Triple Therapy In First 12 Months (cOR)	P Value
Male Sex	0.62	0.35	0.87	0.73
White Ethnicity	0.49	0.27	0.28	0.05
Mother had Prenatal Care	1.47	0.20	1.41	0.17
Not Low Birth weight	0.63	0.36	0.81	0.64
Full-Term Pregnancy	0.84	0.57	0.86	0.56
Immune Status at Treatment Initiation	1.61	0.12	1.82	0.03

Table 3: Sample Characteristics and Associations With C Diagnosis With 36 Months

	Associations With C Diagnosis (cOR)	P Value	Associations With C Diagnosis or Death (cOR)	P Value
Male Sex	0.86	0.60	0.87	0.62
White Ethnicity	2.39	< 0.01	2.01	0.02
Mother had Prenatal Care	1.00	0.98	0.96	0.82
Not Low Birth weight	1.23	0.51	0.98	0.95
Full-Term Pregnancy	0.87	0.43	0.83	0.31
Immune Status at Treatment Initiation	2.17	< 0.01	1.90	< 0.01

Table 4: Comparison of Estimates from G-Comp and Traditional Techniques Therapy Initiation Under 6 Months or Under 12 Months (36 Months of Follow-up)

	Triple Therapy Initiated First 6 Months				**Triple Therapy Initiated First 12 Months**			
Outcome	Marginal Additive Difference $\Psi AD(p_0)(tk)$	Marginal log Hazard Ratio $\Psi HZ(p_0)(tk)$	Mean Marginal Log HR over tk	Cox PH Model	Marginal Additive Difference $\Psi AD(p_0)(tk)$	Marginal log Hazard Ratio $\Psi HZ(p_0)(tk)$	Mean Marginal Log HR over tk	Cox PH Model
	Overall				Overall			
C Diagnosis	0.120 (-0.127, 0.270)[1]	-0.466 (-1.457, 0.397)[1]	-0.470 (-1.468, 1.353)[1]	-0.476 (-1.486, 0.534)[1]	0.087 (-0.065, 0.151)[1]	-0.321 (-0.588, 0.212)[1]	-0.324 (-0.617, 0.217)[1]	-0.407 (-0.876, 0.062)[1]
C Diagnosis or Death	0.108 (-0.110, 0.311)	-0.369 (-1.588, 0.300)	-0.369 (-1.634, 0.822)	-0.346 (-1.249, 0.558)	0.055 (-0.150, 0.158)	-0.180 (-0.599, 0.445)	-0.179 (-0.625, 0.468)	-0.431 (-0.875, 0.012)
	Asymptomatic				Asymptomatic			
C Diagnosis	0.115 (-0.327-0.446)	-0.429 (-16.23, 0.948)	-0.429 (-16.27, 1.438)	-0.712 (-1.445, 0.022)	0.030 (-0.250, 0.225)[2]	-0.106 (-1.054, 0.739)[2]	-0.106 (-1.138, 2.105)[2]	-0.516 (-1.134, 0.102)[2]
C Diagnosis or Death	0.140 (-0.380-0.490)[2]	-0.514 (-17.72, 1.086)[2]	-0.515 (-17.76, 1.316)[2]	-0.570 (-1.231, 0.091)[2]	0.061 (-0.208, 0.267)[2]	-0.205 (-1.205, 0.601)[2]	-0.206 (-1.279, 1.906)[2]	-0.471 (-0.523, 0.004)[2]
	Symptomatic				Symptomatic			
C Diagnosis	0.146 (-0.214-0.355)[3]	-0.563 (-14.34-0.653)[3]	-0.570 (-14.42, 0.755)[3]	-0.634 (-2.041, 0.774)[3]	0.152 (-0.153-0.240)[3]	-0.587 (-1.217-0.480)[3]	-0.594 (-15.65, 1.529)[3]	-0.068 (-0.610, 0.474)[3]
C Diagnosis or Death	0.062 (-0.236-0.403)	-0.204 (-14.75-0.683)	-0.204 (-14.73, 0.766)	-0.291 (-0.951, 0.369)	0.052 (-0.255-0.319)[3]	-0.168 (-1.526-0.709)[3]	-0.171 (-15.36, 2.807)[3]	-0.214 (-0.749, 0.320)[3]

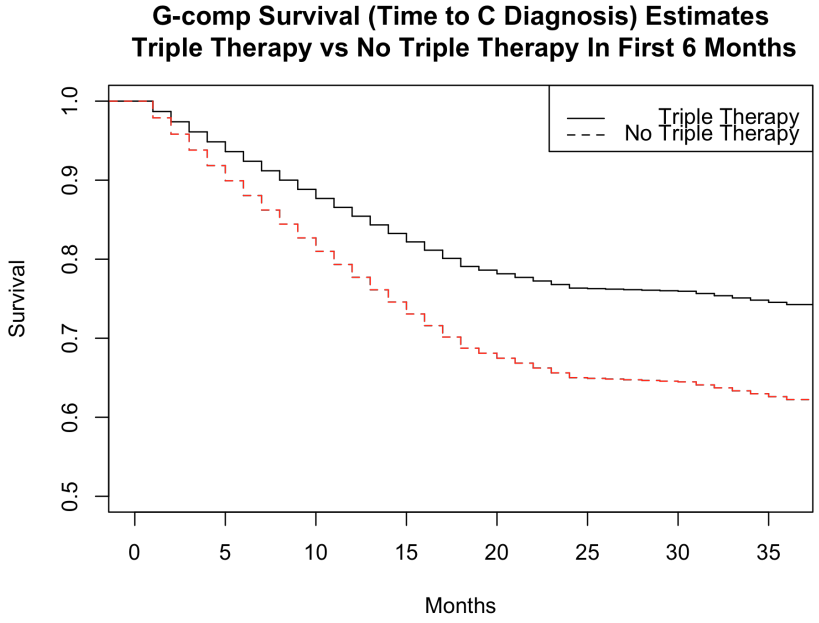
1 Adjusted for sex, race, and pregnancy term

2 Adjusted for pregnancy term

3 Adjusted for an interaction between prenatal care and a time indicator variable for 25-30 months and an interaction between sex and pregnancy term

Figures

Figure 1:



References

1. De Cock KM, Fowler MG, Mercier E, de Vincenzi I, Saba J, Hoff E, Alnwick DJ, Rogers M, Shaffer N. Prevention of mother-to-child HIV transmission in resource-poor countries: translating research into policy and practice. Jama 2000;283(9):1175.

2. HIV/AIDS Update [Internet]. [Access December 1, 2009];Available from: http://www.unaids.org/

3. UNAIDS/World Health Organization. HIV/AIDS Update. Available from: http://data.unaid- s.org/pub/EPISlides/2007/2007_epiupdate_en.pdf

4. World Health Organization. Report of the WHO Technical Reference Group, Paediatric HIV/ART Care Guideline Group Meeting. Geneva, Switzerland: 2008.

5. Maldonado Y, Wang N, Caldwell B. Factors associated with early clinical recognition of children with perinatal human immunodeficiency virus infection. Northern California Pediatric HIV Consortium. J Infect Dis 1995;171:689-692.

6. Shah B, Laupacis A, Hux J, Austin P. Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review. Journal of Clinical Epidemiology 2005;58(6):550-559.

7. World Health Organization. Guidelines for Anitretroviral Therapy for HIV Infection in Adults and Adolescents: Recommendations for a public health approach. 2006.

8. BARRE-SINOUSSI F, Chermann C, Rey F, Nugeyre M. Isolation of a T-Lymphotropic Retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science 1983;220(4599):868-871.

9. Gallo R, Sarin P, Gelmann E, Robert-Guroff M. Isolation of Human T-Cell Leukemia Virus in Acquired Immune Deficiency Syndrome (AIDS). Science 1983;220(4599):865-867.

10. Rutherford GW, Oliva GE, Grossman M, Green JR, Wara DW, Shaw NS, Echenberg DF, Wofsy CB, Weinstein DH, Stroud F, others. Guidelines for the control of perinatally transmitted human immunodeficiency virus infection and care of infected mothers, infants and children. Western Journal of Medicine 1987;147(1):104.

11. Schulte J, Dominguez K, Sukalac T, et al. Declines in low birth weight and preterm birth among infants who were born to HIV-infected women during an era of increased use of maternal antiretroviral drugs: pediatric spectrum of HIV disease, 1989-2004. Pediatrics 2007;119:e900-e906.

12. Taha TE, Kumwenda NI, Hoover DR, Kafulafula G, Fiscus SA, Nkhoma C, Chen S, Broadhead RL. The impact of breastfeeding on the health of HIV-positive mothers and their children in sub-Saharan Africa. Bulletin of the World Health Organization 2006;84:546–554.

13. Adair T. UNMET NEED FOR CONTRACEPTION AMONG HIV-POSITIVE WOMEN IN LESOTHO AND IMPLICATIONS FOR MOTHER-TO-CHILD TRANSMISSION. J. Biosoc. Sci 2008;41(02):269.

14. Embree J, Bwayo J, Nagelkerke N, Njenga S, Nyange P, Ndinya-Achola J, Pamba H, Plummer F. Lymphocyte subsets in human immunodeficiency virus type 1-infected and uninfected children in Nairobi. The Pediatric infectious disease journal 2001;20(4):397.

15. Petru A. [February 2009];

16. AYOUBA A, NERRIENET E, MENU E, LOBE MM, THONNON J, LEKE RJ, BARRE-SINOUSSI F, MARTIN P, CUNIN P. Mother-to-child transmission of human immunodeficiency virus type 1 in relation to the season in Yaounde, Cameroon. The American journal of tropical medicine and hygiene 2003;69(4):447.

17. Mphatswe W, Blanckenberg N, Tudor-Williams G, Prendergast A, Thobakgale C, Mkhwanazi N, McCarthy N, Walker BD, Kiepiela P, Goulder P. High frequency of rapid immunological progression in African infants infected in the era of perinatal HIV prophylaxis. Aids 2007;21(10):1253.

18. Guay LA, Musoke P, Fleming T, Bagenda D, Allen M, Nakabiito C, Sherman J, Bakaki P, Ducar C, Deseyve M, others. Intrapartum and neonatal single-dose nevirapine compared with zidovudine for prevention of mother-to-child transmission of HIV-1 in Kampala, Uganda: HIVNET 012 randomised trial. The Lancet 1999;354(9181):795–802.

19. Global Partners Forum. Achieving Universal Access to Comprehensive PMTCT Services. [Internet]. [date unknown];Available from: http://www.procaare.org/archive/ procaare/200711/msg00013.php

20. Reynolds HW, Janowitz B, Homan R, Johnson L. The Value of Contraception to Prevent Perinatal HIV Transmission. Sexually Transmitted Diseases 2006;33(6):350-356.

21. Sweat MD, O'Reilly KR, Schmid GP, Denison J, de Zoysa I. Cost-effectiveness of nevirapine to prevent mother-to-child HIV transmission in eight African countries. AIDS 2004;18(12):1661-1671.

22. Rutenberg N, Baek C. Field experiences integrating family planning into programs to prevent mother-to-child transmission of HIV. Studies in Family Planning 2005;36(3):235–245.

23. De Martino M, Tovo PA, Balducci M, Galli L, Gabiano C, Rezza G, Pezzotti P, others. Reduction in mortality with availability of antiretroviral therapy for children with perinatal HIV-1 infection. Jama 2000;284(2):190.

24. The European Collaborative Study. Fluctuations in Symptoms in Human Immunodeficiency Virus-Infected Children: The First 10 Years of Life. PEDIATRICS 2001;108(1):116-122.

25. Blanche S, Newell ML, Mayaux MJ, et al. Morbidity and Mortality in European Children Vertically Infected by HIV-1. The French Pediatric HIV Infection Study Group and European Collaborative Study. J Acquir Immune Defic Syndr Hum Retrovirol. 1997; 14(5): 442-50.

26. Gortmaker SL, Hughes M, Cervia J, Brady M, Johnson GM, Seage III GR, Song LY, Dankner WM, Oleske JM, others. Effect of combination therapy including protease inhibitors on mortality among children and adolescents infected with HIV-1. The New England Journal of Medicine 2001;345(21):1522.

27. Patel K, Hernan MA, Williams PL, Seeger JD, McIntosh K, Dyke RBV, Seage III GR, Pediatric AIDS Clinical Trials Group 219/219C Study Team. Long□Term Effectiveness of Highly Active Antiretroviral Therapy on the Survival of Children and Adolescents with HIV Infection: A 10□Year Follow□Up Study. CLIN INFECT DIS 2008;46(4):507-515.

28. Gibb DM, Duong T, Tookey PA, Sharland M, Tudor-Williams G, Novelli V, Butler K, Riordan A, Farrelly L, Masters J, others. Decline in mortality, AIDS, and hospital admissions in perinatally HIV-1 infected children in the United Kingdom and Ireland. Bmj 2003;327(7422):1019.

29. Borleffs J, Boer K, Boucher C. Antiretroviral Treatment in the Netherlands [Internet]. [date unknown].Available from: www.cbo.nl

30. Centers for Disease Control and Prevention. Guidelines for the use of antiretroviral agents in pediatric HIV infection. [Internet]. [date unknown];Available from: www.aidsinfo.nih.gov

31. De Martino M, Tovo P, Giaquinto C, de Rossi A, Galli L, Gabiano C. Italian Guidelines for antiretroviral therapy in children with human immunodeficiency virus-type 1 infection. Acta Paediatrica 1999;88:228-32.

32. Sharland M, Blanche S, Castelli G, Ramos J, Gibb DM, others. PENTA guidelines for the use of antiretroviral therapy, 2004. HIV medicine 2004;5:61.

33. Draenert R, Goebel F. Recommendations and prospects of antiretroviral therapy. Dtsch Med Wochenschr 2001;126(18):539-43.

34. Paediatrica A. Management of vertically HIV-infected children in Europe. Acta Paediatrica;92(2):246–250.

35. Violari A, Cotton MF, Gibb DM, Babiker AG, Steyn J, Madhi SA, Jean-Philippe P, McIntyre JA. Early antiretroviral therapy and mortality among HIV-infected infants. New England Journal of Medicine 2008;359(21):2233.

36. Ananworanich J, Kosalaraksa P, Siangphoe U, Engchanil C, Pancharoen C, Lumbiganon P, Intasan J, Apateerapong W, Chuenyam T, Ubolyam S, others. A feasibility study of immediate versus deferred antiretroviral therapy in children with HIV infection. AIDS Research and Therapy 2008;5(1):24.

37. Five year follow up of vertically HIV infected children in a randomised double blind controlled trial of immediate versus deferred zidovudine: the PENTA 1 trial. Archives of Disease in Childhood 2001;84(3):230-236.

38. Membership I, In VC, Falsey AR, Walsh EE, Capellan J, Gravenstein S, Zambon M, Yau E, Gorse GJ, Edelman R, others. CD4 cell response to antiretroviral therapy in children with vertically acquired HIV infection: is it associated with age at initiation? The Journal of Infectious Diseases 2006;193(7):954–962.

39. Chiappini E, Galli L, Tovo P. Five year follow up of children with perinatal HIV-1 infection receiving early highly active antiretroviral therapy. BMC Infectious Diseases 2009;9:140.

40. Keating MJ, Flinn I, Jain V, Binet JL, Hillmen P, Byrd J, Albitar M, Brettman L, Santabarbara P, Wacker B, others. Therapeutic role of alemtuzumab (Campath-1H) in patients who have failed fludarabine: results of a large international study. Blood 2002;99(10):3554.

41. Age- and Time-related Changes in Extracellular Viral Load in Children Vertically Infected by Human Immunodeficiency Virus. Pediatr Infect Dis. 1996; 15(12): 1087-91.

42. Palumbo PE, Kwok S, Waters S, Wesley Y, Lewis D, McKinney N, Bardeguez A, Connor EM, Oleske JM. Viral measurement by polymerase chain reaction-based assays in human immunodeficiency virus-infected infants. The Journal of pediatrics 1995;126(4):592–595.

43. Shearer WT, Quinn TC, LaRussa P, Lew JF, Mofenson L, Almy S, Rich K, Handelsman E, Diaz C, Pagano M, others. Viral load and disease progression in infants infected with human immunodeficiency virus type 1. The New England journal of medicine 1997;336(19):1337.

44. Burns D, Mofenson L. Paediatric HIV-1 Infection. The Lancet 1999;354((Supplement II)):1.

45. de Martino M, Tovo PA, Galli L, Gabiano C, Chiarelli F, Zappa M, Gattinara GC, Bassetti D, Giacomet V, Chiappini E, others. Puberty in perinatal HIV-1 infection: a multicentre longitudinal study of 212 children. Aids 2001;15(12):1527.

46. Macmillan C, Magder L, Browers P, et al. Head growth and neurodevelopment of infants born to HIV-1-infected drug-using women. Neurology 2001;57:1402-1411.

47. Dubé MP, Sprecher D, Henry WK, Aberg JA, Torriani FJ, Hodis HN, Schouten J, Levin J, Myers G, Zackin R, others. Preliminary guidelines for the evaluation and management of dyslipidemia in adults infected with human immunodeficiency virus and receiving antiretroviral therapy: recommendations of the Adult AIDS Clinical Trial Group Cardiovascular Disease Focus Group. Clinical Infectious Diseases 2000;31(5):1216–1224.

48. Alimenti A, Burdge DR, Ogilvie GS, Money DM, Forbes JC. Lactic acidemia in human immunodeficiency virus-uninfected infants exposed to perinatal antiretroviral therapy. The Pediatric infectious disease journal 2003;22(9):782.

49. Poirier MC, Divi RL, Al-Harthi L, Olivero OA, Nguyen V, Walker B, Landay AL, Walker VE, Charurat M, Blattner WA. Long-term mitochondrial toxicity in HIV-uninfected infants born to HIV-infected mothers. JAIDS Journal of Acquired Immune Deficiency Syndromes 2003;33(2):175.

50. Sharland M, Gattinara di Zub GC, Tomas Ramos J, Blanche S, Gibb DM. PENTA guidelines for the use of antiretroviral therapy in paediatric HIV infection. HIV medicine 2002;3(3):215–226.

51. Patel K, Hernán M, Williams P, Seeger J, McIntosh K, Dyke R, Seage III G, Pediatric AIDS Clinical Trials Group 219/219C Study Team. Long□Term Effects of Highly Active Antiretroviral Therapy on CD4 Cell Evolution among Children and Adolescents Infected with HIV: 5 Years and Counting. CLIN INFECT DIS 2008;46(11):1751-1760.

52. Virus HI, Addenda OA, Codes I. ICD-9-CM.

53. Rubin D. Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies. 1990;

54. Robins J. Marginal Structural Models versus structural nested models as tools for causal inference. In: Statistical Models in Epidemiology: The Environment and Clinical Trials. New York, NY: Springer-Verlag; 1992 p. 95-134.

55. Greenland S. An introduction to instrumental variables for epidemiologists. International Journal of Epidemiology 2000;29:722-729.

56. Jewell N. Statistics for Epidemiology. Chapman and Hall; 2004.

57. Hogan JW, Lancaster T. Instrumental variables and inverse probability weighting for causal inference from longitudinal observational studies. Statistical Methods in Medical Research 2004;13(1):17-48.

58. Kleinbaum D, Klein M. Survival Analysis. Second. Springer-Verlag; 2005.

59. Hernan M, Brumback B, Robins J. Marginal Structural Models to estimate the joint causal effect of nonrandomized treatments. Journal of the American Statistical Association 2001;96(454):440-448.

60. Stitleman O. Examining the Causal Effect of Non-Nucleoside Reverse Transcriptase Inhibitors Among cART-Treated Adults in Botswana Using Targeted Maximum Likelihood Estimation. (In Press) [2009];

61. The Working Group on Antiretroviral Therapy and Medical Management of HIV Infected Children. Guidelines for the use of antiretroviral agents in pediatric HIV infection. [Internet]. US Department of Health and Human Services;.Available from: http://aidsinfo.nih.gov/ContentFiles/PediatricGuidelines.pdf

62. Sinisi S, van der Laan MJ. Loss-based cross-validated deletion/substitution/addition algorithms in estimation. UC Berkeley Division of Biostatistics Working Paper Series 2004;

63. Verweel G, Saavedra-Lozano J, van Rossum AMC, Ramilo O, de Groot R. Initiating Highly Active Antiretroviral Therapy in Human Immunodeficiency Virus Type 1-Infected Children in Europe and the United States. The Pediatric Infectious Disease Journal 2006;25(11):987-994.

64. HIV Paediatric Prognostic Markers Collaborative Study Group. Short-term risk of disease progression in HIV-1-infected children receiving no antiretroviral therapy or zidovudine monotherapy: a meta-analysis. Lancet 2003;362:1605-11.

65. Campbell SE, Stone WL, Whaley SG, Qui M, Krishnan K. Gamma tocopherol upregulates peroxisome proliferator activated receptor (PPAR) gamma expression in SW 480 human colon cancer cell lines. BMC Cancer 2003;3(1):25.

Appendices

The CDC's clinical categories of HIV disease among children helps determine the progress of the disease and establish immune suppression. Though the CDC's treatment guidelines include all HIV positive children under 12 months, previous algorithms were employed to determine treatment eligibility. Essentially, severe disease was determined by a combination of clinical presentations and immunologic measurements (CD4 count or preferably CD4%). These criteria are outlined in Attachment B. In contrast, the Pediatric European Network for the Treatment of AIDS (PENTA) group's treatment recommendations are less aggressive; essentially, treatment is recommended among infants if they have a C diagnosis or CD4% less than 20% (see Appendix C).(50)

Appendix A: WHO's Treatment Initiation Guidelines Among HIV+ Children (pre-2008)

Immunological marker[a]	Age-specific recommendation to initiate ART[b] [A (I)]*			
	≤11 months	12 months to 35 months	36 months to 59 months	≥5 years
%CD4+[c]	<25%	<20%	<15%	<15%
CD4 count[c]	<1500 cells/mm^3	<750 cells/mm^3	<350 cells/mm^3	<200 cells/mm^3

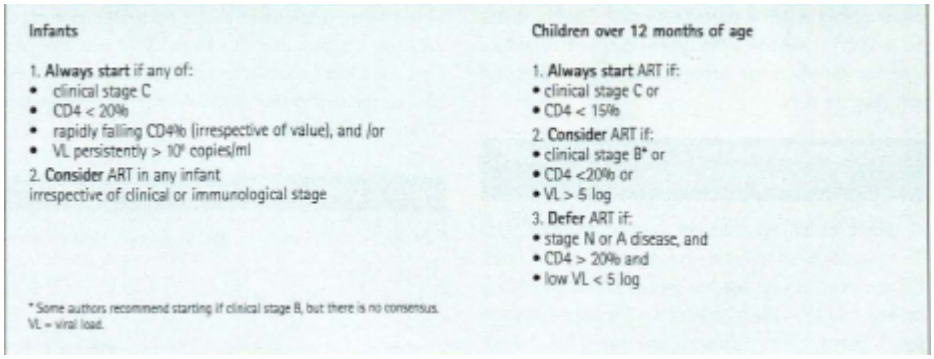
Appendix B: CDC Definition of Clinical HIV Disease Progression Among Children

	Clinical Categories			
Immunologic Categories	**N: No Signs/ Symptoms**	**A: Mild Signs/ Symptoms**	**B: Moderate Signs/ Symptoms**	**C: Severe Signs/ Symptoms**
1. No evidence of suppression <12 mo, ≥1,500 cells/µl ≥25% 1–5 y, ≥1,000 cells/µl ≥25% 6–12 y, ≥500 cells/µl ≥25%	N1	A1	B1	C1
2. Evidence of moderate suppression <12 mo, 750–1,499 cells/µl 15–24% 1–5 y, 500–999 cells/µl 15–24% 6–12 y, 200–499 cells/µl 15–24%	N2	A2	B2	C2
3. Severe suppression <12 mo, <750 cells/µl <15% 1–5 y, <500 cells/µl <15% 6–12 y, <200 cells/µl <15%	N3	A3	B3	C3

* If infection status is unknown but infant is exposed, use E (eg, EN2)

Modified from Centers for Disease Control and Prevention. 1994 revised classification system for human immunodeficiency virus infection in children less than 13 years of age. MMWR Morb Mortal Wkly Rep. 1994;43(No. RR-12):1–19.

Appendix C: PENTA's Treatment Initiation Guidelines For HIV+ Children (2002)



Infants

1. **Always start** if any of:
- clinical stage C
- CD4 < 20%
- rapidly falling CD4% (irrespective of value), and /or
- VL persistently > 10^5 copies/ml

2. **Consider** ART in any infant irrespective of clinical or immunological stage

Children over 12 months of age

1. **Always start** ART if:
- clinical stage C or
- CD4 < 15%

2. **Consider** ART if:
- clinical stage B* or
- CD4 <20% or
- VL > 5 log

3. **Defer** ART if:
- stage N or A disease, and
- CD4 > 20% and
- low VL < 5 log

* Some authors recommend starting if clinical stage B, but there is no consensus.
VL = viral load.

Appendix D:
Two separate, primary analyses were explored regarding timing of initiation of HAART—a) starting triple therapy in the first six months of life, b) starting triple therapy in the first twelve months of life. However, since the early 1990s as researchers have learned more about the spread of the infection within children, the majority of HIV infections in children are attributed to mother-to-child-transmission (MTCT). Specifically, there are three avenues of infection via MTCT: in utero; at delivery; and through breastfeeding. It is estimated that the probability of infection in utero and at time of delivery is approximately 15-30%, though breastfeeding from 18-24 months the overall probability of vertical transmission increases to about 30-45%.(1)

In order to establish t=0, that is the first day of follow-up for each child, I employed an algorithm to include only children who were assumed to be infected in utero or delivery. In turn, I excluded n=60 children who were likely infected postnatally via breastfeeding, which could have occurred any time throughout breastfeeding. This approach included children whose mothers were assumed to be knowledgeable about their infection as breastfeeding was likely discouraged among these women. To identify these children, an algorithm was applied that identified children whose mothers showed HIV symptoms during their pregnancy, symptoms during their delivery, were known to have taken HIV medication, or were known to have received prenatal care, as mandatory HIV testing for pregnant women began in 1987.

Appendix E:
For my subgroup analyses of children who were asymptomatic and symptomatic at time of treatment initiation, I constructed an algorithm to identify their disease status. The definition of asymptomatic children was adapted from the CDC's definition of severe/moderate/mild immune suppression among children (see Appendix B). In short, somewhat similar to PENTA's previous treatment initiation guidelines (see Appendix C), asymptomatic children were described as not having a C diagnosis and not having a CD4% below 15%. In the absence of CD4% data, the CD4 count as it relates to the immune competence age-specific threshold was used. To ensure children who began ART were asymptomatic, another algorithm was applied to identify children who were diagnosed with a C diagnosis 4 weeks or more after the initial treatment. Previously, it

has been shown that at least four weeks of ARV treatment are needed to have any clinical effectiveness.(61) Additionally, this algorithm identified the CD4% or CD4 counts within four weeks of ARV treatment initiation to ensure that the immunological data (CD4% and CD4 count) at (or near) treatment initiation were likely unaffected by ARV initiation.

B.3 The Effect of Highly Active Antiretroviral Therapy Use Among HIV Positive Children on the Hazard of AIDS or Death Using Calendar Year as an Instrumental Variable

This study will be submitted for publication in the journal AIDS in Spring 2010.

The Effect of Highly Active Antiretroviral Therapy Use Among HIV Positive Children on the Hazard of AIDS Using Calendar Year as an Instrumental Variable

Word count exclusive of abstract, tables and figures: 3438 (needs to be 3500)

Word count of abstract: 262

Abstract

Objective: In the absence of randomized data, epidemiologists are tasked with asserting causal inference from observational data. Previously calendar periods have been used as a proxy for Highly Active Antiretroviral Therapy (HAART) use, first introduced in the U.S. in mid-1996 (1-12). This approach, referred to as an instrumental variable analysis, can be biased because of misclassification of HAART use, also known as non-compliance adjustment (1).

Design: Retrospective clinical cohort.

Methods: We performed an adapted instrumental variable analysis of 267 perinatally HIV-infected children living in Northern California from 1988 to 2009 to estimate the causal effect of HAART on the hazard of progression to CDC Category C diagnosis.

Results: During 61,860 person-days, 100 HIV-positive children received their initial C diagnosis. The intention to treat (ITT) rate ratio of C diagnosis comparing the pre-HAART and HAART eras was estimated at 2.74 (95% CI 1.50 - 5.01). An instrumental variable estimator was used to adjust for HAART use misclassification, yielding an instrumental variable rate ratio of 4.99 (95% CI 2.73 - 9.12). A secondary analysis followed 65 HIV-positive children previously diagnosed with a B illness until they received their first C diagnosis. The ITT rate ratio was estimated at 3.08 (95% CI 1.52 – 6.23) while the instrumental variable estimator yielded a rate ratio of 4.74 (95% CI 2.34 - 9.58). Weighting by the inverse probability of calendar era given selected covariates was performed, which did not significantly alter the results.

Conclusions: Noncompliance adjustments in instrumental variable analyses may help bridge the gap in understanding the evidence of HAART's effectiveness from both randomized controlled trial.

Introduction

Following its advent in 1996, highly active antiretroviral therapy (HAART) for the pediatric HIV-infected population has proved to be a significant factor in delaying time to acquired immune deficiency syndrome (AIDS) and death (13-17).

The gold standard in research, results from randomized controlled trials (RCTs) are often the first steps in estimating treatment effects of specific therapies. Research from an RCT has established HAART's protective effect on time to AIDS or death before researchers had the ability to estimate HAART's population-level effect (13).

The protective effect of HAART on delaying time to an AIDS-defining illness and/or death is well-established in observational study findings in several study settings and countries and support findings from clinical trial research (14-17). However, no study has been published that applied an instrumental variable analysis on a pediatric population. Furthermore, results from observational studies have allowed researchers to estimate HAART's effect on a population level (7).

Data from RCTs are most often analyzed using the intention to treat (ITT) principle. Under this rule, once person A is randomized to a treatment X, he should be included in any future analyses comparing treatment assignment arms as if he actually received treatment X. In this respect, ITT is actually determining the effect of assigning person A to treatment X. In a well-run RCT in which few participants are censored or change treatment assignments, the causal effect of the treatment and ITT effects are similar. Alternatively, if randomization failed or treatment assignments were rarely followed, non-compliance adjustments are needed to ensure the results are a true reflection of the causal effect of treatment X (18-20). In the absence of a non-compliance adjustment, the estimates produced using an ITT approach will likely be biased toward the null hypothesis if 100% compliance is not realized.

Though results seen in RCTs, with strict random allocation principles, are often supported by observational data, confounding biases are a persistent, inherent problem in observational studies. For example, while an RCT randomizes study participants to receive treatments, the physicians or the participants themselves select receipt of treatments in an observational study. Therefore, an argument for causality is difficult in observational studies because the effect seen could either be a result of the treatment or it could be a result of the reason for selecting the treatment (21).

Though RCTs provide the best epidemiologic scenario for estimating the causal effect of therapy on disease progression, often researchers are somewhat restricted to the analysis of observational data because of limitations in funding, recruiting adequate sample sizes, as well as the ethical concerns inherent with implementing

RCTs. Thus, instrumental variables (IV) was a method developed to deal with the difficulties with asserting causality in observational data. Initially implemented in econometric theory(22), IV account for the lack of randomization and sparse data found in the economic literature. Used initially in HIV/AIDS data in 2001, Tarwater et al used calendar year as an external time-dependent variable in assessing the effect of HAART on AIDS diagnoses (9). Of causal inference techniques, IV are underutilized as a means of accounting for biases commonly found in observational data (23).

Like the principle of random allocation to treatments in RCTs, traditionally IV only affect the outcome through their effect on the treatment or exposure alone (21). Moreover, referring to two principle concepts--the counterfactual framework and the randomization assumption, all counterfactual observations are independent of the process of treatment allocation. Additionally, the variation in the identified instrument is assumed to be substantial enough to cause variation in the treatment. The first assumption is particularly important as if it is rejected and the instrument is directly related to the outcome, the results will be biased (24). Though this assumption may not be directly evaluable, the use of directed acyclic graphs (DAGs) can help justify the use of a specific IV.

Figure 1 illustrates that calendar year is only related to the outcome (death) through the exposure (HAART), making calendar year an ideal IV. As a result, the researcher will be able to estimate how much the variation in the antiretroviral regimen is explained by the calendar year. Detels et al has previously shown that confounders of HAART's effect, including HIV-related, non-HAART therapies and use of health care, are not supported as neither of these factors vary significantly across calendar periods (7). Regarding the assumption of assumed, incurred variation, the results from more traditional approaches such as ordinary least squares (OLS) will be similar to IV results if the variation seen in the instrument used does not create variation in the treatment variable (24).

An important contribution of IVs is their ability to control known as well as unknown confounding. The details of individual health characteristics and disease severity are usually of no consequence if an RCT is performed because their influence would be mitigated through the random allocation assumption. In observational studies, however, these particular details can incur confounding bias on the estimated effect (e.g. ART and death).

While IV analysis is not a panacea for inherent problems in the HIV/AIDS observational data, there are benefits to using the HAART era (post HAART introduction in 1996) as a proxy for actual HAART use. In this approach, the confounding bias seen in observational studies where there are differences in underlying health conditions between the treated and untreated populations is minimized (24). In situations where there is no misclassification of HAART exposure, calendar year performs well as a proxy for actual HAART use. However information bias can be introduced if the calendar period, pre-1998 and 1998

onwards in the present case, is not representative of actual HAART exposure (1). Furthermore, some covariates may have associations with calendar period *and* the outcome of interest, violating a principle concept of IV—their independence of the outcome given treatment and covariates that affect both treatment and the outcome.

To account for possible information bias introduced by calendar year, noncompliance correction in RCTs has been previously adapted for use in observational HIV/AIDS data(1). In clinical trials, this correction method is useful in situations where randomization fails(20). Unfortunately, often treatment contamination (the use of the intervention among controls) remains a possibility in RCTs, especially in prevention trials, or is an inherent attribute of the trial design(1). As a result, baseline risks for compliers and non-compliers may be different, negating the benefits of randomization. As opposed to an ITT analysis, which analyzes subjects as if they actually received the intervention they were randomized to receive, non-compliance methods have been developed for these circumstances of contamination.

To account for the potential for misclassification of HAART exposure and to adjust for covariates that may be associated with calendar period and AIDS or death, we employed methods (1) to estimate the effect of HAART on AIDS or death in a population of perinatally HIV-infected children in Northern California.

Methods

Study Population

The Northern California Pediatric Spectrum of Disease (PSD) project is a prospective, longitudinal, multi-center pediatric HIV surveillance project. The surveillance area includes 12 counties in Northern California, with a total population of approximately 10 million. All HIV-infected children included in this study were identified through the PSD surveillance system, which included hospital surveillance and record matching, as described previously (15). Study nurses went to each study site in consecutive 6 month intervals. At these visits, chart abstractions performed on all perinatally exposed infants and these infants were then followed longitudinally. All children born to HIV-infected mothers were followed longitudinally until definitive HIV serostatus was determined based on case definitions from the US Centers for Disease Control and Prevention (CDC) (25). Thereafter, cumulative chart updates were done on known HIV-infected children at minimum 6 month intervals. XX (number of infants) perinatally infected infants enrolled from 1988 through 2009 who reached the primary endpoint of a 1994 CDC Revised Classification System C Diagnosis(25) are included in the present analysis.

Exposure Assessment

By current guidelines, antiretroviral use is optimal with three drugs from at least two classes(26). Triple antiretroviral therapy was defined as 3 antiretroviral agents of any class, so long as there were two classes represented. All treatment decisions were made at the discretion of the treating physician. Treatment regimens were recorded at each study visit and entered into a standardized database.

To estimate each child's person-time within calendar periods, we created an indicator variable for HAART calendar eras: pre-HAART (before 1998) and HAART (1998 and beyond). As defined by Greenland, (22) the properties of an IV are assumed to be satisfied in the present analysis. Specifically, calendar period meets three conditions: 1) is independent of unmeasured confounding between HAART and outcome; 2) is associated with HAART; 3) is independent of the outcome given HAART and unmeasured confounding between HAART and the outcome.

Adapting these noncompliance methods in RCTs for use in observational studies, requires estimating the rate rather than the risk (20). Additionally, there remains the possibility that condition 3, described above, may be too restrictive as there may be some covariates, V, that are related with both calendar period, Z, and the outcome, Y. Potential examples of these variables are length of infection, age at seroconversion, or race/ethnicity. Therefore, we attempted to remove these possible associations by creating a weighted pseudo-population by using inverse probability weighting (27). As a result, the new observations in this pseudo-population are now weighted by the inverse of the probability of calendar period given V.

Endpoint Ascertainment

Two populations were defined by outcome of interest. To explore the dynamics of disease progression, we followed children from their first day of follow up until their first Category C diagnosis. We also looked at a subgroup of children from their first date of a Category B diagnosis and followed them until their first Category C diagnosis.

While children vertically infected with HIV can seroconvert *in utero*, at delivery, or post-partum, we assumed all children were infected at birth. The presence of a CDC Category C diagnosis determined by a clinician and noted within the child's medical records. The CDC 1994 Revised Classification System system(insert ref), was used to identify children who reached Category C diagnosis. See Appendix A for details. Censoring occurred at three potential time points—time of C diagnosis, time of death, or the end of the study period.

Statistical Methods

HAART use is indexed by subscript x, where 1 is HAART use and 0 non-HAART use. When a child experiences the outcome, a diagnosis of an AIDS defining illness or death, the script $D_{ijxz} = 1$ is used to indicate that child i experienced the outcome between visits j and $j + 1$ during calendar period z while using therapy x. $D_{ijxz} = 0$ indicates that child did not experience the outcome. The number of person-days that each child i contributed between visits $j - 1$ and j while using therapy x during calendar period z is indicated by T_{ijxz}. Identified covariates for each child i at visit j, variables include both time-varying and time-fixed alike, are included in vector V_{ij}. Among the possible covariates included in V_{ij} are race/ethnicity, age at randomization, age (time since randomization).

In a traditional IV approach, observational data would be analyzed with a standard ITT approach—comparing rates between calendar periods (before/during the HAART era), regardless of actual HAART use. Between calendar periods, we compared rates of therapy "compliers": non-HAART users in pre-HAARTera and those who used HAART in the HAART-era (19). See Appendix B for statistical details.Therapy "non-compliers" refer to those who use HAART during the pre-HAART era, or do not receive HAART in the HAART-era. In the absence of contamination, or non-compliers, the IV estimator equals the ITT estimator.

Inverse probability of calendar period weights, W_{ij+Z} , were estimated to adjust for V_{ij}. Specifically, $W_{ij+z} = P(Z=z)/P(Z=z|V_{ij} = v)$ for $i = 1$ to the number of children who experienced event, $j = 1$ to J_i, where $z = 0$ or 1. By reintroducing the observed distribution of Z into the weights, maximum efficiency and stabilization of weights are realized(28). This stabilization is accomplished by the numerator in the weights—an estimate of the probability of being in the same calendar period as what is observed. In contrast, the denominator represents the probability of being in the same calendar period as what is observed, given the covariates.

To select the covariates V_{ij} from a set of potentially influential variables associated with both AIDS, death and calendar period, we employed super learner software -- Deletion/Substitution/Addition (DSA). By using this data-adaptive machine learning algorithm and its cross-validation based on likelihood, we avoid the problems inherent with traditional approaches and model building. Specifically, DSA was used to search through function forms using deletion, substitution, and addition actions as previously described (29). Similar to the covariates used by Cain et al, the pool of covariates DSA selected from included age (i.e. time since seroconversion), age at randomization, and race/ethnicity (defined as White or non-White)(1).

The weighted IV estimator of the causal rate ratio among compliers is described in Appendix B. Confidence intervals for unweighted ITT and IV estimates were calculated (30). The 95% confidence intervals for the weighted ITT and IV estimates were estimated by bootstrap.

Results

CDC Category C Diagnosis Among Children with Previous Category B Diagnosis

The study population (n=65) was 55.4% males of non-White ethnicity (61.5%) and over half (50.8%) of their mothers received prenatal care (data not shown).

Among children with a previous Category B diagnosis, the distribution of first Category C diagnosis events, person-days, and rates by calendar period and HAART use are described in Table 1. In the pre-HAART era there were 2 of 56 misclassified events (2.1%) as well as 3,056 of 34,032 (6.9%) misclassified person-days. During the HAART era, 3 out of 9 events (62.5%) and 8,230 out of 16,847 person-days

(48.9%) were misclassified as the participants did not use HAART during this period. The rate of AIDS progression was estimated at 1.45 events per 1,000 person-days for the children using non-HAART regimens. For children using HAART, the rate of AIDS progression was estimated at 0.69 events per 1,000 person-days. Overall, the rate of AIDS progression or death was 1.28 events per 1,000 person-days.

The unweighted ITT rate ratio was estimated at uRR_{ITT} = 3.08 (95% CI 1.52, 6.23) when comparing the pre-HAART era with the HAART era. The ITT rate difference was estimated at uRD_{ITT} = 1.12 (95% CI 0.57 - 1.67). The instrumental variable rate ratio using the unweighted data was estimated at uRR_{IV} = 4.74 (95% CI 2.34,9.58) when comparing the pre-HAART era with the HAART era.

When considering the weighted data, the super-learner, DSA, only selected one covariate to include in the most efficient vector of variables for V_{ij}--race/ethnicity for all analyses. The weighted ITT rate ratio was estimated at wRR_{ITT} = 3.01 (95% CI 1.01-6.87) when comparing the pre-HAART era with the HAART era. The ITT rate difference was estimated at wRD_{ITT} = 1.07 (95% CI 0.01-1.80). The IV rate ratio using the weighted data was estimated at wRR_{IV} = 4.64 (95% CI 1.71-12.95) when comparing the pre-HAART era with the HAART era.

C Diagnosis Among All Children

The study population (n=100) was equally distributed by gender, non-White ethnicity (65.0%) and at least 48 percent of mothers received prenatal care (data not shown).

The distribution of first Category C diagnosis events (regardless of previous Category B diagnosis), person-days, and rates by calendar period and HAART use are described in Table 3. Using 1998 as a cut-off, 100 AIDS events occurred over 61,860 person-days. In the pre-HAART era there were 2 misclassified events out of 88 (2.3%), as well as 3,099 of 45,013 (6.9%) misclassified person-days. During the HAART era, 6 out of 12 events (50.0%) and 8,230 out 16,847 person-days (48.9%) were misclassified as the participants were not observed using HAART during this period. The rate of AIDS progression was estimated at 1.83 events per 1,000 person-days for the children using non-HAART regimens. For children using HAART, the rate of AIDS progression was estimated at 0.68 events per 1,000 person-days. Overall, the rate of AIDS progression was 1.62 events per 1,000 person-days.

The distribution of events and person-days by calendar period for the weighted and unweighted data is described in Table 4. When considering the unweighted data, the ITT rate ratio was estimated at uRR_{ITT} = 2.74 (95% CI 1.50 – 5.01) when comparing the pre-HAART era with the HAART era. The ITT rate difference was estimated at uRD_{ITT} = 1.24 (95% CI 0.67-1.81). The IV rate ratio using the unweighted data was estimated at uRR_{IV} = 4.99 (95% CI 2.73 – 9.12) when comparing the pre-HAART era with the HAART era. The weighted ITT rate ratio was estimated at wRR_{ITT} = 2.71 (95% CI 1.11-6.45) when comparing the pre-HAART

era with the HAART era. The ITT rate difference was estimated at wRR_{ITT} = 1.22 (95% CI 0.16-2.13). The IV rate ratio using the weighted data was estimated at wRR_{IV} = 4.94 (95% CI 2.37-16.85) when comparing the pre-HAART era with the HAART era.

Discussion

In the analyses exploring the effect of HAART on Category C diagnoses among children previously diagnosed with a Category B diagnosis suggest that the ITT rate ratios are biased toward the null when exposure misclassification is not considered. Using an IV estimator, exposure to non-HAARTregimens increased the hazard of a Category C diagnosis 4.74 times when compared to HAART exposure. This IV estimate is 54% (4.74/3.08) higher than the result one would see from a traditional ITT approach using calendar period. We weighted the number of events and amount of person-days by inverse probability of calendar period given race/ethnicity. In turn, the weighted rate ratios were estimated, which were expanded from basic instrumental variable methods,(20) after adjusting for measured covariates. Similar biases toward the null were also seen for the analyses using an earlier defined instrument and for the analyses among all children regardless of previous B diagnosis.

In the present analysis, we assume that calendar period is an appropriate instrument for HAART use. This assumption is based on three key characteristics of the IV: 1) is independent of unmeasured confounding between HAART and outcome; 2) is associated with HAART; 3) is independent of the outcome given HAART and unmeasured confounding between HAART and the outcome. The second principle is well established from previous research (1; 5; 7). However, we were unable to test the first and third principle with our data. In an attempt to address the third principle, we employed inverse probability of calendar period weights which, the assumption that calendar year is independent of the outcome given HAART (actual use) and measured confounding (indications for actual HAART use), as previously described(1). Thus, this IV estimator analysis assumes exchangeability between calendar eras (20).

Our results are contingent on the assumption that the model for weights has within it all possible determinants of calendar era and Category C diagnoses. The algorithm DSA selected only race/ethnicity for the final model, ignoring age at seroconversion and time since seroconversion.

The impact of the choice of calendar period on the estimates is not negligible. We chose 1998 as a cut-off for the first calendar period as HAART was only first introduced in mid-1996. To allow for the possibility that the use of HAART was not widely available until later, we explored 1997 as a cut-off in a second analysis. Cain et al performed a similar analysis using 1996 and 1998 as separate instrumental variable cut-offs(1). They found the unweighted and weighted ITT estimates 3% and 8% higher when using the 1998 calendar year cut-off than using the 1996 calendar year cut-off. The researchers also found that the 1998 unweighted and

weighted instrumental variables estimates were 5% and 4% higher than the 1996 instrumental variable estimates. When using the 1998 calendar period cut-off for our present analysis of risk of first Category C diagnosis among all children, the unweighted IV estimates are 28% higher than the unweighted IV estimates we would have calculated using the 1997 calendar period (uRR_{IV}=4.99 and uRR_{IV}=3.89, respectively). In contrast, the unweighted ITT estimates are 5% higher when using the 1998 cut-off rather than the 1997 cut-off (data not shown).

Despite efforts to estimate the effect of HAART using an adapted IV approach, the present study has limitations. Children in this population may have already been treated with HAART before entering the cohort if they moved from another state or region of California. This would increase misclassification of therapy exposure if it occured during the pre-HAART era. Additionally, there were only 8 Category C events among children treated with HAART, which could have increased the variability in the weighted estimates. The small number of events may have also influenced the complier average causal effect of HAART.

Tables

Table 1: Distribution of Events (first Category C diagnosis), Person-Days, and Rates by Calendar Period and HAART Use Among Children With a Previous Category B Diagnosis

Calendar Period (Before 1998/ 1998 and After)	No. of events of first C Diagnosis	No. of Person Days	Rate[1]
	Non-HAART Use/Regimen?		
Pre-HAART	54	30,976	1.74
HAART	3	8,230	0.36
Total	57	39,206	1.45
	HAART Use/Regimen?		
Pre-HAART	2	3,056	0.65
HAART	6	8,617	0.70
Total	8	11,673	0.69
	Total		
Pre-HAART	56	34,032	1.65
HAART	9	16,847	0.53
Total	65	50,879	1.28

1. Events per 1,000 person-days

Table 2: Distribution of Events (first Category C diagnosis) and Person-Days by Calendar Period Among Children With a Previous Category B Diagnosis

Calendar Period (Before 1998/ 1998 and After)	No. of events of first C Diagnosis	No. of Person-Days	Rate[1]	**Intent To Treat**				**Instrumental Variable**	
				Rate Difference	95% CI	Rate Ratio	95% CI	Rate Ratio	95% CI
				Unweighted					
Pre-HAART	56	34,032	1.65	1.12	(0.57, 1.67)	3.08	(1.52, 6.23)	4.74	(2.34, 9.58)
HAART	9	16,847	0.53	0		1		1	
Total	65	50,879	1.28						
				Weighted					
Pre-HAART	54.21	33,724	1.61	1.07	(0.01, 1.80)	3.01	(1.01, 6.87)	4.64	(1.71, 13.0)
HAART	7.24	13,549	0.53	0		1		1	
Total	61.45	47,273	1.30						

1. Events per 1,000 person-days

Table 3: Distribution of Events (first Category C diagnosis), Person-Days, and Rates by Calendar Period and HAART Use

Calendar Period (Before 1998/ 1998 and After)	No. of events of first C Diagnosis	No. of Person Days	Rate[1]
	Non-HAART Use/Regimen?		
Pre-HAART	86	41,914	2.05
HAART	6	8,230	0.73
Total	92	50,144	1.83
	HAART Use/Regimen?		
Pre-HAART	2	3,099	0.64
HAART	6	8,617	0.70
Total	8	11,716	0.68
	Total		
Pre-HAART	88	45,013	1.95
HAART	12	16,847	0.71
Total	100	61,860	1.62

1. Events per 1,000 person-days

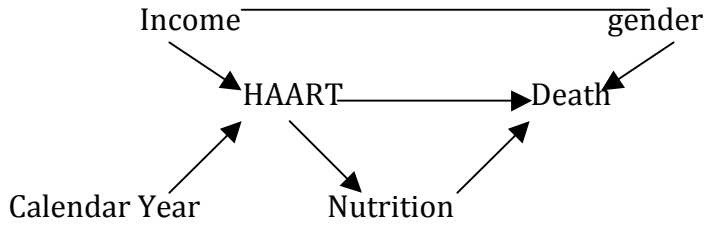
Table 4: Distribution of Events (first Category C diagnosis), and Person-Days by Calendar Period

Calendar Period (Before 1998/ 1998 and After)	No. of events of first C Diagnosis	No. of Person-Days	Rate[1]	Intent To Treat				Instrumental Variable	
				Rate Difference	95% CI	Rate Ratio	95% CI	Rate Ratio	95% CI
				Unweighted					
Pre-HAART	88	45,013	1.95	1.24	(0.67, 1.81)	2.74	(1.5, 5.0)	4.99	(2.7, 9.1)
HAART	12	16,847	0.71	0		1		1	
Total	100	61,860	1.62						
				Weighted					
Pre-HAART	86.55	44,758	1.93	1.22	(0.16, 2.13)	2.71	(1.1, 6.5)	4.94	(2.4, 16.9)
HAART	9.27	13,012	0.71	0		1		1	
Total	95.82	57,770	1.66						

1. Events per 1,000 person-days

Figures

Figure 1. Traditional instrumental variable approach



References

1. Cain LE, Cole SR, Greenland S, Brown TT, Chmiel JS, Kingsley L, Detels R. Effect of Highly Active Antiretroviral Therapy on Incident AIDS Using Calendar Period as an Instrumental Variable. American Journal of Epidemiology 2009;

2. Cain LE, Cole SR, Chmiel JS, Margolick JB, Rinaldo Jr CR, Detels R. Effect of highly active antiretroviral therapy on multiple AIDS-defining illnesses among male HIV seroconverters. American journal of epidemiology 2006;163(4):310.

3. Mocroft A, Ledergerber B, Katlama C, Kirk O, Reiss P, Monforte AA, Knysz B, Dietrich M, Phillips AN, Lundgren JD. Decline in the AIDS and death rates in the EuroSIDA study: an observational study. The Lancet 2003;362(9377):22–29.

4. Palella FJ, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, Aschman DJ, Holmberg SD. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. New England Journal of Medicine 1998;338(13):853.

5. Porter K, Babiker A, Bhaskaran K. Determinants of survival following HIV-1 seroconversion after the introduction of HAART. Lancet 2003;362(9392):1267-1274.

6. Gange SJ, Barron Y, Greenblatt RM, Anastos K, Minkoff H, Young M, Kovacs A, Cohen M, Meyer III WA, Munoz A. Effectiveness of highly active antiretroviral therapy among HIV-1 infected women. British Medical Journal 2002;56(2):153.

7. Detels R, Munoz A, McFarlane G, Kingsley LA, Margolick JB, Giorgi J, Schrager LK, Phair JP, others. Effectiveness of potent antiretroviral therapy on time to AIDS and death in men with known HIV infection duration. Jama 1998;280(17):1497.

8. Ahdieh-Grant L, Li R, Levine AM, Massad LS, Strickler HD, Minkoff H, Moxley M, Palefsky J, Sacks H, Burk RD, others. Highly active antiretroviral therapy and cervical squamous intraepithelial lesions in human immunodeficiency virus-positive women. JNCI Journal of the National Cancer Institute 2004;96(14):1070.

9. Tarwater PM, Mellors J, Gore ME, Margolick JB, Phair J, Detels R, Munoz A. Methods to assess population effectiveness of therapies in human immunodeficiency virus incident and prevalent cohorts. American journal of epidemiology 2001;154(7):675.

10. Moore RD, Chaisson RE. Natural history of HIV infection in the_era of combination antiretroviral therapy. Aids 1999;13(14):1933.

11. Schneider MF, Gange SJ, Williams CM, Anastos K, Greenblatt RM, Kingsley L, Detels R, Mu\ noz A. Patterns of the hazard of death after AIDS through the evolution of antiretroviral therapy: 1984-2004. Aids 2005;19(17):2009.

12. Tassie JM, Grabar S, Lancar R, Deloumeaux J, Bentata M, Costagliola D, others. Time to AIDS from 1992 to 1999 in HIV-1-infected subjects with known date of infection. JAIDS Journal of Acquired Immune Deficiency Syndromes 2002;30(1):81.

13. Violari A, Cotton MF, Gibb DM, Babiker AG, Steyn J, Madhi SA, Jean-Philippe P, McIntyre JA. Early antiretroviral therapy and mortality among HIV-infected infants. New England Journal of Medicine 2008;359(21):2233.

14. Judd A, Doerholt K, Tookey PA, Sharland M, Riordan A, Menson E, Novelli V, Lyall EGH, Masters J, Tudor‐Williams G, Duong T, Gibb DM, Collaborative HIV Paediatric Study (CHIPS) and the National Study of HIV in Pregnancy and Childhood (NSHPC). Morbidity, Mortality, and Response to Treatment by Children in the United Kingdom and Ireland with Perinatally Acquired HIV Infection during 1996–2006: Planning for Teenage and Adult Care. CLIN INFECT DIS 2007;45(7):918-924.

15. Berk DR, Falkovitz-Halpern MS, Hill DW, Albin C, Arrieta A, Bork JM, Cohan D, Nilson B, Petru A, Ruiz J, others. Temporal trends in early clinical manifestations of perinatal HIV infection in a population-based cohort. JAMA 2005;293(18):2221.

16. De Martino M, Tovo PA, Balducci M, Galli L, Gabiano C, Rezza G, Pezzotti P, others. Reduction in mortality with availability of antiretroviral therapy for children with perinatal HIV-1 infection. Jama 2000;284(2):190.

17. Gortmaker SL, Hughes M, Cervia J, Brady M, Johnson GM, Seage III GR, Song LY, Dankner WM, Oleske JM, others. Effect of combination therapy including protease inhibitors on mortality among children and adolescents infected with HIV-1. The New England Journal of Medicine 2001;345(21):1522.

18. Nagelkerke N, Fidler V, Bernsen R. Estimating treatment effects in randomized clinical trials in the presence of non-compliance. STATISTICS IN MEDICINE 2000;19(14):1849-1864.

19. Cole SR, Chu H. Effect of acyclovir on herpetic ocular recurrence using a structural nested model. Contemporary clinical trials 2005;26(3):300–310.

1. Cuzick J, Edwards R, Segnan N. ADJUSTING FOR NON-COMPLIANCE AND CONTAMINATION IN RANDOMIZED CLINICAL TRIALS JACK CUZICK*, ROBERT EDWARDS AND NEREO SEGNAN 2. STATISTICS IN MEDICINE 1997;16:1017–1029.

21. Bond SJ, White IR, Sarah Walker A. Instrumental variables and interactions in the causal analysis of a complex clinical trial. Statist. Med. 2007;26(7):1473-1496.

22. Riersol E. Confluence analysis by means of lag moments and other methods of confluence analysis. Econometrica 1941;9:1-24.

23. Anglemyer A, Colford J, Hubbard A. Review of Causal Inference Methods Used in HIV/AIDS Epidemiologic Studies. In Press. [date unknown];

24. Newhouse JP, McClellan M. Econometrics in outcomes research: the use of instrumental variables. Annual Review of Public Health 1998;19(1):17–34.

25. Caldwell B, Oxtoby M, Simonds R. 1994 Revised Classification System for Human Immunodeficiency Virus Infection in Children Less Than 13 Years of Age. Morbidity and Mortality Weekly Reports 1994;43(RR-12):1-10.

26. Working Group on Antiretrotrival Therapy and Medical Management of HIV Infected Children. Guidelines for the use of antiretroviral agents in pediatric HIV infection. National Institutes of Health 2009 Feb;:1-139.

27. Robins JM, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000;11(5):550–560.

28. Hernan M, Brumback B, Robins J. Marginal Structural Models to estimate the joint causal effect of nonrandomized treatments. Journal of the American Statistical Association 2001;96(454):440-448.

29. Sinisi S, van der Laan MJ. Loss-based cross-validated deletion/substitution/addition algorithms in estimation. UC Berkeley Division of Biostatistics Working Paper Series 2004;

30. Rothman K, Greenland S. Modern Epidemiology. 2nd ed. Philadelphia, PA: Lippincott-Raven; 1998.

Appendix A

Centers for Disease Prevention and Control Definition of Clinical HIV Disease Progression Among Children: Category C Severely Symptomatic

-- Serious bacterial infections, multiple or recurrent (i.e., any combination of at least two culture-confirmed infections within a 2-year period), of the following types: septicemia, pneumonia, meningitis, bone or joint infection, or abscess of an internal organ or body cavity (excluding otitis media, superficial skin or mucosal abscesses, and indwelling catheter-related infections)
-- Candidiasis, esophageal or pulmonary (bronchi, trachea, lungs)
-- Coccidioidomycosis, disseminated (at site other than or in addition to lungs or cervical or hilar lymph nodes)
-- Cryptococcosis, extrapulmonary
-- Cryptosporidiosis or isosporiasis with diarrhea persisting >1 month
-- Cytomegalovirus disease with onset of symptoms at age >1 month (at a site other than liver, spleen, or lymph nodes)
-- Encephalopathy (at least one of the following progressive findings present for at least 2 months in the absence of a concurrent illness other than HIV infection that could explain the findings): a) failure to attain or loss of developmental milestones or loss of intellectual ability, verified by standard developmental scale or neuropsychological tests; b) impaired brain growth or acquired microcephaly demonstrated by head circumference measurements or brain atrophy demonstrated by computerized tomography or magnetic resonance imaging (serial imaging is required for children <2 years of age); c)acquired symmetric motor deficit manifested by two or more of the following: paresis, pathologic reflexes, ataxia, or gait disturbance
-- Herpes simplex virus infection causing a mucocutaneous ulcer that persists for >1 month; or bronchitis, pneumonitis, or esophagitis for any duration affecting a child >1 month of age
-- Histoplasmosis, disseminated (at a site other than or in addition to lungs or cervical or hilar lymph nodes)
-- Kaposi's sarcoma
-- Lymphoma, primary, in brain
-- Lymphoma, small, noncleaved cell (Burkitt's), or immunoblastic or large cell lymphoma of B-cell or unknown immunologic phenotype
-- Mycobacterium tuberculosis, disseminated or extrapulmonary
-- Mycobacterium, other species or unidentified species, disseminated (at a site other than or in addition to lungs, skin, or cervical or hilar lymph nodes)
-- Mycobacterium avium complex or Mycobacterium kansasii, disseminated (at site other than or in addition to lungs, skin, or cervical or hilar lymph nodes)
-- Pneumocystis carinii pneumonia
-- Progressive multifocal leukoencephalopathy
-- Salmonella (nontyphoid) septicemia, recurrent
-- Toxoplasmosis of the brain with onset at >1 month of age
-- Wasting syndrome in the absence of a concurrent illness other than HIV infection that could explain the following findings: a) persistent weight loss >10% of baseline OR b) downward crossing of at least two of the following percantile lines on the weight-for-age chart (e.g., 95th, 75th, 50th, 25th, 5th) in a child >=1 year of age OR c) <5th percentile on weight-for-height chart on two consecutive measurements, >=30 days apart PLUS a) chronic diarrhea (i.e., at least two loose stools per day for >30 days) OR b) documented fever (for >=30 days, intermittent or constant)

Appendix B

Subscript i indexes the 1 to N=65 children, j indexes the 1 to J_i visits for each child i. The maximum number of visits was 57. HAART use is indexed by subscript x, where 1 is HAART use and 0 non-HAART use. When a child experiences the outcome, a diagnosis of an AIDS defining illness or death, the script $D_{ijxz} = 1$ is used to indicate that child i experienced the outcome between visits j and $j + 1$ during calendar period z while using therapy x. $D_{ijxz} = 0$ indicates that child did not experience the outcome. The number of person-days that each child i contributed between visits $j - 1$ and j while using therapy x during calendar period z is indicated by T_{ijxz}. Identified covariates for each child i at visit j, variables include both time-varying and time-fixed alike, are included in vector V_{ij}.

Let $T_{xz}=\sum_{i=1}^{65}\sum_{j=1}^{Ji} T_{ijxz}$, the total number of person-days contributed while using therapy x during calendar period z summed over all children I and visits.

And let $D_{xz}=\sum_{i=1}^{65}\sum_{j=1}^{Ji} D_{ijxz}$, the total number of events experienced while using therapy x during calendar period z summed over all children I and visits J. As described in Cain et al, let α_{xz} be the conditional probability of using therapy x given calendar period z, as estimated by the proportion of person-days while treated with therapy x during calendar period z. That is to say, $\alpha_{xz} = P(X = x| Z = z) = T_{xz}/T_{+z}$ where

$$T_{+z} = \sum_{x=0}^{1} T_{xz}.$$

In a traditional IV approach, the estimator for ITT of the average causal effect is:

$$\beta_{ITT} = \frac{\alpha_{10} \text{ X } (D_{10}/T_{10}) + \alpha_{00} \text{ X } (D_{00}/T_{00})}{\alpha_{11} \text{ X } (D_{11}/T_{11}) + \alpha_{01} \text{ X } (D_{01}/T_{01})} = \frac{(D_{+0}/T_{+0})}{(D_{+1}/T_{+1})}.$$

Assuming that calendar periods are exchangeable and that calendar period is a valid instrument, the estimator is:

$$\beta_{IV} = \frac{\alpha_{00} \text{ X } (D_{00}/T_{00}) - \alpha_{01} \text{ X } (D_{01}/T_{01})}{\alpha_{11} \text{ X } (D_{11}/T_{11}) - \alpha_{10} \text{ X } (D_{10}/T_{10})} =$$

$$\frac{\beta_{ITT}}{\dfrac{[\alpha_{00} \text{ X } (D_{00}/T_{00}) - \alpha_{01} \text{ X } (D_{01}/T_{01})] \text{ X } [\alpha_{00} \text{ X } (D_{00}/T_{00}) + \alpha_{01} \text{ X } (D_{01}/T_{01})]}{[\alpha_{11} \text{ X } (D_{11}/T_{11}) - \alpha_{10} \text{ X } (D_{10}/T_{10})] \text{ X } [\alpha_{10} \text{ X } (D_{10}/T_{10}) + \alpha_{01} \text{ X } (D_{01}/T_{01})]}}$$

The ITT estimator divided by the estimator of the association between the exposure and the IV, as illustrated in the lower ratio, depicts a traditional β_{IV} analysis. In the absence of contamination or non-compliers, which would occur if no one contributes person-time to the non-HAART calendar period while using HAART and if no one contributes person-time to the HAART calendar period while not using HAART, then $\beta_{IV} = \beta_{ITT}$ and $\alpha_{10} = \alpha_{01} = 0$.

Inverse probability of calendar period weights, W_{ij+Z} , were estimated to adjust for V_{ij}. Specifically, $W_{ij+z} = P(Z=z)/P(Z=z|V_{ij} = v)$ for $i = 1$ to 65, $j = 1$ to J_i, where $\max(J_i) = 57$ and $z = 0$ or 1.

The new weighted instrumental variable estimator of the causal rate ratio among compliers can be written as:

$$ {}_w\beta_{IV} = \frac{{}_w\alpha_{00} \text{ X } ({}_wD_{00}/{}_wT_{00}) - {}_w\alpha_{01} \text{ X } ({}_wD_{01}/{}_wT_{01})}{{}_w\alpha_{11} \text{ X } ({}_wD_{11}/{}_wT_{11}) - {}_w\alpha_{10} \text{ X } ({}_wD_{10}/{}_wT_{10})} $$

where ${}_w\alpha_{xz} = {}_wT_{xz}/{}_wT_{+z}$, ${}_wD_{xz}=\sum_{i=1}^{65}\sum_{j=1}^{Ji} D_{ijxz} \text{ X } W_{ijxz}$, and ${}_wT_{xz}=\sum_{i=1}^{65}\sum_{j=1}^{Ji} T_{ijxz} \text{ X } W_{ijxz}$.