TIME-DEPENDENT CONFOUNDING IN ANTIHYPERTENSIVE DRUG STUDIES

By

TOBIAS GERHARD

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2007

To my parents, Gertrud and Albrecht Gerhard

ACKNOWLEDGMENTS

I want to thank my adviser, Almut Winterstein, for her continuous support. She has been a terrific mentor and has never stopped challenging me to do my best. I would also like to thank my supervisory committee members Julie Johnson, Abraham Hartzema, Carl Pepine, and Jonathan Shuster for their expertise, advice, and encouragement. I thank Yan Gong and Rhonda Cooper-DeHoff for their help with the INVEST dataset. Finally, I thank my fellow graduate students for their support and friendship.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

TIME-DEPENDENT CONFOUNDING IN
ANTIHYPERTENSIVE DRUG STUDIES

By

Tobias Gerhard

May 2007

Chair: Almut Winterstein
Major Department: Pharmacy Health Care Administration

Accurate estimation of blood pressure (BP) effects on the risk for cardiovascular outcomes has important implications for the treatment of hypertension. The extent to which operationalization of BP affects these risk estimates is unclear. Furthermore, the presence of a time-dependent confounder may lead to biased estimates for the risk of BP on cardiovascular outcomes and can not be adjusted for by standard statistical methods. The same bias may occur in the estimation of drug effects in the presence of time-dependent confounding by BP.

To examine the impact of systolic blood pressure (SBP) operationalization on risk estimates for myocardial infarction, stroke, or all cause death (primary outcome) we estimated the hazard ratios of 7 SBP categories for six different Cox proportional hazards models in patients of the International Verapamil-Trandolapril Study (INVEST), a randomized study of 22,576 hypertensive coronary artery disease patients. To test for the presence of time-dependent confounding by antihypertensive treatment (or, alternatively, SBP control), we estimated both standard Cox models and marginal structural Cox models (causal models) for the effect of SBP control (or, respectively, aggressive antihypertensive treatment), adjusting for the number of concurrently used antihypertensive drugs (or, respectively, SPB).

Estimates of the effect of SBP on primary outcome vary significantly depending on the method of SBP operationalization. Some of the operationalization approaches, most notably the use of average SBP, may lead to systematically biased estimates. Causal analyses suggest that time-dependent confounding by SBP may bias estimates of treatment effects (Hazard ratio [HR] standard model: 0.96; 95% confidence interval [CI] 0.87-1.07; HR marginal structural model: 0.81; 95% CI 0.71-0.92), but provides no evidence of time-dependent confounding by treatment in the estimation of risk associated with SBP control (HR standard model: 0.54; 95% CI 0.48-0.60; HR marginal structural model: 0.55; 95% CI 0.50-0.61).

Our results suggest that time-dependent confounding by SBP, leads to an underestimation of the effectiveness of antihypertensive treatment. No evidence for time-dependent confounding of the effect of SBP control by antihypertensive treatment was found, implying, that antihypertensive treatment as modeled in our analysis does not affect cardiovascular outcomes in pathways other than through SBP.

CHAPTER 1
INTRODUCTION

Background

In 2005, chronic diseases, such as cardiovascular disease, diabetes, or cancer were estimated to be responsible for 60% of the total global mortality (35 million deaths).[1] In the United States alone, chronic diseases affect the lives of over 90 million Americans and account for 70% of all deaths.[2] Treatment of patients suffering from chronic diseases occurs over extended periods of time, frequently involves multi-drug regimens, and often relies on surrogates (i.e., intermediate markers of health) to evaluate the effectiveness of treatment in the individual patient.

Clinically relevant outcomes of chronic diseases such as myocardial infarction, stroke, or death often occur only in a proportion of affected patients and often after years or even decades of the disease. As a consequence, immediately observable surrogate measures such as blood pressure, low density lipoprotein (LDL) level, or CD4 cell count (a marker of circulating T helper cells) play an important role in the treatment of patients suffering from chronic disease as they predict the risk for manifestation of adverse outcomes. In addition, surrogate measures typically facilitate shorter clinical trials with smaller sample sizes. Accurate estimation of the association between the surrogate measure and the risk of clinically relevant outcomes over time is a prerequisite for the informed use of a surrogate measure in treatment and research. Furthermore, to avoid confounding, the influence of the surrogate has to be carefully considered in the planning and analysis of observational studies of drug effects, because surrogate measures play an important role in the determination and management of drug therapy.

Need for Study

Association of a Surrogate with Clinical Outcome: Surrogates are Time-Dependent Variables

Surrogate measures are rarely constant over time. Disease progression, life-style modification, pharmacologic and nonpharmacologic treatments may all contribute to changes in a surrogate measure over time. Thus, any analysis that aims to quantify an association of a surrogate with a relevant clinical outcome (i.e., morbidity or mortality) over a prolonged period of time and uses a single, fixed value to represent the true surrogate values over the course of time will lead to misclassification. Inclusion of multiple values of the surrogate over time (i.e., its inclusion as a time-dependent variable) can significantly reduce misclassification bias. However, it is often unclear to what extend bias introduced by such misclassification will alter estimates of the effect of a surrogate on a clinical outcome in practice. Since, time-dependent modeling of surrogates involves more complex statistical methods and requires regularly measured data-points for the surrogate over time, study results are frequently based on single, fixed surrogate values (such as baseline, or average over follow-up).

Another problem closely related to the time-dependent nature of surrogate measures is the potential lag time of a surrogate's effect on the clinical outcome. Depending on the pathophysiological mechanism through which the surrogate affects the clinical outcome, associations between the surrogate and the clinical outcome may be immediate or delayed. This has profound consequences for analysis because it determines whether current or historical values (or a combination of the two) of the surrogate should be used to model the risk for the clinical outcome (e.g., it would affect to what extend BP history, as opposed to current BP values, should be included in risk models for cardiovascular disease).

Association of a Surrogate with Clinical Outcome: Confounding by Treatment

Surrogate measures are routinely used to guide clinical practice. Physicians will for example, increase the dose of an antihypertensive medication or add an additional agent, if a patient's blood pressure is considered uncontrolled, and patients with elevated cholesterol may be treated with increasing doses of statins until recommended LDL levels are achieved. The validity of this approach relies on unbiased estimation of the causal effect of the surrogate on the clinical outcome. Several factors make this estimation difficult. As mentioned above, surrogate measures are not constant over time, and thus, estimation of the causal effect of the surrogate on clinical outcome needs to account for these time-dependent changes. In addition, surrogate measures are rarely observable in untreated patients, because changes in a surrogate are expected to result in changes in clinical outcomes, and thus, once identified, patients with elevated values of a surrogate are routinely receiving treatment. This treatment may confound the estimation of a causal effect between surrogate and clinical outcome in a treated cohort. Specifically, only if the effect of a given drug is entirely mediated by the surrogate, which rarely is the case in practice, will an estimate of the causal effect of the surrogate on a clinical outcome in the presence of the drug be unconfounded. If a drug affects clinical outcome in parts through pathways different from the surrogate measure of interest, the drug acts as a confounder and needs to be controlled for in any analysis that aims to estimate the causal effect of the surrogate on the clinical outcome. However, since drug use is commonly affected by a prior value of the surrogate, and thus is simultaneously a direct cause for subsequent values and a direct result of prior values of the surrogate, assumptions for standard methods of confounder adjustments, such as inclusion in regression models are violated, and such methods fail to produce unbiased, causally interpretable estimates.[3]

In summary, the evaluation of the causal effect of surrogate measures must incorporate changes of the surrogate observed over time and account for time-dependent confounding by treatment. To the best of our knowledge, time-dependent confounding by treatment has never been accounted for in the analysis of surrogate measures.

Estimation of Drug Effectiveness in Observational Research

Pharmacoepidemiological research, particularly when it relates to drugs used in the treatment of chronic diseases, commonly deals with the risks and benefits of drugs used over prolonged periods of time. More often than not, treatment will not be stable over time but rather will doses be adjusted and drugs added or removed from the treatment regimen. Such adjustments of therapy over time do not occur at random and thus, control of factors that influence both treatment changes and treatment outcome (i.e., confounders) is necessary to obtain unbiased estimates of risks and benefits for the various treatment choices. However, as detailed in the previous section, when factors that predict changes in treatment are also affected by the change in the treatment regimen, as is common for surrogate measures of chronic disease states (e.g., BP, HDL/LDL, HbA1C), assumptions for standard methods of confounder adjustments are violated and these methods fail to produce unbiased estimates. Conventional evaluation of drug effects in the presence of a time-dependent confounder therefore may produce biased estimates.

For a number of reasons evidence from randomized clinical trials alone is often insufficient to provide the evidence needed for optimal selection of individual treatment strategies. First, randomized clinical trials are typically conducted over short periods of time and take place in narrow study populations defined by explicit inclusion and exclusion criteria. In contrast, drug use in clinical practice will occur over extended periods of time in less homogeneous populations. Second, and more importantly in the context of this study, the

comparisons made in clinical trials are limited, frequently involving comparisons of single therapeutic agents with each other or placebo. In practice however, many chronic disease patients will require a combination of two or more drugs to be adequately treated. While clinical trials comparing specific combination therapies or flexible treatment strategies are possible, and have been conducted, the number of possible drug (and dose) combinations will likely exceed what can be feasibly tested in a clinical trial setting.

Thus, observational pharmacoepidemiological research, which is able to explore the broad spectrum of treatments occurring in practice over extended periods of time, could play an important role in evaluating the long-term effectiveness of complex multi-drug strategies. However, careful control of treatment decisions that determine the exposure of individual patients to specific drug regimens over the course of a study and that may lead to time-dependent confounding is necessary to avoid biased estimates of regimen effectiveness.

Purpose of the Study

Our study used a dataset from a large international antihypertensive trial to estimate the association of systolic blood pressure (surrogate measure) over time on the risk for cardiovascular morbidity and mortality (clinical outcome) and to assess whether time–dependent changes in treatment confound this association. In addition, this study will evaluate the effects of treatment on clinical outcome, when initiation of treatment is partly conditional on inadequate response to prior treatment and thus, confounded by a surrogate measure. Our study will illustrate problems arising from the presence of time-dependent confounding in such a setting and use newly developed statistical methods to obtain estimates of unbiased effects for both the surrogate and treatment on clinical outcome in the presence of time-dependent confounding.

Specifically, this study will describe blood pressure and antihypertensive drug use patterns for patients participating in the International Verapamil SR/Trandolapril Study (INVEST)[4, 5] over

the time of follow-up. It will then evaluate how systolic blood pressure over time associates with the incidence of primary outcome events (nonfatal myocardial infarction, nonfatal stroke, or all cause death) and compare this time-dependent approach to analytic approaches that use single, fixed SBP estimates (e.g., baseline SBP, average SBP over treatment period). Using marginal structural Cox models, the present study will then assess whether time-dependent treatment with antihypertensive drugs confounds the association of blood pressure control (SBP <140 mm Hg versus ≥140 mm Hg) with clinical outcome and to what extend failure to consider this in traditional methods biases these estimates. Lastly, adjusting for SBP control over time using marginal structural Cox models, this study will derive an unbiased estimate for the effect of antihypertensive therapy (aggressive versus standard antihypertensive therapy) on the clinical outcome. Of note, the necessity to dichotomize the independent variables of interest (SBP control, aggressive antihypertensive therapy) is a limitation inherent in the use of current marginal structural models and will likely produce estimates of association that are of limited clinical utility.

More complex and specific comparisons between individual drug combinations or use of multiple BP categories may be possible and should be addressed in future research. The present study will use hypertension to illustrate the aforementioned problems arising from time-dependent confounding when treatment initiation and choice are affected by the prior surrogate and the surrogate lies on the hypothesized pathway through which the treatment affects the risk for the clinical outcome. Blood pressure is an important and widely used surrogate measure that plays an essential role in the selection and management of antihypertensive treatment and is likely to be the major pathway through which antihypertensive drugs affect the risk of clinical outcome. Causal methods such as marginal structural models may indirectly contribute to better

answer the question to which extent the effects of specific antihypertensive drugs and drug classes are mediated by blood pressure and may ultimately contribute to the identification of optimized combination therapies.

The INVEST cohort provides the opportunity to investigate the independent effects of antihypertensive drug use and a surrogate measure (blood pressure) on a clinical outcome (INVEST primary outcome) in a rich and validated clinical dataset with independently adjudicated outcomes. Although INVEST is a randomized controlled trial, evaluation of individual steps of the treatment strategies negates randomization and thus, requires an epidemiologic analysis approach similar to an observational study. Its large sample size and high level of data quality make the INVEST an appropriate setting for the simultaneous evaluation of time-dependent treatment and time-dependent surrogate measure.

Research Questions and Hypotheses

The first research question aims to evaluate whether the ability to predict adverse cardiovascular outcomes in the INVEST is increased when blood pressure is operationalized as a multi-category time-dependent variable, as compared to a single, fixed estimate (such as baseline or average over follow-up). However, the estimation of the association of systolic blood pressure presented in research question 1 does not control for concurrent use of antihypertensive drugs. Thus, research question 2 will estimate the effect of blood pressure control on clinical outcome over time controlling for time-dependent treatment (operationalized as the number of antihypertensive INVEST study drugs as well as the total number of antihypertensive drugs), while research question 3 evaluates whether the concurrent use of antihypertensive drugs confounds the association of systolic blood pressure over time with primary outcome. Lastly, the fourth and fifth research questions address the problem of estimating the effectiveness of a drug or treatment strategy (in our study, aggressive versus standard antihypertensive therapy) in the

presence of time-dependent confounding by a surrogate measure (in our study, SBP control). The a priori significance level for all research questions is set at 0.05.

- **Research Question 1:** Does time-dependent operationalization of systolic blood pressure increase the ability to predict primary outcome events in the INVEST as compared to the use of single, fixed blood pressure values?

- **Hypothesis for Research Question 1:** The null hypothesis for research question 1 is that the generalized R^2 for a model that updates the systolic blood pressure at each visit is not different from models with fixed systolic blood pressure values, specifically baseline systolic blood pressure, and average systolic blood pressure over follow-up. The alternative hypothesis is that the generalized R^2 is larger.

- **Research Question 2:** In the INVEST, does systolic blood pressure control over follow-up affect the risk of primary outcome controlling for time-dependent confounding by concurrent antihypertensive drug use?

- **Hypothesis for Research Question 2:** The null hypothesis for research question 2 is that the hazard ratio for systolic blood pressure control is not significantly different from 1.0. The alternative hypothesis is that the hazard ratio is significantly different from 1.0.

- **Research Question 3:** In the INVEST, does time-dependent treatment (i.e., the number of antihypertensive INVEST study drugs as well as the total number of antihypertensive drugs) confound the effect of systolic blood pressure control over the follow-up period.

- **Hypothesis for Research Question 3:** The null hypothesis for research question 3 is that the hazard ratio for SBP control obtained from a marginal structural Cox proportional hazards model that incorporates time-dependent treatment, is not significantly different from the hazard ratio obtained from a standard time-dependent Cox proportional hazards model that does not control for treatment after baseline.

- **Research Question 4:** In the INVEST, adjusting for time-dependent systolic blood pressure control, is there a difference in risk of primary outcome between patients receiving aggressive antihypertensive therapy compared to standard antihypertensive therapy over follow-up?

- **Hypothesis for Research Question 4:** The null hypothesis for research question 4 is that the hazard ratio for patients receiving aggressive antihypertensive therapy (three or more concurrent total antihypertensive drugs) versus standard antihypertensive therapy (less than three concurrent total antihypertensive drugs) is not significantly different from 1.0. The alternative hypothesis is that the hazard ratio is significantly different from 1.0.

- **Research Question 5:** In the INVEST, does time-dependent systolic blood pressure control confound the effect of treatment (aggressive versus standard antihypertensive therapy) over the follow-up period?

- **Hypothesis for Research Question 5:** The null hypothesis for research question 5 is that the hazard ratio for patients receiving aggressive antihypertensive therapy (three or more total antihypertensive drugs) versus standard antihypertensive therapy (less than three total antihypertensive drugs) obtained from a marginal structural Cox proportional hazards model that controls for time-dependent confounding by systolic blood pressure control, is not significantly different from the hazard ratio obtained from a standard time-dependent Cox proportional hazards model that does not control for systolic blood pressure over follow-up.

CHAPTER 2
LITERATURE REVIEW

Surrogates

A surrogate is defined as a laboratory measurement or physical sign used as a substitute for a clinically meaningful endpoint that measures directly how a patient feels, functions or survives. Changes induced by a therapy on a surrogate measure are expected to reflect changes in a clinically meaningful endpoint.[6] Although drug therapy is ultimately aimed at affecting clinically meaningful endpoints, the use of surrogate measures offers important advantages. When clinical endpoints are rare or manifest after substantial periods of time, as is often the case in chronic disease states, the use of surrogate measures allows shorter, smaller, and less costly clinical trials, which in turn allow more rapid approval of new therapies.[7-9] Patient advocacy groups, interested in the rapid availability of new and promising therapies, as well as drug manufacturers, who save costs and patent life of their products, consequently support the use of surrogates in Phase III trials. The Food and Drug Administration (FDA) has responded to these demands by allowing the use of surrogate measures to demonstrate the efficacy of new drug products as part of its accelerated approval process for serious or life-threatening illness.[10] However, due to the concern that changes in the surrogate may not translate into changes in the clinical outcome, the use of surrogates is not without controversy.[9, 11-14]

In addition to their function in the evaluation of new therapies, surrogate measures play an important role in the evaluation of treatment response in individual patients and the modification of individual pharmacotherapy, often following guidelines that recommend treatment towards specific target levels of the surrogate measure. However, the use of surrogate measures is problematic when a treatment also affects the clinical endpoints in ways not mediated by the surrogate, since such effects are not captured by the observed changes in the surrogate. Prentice

formalized the definition of a valid surrogate measure by requiring two sufficient conditions, (1) the surrogate must correlate with the true clinical endpoint, and (2) the surrogate must fully capture the treatments net effect (the aggregate effect accounting for all causal effects of the treatment on the true clinical outcome)[8]. Figures 2-1 to 2-3 illustrate Prentice's conditions and potential deviations that may compromise the validity of surrogate measures. Figure 2-1 depicts a situation where Prentice's conditions are met. In contrast, Figure 2-2 illustrates a scenario in which use of a surrogate would completely fail to predict the effect of an intervention on the true clinical outcome. While the surrogate is correlated with the true clinical outcome, and thus, satisfies the "correlation" condition, it does not lie on the biological pathway by which the disease causes the clinical outcome. In consequence an intervention affecting the surrogate will have no effect on the clinical outcome. A situation somewhere between the scenarios depicted in the previous two figures may give a more accurate reflection of reality. A disease may affect the clinical outcome through more than one biologic pathway. If an intervention affects more than one biologic pathway (arrows A and B in Figure 2-3) and the surrogate lies only in one of these pathways, then only a part of the effects of the intervention will be captured by its effect on the surrogate. In addition, the intervention may causally affect the clinical outcome unrelated to the disease (arrow C in Figure 2-3), for example, through adverse effects.

While Prentice's conditions define a surrogate in absolute terms, Figure 2-3 illustrates a more common scenario where the second condition is not completely satisfied, but rather partially met. The statistical literature has approached this problem by introducing the proportion of treatment effect (PTE) explained by a surrogate marker which allows a more subtle evaluation of a surrogate's validity for a specific intervention.[15]

Importantly, when Prentice's conditions are not fully met, the evaluation of the effect of a specific treatment on a clinical outcome by assessing the treatment effect on the surrogate measure (be it in an aggregate form in the context of a clinical trial, or on an individual level when assessing response to treatment) will not capture the full effect of the treatment on the clinical outcome and thus, be biased.

In summary, the validation of surrogate measures is a challenging task. First, it requires the understanding of the biological pathway through which the surrogate affects the clinical outcome. Only in the second step follows the statistical evaluation. The validation of a surrogate for a specific treatment requires larger sample sizes than are needed to determine the effect of the treatment on the clinical outcome. Therefore meta-analysis of large clinical trials that document the effects of treatment on both surrogate and clinical outcomes are usually necessary. Since the validity of a surrogate is treatment specific, validation should be repeated for different classes of drugs in the same disease state.

The Epidemiology of Blood Pressure Control and Cardiovascular Outcomes

Blood pressure (BP) is a strong independent predictor of adverse cardiovascular (CV) outcomes and its control is one of the central goals in the prevention and treatment of cardiovascular disease.[16, 17] Blood pressure is currently classified into normal (systolic blood pressure [SBP]/diastolic blood pressure [DBP] < 120/80 mm Hg), prehypertension (120/80 mm Hg < SBP/DBP < 140/90 mm Hg), stage 1 (140/90 mm Hg < SBP/DBP < 160/100 mm Hg) and stage 2 (SBP/DBP > 160/100 mm Hg) hypertension. The prevalence of hypertension in the United States between 1989 and 1991 has been estimated to reach almost 25%, and increases sharply with advancing age.[18] Mortality from stroke and ischemic heart disease (IHD) increases with higher blood pressure levels starting from 115mm Hg SBP and 75 mm Hg DBP, respectively (Figure 2-4).[17]

While DBP is a more important risk factor for cardiovascular disease than SBP before the age of 50, the importance reverses in patients older than 50 years of age.[19] Treatment of hypertension ultimately aims at the reduction of circulatory and renal mortality and includes lifestyle modifications as well as pharmacologic treatment. A large number of drugs from multiple drug classes are currently approved for the treatment of hypertension and more than two-thirds of treated hypertensive patients require two or more antihypertensive drugs to reach blood pressure control.[16] Initial antihypertensive drug choice and following management are influenced by the presence of secondary diagnoses with compelling advantages of specific antihypertensive drug classes in regards to efficacy, tolerability, and blood pressure response. For most patients without comorbidities, thiazide-type diuretics are recommended as first line treatment. [16] An algorithm for treatment of hypertension is shown in Figure 2-5.

However, a number of questions regarding the optimal therapy of hypertension remain controversially discussed. Arguably the most important issue is whether differences exist in the beneficial effects on adverse cardiovascular outcomes between the major antihypertensive drug classes. Closely related to this problem is the question that if a comparison between two drug classes results in differences in cardiovascular outcomes, then are these differences fully accounted for by the level of achieved blood pressure reduction or do non blood pressure mediated effects play a role in the effectiveness of antihypertensive drugs? In other words: does it matter how blood pressure reduction is achieved? Over the last decades a plethora of antihypertensive drug trials have been conducted comparing various drugs from the major antihypertensive drug classes with placebo or active control treatments. Two recent meta-analyses have aggregated data from 29 trials with 162,341 patients[20], and 42 trials with 192,478 patients[21], respectively. Both studies found no differences in the reduction of all cause or

cardiovascular mortality between the major antihypertensive drug classes. However, some differences were shown in the effects on specific cardiovascular outcomes, most notably heart failure, with diuretics presenting the most beneficial effects. One of the studies also reported that, with the exception of heart failure, differences in achieved blood pressure between trials randomized to the major antihypertensive drug classes were proportional to differences in risk of cardiovascular outcomes.[20] While some argue that it is 'relatively unimportant' which specific agents are used to achieve blood pressure control[22], the differences in the effectiveness on specific cardiovascular outcomes between drug classes suggest the existence of drug-specific mechanisms that affect cardiovascular outcomes independent of blood pressure.

As mentioned earlier, the majority of hypertensive patients will require two or more antihypertensive drugs to achieve blood pressure control according to current guidelines. While the comparative effectiveness of specific antihypertensive drugs is still not conclusively established, much less is known regarding the comparative effectiveness of different combination therapies. The underlying question is whether synergistic effects exist for specific antihypertensive drug combinations or whether only achievable blood pressure reduction, tolerability, and cost should determine treatment. The comparative effectiveness is considerably harder to assess for combination therapy than for monotherapy because of the large number of antihypertensive drugs and drug classes (and thus, the large number of possible comparisons). Additionally, benefits may be associated with the more rapid control of blood pressure through immediate initiation of combination therapy versus initial treatment with monotherapy followed by additional antihypertensive drugs if blood pressure control has not been achieved.

Lastly, it is not clear what the ideal blood pressure goal should be and when to begin treatment. For individuals with uncomplicated hypertension, current guidelines prescribe

initiation of antihypertensive pharmacotherapy, if lifestyle modifications alone do not lower blood pressure below 140/90 mm Hg [SBP/DBP] (lower recommendations apply to individuals with specific comorbidities).[16] However, epidemiological evidence suggests a more than twofold difference in cardiovascular risk for blood pressure values of 130-139/85-89 mm Hg as (currently defined as prehypertension) compared to values below 120/80 mm Hg. Thus, additional benefits may be achievable by lowering treatment goals.

Recently, there has been considerable controversy about the feasibility to treat prehypertension.[23] A recent randomized controlled trial demonstrated that treatment of prehypertensive patients delayed progression to stage I hypertension, but to date no data is available for the effect of such treatment on cardiovascular morbidity and mortality.[24]

Time-Dependent Confounding in Pharmacoepidemiology

Confounding occurs when the measure of the effect of an exposure is distorted because of an association of the exposure with other factors that influence the outcome under study. Its control is one of the central issues in pharmacoepidemiological research. It is important to distinguish between measured and unmeasured confounding. The present study will focus on measured confounding. Measured confounding may be addressed by restriction or matching within the design of a study or by stratification or multivariate regression within the analysis stage of a study[25]. Traditionally such methods would use variables at the beginning of the exposure period and then follow patients over time. Through the increasing use of time-dependent methods in which exposure status can vary over the follow-up period in recent years, the problem of time-dependent confounding has become apparent. Time-dependent confounding occurs when a covariate predicts future treatment and future outcome and is itself predicted by past treatment (Figure 2-6). This poses unique problems because standard methods of confounder adjustment do not suffice to produce unbiased estimates. To illustrate why standard

models fail to produce unbiased estimates under the aforementioned conditions, consider the following example. In Figure 2-6, if the time dependent confounder L is not controlled for in the analysis, then L_1 confounds the association of treatment A_1 with outcome Y, because it simultaneously affects both A_1 and Y. Thus, any estimate of the association of A with Y would be biased if L is not controlled for. However, if L is controlled in the analysis, then L_1, a variable in the causal path of A_0 on Y, is blocked, again, resulting in a biased estimate of the association of A with Y.

Marginal Structural Models

Marginal structural models (MSMs), first introduced by Robins, Hernan, and Brumback, aim to produce unbiased estimates in the presence of time-dependent confounding.[3] MSMs use inverse probability of treatment weights (IPTWs) and inverse probability of censoring weights (IPCWs) to create a pseudo-population in which treatment is unconfounded and no censoring occurs.[26] MSMs are fitted in a two stage process. The first step estimates the individual IPTWs and IPCWs. The IPTWs are based on each subject's probability of having their own treatment history at each time point given the subject's covariates (with the time-dependent confounder as one of the covariates). The IPCWs are similarly estimated based on each subjects probability at each time point to be censored based on his covariates. The second step uses the IPTWs and IPCWs as weights in a regression model of the effect of the treatment on the outcome. Because of the weighting, the regression now takes place in the pseudo-population and—if all assumptions are met—results in a causal estimate of the treatment's effect on the study outcome. The method assumes no unmeasured confounding factors and correct model specification for both the weights and the final regression model.

Marginal structural models have been used in a number of disease states to obtain causal estimates of the effect of treatments in the presence of time-dependent confounding. In a recent

observational study that aimed to estimate the causal effect of treatment with zidovudine on the survival of HIV-positive men, the inclusion of CD4 cell count into standard models was prohibited because it simultaneously predicted initiation of zidovudine, was part of the pathway through which zidovudine is hypothesized to work, and was a risk factor for the study outcome.[26] While a standard time-dependent Cox model, adjusted for baseline covariates but not for CD4 cell count resulted in a hazard ratio of 2.3 (95%CI 1.9-2.8), the marginal structural Cox model showed a hazard ratio of 0.7 (0.6-1.0), revealing the beneficial effect of the treatment. Similar results have been obtained for treatment with methotrexate in patients with rheumatoid arthritis,[27] or the effect of aspirin on cardiovascular mortality.[28]

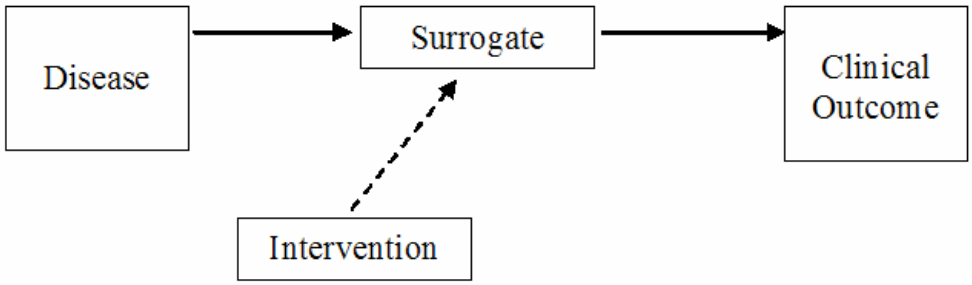

Figure 2-1. Prentice criteria satisfied (adapted from Fleming et al.)[13]

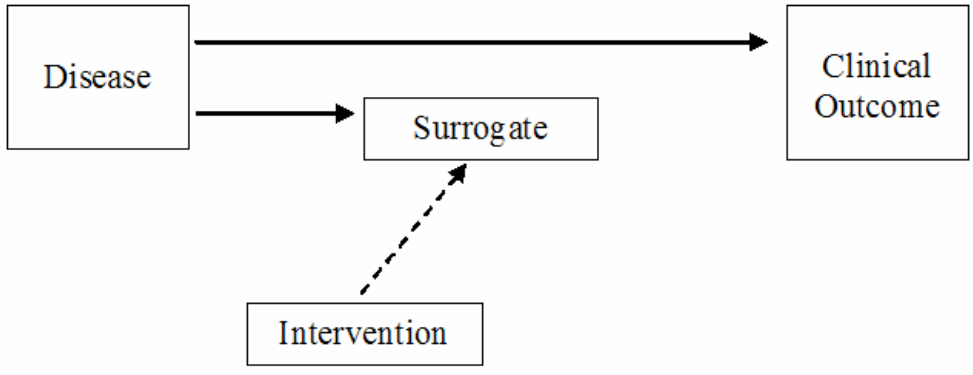

Figure 2-2. The surrogate is correlated with the clinical outcome but captures no treatment effect (adapted from Fleming et al.)[13]

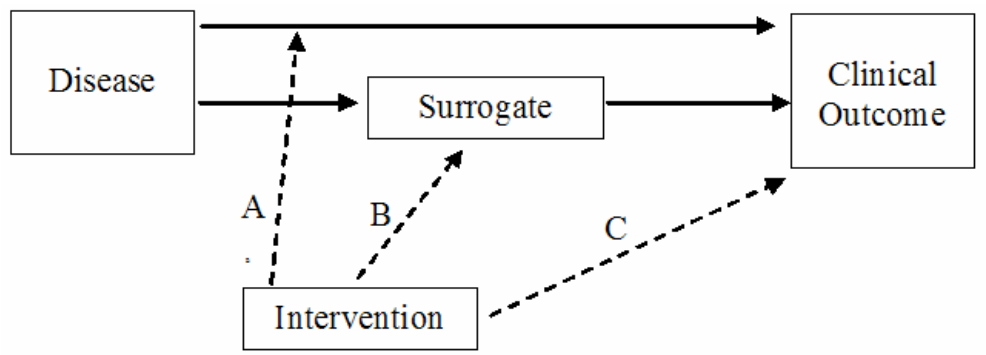

Figure 2-3. Net effect of treatment is only partially captured by the surrogate (adapted from Fleming et al.)[13]

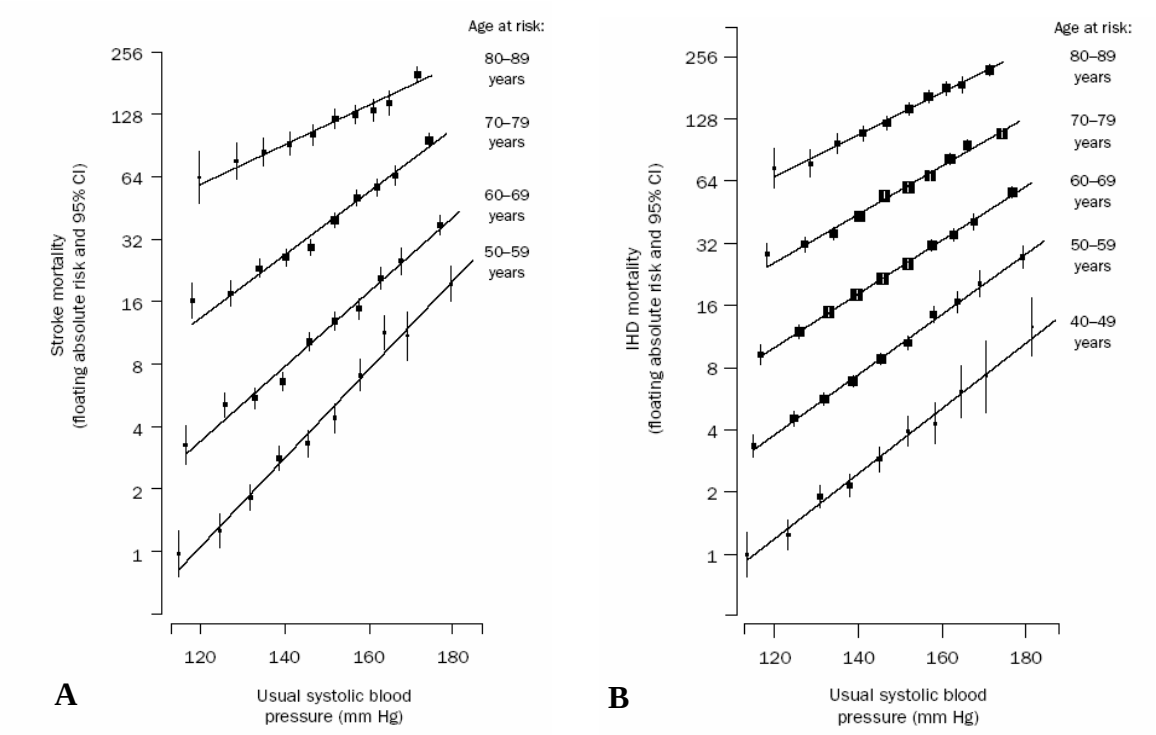

Figure 2-4. Mortality from stroke (A) and ischemic heart disease (B) in each decade of age versus usual systolic blood pressure at the start of that decade. Reprinted with permission from Lewington S, Clarke R, Qizilbash N, Peto R, Collins R. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. Lancet. Dec 14 2002;360(9349):1903-1913 (Figures 2 and 4, pages 1906 and 1908).[17]

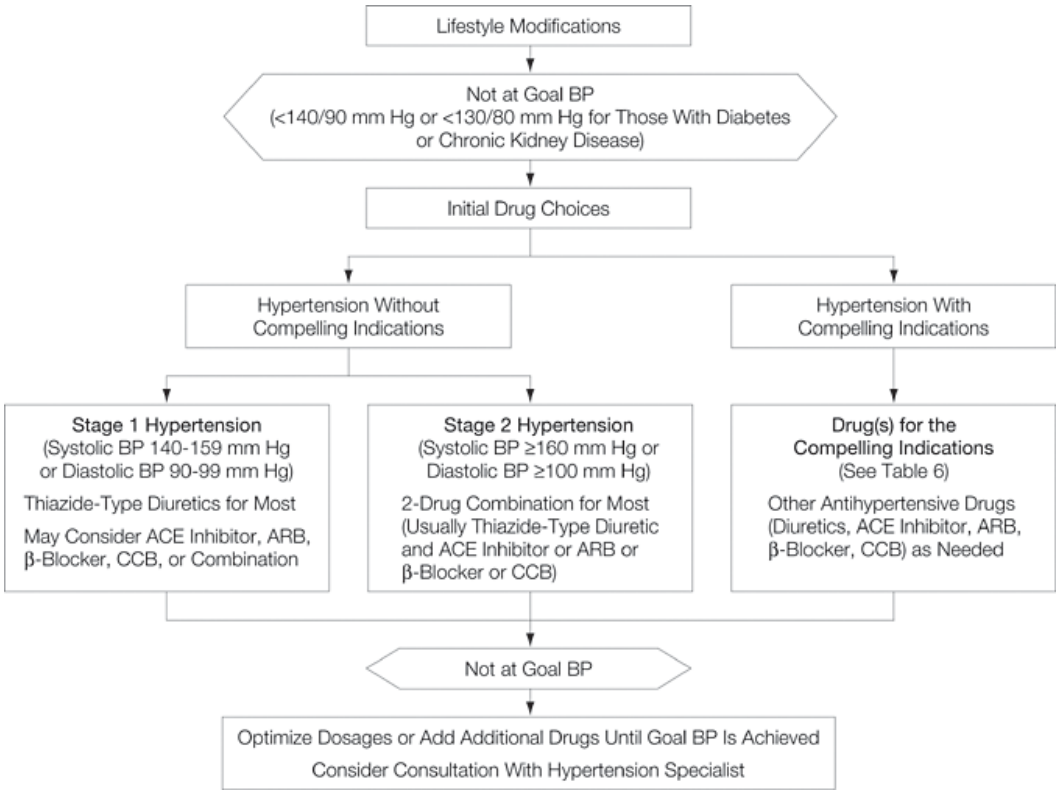

Figure 2-5. Algorithm for the treatment of hypertension. Reprinted with permission from Chobanian AV, Bakris GL, Black HR, et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. Jama. May 21 2003;289(19):2560-2572 (Figure 1, page 2564).[16]

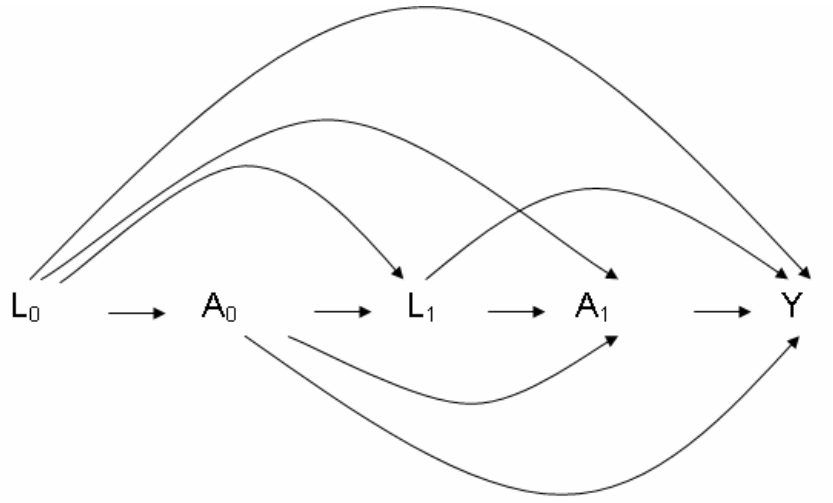

Figure 2-6. Directed acyclic graph for time-dependent confounding.
⟶, causal effect; L_0, vector of measured confounders at time 0; L_1, vector of measured confounders at time 0; A_0, treatment at time 0; A_1, treatment at time 1; Y, outcome of interest.

CHAPTER 3
METHODS

The methods for this study are presented in four parts: (1) a description of the dataset including operationalization of key variables, (2) a section detailing descriptive statistics, (3) a section describing several statistical models used to determine the association between systolic blood pressure over time and the risk of a primary outcome event, and (4) a section describing the methods used to determine the relation between time-dependent treatment and the risk of a primary outcome event. Sections 3 and 4 include both standard methods (Cox-regression with and without time-dependent covariates) and novel, causal methods (marginal structural Cox regression) to address potential bias introduced by time-dependent confounding.

The INVEST and the INVEST Dataset

The International Verapamil-Trandolapril SR Study (INVEST) was a large, international, randomized controlled antihypertensive trial involving patients with hypertension and coronary artery disease from 862 sites in 14 countries.[4] After an extensive cardiovascular history and physical exam the INVEST randomly assigned 22,576 CAD patients ≥50 years old to either a verapamil SR- or an atenolol-based multidrug antihypertensive strategy. Trandolapril and hydrochlorothiazide (HCTZ) were specified as added agents, if needed for blood pressure control, with trandolapril added first in the verapamil SR strategy and HCTZ added first in the atenolol strategy. In both strategies, trandolapril was recommended for patients with heart failure, diabetes, or renal impairment (Figure 3-1). Between 1997 and 2003, 61,835 patient-years follow-up were accumulated and each strategy provided excellent BP control (>70% of patients achieved BP <140/90 mm Hg) without differences in BP between the strategies. The strategies were equivalent in preventing the primary outcome defined as all-cause death, nonfatal myocardial infarction (MI), or nonfatal stroke. All components of the primary outcome (defined

as first occurrence of all-cause death, nonfatal MI, or nonfatal stroke) were fully adjudicated by an independent adjudication committee. Further details on the design and results have been published.[4, 5] The INVEST and all subsequent studies including the study at hand were approved by the institutional review board (IRB) of the University of Florida, which acted as the central IRB for all participating sites.

During the INVEST patients had scheduled visits every six weeks for the first six months and every six months thereafter. At each visit, patients were assessed for occurrence of symptoms, adverse events, and response to treatment. SBP and DBP were measured twice at each visit (at least two minutes apart) with a standard mercury sphygmomanometer in a sitting position. In a given patient throughout the trial all measurements were taken on the same arm, and, when possible, approximately the same time of day to minimize measurement error. In addition, all antihypertensive drug use was recorded at each visit. Throughout the remainder of the manuscript we refer to the antihypertensive drugs included in either of the INVEST treatment strategies (Atenolol, Verapamil, HCTZ, and Trandolapril) as *study drugs*, and all other antihypertensive drugs as *nonstudy drugs*. The term *total antihypertensive drugs* refers to both study and nonstudy antihypertensive drug use.

Follow-up continued until a patient was lost to follow up, died, or the end of the study. In the online data acquisition system, protocol visits were numbered consecutively from 1 (baseline) to 14 (maximum follow up of 5 years). Visits outside of the protocol schedule were also recorded and numbered 0. These visits were only included in the analysis if a protocol visit was not observed but an unscheduled visit was recorded in a time interval close to the omitted protocol visit. If patients did not return for one or multiple protocol visits and did not have suitable non-schedule visits to replace the unobserved protocol visit(s), values (e.g., SBP,

antihypertensive drug use) from the last observed visit were carried forward. If a patient was lost to follow-up (i.e., does not have another observed visit or final assessment), the patient was censored at the time of the last observed visit. For patients who experienced an event on the day of the recorded visit, BP and treatment measures from the last recorded visit before the event were used instead of the measures from the event visit to avoid the possibility that the observed measures on the event date were affected by the event (reverse causation).

Descriptive Statistics

The following basic descriptive statistics were computed at each visit:

- Number and percentage of patients on each INVEST study drug
- Number and percentage of patients by number of INVEST study drugs and number of total antihypertensive drugs used
- Mean SBP, and percentage of patients in 10 mm Hg SBP categories
- Change in number and percentage of patients between number of INVEST study, and total antihypertensive drugs between visits
- Change in percentage of patients within 10 mm Hg categories between visits

Blood Pressure and CV Outcomes

The association of systolic blood pressure with the risk of primary outcome event was first assessed unadjusted for time-dependent antihypertensive treatment using Poisson- and Cox proportional hazards regression.

Incidence Rates by Categories of Systolic Blood Pressure

Incidence of primary outcome events per category of systolic blood pressure was expressed as number of primary outcome events per 1000 patient years of follow-up. Blood pressure categories were defined as <110, 110-119, 120-129, 130-139, 140-149, 150-159, and ≥160 mm Hg. Adjusted incidence rates were calculated using Poisson regression. The model adjusted for following baseline covariates that include predictors of CV-outcomes in the INVEST[29], as well

as basic demographic variables: sex, ethnicity, age, residency (US vs. non-US), smoking status, history of heart failure, history of diabetes, history of renal impairment, prior stroke or transient ischemic attack, prior myocardial infarction, history of peripheral vascular disease, and prior coronary revascularization. Adjusted incidence rates are presented for female sex, White ethnicity, age at baseline between 60 and 70 years, US residency, and in absence of other risk factors (these values reflect the median values of the included variables).

Cox Proportional Hazards Models

The potential association of systolic blood pressure with the risk of primary outcome event unadjusted for treatment (other than at baseline) was assessed using standard and time-dependent Cox models. The models included the following static covariates measured at baseline (sex, ethnicity, age, residency (US vs. non-US), smoking status, history of heart failure, history of diabetes, history of renal impairment, prior stroke or transient ischemic attack, prior myocardial infarction, history of peripheral vascular disease, and prior coronary revascularization) as well as SBP. All Cox models categorized SBP in seven 10 mm Hg categories as defined above and used SBP 130 to 139 mm Hg as reference category. SBP categories were operationalized either as static variables (Equation 3-1) or as a time-dependent variables (Equation 3-2), depending on the respective model. The Cox models were specified as follows:

Cox Proportional Hazards model:

$$\lambda_i(t) = \lambda_0(t)\exp\{\beta_1 x_{i1} + (\beta_2 x_{i2} + ... + \beta_k x_{ik})\} \quad (3\text{-}1)$$

Cox model with time-dependent covariate:

$$\lambda_i(t) = \lambda_0(t)\exp\{\beta_1 x_{i1}(t) + (\beta_2 x_{i2} + ... + \beta_k x_{ik})\} \quad (3\text{-}2)$$

- $\lambda_i(t)$ = individual i's hazard to experience an event at time t
- $\lambda_0(t)$ = baseline hazard function at time t

- β_1 = association parameter for the SBP category (in the actual model, there are six parameters, one for each SBP category dummy variable)
- x_{i1} = individual i's SBP category (in the actual model, six dummy variables are used)
- $\beta_2 - \beta_k$ = association parameters for individual i's k-1 static covariates
- $x_{i2} - x_{ik}$ = k-1 static covariates for individual i
- $x_{i1}(t)$ = individual i's SBP category at time t (six dummy variables)

Six Cox models using different operationalizations of systolic blood pressure were evaluated: (1) baseline, (2) average over follow-up (fixed average), (3) average over follow-up weighted by follow-up time, (4) time-dependent using values from the previous visit (updated previous), (5) time-dependent using values from the next visit (updated next) and (6) time-dependent using an average updated at every visit (updated mean). Models 1 to 3 used a single, static value of SBP over the time of follow up. In contrast, models 4 to 6 are time-dependent and SBP values were updated at each observation (during the remainder of the study we refer to these models as 'updated models'). Figure 3-2 shows how these different models conceptionalize SBP over follow-up for a sample INVEST patient. The sample patient has a baseline SBP of 160 mm Hg (visit 1), observed scheduled visits 2, 3, and 5 (with measured SBPs of 140, 130, and 155 mm Hg, respectively), and experienced an event after 32 weeks (before scheduled visit 6). Importantly, Figure 3-2 shows continuous SBP values for each model, while the Cox models utilized categorized data as described above (i.e., for the Cox models the resulting SBP values are converted into dummy variables representing the 7 SBP categories).

The *baseline model* simply used the SBP observed at baseline to represent the patient's SBP throughout follow-up. Like the *baseline model*, the *average model* used a static SBP value to represent the patient's SBP over the entire follow-up, however, instead of the baseline SBP, it

used the average SBP calculated over the observed follow up period. The average was calculated as follows:

$$\overline{SBP} = \frac{\sum_{i=1}^{n-1} SBP_i t_i}{\sum_{n=1}^{n-1} t_i} \quad (3\text{-}3)$$

- $\overline{SBP}$ = average SBP over follow-up
- SBP_i = SBP at visit i
- n = total number of observed visits
- t_i = time between visit i and visit i+1

Note that, because visit 4 was not observed, the calculation assigned the SBP value observed at visit 3 to the entire time period between visits 3 and 5 (see Figure 3-2 for a numerical example of the calculation). The *time-weighted average model* used the same fixed average value calculated in the equation above, but weighted each individual's observation by his or her respective total follow-up time. This model thus weighted a subjects' contribution according to the total follow up-time the subject provided. For each time period between two observed visits, the *updated previous model* assigned the SBP value measured at the visit at the beginning of the respective time-period (i.e., it carries the value forward), while the *updated next model* assigned the SBP value from the visit that marked the end of the time period. Since no observations existed after an event is observed, the updated next model used the last available SBP measurement before the event for the time period from the last observed visit before the event up to the event (i.e., it used the same value as the updated previous model). Lastly, the *updated mean model* used an SBP average calculated as in equation 3-3, but instead of calculating a single average at the end of follow-up (as in the average, and time-weighted average models),

calculated a new (updated) average at each observed visit. Note that the updated mean SBP at the end of follow-up is equivalent to the average SBP over the entire follow-up.

A generalized R^2 was calculated for each of the six models and used to assess and compare the strength of association of the predictor variables with the outcome.[30]

$$R^2 = 1 - \exp\left(-\frac{G^2}{n}\right) \quad (3\text{-}4)$$

- G^2 = likelihood-ratio chi-squared statistic for testing the null hypothesis that all covariates have coefficients of 0
- n = sample size

Marginal Structural Cox Model

A marginal structural Cox model was used to estimate the effect of SBP control (SBP less than 140 mm Hg) over the course of follow-up on primary outcome controlling for potential time-dependent treatment. Because SBP control, the independent variable of interest, is a binary variable (a requirement of the marginal structural model) all patients with SBP of less than 110 at any visit were excluded. This was necessary because a previous report[31] and preliminary data from our analysis showed a J-shaped relationship between SBP and the risk for cardiovascular outcomes, with substantially increased risk for cardiovascular outcomes associated with SBP of less than 110 mm Hg. Thus, if patients with such low SBP (that would be included in the category of less than 140 mm Hg) were not excluded, the estimate of the benefit of controlled SBP would be skewed towards the null.

The remainder of this section describes the estimation of the marginal structural Cox model. First, the stabilized inverse probability of treatment and inverse probability of censoring weights were estimated (Equations 3-5 to 3-7). Stabilized weights have been shown to produce more narrow confidence intervals with better coverage rates. Note that in this instance *treatment*

refers to having controlled versus uncontrolled systolic blood pressure. In addition, the method requires the intend-to-treat like assumption that once treatment is initiated (here, SBP control is reached), patients remain on it until the end of their follow-up. Thus, the datasets used for the marginal structural Cox models are adjusted accordingly and all observations after treatment initiation are—regardless of observed exposure status—recoded as exposed to treatment.

Inverse probability of treatment weight (IPTW):

$$w_i(t) = \prod_{k=0}^{t} \frac{1}{pr[A(k) = a_i(k) \mid \overline{A}(k-1) = \overline{a}_i(k-1), \overline{L}(k) = \overline{l}_i(k)]} \tag{3-5}$$

Stabilized IPTW:

$$sw_i(t) = \prod_{k=0}^{t} \frac{pr[A(k) = a_i(k) \mid \overline{A}(k-1) = \overline{a}_i(k-1), V = v_i]}{pr[A(k) = a_i(k) \mid \overline{A}(k-1) = \overline{a}_i(k-1), \overline{L}(k) = \overline{l}_i(k)]} \tag{3-6}$$

Stabilized Inverse probability of censoring weight (IPCW):

$$sw_i^{\dagger}(t) = \prod_{k=0}^{t} \frac{pr[C(k) = 0 \mid \overline{C}(k-1) = 0, \overline{A}(k-1) = \overline{a}_i(k-1), V = v_i]}{pr[C(k) = 0 \mid \overline{C}(k-1) = 0, \overline{A}(k-1) = \overline{a}_i(k-1), \overline{L}(k-1) = \overline{l}_i(k-1)]} \tag{3-7}$$

Model parameters (uppercase letters represent random variables, lowercase letters denote specific realizations of that random variable):

- $w_i(t) =$ probability of individual *i* to have experienced his or her own observed treatment history from time 0 to time *t*
- $sw_i(t) =$ stabilized form of $w_i(t)$
- A(k) = 1 if SBP < 140 mm Hg, 0 otherwise
- L(k) = vector of all measured risk factors for Y at time k (number of antihypertensive drugs)
- Y = 1 if the INVEST primary outcome occurred
- V = vector of all baseline risk factors for Y
- $sw_i^{\dagger}(t)$ = stabilized weight for the probability of censoring for individual *i*

- C(k) =1 if a subject was lost to follow up by time k

The IPTWs were estimated using a pooled logistic regression model for the probability of having controlled systolic blood pressure at visit (k) conditional on baseline covariates (all measured baseline variables were included) and antihypertensive treatment (number of both study and total antihypertensive drugs) at baseline and visit (k-1). The IPCWs were estimated in the same fashion using a pooled logistic regression model for the probability of being censored at visit (k). Second, combined stabilized weights $sw_i(t)$ x $sw_i^{\dagger}(t)$ were calculated for each patient visit.

Lastly, the combined stabilized weights were used in a weighted Cox proportional hazards model. To overcome computational limitations of standard software (most available programs including SAS do not allow subject specific time-varying weights), the Cox proportional hazards model was estimated by fitting a pooled logistic regression that included the weights, baseline covariates and the time-dependent systolic blood pressure control variable.[26, 32] The model was specified as follows:

$$\text{logit pr}[D(t)=1 \mid D(t-1)=0, \bar{A}(t-1), V] = \beta_0(t) + \beta_1 A(t-1) + \beta_2 V \qquad (3\text{-}8)$$

$D(t)$ = 1 if the subject had an event in month t and $D(t)$ = 0 otherwise.

To assess whether antihypertensive treatment acted as a time-dependent confounder, we compared the hazard ratio for SBP control obtained from the marginal structural Cox model (e^{β_1} from equation 3-8 with each patient visit weighted by the combined stabilized weights) with the estimate obtained from a standard time-dependent Cox model (i.e., a model that did not adjust for time-dependent confounding). The hazard ratio for SBP control in the standard time-dependent Cox model was estimated simply by using equation 3-8 without the combined weights

(e^{β_1} from equation 3-8). Statistical significance of the difference was assessed by comparing the 95% confidence intervals of both estimates.

Antihypertensive Treatment and CV Outcomes

The time-dependent effect of treatment on the risk of primary outcome event was assessed both adjusted and unadjusted for time-dependent SBP, using Cox proportional hazards regression with and without combined stabilized weights (i.e., using a marginal structural- as well as a standard time-dependent Cox model as in the section above).

Marginal Structural Cox Models

A marginal structural Cox model similar to the one described before was used to estimate the effect of time-dependent treatment (aggressive antihypertensive treatment versus conventional antihypertensive treatment) on primary outcome controlling for SBP at each visit. Aggressive antihypertensive treatment was defined as being simultaneously exposed to three or more total antihypertensive drugs. To assess the sensitivity of the results to this rather arbitrary definition (that is necessitated by the method's restriction to a binary independent variable), the analyses were also conducted using the four following additional definitions for aggressive treatment: (1) more than one total antihypertensive drug, (2) more than three total antihypertensive drugs, (3) more than one antihypertensive study drug, and (4) more than two antihypertensive study drugs. Because of the U-shaped relationship between SPB and the risk for cardiovascular outcomes in INVEST, time dependent SBP was categorized into three categories, low (<120 mm Hg), normal (120 mm Hg to <140 mm Hg), and high (≥ 140 mm Hg).

As in the model for SBP control, the stabilized inverse probability of treatment and inverse probability of censoring weights were estimated (Equations 3-6 and 3-7). The IPTWs were estimated using a pooled logistic regression model for the probability of being exposed to aggressive versus standard antihypertensive therapy at visit (k) conditional on baseline covariates

(all measured baseline variables were included) and systolic blood pressure (low, normal, or high) at baseline, visit (k) and visit (k-1). The IPCWs were estimated in the same fashion using a pooled logistic regression model for the probability of being censored at visit (k).

Equations 3-5 to 3-8 apply as before with treatment now defined as aggressive antihypertensive treatment and SBP acting as a potential time-dependent confounder, thus:

- A(k) = 1 if the number of total antihypertensive drugs was greater than 2, 0 otherwise
- L(k) = vector of all measured risk factors for the outcome at visit k (including the three previously defined SBP categories)

As in the previous section, combined stabilized weights was computed and used in a weighted Cox proportional hazards model that is estimated through pooled logistic regression including the combined weights, baseline covariates, and the time-dependent treatment variable (aggressive antihypertensive treatment).

Cox Proportional Hazards Models

In addition to the marginal structural Cox model, a standard time-dependent Cox model was estimated analogous using the same pooled logistic regression model that was used in the marginal structural Cox model above (equation 3-8) but without weighting. The hazard ratio for SBP control obtained by the standard time-dependent Cox model was then compared to the one obtained from the marginal structural Cox model to determine whether a confounding effect of SBP on the hazard ratio for aggressive versus standard antihypertensive treatment exists.

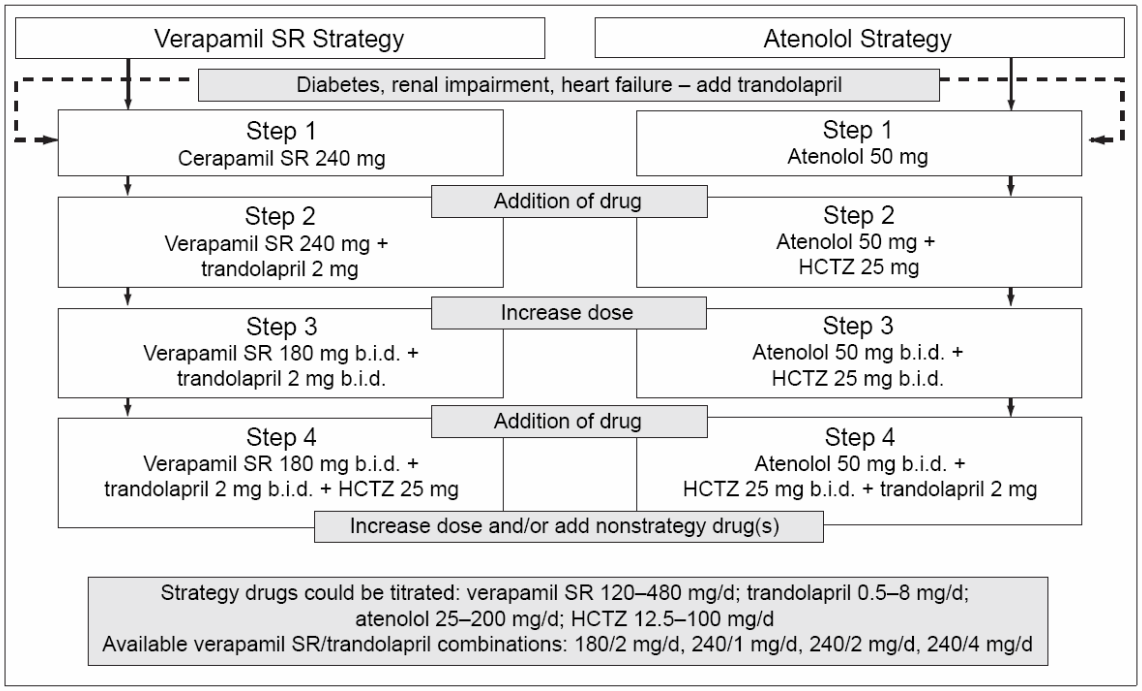

Figure 3-1. Treatment strategies in the INVEST. Reprinted with permission from Elliott WJ, Hewkin AC, Kupfer S, Cooper-DeHoff R, Pepine CJ. A drug dose model for predicting clinical outcomes in hypertensive coronary disease patients. J Clin Hypertens (Greenwich). Nov 2005;7(11):654-663 (Figure 1, page 656).[33]

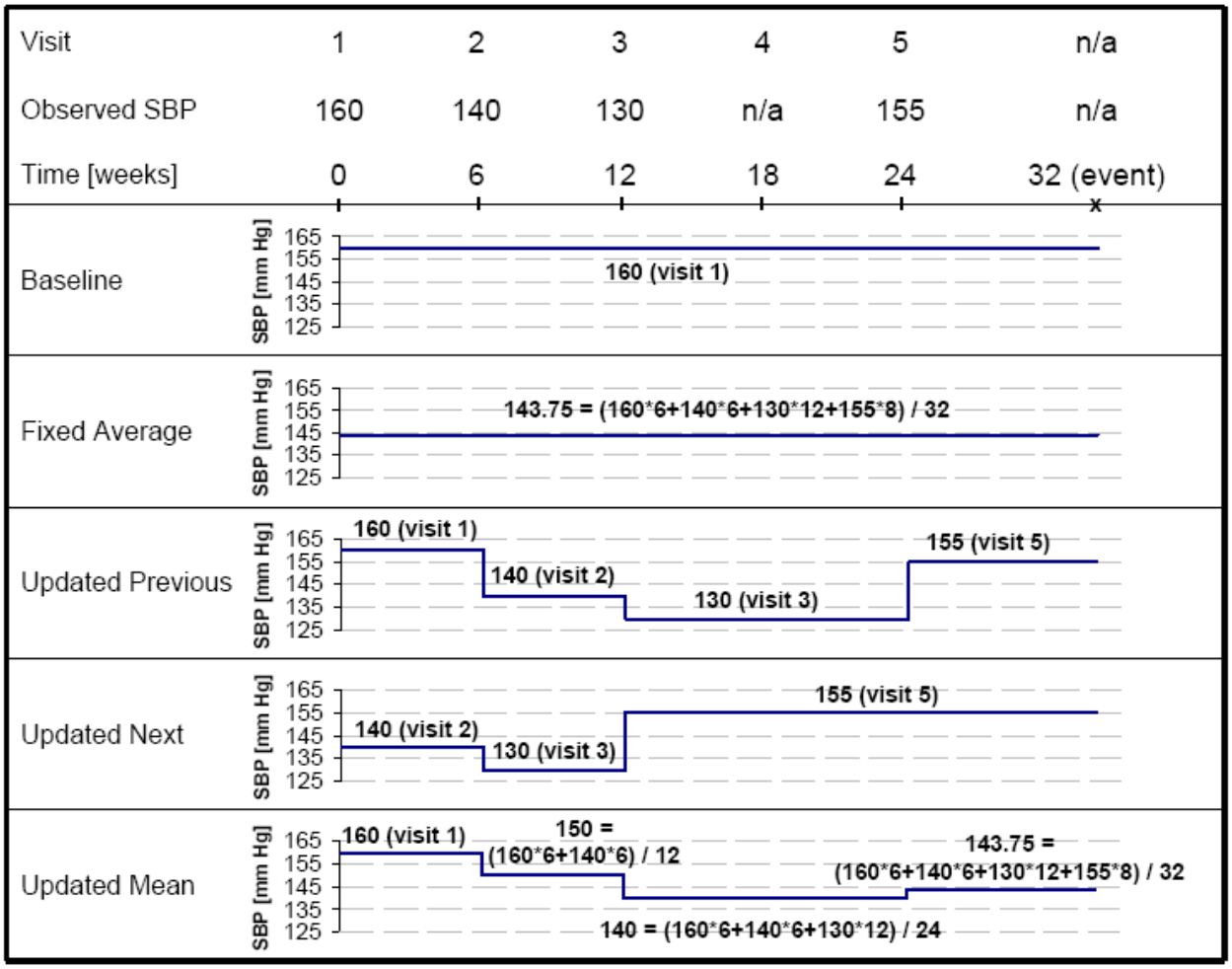

Figure 3-2. Operationalization of SBP: Examples for a sample patient.

CHAPTER 4
RESULTS

A total of 22,576 patients satisfied all requirements for inclusion into the original INVEST analysis.[4] Total follow-up time for the cohort was 61,845 patient years with 2,269 patients experiencing a primary outcome event during this period. For the present study, 906 patient years of follow-up were excluded from the analysis because they accrued after the occurrence of a nonfatal primary outcome event and thus, a total of 60,939 patient years of follow-up remained available for analysis.

Descriptives

Baseline characteristics of the INVEST cohort relevant to the present study are presented in Table 1. Briefly, the cohort had a mean age of 66.1 (±9.8) years and included slightly more women than men. The majority of patients were White, followed by large proportions of Hispanics and Blacks. Considerable proportions of patients had a history of cardiovascular events, or conditions recognized as cardiovascular risk factors. No breakdown by INVEST treatment strategy is provided since randomization is not relevant to the analyses presented in this study.

Average follow-up time to primary outcome event or censoring (end of follow-up or loss to follow up) was 2.7 (±0.9) years, ranging from 1 day (a patient who experienced a PO event on the day of the first visit) to a maximum of 5.4 years. Of a maximum of 14 possible physician visits designated for data collection and treatment adjustments, the average number of visits during INVEST follow up (including last encounter) was 7.3 (±2.7). After missed visits prior to censoring were imputed by carrying forward values from the last observed visit, this number increased to 9.5 (±1.8).

The number of patients at risk at each visit is depicted in Figure 4-1. All figures display follow-up over only 48 months (visits 1 to 12) due to the low numbers of uncensored patients at the last two visits (2583 patients after 54 months, and 792 patients after 60 months).

However, all analyses are conducted using data of all 60 months (visits 1to14). Ninety five percent of patients remained in the trial at 24 months, slightly less than 50% at 36 months and only about 12% at 48 months. Depending on the visit number, between 64% and 85% of uncensored patients had an observed visit. Visit four (week 18) shows the lowest percentage of observed visits, with only 64% of uncensored patients assessed at this visit. For the remainder of this study, all results are reported for the imputed data (with values carried forward for missed visits as described above) unless noted otherwise.

Blood Pressure

Mean SBP at the baseline visit was 150.9 mm Hg (±19.5). After 24 months, the mean SBP was reduced by 17.2 mm Hg to 133.7 mm Hg (±16.8). This reduction occurred mainly early in the trial with 52%, and 80% of the reduction observed at the six week and 12-week visits, respectively (Figure 4-2). When only observed visits were evaluated, SBP values at each visit were on average about 2.5 mm Hg lower than in the imputed dataset. For the majority of the analyses in this study, SBP was categorized into 7 categories, each spanning 10 mm Hg and ranging from smaller than 110 mm Hg to greater or equal to 160 mm Hg (Figure 4-3). At baseline 75% of patients had uncontrolled SBP (≥140 mm Hg) and 33% had a SBP of greater or equal to 160 mm Hg. Forty-four and 31% of patients showed uncontrolled SBP after three and 24 months, respectively, with 15% and 9% of patients in the range of greater than 160 mm Hg.

Figure 4-4 depicts the proportions of patients within each SBP category and at each visit (starting with visit two) who where not in the same category at the visit before (i.e., had changed SBP category between visits). The early drop in mean SBP shown in Figure 4-2 is mirrored at

the first data points (six week visit) in Figure 4-4. Between 70% and 85% of patients in the lowest four SBP categories had not been in the same category at baseline. With the exception of the highest SBP category that throughout follow-up at each visit included between 25% and 40% of patients who were in lower SBP categories at the visit before, all other categories show between 40% and slightly above 60% of patients with changed SBP category at each visit. Overall, at any given visit, only about half of all patients did not change blood pressure categories.

Antihypertensive Drugs

At the baseline visit, patients received on average 2.9 antihypertensive drugs, 1.5 of which were INVEST study drugs (Figure 4-5). After six months of follow-up the number of study drugs had increased to 1.9 while the total number of drugs had decreased to 2.6. These numbers remained relatively stable throughout month 30, after which the total number of drugs and the number of study drugs dropped to about 2 and 1.5, respectively. The proportions of patients receiving each individual study drug over follow-up are depicted in Figure 4-6. At baseline, the first line study drugs verapamil and atenolol were utilized by 50% and 45% of patients, with 20% and 33%, respectively, receiving the add-on drugs HCTZ and trandolapril. Over the follow-up period the percentage of patients on first line study drugs continuously dropped to 40% and 30% of patients on verapamil and atenolol after 24 months, respectively, and 38% and 28% after 48 months. Over the initial six months of follow-up, the proportion of patients on add-on study drugs increased to 42% and 44% for HCTZ and trandolapril, respectively. Starting at week 12, trandolapril was the most commonly used study drug within INVEST, with atenolol becoming the second most commonly used study drug after one year. From 30 months to the end of follow-up, the proportion of patients on study add-on drugs dropped to about 40% to 45% for HCTZ and trandolapril, after remaining relatively constant at 47% to 55% from 12 to 30 months.

Proportions of INVEST patients categorized by the number of study and total antihypertensive drugs are shown in Figures 4-7 and 4-8. At baseline 62% of patients received one, 28% two, and the remaining 10 percent three study antihypertensive drugs. After the first six months, the number of patients on one study drug had more than halved to 29%, while 41% and 25% of patients were placed on two and three study drugs, respectively, and 5% of patients did not receive any study antihypertensive treatment. Over the following months, the proportion of patients on one study drug continued to decrease to 19% at month 30, with 20% having no study antihypertensive drug at this time point. Over the same time period the proportion of patients on three study drugs increased to 30% while the proportion of patients on two study drugs dropped to 31%. After 30 months the percentages of patients on study drugs continued to drop slightly with the proportion of patients without study drugs increasing to 34% by month 48.

Similar trends can be observed when examining the data for all antihypertensive drugs. From baseline to 24 months most patients (between 33% and 39%) received two antihypertensive drugs followed by patients on three antihypertensive drugs. From month 30 to the end of follow-up, patients on three antihypertensive drugs contributed the largest proportion with 27% to 30%. Starting at month 36 and throughout the remainder of follow-up about one quarter of patients did not receive any antihypertensive drugs. Throughout the follow-up period a stable proportion of between 5% and 10% of patients each received five and more than five antihypertensive drugs.

Compared to the high proportion of patients who experienced changes in their SBP categories between visits (Figure 4-4), antihypertensive drug use was relatively stable. Figure 4-9 shows by visit and starting at the second visit the proportion of patients who had a change in the number of prescribed antihypertensive study drugs from the prior visit. After 18 months this

proportion stabilized at around 10% for patients on one to three study drugs, but ranging from 30% to 50% for patients who did not have a prescription for a study drug.

Compared to the changes reported for the number of antihypertensive study drugs (Figure 4-9), changes in the number of total antihypertensive drugs were more pronounced. Figure 4-10 shows by visit and starting at the second visit the proportion of patients who had a change in the number of prescribed total antihypertensive drugs compared to the prior visit. Patients with no antihypertensive drugs are most likely to report a change from the visit before (depending on the visit, 50% to 90% of patients who did not take any antihypertensive drugs were on at least one antihypertensive drug at the prior visit. After visit two and for those patients who had a prescription for at least one antihypertensive drug, the proportion who reports changes in the total number of antihypertensive drugs ranges from 10% to 45% with patients on five total antihypertensive drugs generally reporting the highest proportion of change.

Primary Outcome Events

Figure 4-11 displays a constant slope for the cumulative number of INVEST primary outcome events over the initial three years of follow-up, followed by a slightly steeper slope for the remaining years, indicating that the event rate in INVEST is initially constant with only a small increase after more than three years of follow-up.

Hazard Ratios

The following section shows plots of hazard ratios (HR) by categories of SBP, obtained by conventional and time-dependent Cox PH regression. All hazard ratios show the hazard for primary outcome event in a specific SBP category compared to the hazard for a primary outcome event in the category of SBP 130-140 mm Hg (reference category). Operationalization of SBP was varied and included baseline, mean over follow-up, mean over follow-up weighted by time

of follow-up, actual value at prior visit (time-dependent), actual value at the following visit (time-dependent), and updated mean over follow up (time-dependent).

Baseline SBP Model

Figure 4-12 shows the hazard ratios and 95% confidence intervals for a primary outcome event by categories of baseline SBP. The relationship between baseline SBP and the hazard for an outcome event follows a reverse J-shape with a nadir at the reference category of 130 to < 140 mm Hg. The lowest and two highest SBP categories show significantly increased risk, while the remaining three categories show no difference. The distribution of patients within each SBP category was as follows: Almost 33% of patients had a baseline SBP of greater than 160 mm Hg with relatively few patients contributing to categories of 110 and 110 to 120 mm Hg (3%) and smaller than 110 mm Hg (0.9%). The remaining categories of baseline SBP included between 8% (SBP 120 to <130 mm Hg) and 22% (SBP 140 to <150 mm Hg) of the 22576 INVEST patients.

Average SBP Model

Calculation of the average SBP over follow up for each subject resulted in 7259 patients (32%) with an average between 130 mm Hg and 140 mm Hg, followed by 6403 subjects (28%) with an average follow-up SBP between 120 mm Hg and 130 mm Hg. The two extreme SBP categories of less than 110 mm Hg and greater than 160 mm Hg, included fewer subjects with 190 (.8%) and 1500 subjects (6.6%), respectively. The relationship between risk for primary outcome and average SBP categories over follow-up follows a J-curve with the nadir at 120 mm Hg to 130 mm Hg (Figure 4-13). Patients with an average SBP over follow-up between 120 mm Hg and 130 mm Hg show no difference in the risk for an outcome event (HR 0.96, 95% CI 0.85-1.09) compared to the reference category, while all remaining categories show significantly higher risk. Different from the hazard ratios obtained for baseline SBP the increase in the hazard

is larger for the highest two categories of average SBP (with a maximum increase of 180%, 95% CI 44%-221%, at SBP greater or equal 160 mm Hg) than for the lowest category (80%, 95% CI 25%-158%). In comparison, the estimates for SBP of greater or equal 160 mm Hg and lower than 110 mm Hg were a 17% increase (95% CI, 2-33%) and a 64% increase (95% CI, 17-130%) when the baseline SBP was used.

Average SBP Weighted by Follow-up Time Model

Individual follow-up time in INVEST was used to weight each individual's average follow-up SBP (Figure 4-14) to avoid over-representation of patients who contributed little follow-up time due to an early event or loss to follow-up early in the trial. The resulting curve shows a similar shape as Figure 4-13 but is generally flatter. Its maximum hazard ratio occurs at an average SBP greater or equal to 160 mm Hg and with 2.11, 95% CI 1.76 – 2.54, is about one third lower than the hazard ratio for the un-weighted average at the same SBP category (HR 2.80, 95% CI 2.44-3.21). The risk at the lowest SBP category of less than 110 mm Hg is not different from the risk at the next higher category and shows a very wide confidence interval due to the low number of patients with low average SBP.

Time-Dependent SBP Models

Figure 4-15 shows the relationship between the SBP category observed at the last recorded visit of a given time interval in the Cox model and the risk for a primary outcome event. SBP categories were continuously updated over the course of follow-up. Figure 4-15 shows a flat V-shape with its nadir at SBP between 130 mm Hg and 140 mm Hg. All but the category between 120 to 130 mm Hg show a significantly increased risk for a primary outcome event compared to the reference category. The lowest and the two highest SBP categories show slightly over 50% increase in risk while the remaining two categories are at about 20%.

In contrast to Figure 4-15, Figure 4-16 shows hazard ratios for the primary outcome by category of time-dependent SBP from the next observed visit. Here, SBP for each given time interval is taken from the first observed visit after this time interval. One exception was made: For the time period immediately preceding an event, the SBP from the last observed visit was used since there was no observation available after the patient was censored. Figure 4-16 shows a J-shaped curve with nadir for the hazard ratio at 120 mm Hg to 130 mm Hg, significantly lower than the reference category (HR 0.73, 95% CI 0.65-0.83). Departing from the curve's J-shape the two lowest SBP categories show almost identical hazard ratios.

Updated Mean SBP Model

The hazard ratios for Figure 4-17 were derived calculating at each observed visit an updated mean SBP that was used as a time-dependent covariate in the model. This produced a nearly symmetric V-shaped curve with its nadir at SBP 130-140 mm Hg. With the exception of updated mean SBP values between 120 mm Hg and 130 mm Hg (HR 1.09, 95% CI 0.96-1.23), all remaining SBP categories showed significantly higher hazard ratios than the reference category. The lowest and highest categories showed the greatest increase in risk with hazard ratios of 1.61 (95% CI, 1.12-2.32), and 1.80 (95% CI, 1.56 – 2.06), respectively. Compared to the analysis using the unweighted average, the highest three categories using the updated mean showed between 13% (SBP between 140 mm Hg and 150 mm Hg) and 35% (SBP larger than 160 mm Hg) lower risk for a primary outcome event. Compared to the time dependent SBP model (prior visit) the results of the updated mean model show only small differences ranging from a 5% lower estimate (SBP between 150 mm Hg and 160 mm Hg) to a 25% overestimation (SBP between 110 mm Hg and 120 mm Hg) with the lowest and highest SBP category within 3% and 13% respectively.

Incidence

While the previous section presented numerous estimates of the hazard for an outcome event in a specific SBP category relative to the reference category of 130 to 140 mm Hg, absolute risks for patients in specific SBP categories have thus far not been presented. The following section shows both unadjusted and adjusted incidence rates presented as the number of primary outcome events per 1000 person years spent in the seven respective SBP categories.

Figure 4-18 presents unadjusted incidence rates for the primary outcome events by category of SBP. Each subject's follow-up period could contribute person time to multiple SBP categories. Person time was accrued for the last observed SBP category until a change in SBP category, the occurrence of an event, or censoring. Likewise, events were attributed to the SBP category observed at the last visit prior to the event. Most person years of follow-up were accrued in the category of 130-140 mm Hg (16188 person years), while the category of <110 mm Hg contributed the lowest number (1678 person years). The crude incidence of the INVEST primary outcome by SBP category was V-shaped and consequently highest in the extremely low and high SBP categories <110 mm Hg (53.6 primary outcome events/1000 person years) and ≥160 mm Hg (53.8 primary outcome events/1000 person years). The crude incidence rates in the categories of 130 to 140 mm Hg, and 120 to130 mm Hg were more than 40% lower with 30.6 primary outcome events/1000 person years for these SBP categories.

The respective adjusted incidence rates, calculated by Poisson regression for White, US, female patients without co-morbidities and between the ages of 60 and 70 years, displayed in Figure 4-19 show a similar V-shape and range from 14.1 primary outcome events/1000 person years for the SBP category between 130 to140 mm Hg to 18.8 primary outcome events/1000 patient years for the category greater than 160 mm Hg.

Comparisons

The two previous sections present several approaches to determine relative and absolute risks to experience an outcome event for patients of the INVEST depending on their SBP category. The following section compares the results of these approaches and provides a brief description of the assumptions implied in each of the modeling approaches.

Table 4-2 presents a summary of the results obtained by the modeling approaches of the previous two sections. For each model, Table 4-2 shows the hazard ratios (relative risks for the Poisson regression model) for the lowest SBP category (<110 mm Hg), the category with the lowest risk for an outcome event, and the highest SBP category (≥160 mm Hg). Hazard ratios and relative risks for the lowest and highest SBP category are presented both in comparison to the reference category (130 to140 mm Hg) and the SBP category with the lowest risk (in parenthesis) if that is not the reference category in the respective model. Additionally, Table 4-2 shows the generalized R^2 for each model. Compared to the reference category, the hazard ratios/relative risks for the SBP category of lower than 110 mm Hg range from 1.16 (time-dependent next) to 1.80 (average over follow up). At the same time hazard ratios/ relative risks for the highest SBP category (>160 mm Hg) range from 1.17 (baseline) to 2.80 (average). Notably, the most extreme estimates for both the lowest and highest SBP categories were obtained from the average SBP model. Generalized R^2 values for the presented models were small, ranging from 0.039 to 0.070, with little difference between models.

Average SBP and Bias

Table 4-3 illustrates in a much simplified scenario that the use of an average of a measure (e.g., controlled versus uncontrolled SBP averaged over follow-up) can lead to biased estimates if (1) the duration follow-up time is not fixed but rather determined by the occurrence of an event and (2) there is a directed change of the measure over the course of follow-up (i.e., mean SBP is

lowered during the course of INVEST, especially early). Of note, conditions (1) and (2) are typically met in the context of clinical trial data that treat a surrogate over follow-up and have a primary outcome at which patients are censored, such as the INVEST data. It also applies to observational studies where a surrogate shows a time-dependent upward or downward trend. Notably a directed change in a mean over follow-up can also occur as the result of regression to the mean if inclusion of subjects in the study is dependent on extreme values of the surrogate.

The simplified scenarios (A to F) below make the following assumptions:

- Patients start with either controlled or uncontrolled blood pressure
- Duration of follow-up is three years
- Measures are taken after the first and third years
- The event rate is constant at 10% per year and independent of blood pressure control
- Events occur at the time of measurement
- At the end of follow up the average is calculated as high and low, respectively, for those patients who remained in their BP category over follow up and for those who changed BP category after the first year as the value of the second (longer) measurement period.

Scenarios A to F show the effects of varying longitudinal changes in BP (A to D), and different patient distributions between controlled and uncontrolled BP at baseline (E to F):

- **Scenario A:** 100 patients each start with controlled and uncontrolled BP respectively. Patients do not change BP categories throughout follow-up.
- **Scenario B:** 100 patients each start with controlled and uncontrolled BP respectively. 30% of patients with uncontrolled blood pressure will be controlled after the first follow-up year. Patients with controlled BP will continue to have controlled BP.
- **Scenario C:** 100 patients each start with controlled and uncontrolled BP respectively. 60% of patients with uncontrolled blood pressure will be controlled after the first follow-up year. Patients with controlled BP will continue to have controlled BP.

- **Scenario D:** 100 patients each start with controlled and uncontrolled BP respectively. 60% of patients with uncontrolled blood pressure will be controlled after the first follow-up year. 60% of patients with controlled blood pressure will have uncontrolled BP after the first follow-up year and over the remainder of follow-up.
- **Scenario E:** Equivalent to scenario D except that 200 patients have uncontrolled BP and 100 patients have controlled BP at baseline.
- **Scenario F:** Equivalent to scenario D except that 100 patients have uncontrolled BP and no patients have controlled BP at baseline.

Scenario A (no changes in BP categories over follow-up) results, as expected, in an event rate of 10 events per 100 patient years. Overall, 28 events occurred, 10 (35.7%) of which occurred after the first year. The number is slightly higher than one third because only 90 of the original 100 patients were still at risk for an event in the second, two-year long, follow-up period. Scenario B, which has 30% of patients switch from uncontrolled to controlled BP after one year, deviates from the expected event rates. It results in an event rate of 11.6% for patients with average uncontrolled BP over follow-up and 9.2% for patients with average controlled BP. The percentage of all events of the categories stemming from patients whose average BP was based on only the first year of follow-up is about 10 percentage points higher, at 45.5%, as expected for patients with uncontrolled BP and about 6% lower than expected for those who average controlled BP (29.4%). Of note, the number of events after one year of follow up is constant throughout all four scenarios at 10 each for controlled and uncontrolled patients, respectively. Scenario C, which doubles the percentage of patients who, starting as uncontrolled, achieve BP control after the one year visit, results in an even larger deviation from the true event rate. This scenario results in a 16% event rate for those patients averaging uncontrolled BP and an 8.7% event rate for those averaging controlled BP. Simultaneously, the percentage of all events per category contributed by patients whose average was only based on the first year of follow-up diverged further from the expected 35.7%, with 62.5% of all events within patients

whose average BP over follow-up was uncontrolled resulting from patients who had an event at the end of the first year, and only 25% of all events in patients with an average controlled BP attributable to patients who had their event after one year. Model D demonstrates, that when patients switch BP categories evenly, that is when as many patients become uncontrolled after one year as achieve BP control, calculated event rates—as in scenario A where patients do not switch between categories at all—match expected event rates. Scenario E demonstrates that not the proportional change between categories is responsible for the proportion of events that occur in the first year but rather the absolute number of patients with a change in BP categories. While the proportions of patients who switch categories in scenario E are identical (60% in each group), the absolute number of patients who switch from uncontrolled to controlled is double the number of patients that switch from controlled to uncontrolled since twice as many patients started follow-up in the uncontrolled group. Overall, model E results in 45.5% of all events within the uncontrolled group from the first year of follow-up. Lastly, Scenario F, in which all patients start with uncontrolled BP and thus, change is by definition directed for the first year of follow up, results in 62.5% of all events within the uncontrolled group resulting from the first visit compared to 0% of all events in the controlled group from the same time period (since no patients with uncontrolled BP were at risk during the first year).

To show the consequences of using the average BP control over follow-up as the predictor of cardiovascular risk, we calculated the relative risks to experience an event comparing patients with an uncontrolled average over follow-up to patients with a controlled average over follow up. The relative risks were calculated by dividing the absolute risk (event rate per patient year) of uncontrolled patients over the absolute risk of patients with controlled SBP over follow-up. Absolute risk estimates were obtained by dividing the number of events that occurred in each

group by the person years of follow-up in the respective BP category based on the average BP over follow up (e.g., a patient who was uncontrolled during year one and controlled for the remaining two years contributes all 3 years of follow-up to the uncontrolled SBP group). Of note, the true relative risk is 1.0, since the event rates used for the simulation were identical between groups at 10% per patient per year.

Scenarios A and D, the scenarios that show equal numbers of patients in both average groups after follow-up (i.e., lack directed change of the measure) accurately show relative risks of 1.0. In contrast, scenarios B and C, that both simulate a directed change of patients from uncontrolled to controlled average BP over follow-up as a result of higher rates of change between groups from uncontrolled to controlled BP, show relative risks of 1.26 and 1.84, respectively. However, scenario E in which patients change between groups in equal proportions as in scenario D, but shows a directed net change of subjects resulting from regression to the mean (since twice as many patients start follow-up in the uncontrolled group with both groups having equal rates of change), estimates a relative risk of 1.59, similar to scenarios B and C. Lastly, scenario F, that at baseline only includes patients with uncontrolled BP and thus shows the strongest possible directed change of patients between groups, estimates the largest relative risk with 2.39. Scenarios A to F show that even with identical true event rates per year of follow-up, the mere fact that a directed change in BP control exists, leads to an overestimation of risk in the categories that lose patients and to an underestimation of risk in those categories that gain patients.

In contrast, Table 4-4 shows the results of the equivalent six simulations, when the denominator (person time) of the absolute risk estimates for each BP category is time-dependent. In these simulations, a patient who started with uncontrolled BP at baseline and achieved blood

pressure control for years two and three, contributes one person year to the total person time in uncontrolled BP and two person years to the total person time in controlled BP (in the average calculation the same patient contributes all three years to the total person time in controlled BP since his average over the entire follow-up is controlled). Notably, when using this updated mean SBP, the simulations under all scenarios correctly estimate the true relative risk of 1.0.

The results of these simple simulations help explain the observed differences between the Cox proportional hazards regressions with average SBP over follow-up (Figure 4-15) and updated mean SBP (Figure 4-19). In the simulation we observe a correlation between the magnitude of the observed bias in Scenarios B and C and the proportion of total events within each category that resulted from the patients with only one year of follow-up at their event time. The higher this proportion, the higher was the observed relative risk (i.e., bias, since the true RR is 1.0). Figure 4-20 illustrates within INVEST the calculation of bias introduced by the Cox model using the average SBP, expressed as the percent difference from the hazard ratio obtained by the time-dependent Cox model using the updated mean SBP. The largest bias exits in the highest two SBP categories where the Cox model using the average over follow-up overestimates the hazard ratios by 34% and 56%. Figure 4-21 shows for INVEST within categories of average blood pressure over follow-up the proportion of total events that were contributed by patients based on their number of observed visits at the time of the event. Figure 4-21 shows that patients in the highest two and the lowest average SBP categories were considerably more likely to have their average based on only the baseline visit (i.e., their event occurred before the second visit).

Figure 4-22 combines data generated by Figures 4-20 and 4-21 and plots the bias of the Cox model using the average SBP (Figure 4-20) and the proportion of events resulting from

patients with only two observed visits (Figure 4-21) in the same graph by SBP category. As observed in the simulation, the bias resulting from using an average over follow-up correlates with the percentage of events after the first follow-up period, strongly suggesting, that the higher hazard ratios obtained from the Cox model using average SBP over-follow up are indeed resulting from a bias analogous to the bias simulated in scenarios B, C, E, and F of Table 4-2.

Of note, the time-weighted average is equally affected by this suggested bias. However, the impact of the bias is likely attenuated since observations who contribute strongest to the bias (patients who experience events early) receive small weights compared to patients with longer follow-up times.

Marginal Structural Models

The first part of this section intends to expand on the analyses of the prior sections by using a marginal structural Cox model to estimate the effect of SBP on the risk for cardiovascular events controlling for potential time-dependent confounding by antihypertensive drug use. However, instead of the seven SBP categories used in the previously presented models, a binary operationalization of SBP (controlled vs. uncontrolled, with controlled defined as SBP< 140 mm Hg) was used to accommodate current limitations in the method. Note that 2038 patients with extremely low SBP values (< 110 mm Hg at any visit during follow-up) were excluded to allow the use of a binary SBP variable (required by the marginal structural Cox model) even though the relationship between SBP and cardiovascular outcomes follows a U shape (Figures 4-12 to 4-19). Time-dependent use of antihypertensive drugs was operationalized by the number of total antihypertensive drugs and the number of study antihypertensive drugs (defined as antihypertensive drugs that were part of the INVEST protocol) at baseline, each follow-up visit, as well as lagged drug variables that, at each visit other than baseline, represent the values of the respective prior visit. The second part of the section aims to estimate the effect of

antihypertensive drug use controlling for time-dependent confounding by SBP. Antihypertensive drug use was operationalized as aggressive (more than two total antihypertensive drugs) versus standard (two or less total antihypertensive drugs) antihypertensive therapy, with additional definitions used in a sensitivity analysis. SBP was operationalized as low (lower than 120 mm Hg); normal (120 mm Hg to lower than 140 mm Hg); and high (equal or higher than 140 mm Hg).

Effect of Systolic Blood Pressure Control

Table 4-5 shows the result of the inverse probability of treatment weighted Cox model (marginal structural Cox model) compared to the equivalent standard time-dependent Cox model. Both models show a significantly reduced hazard for patients with controlled SBP compared to patients with uncontrolled SBP. The standard model estimates a 46% (95% CI 40%-52%) reduction in the hazard for a cardiovascular event while the inverse probability of treatment weighted model estimates an almost identical reduction of 45% (95% CI 39%-50%). There was no significant difference between the estimates of both models. A significant difference between the models would support the presence of a time-dependent confounding effect of antihypertensive treatment (i.e., a non SBP mediated effect of aggressive antihypertensive treatment as defined).

Effects of Antihypertensive Drugs

Tables 4-6 provides an inverse probability of treatment weighted estimate for the effect of aggressive (more than two concurrent total antihypertensive drugs) versus standard antihypertensive treatment (two or less concurrent total antihypertensive drugs), compared with estimates from the equivalent standard time-dependent Cox model. The standard Cox model estimated no difference in cardiovascular risk between the treatment with aggressive and standard antihypertensive therapy (HR 0.96, 95% CI 0.87-1.07). In contrast, the inverse

probability of treatment weighted Cox model estimated a significant 19% (95% CI 8%-29%) reduction in the hazard of an outcome event for patients treated with aggressive antihypertensive therapy. However, the two results have overlapping 95% confidence intervals and thus, are not significantly different from each other. A significant difference between the two estimates would suggest that time-dependent confounding by SBP control substantially alters the estimate of the effect of receiving aggressive antihypertensive treatment.

Table 4-6 shows the results of a sensitivity analysis conducted to assess whether and by how much the results were influenced by the definition of aggressive antihypertensive therapy. As in the original analysis (Table 4-6), the effect of aggressive versus standard antihypertensive treatment was assessed by estimating both a standard Cox model and a marginal structural Cox model. Four additional definitions of aggressive antihypertensive therapy were assessed, (1) more than one total antihypertensive drug, (2) more than three total antihypertensive drugs, (3) more than one antihypertensive study drug, and (4) more than two antihypertensive study drugs. Regardless of the definition used for aggressive antihypertensive treatment, all four comparisons estimated a larger beneficial effect for aggressive antihypertensive therapy from the marginal structural Cox model compared to the standard model.

As in the primary analysis, the difference between the models failed to reach significance in three of the four models. The analyses that define aggressive antihypertensive treatment as the concomitant use of more than three total antihypertensive drugs found the hazard ratio estimated from the marginal structural Cox model (HR 0.79, 95% CI 0.71-0.89) to be significantly lower than the one estimated by the standard model (HR 0.98, 95% CI 0.90-1.07). The analysis defining aggressive antihypertensive therapy as the concomitant use of more than one study antihypertensive drug found a borderline significant difference between the models with an HR

estimate of 0.61 (95% CI 0.54-0.68) for the marginal structural Cox model and 0.73 (95% CI 0.67-0.80) for the standard time-dependent Cox model, respectively.

Table 4-1. Composition of the INVEST Cohort at Baseline

Variable	N=22576
Demographic	
Age, mean (SD), years	66.1 (9.8)
Women	11770 (52.1)
Race/Ethnicity	
White	10925 (48.4)
Black	3029 (13.4)
Hispanic	8045 (35.6)
Other	577 (2.6)
Calcium Antagonist Strategy	11267 (49.9)
Condition	
Myocardial infarction	7218 (32.0)
Stroke/transient ischemic attack	1629 (7.2)
Congestive Heart Failure	1256 (5.6)
Diabetes	6400 (28.4)
Renal impairment	424 (1.9)
Peripheral vascular disease	2699 (12.0)
CABG	6166 (27.3)
Smoking (ever)	10454 (46.3)

Abbreviations: SD, Standard deviation; CABG, Coronary artery bypass graft.
Unless indicated otherwise, values express numbers (percentage).

Table 4-2. Comparison of models

Model	Hazard Ratio at SBP <110 mm Hg	Minimum Hazard Ratio (at SBP category)	Hazard Ratio at SBP ≥160 mm Hg	Generalized R^2
Baseline	1.64	1.00 (130 to <140 mm Hg)	1.17	0.059
Average over follow-up	1.80 (1.88*)	0.96 (120 to <130 mm Hg)	2.80 (2.92*)	0.070
Time-weighted average	1.41	1.00 (130 to <140 mm Hg)	2.54	0.039
Time-dependent (prior)	1.57	1.00 (130 to <140 mm Hg)	1.59	0.061
Time-dependent (next)	1.16 (1.59*)	0.73 (120 to <130 mm Hg)	2.26 (3.10*)	0.071
Updated mean	1.61	1.00 (130 to <140 mm Hg)	1.80	0.062
RR from Poisson	1.50**	1.00** (130 to <140 mm Hg)	1.57**	n/a

*Compared to the lowest risk category, ** Relative risk

Table 4-3. Simulation of six scenarios using average over follow-up

BP Status	1st Year of Follow-up				2nd & 3rd Years		Average Over Follow-up				
	N (baseline)	Events	Change in BP status	No change in Bp status	N (year 2)	Events	Total Person Time*	Total Events	Event rate per patient year	Percentage of events from 1st year	Relative Risk
					Scenario A						
Uncontrolled	100	10	0	90	90	18	280	28	10%	35.7	1.0
Controlled	100	10	0	90	90	18	280	28	10%	35.7	1.0
					Scenario B						
Uncontrolled	100	10	30	60	60	12	190	22	11.6%	45.5	1.26
Controlled	100	10	0	90	120	24	370	34	9.2%	29.4	1.0
					Scenario C						
Uncontrolled	100	10	60	30	30	6	100	16	16%	62.5	1.84
Controlled	100	10	0	90	150	30	460	40	8.7%	25	1.0
					Scenario D						
Uncontrolled	100	10	60	30	90	18	280	28	10%	35.7	1.0
Controlled	100	10	60	30	90	18	280	28	10%	35.7	1.0
					Scenario E						
Uncontrolled	200	20	120	60	120	24	380	44	13.8%	45.5	1.59
Controlled	100	10	60	30	150	30	460	40	8.7%	25	1.0
					Scenario F						
Uncontrolled	100	10	60	30	30	6	100	16	16%	62.5	2.39
Controlled	0	0	0	0	60	12	180	12	6.7%	0	1.0

*in average over follow-up controlled vs. uncontrolled

Table 4-4. Simulation of six scenarios using updated mean

BP Status	1st Year of Follow-up				2nd & 3rd Years		Average Over Follow-up				
	N (baseline)	Events	Change in BP status	No change in Bp status	N (year 2)	Events	Total Person Time*	Total Events	Event rate per patient year	Percentage of events from 1st year	Relative Risk
					Scenario A						
Uncontrolled	100	10	0	90	90	18	280	28	10%	35.7	1.0
Controlled	100	10	0	90	90	18	280	28	10%	35.7	1.0
					Scenario B						
Uncontrolled	100	10	30	60	60	12	220	22	11.6%	45.5	1.0
Controlled	100	10	0	90	120	24	340	34	9.2%	29.4	1.0
					Scenario C						
Uncontrolled	100	10	60	30	30	6	160	16	16%	62.5	1.0
Controlled	100	10	0	90	150	30	400	40	8.7%	25	1.0
					Scenario D						
Uncontrolled	100	10	60	30	90	18	280	28	10%	35.7	1.0
Controlled	100	10	60	30	90	18	280	28	10%	35.7	1.0
					Scenario E						
Uncontrolled	200	20	120	60	120	24	440	44	13.8%	45.5	1.0
Controlled	100	10	60	30	150	30	400	40	8.7%	25	1.0
					Scenario F						
Uncontrolled	100	10	60	30	30	6	160	16	16%	62.5	1.0
Controlled	0	0	0	0	60	12	120	12	6.7%	0	1.0

* updated mean controlled vs. uncontrolled

Table 4-5. Inverse probability of treatment weighted estimates for the causal effect of controlled SBP on primary INVEST primary outcome event

	Hazard Ratio	95% Confidence Interval
Standard Cox model	0.54	0.48-0.60
Marginal Structural Cox Model	0.55	0.50-0.61

SBP control is defined as SBP < 140 mm Hg; patients who had SBP of <110 mm Hg at any visit were excluded from the analysis (n=2038)

Table 4-6. Inverse probability of treatment weighted estimates for the effect of receiving more than two total antihypertensive drugs on primary INVEST primary outcome event

	Hazard Ratio	95% Confidence Interval
Standard Cox model	0.96	0.87-1.07
Marginal Structural Cox Model	0.81	0.71-0.92

Table 4-7. Inverse probability of treatment weighted estimates for the causal effect of receiving various numbers of total antihypertensive drugs on INVEST primary outcome event

	Hazard Ratio	95% Confidence Interval
> 1 total antihypertensive drug		
Standard Cox model	0.89	0.70-1.13
Marginal structural Cox model	0.83	0.64-1.06
> 3 total antihypertensive drugs		
Standard Cox model	0.98	0.90-1.07
Marginal structural Cox model	0.79	0.71-0.89
> 1 study antihypertensive drug		
Standard Cox model	0.73	0.67-0.80
Marginal structural Cox model	0.61	0.54-0.68
> 2 study antihypertensive drugs		
Standard Cox model	0.78	0.71-0.85
Marginal structural Cox model	0.72	0.64-0.81

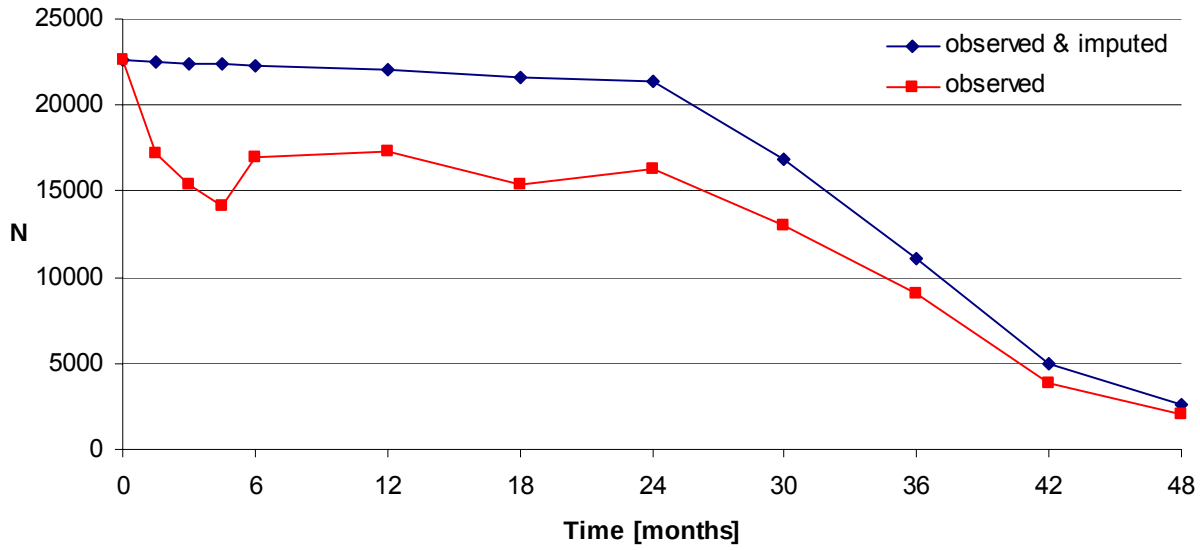

Figure 4-1. Patients remaining in the study at each visit

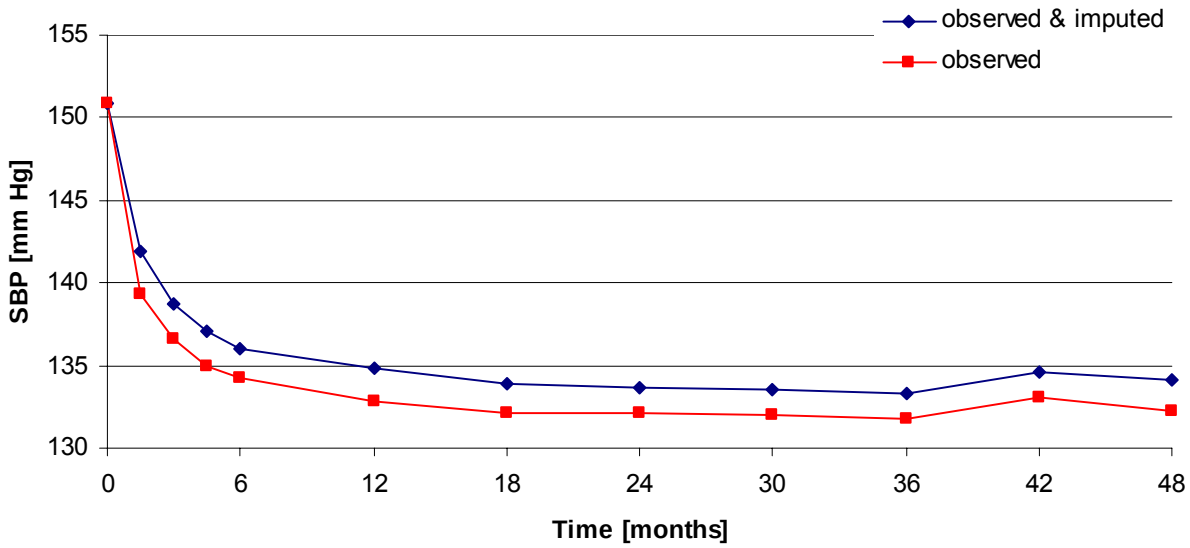

Figure 4-2. Mean systolic blood pressure over follow-up (observed vs. imputed data)

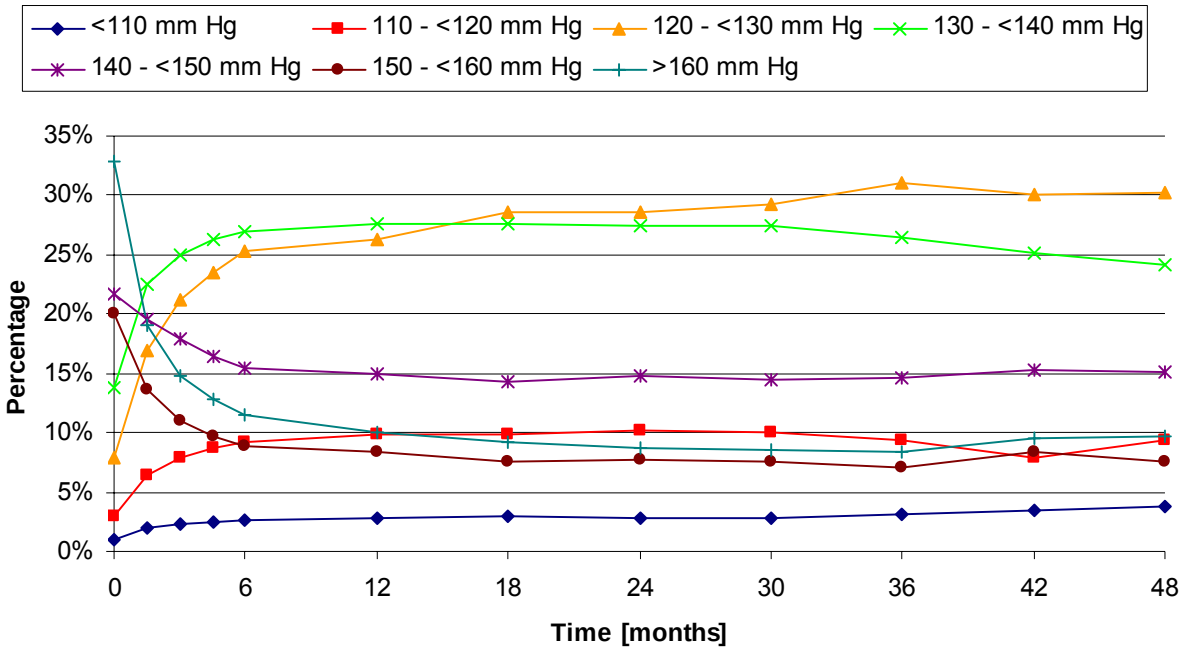

Figure 4-3. Percentage of patients within each SBP category over follow-up

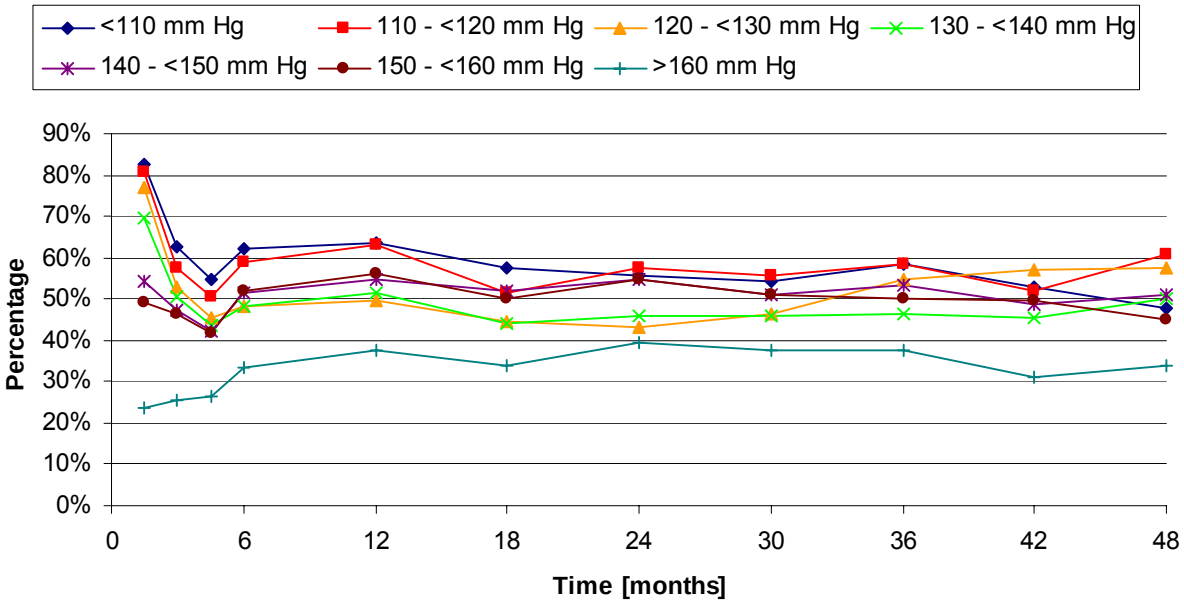

Figure 4-4. Percentage of patients within each SBP category who were not within the same SBP category at the prior visit

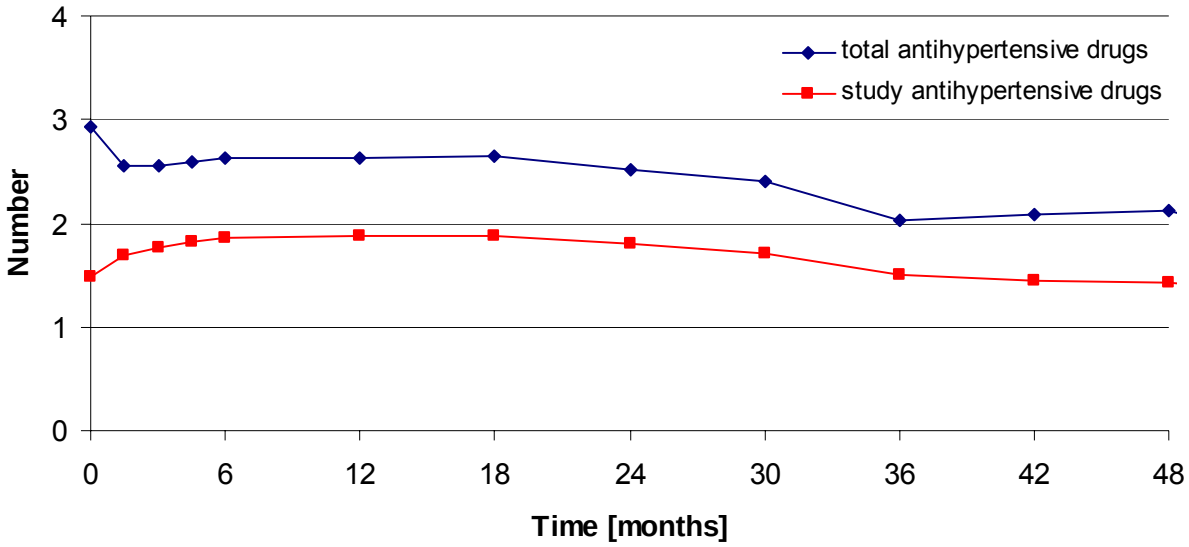

Figure 4-5. Number of total antihypertensive drugs and antihypertensive study drugs over follow-up

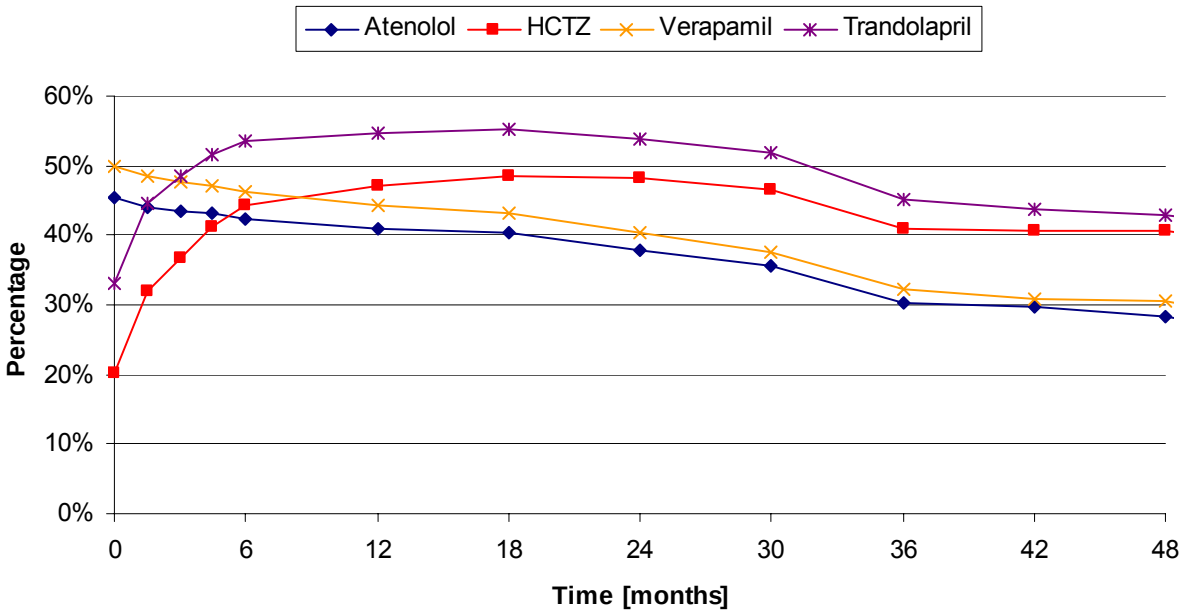

Figure 4-6. Percentage of patients on each individual study drug over follow-up

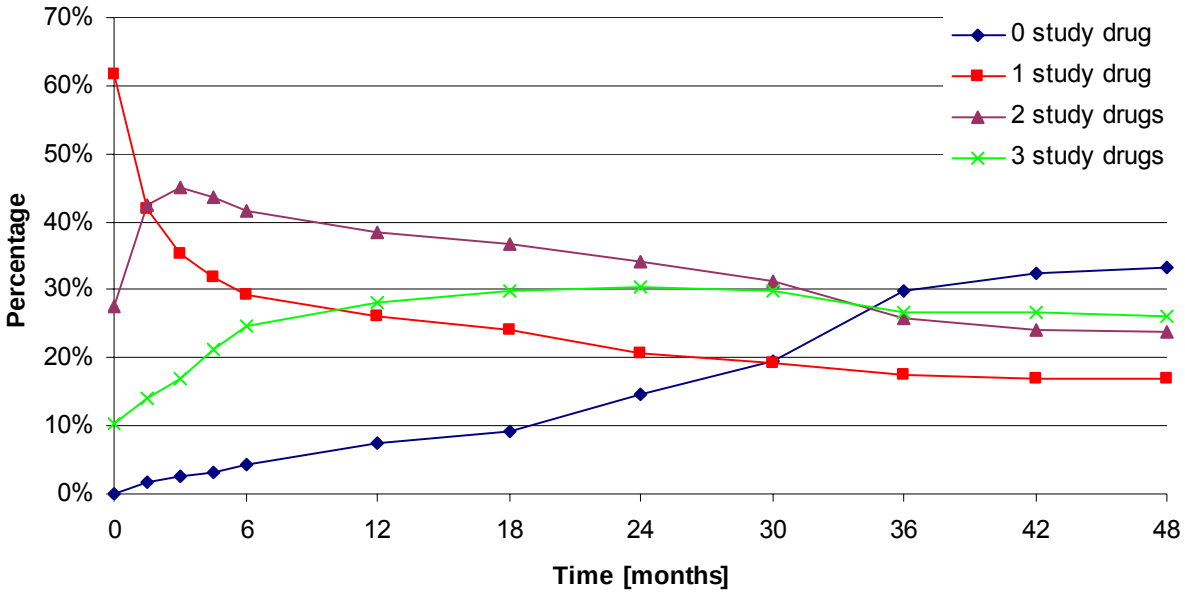

Figure 4-7. Number of INVEST study drugs over follow-up

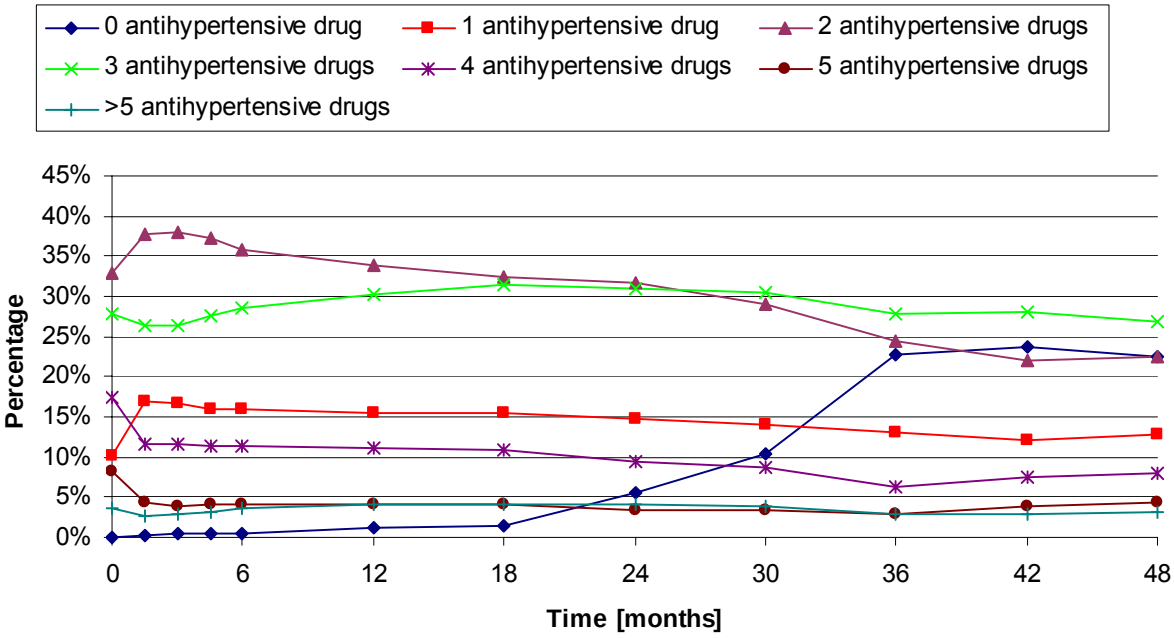

Figure 4-8. Number of total antihypertensive drugs over follow-up

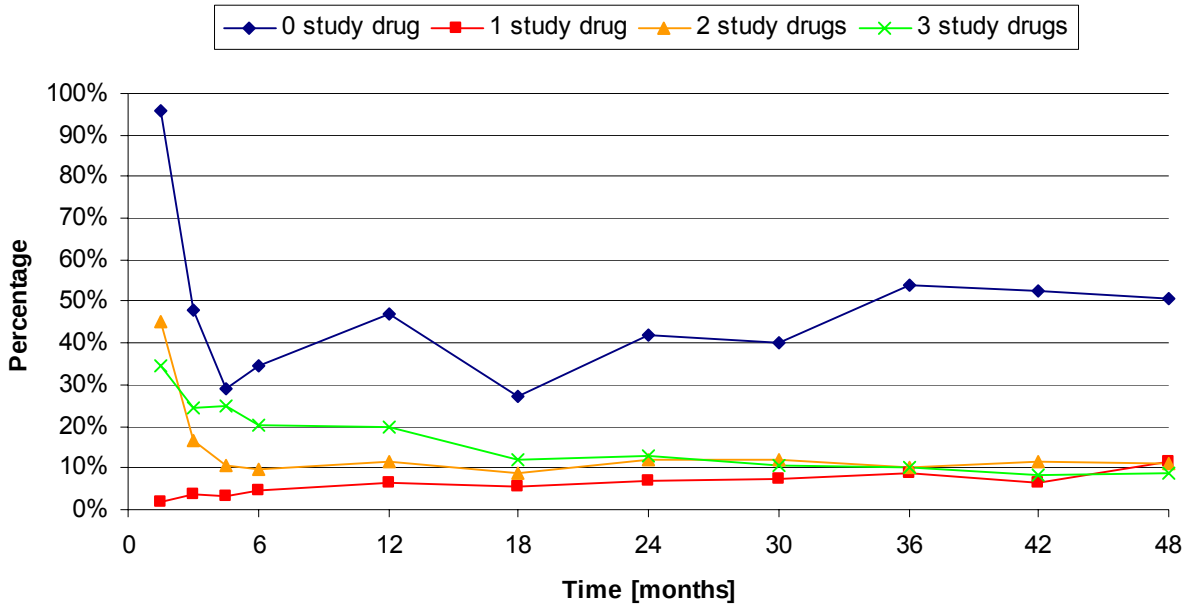

Figure 4-9. Percentage of patients on a specific number of antihypertensive study drugs who were not on the same number of antihypertensive drugs at the prior visit

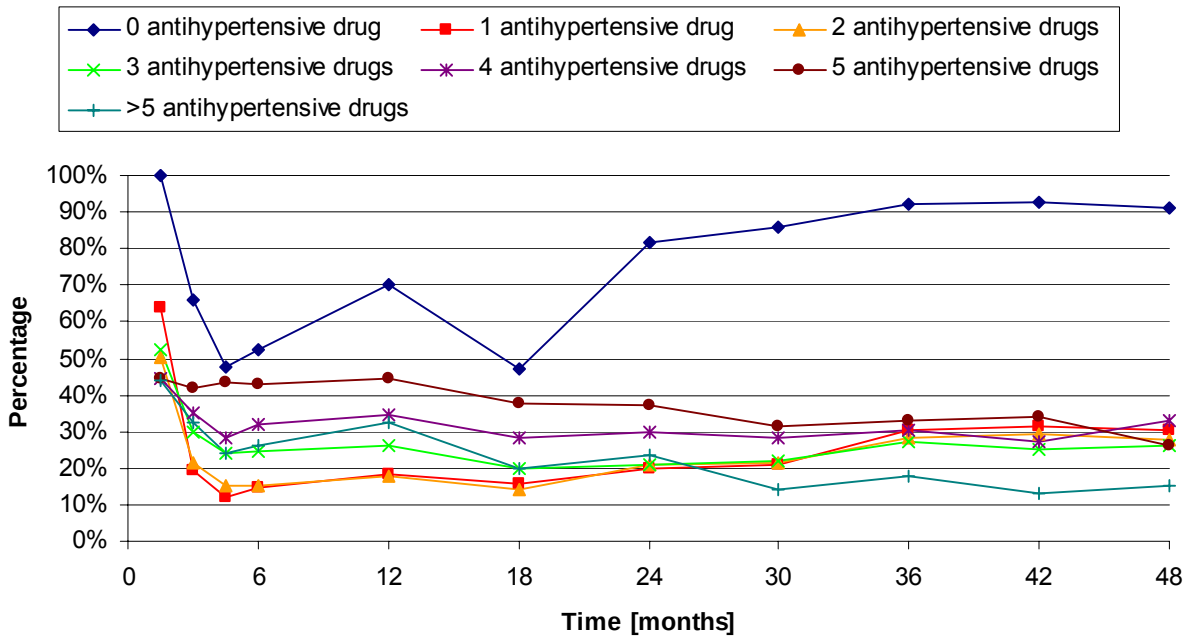

Figure 4-10. Percentage of patients on a specific number of antihypertensive drugs who were not on the same number of antihypertensive drugs at the prior visit

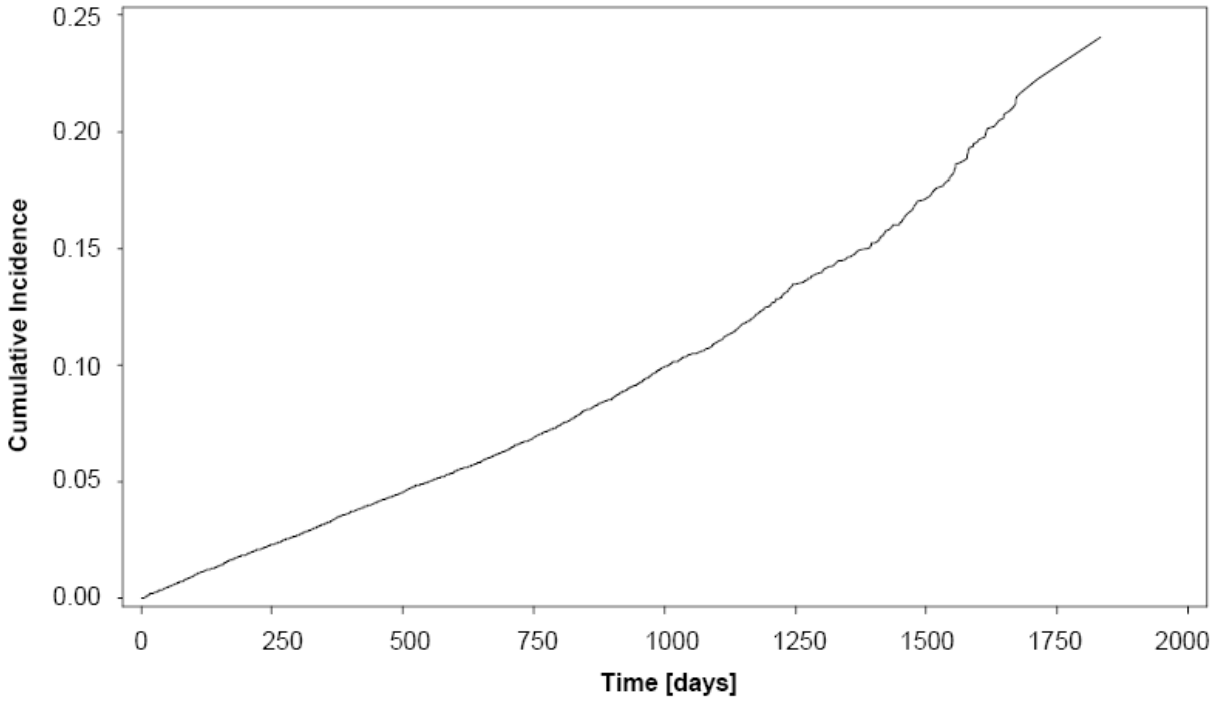

Figure 4-11. Cumulative incidence of the primary outcome event over follow-up

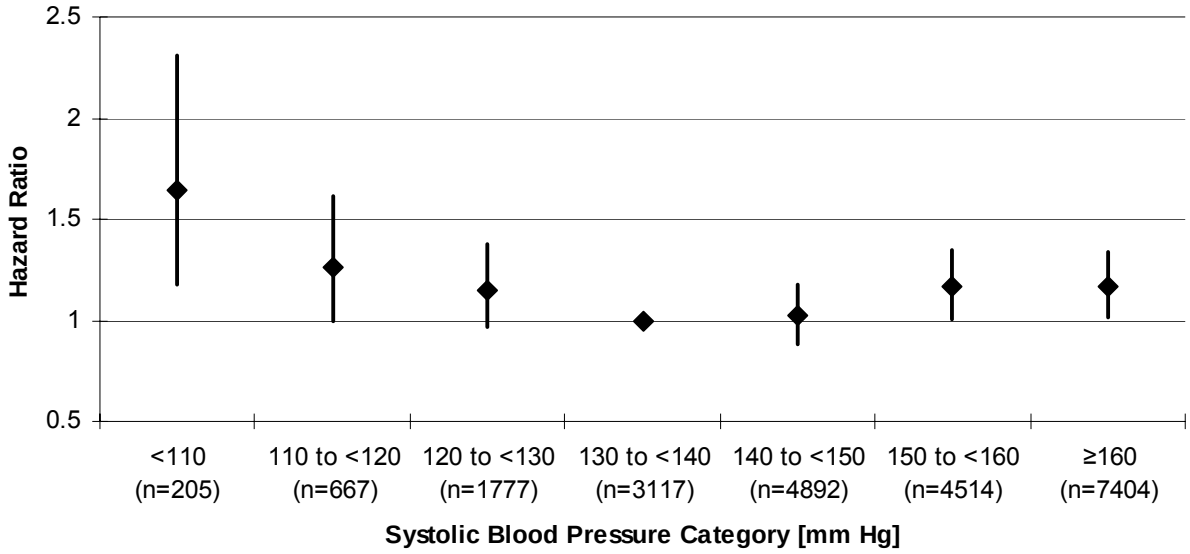

Figure 4-12. Hazard ratios for an INVEST primary outcome event by categories of baseline systolic blood pressure.

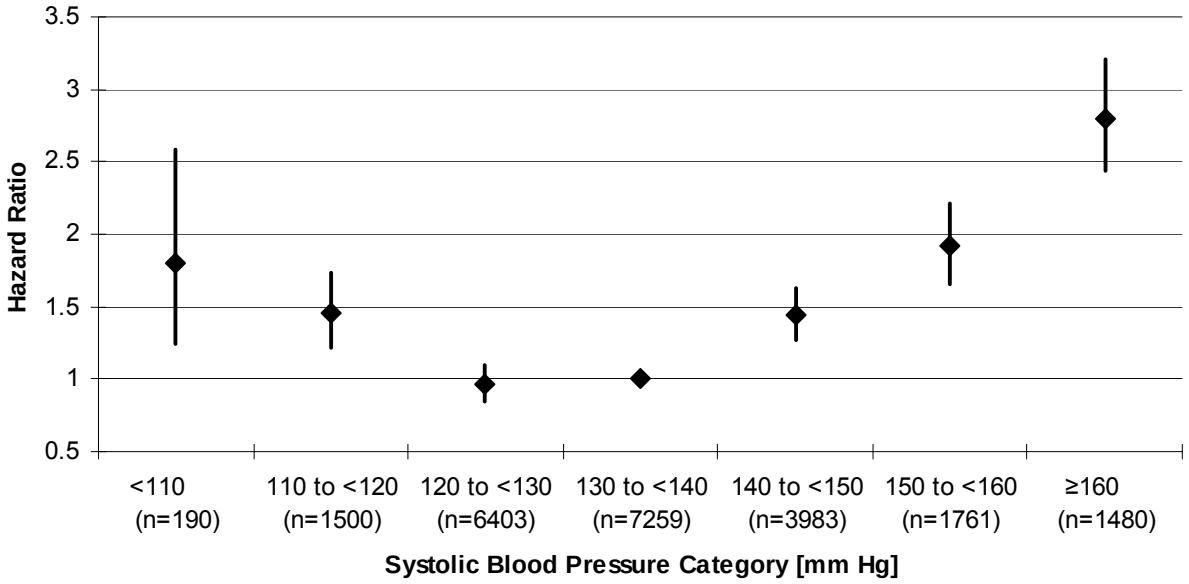

Figure 4-13. Hazard ratios for an INVEST primary outcome event by categories of average systolic blood pressure over follow-up.

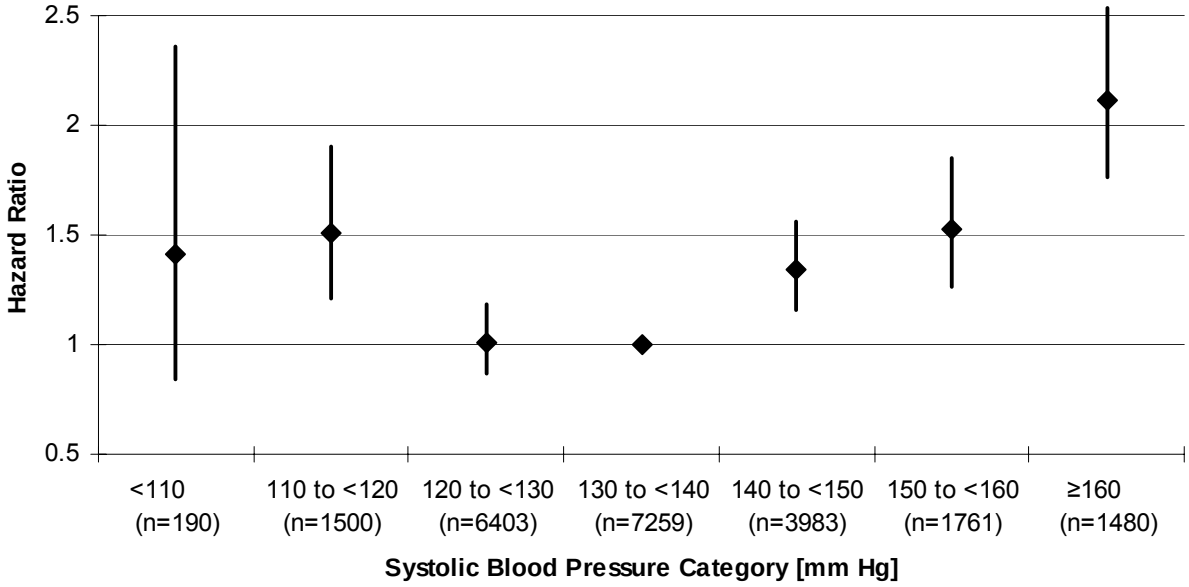

Figure 4-14. Hazard ratios for an INVEST primary outcome event by categories of average systolic blood pressure over follow-up, weighted by time of follow-up.

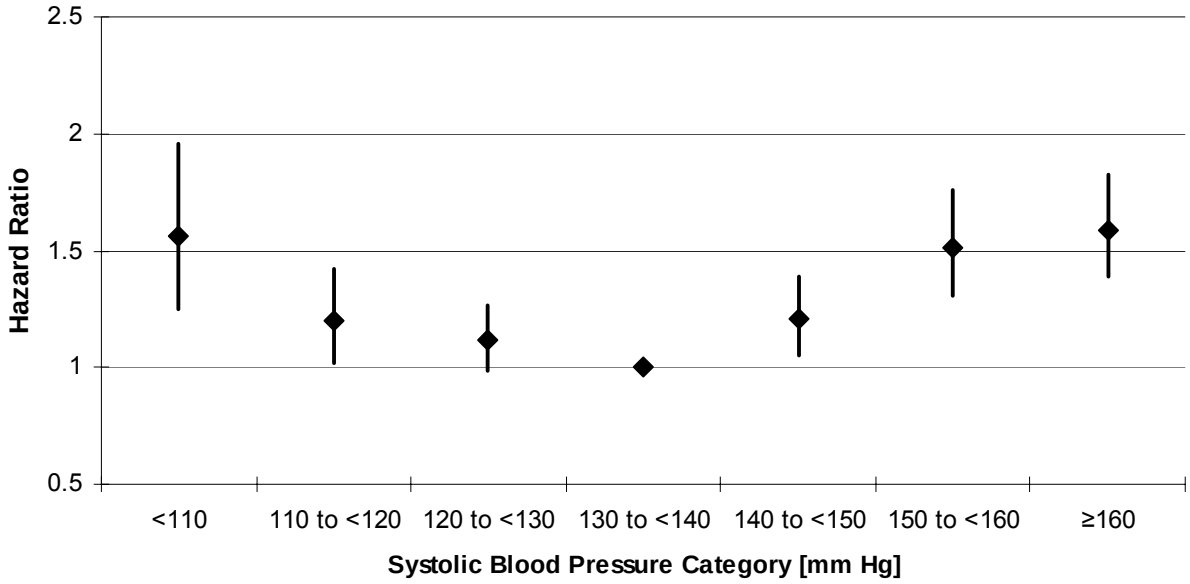

Figure 4-15. Hazard ratios for an INVEST primary outcome event by categories of systolic blood pressure (updated; carried forward from last observed visit)

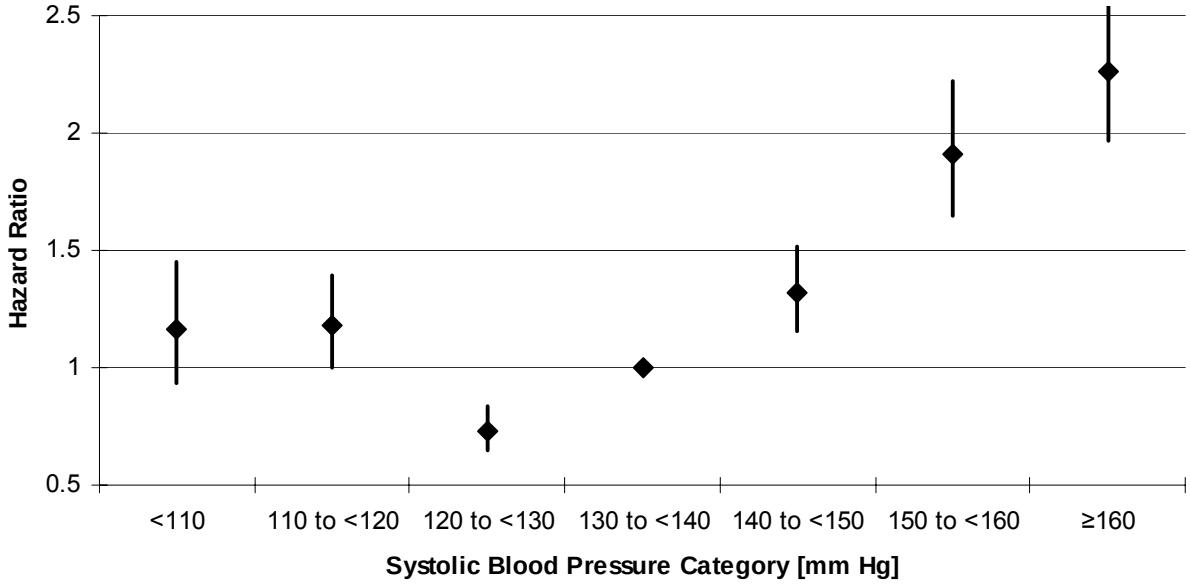

Figure 4-16. Hazard ratios for an INVEST primary outcome event by categories of systolic blood pressure (updated; from next observed visit)

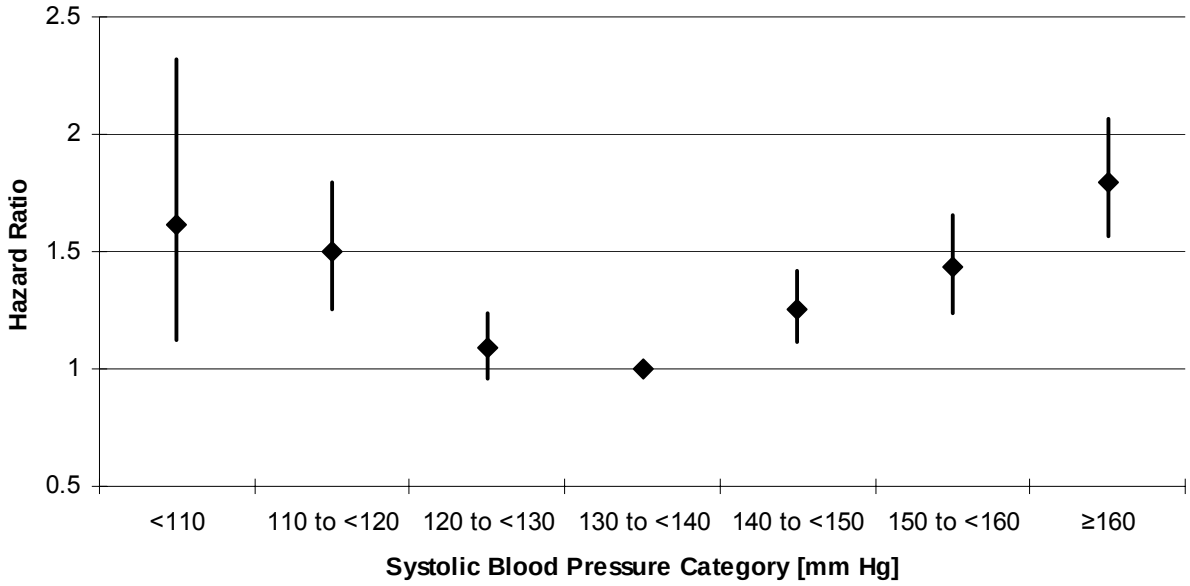

Figure 4-17. Hazard ratios for an INVEST primary outcome event by categories of updated mean systolic blood pressure (time-dependent; updated at each visit)

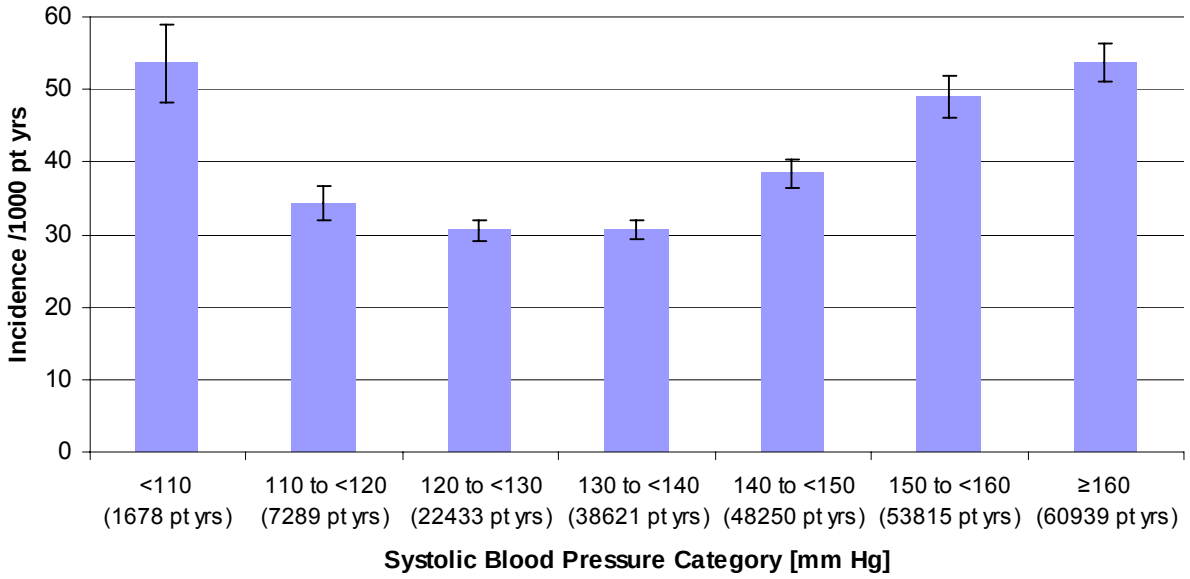

Figure 4-18. Crude incidence of primary outcome events by SBP category

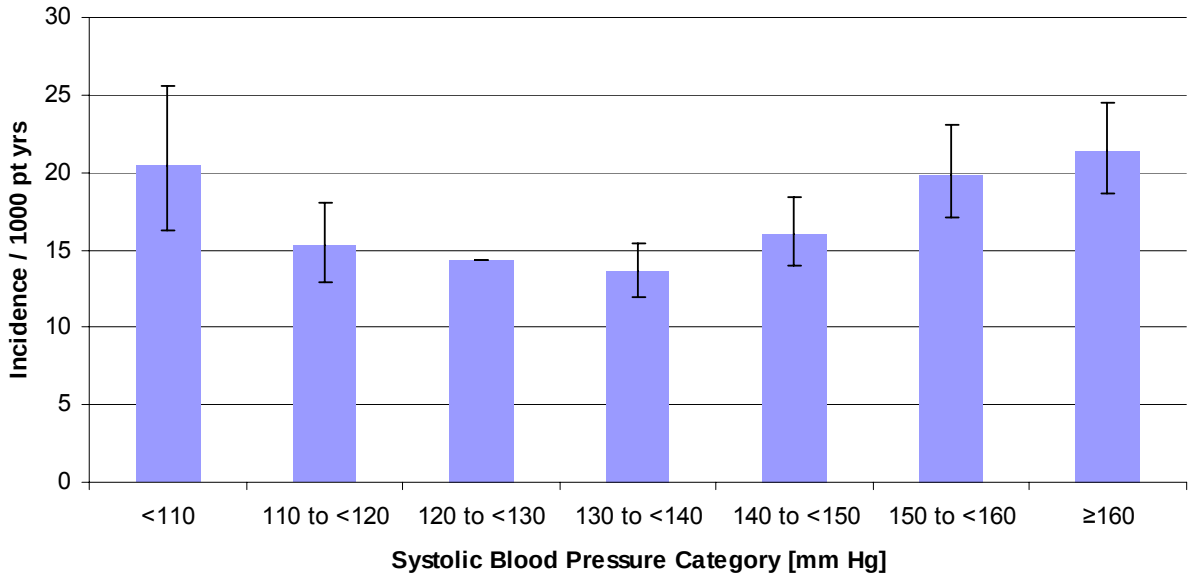

Figure 4-19. Adjusted incidence of primary outcome events for White, female, US patients between the ages of 60 to 70 years by SBP category

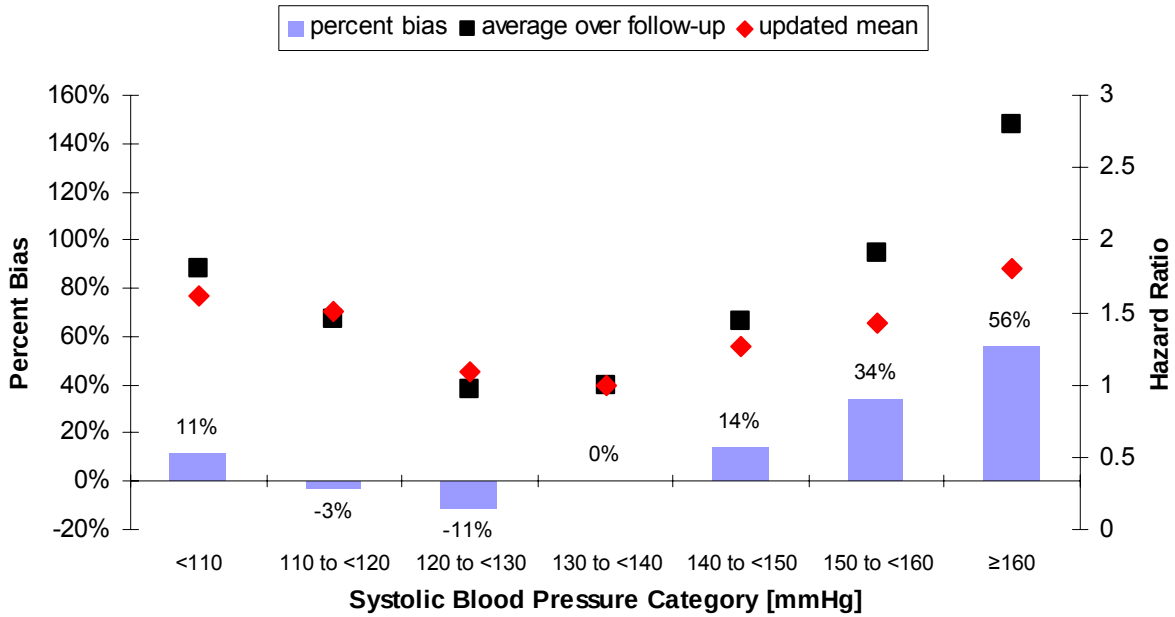

Figure 4-20. Bias of outcome event hazard ratios obtained average SBP compared to updated mean SBP

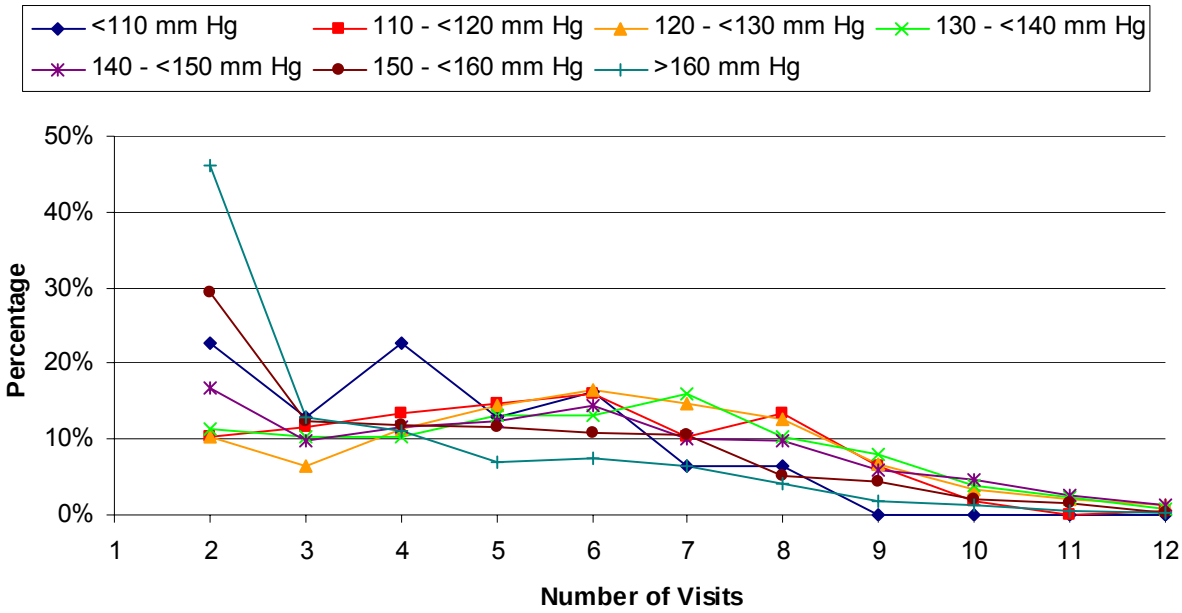

Figure 4-21. Proportion of events within category of average SBP by number of observed visits at the occurrence of the event

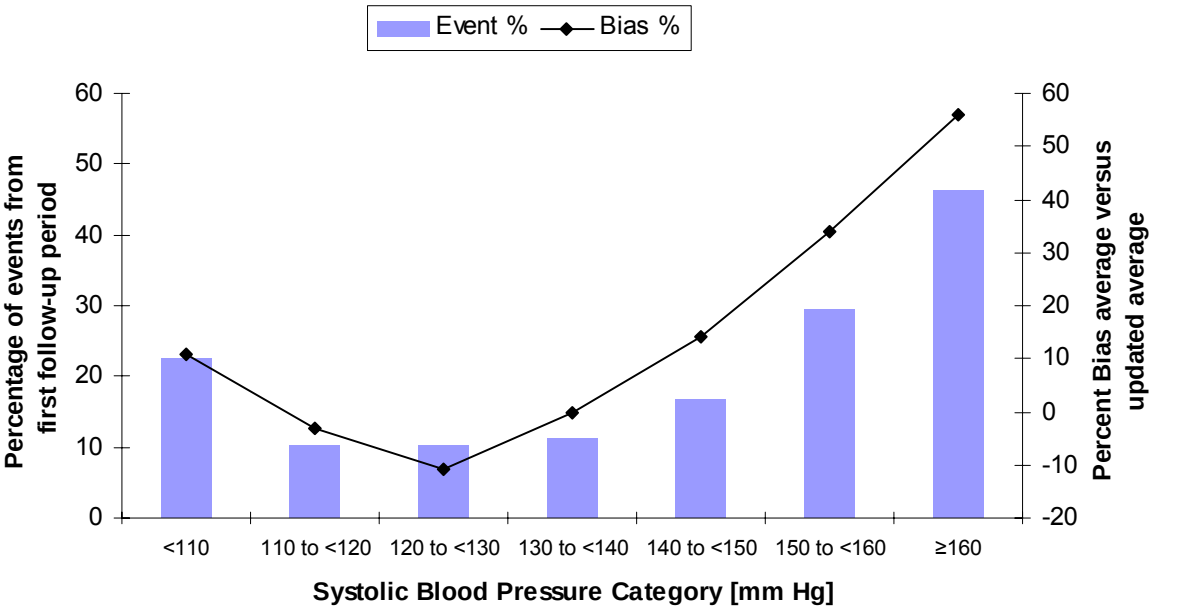

Figure 4-22. Bias and timing of events by category of SBP

CHAPTER 5
DISCUSSION

The present study shows that, in the International Verapamil SR/Trandolapril Study (INVEST), estimates for the effect of SBP on cardiovascular morbidity and mortality vary significantly depending on the method of SBP operationalization. It further demonstrates that using the average SBP over follow-up as a predictor of cardiovascular events systematically overestimates the risk associated with extreme SBP categories. Causal analyses suggest that time-dependent confounding by SBP may bias estimates of treatment effects, but provides no evidence of time-dependent confounding by treatment in the estimation of risk associated with SBP control. However these analyses are considerably restricted by limitations imposed by the statistical method, which necessitated significant simplifications of the data. Consequently, these results are only exploratory in nature. Lastly, our study provides detailed longitudinal descriptions of SBP and antihypertensive treatment patterns over the course of INVEST, which facilitate better understanding and interpretation of the presented inferential analyses.

Descriptive Analyses: Antihypertensive Treatment and SBP in the INVEST

Several observations from the descriptive analyses deserve note. First, when choosing specific time points of follow-up to report BP or other variables (e.g., percentage of patients with BP control at month 24),[4] a proportion of patients (those who are uncensored but do not have an observation at the time-point) is not included in the estimate. Figure 4-1 shows that for the first 24 months of follow-up this proportion is rather significant ranging from about 20% to more than 35% of patients depending on the specific visit chosen to report. As shown in Figure 4-2 there may be differences between results reported for patients with an observation compared to the entire uncensored cohort (with imputed values carried forward from the last observation for patients who do not have an observation at the time-point of interest). A decision between the

two approaches resembles a trade off between a possible selection bias that occurs if the actually observed patients at a specific point of follow-up differ systematically from the uncensored patients who lack this observation, and a measurement bias that would occur if the imputed data for patients who lack the observation of interest is systematically different from the actual, unmeasured value for the variable of interest (e.g., if there is reduction in mean SBP for the cohort including those lacking specific observations, then imputation of missing values from the last observation would systematically overestimate the true SBP value). If feasible, both approaches should be presented to best reflect the true situation at specific points of follow-up. Another consideration may involve choosing the descriptive approach that is in line with the inferential statistical methods to be used. In the present study, all inferential analyses implicitly or explicitly use imputed data and thus the majority of descriptive information was presented for these data.

Second, after mean SBP for the INVEST cohort dropped about 15 mm Hg over the first six months of follow up, it remained rather stable throughout the remaining 42 months, never diverging more than two to three mm Hg from the value observed at six months (Figure 4-2). Consequently, the proportion of patients within the respective SBP categories was stable after the initial six months. However, the proportion of individual patients who experienced a change in SBP categories between visits was continuously high and remained between 30% and 60%, depending on SBP category, throughout the entire follow-up. Thus, the rather stable SBP displayed by the INVEST cohort as a whole after the initial six month of follow-up, was not the result of equally stable SBP on the individual level but rather the result of a constant undirected change (steady state) between categories. This observation likely reflects natural variation as well as measurement error in the assessment of SBP and shows that careful examination of the

data is necessary to distinguish between a stable mean over follow-up that is based on a stable value on the patient level and significant undirected individual variation as observed in the INVEST. This distinction has profound consequences for the choice of an appropriate method of analysis. Standard regression methods using fixed values for SBP may be used when there is little individual variation or measurement error, while time-dependent methods are likely more appropriate to capture and incorporate short term individual variation and its potential effects on outcomes.

Third, descriptive analyses as presented in the first section of Chapter 4 illustrate the limitation of intend to treat (ITT) analyses in randomized controlled trials. While ITT is necessary to preserve randomization, it does introduce misclassification (e.g., at the 24 month visit less than 80% of uncensored patients were still receiving one of the first line study drugs atenolol or verapamil (Figure 4-6)). The widespread utilization of nonstudy antihypertensive drugs (more than a quarter of all antihypertensive medication in INVEST were nonstudy drugs) will likely also result in the attenuation of differences between treatment strategies. Thus, an as treated analysis as performed in the present study should be considered complementary to the original ITT analysis, especially when the nature of the results suggesting noninferiority are considered.

Operationalization of SBP

The present study confirms findings from a prior report which suggested that in the INVEST both low and high blood pressures are associated with an increase in the risk for the primary outcome.[31] While varying in the magnitude of the risk associated with high and low SBP categories, all SBP models presented in our study support this observation. The previously published report, which modeled the average systolic and diastolic BP over follow-up, suggests that the relationship between BP and the INVEST primary outcome event follows a J-shape with

its nadir for SBP at 120 to 130 mm Hg. However, it is important to note that our study presents evidence suggesting that the use of average BPs over follow-up may lead to biased estimates of association. Depending on the SBP model used, our study finds the relationship between SBP and cardiovascular outcomes varying between a J, V, and a reverse-J-shape with the nadir at either 120 to 130 mm Hg, or 130 to 140 mm Hg. These findings differ significantly from other reports that have consistently suggested a log linear increase of cardiovascular risk with increasing SBP (Figure 2-4). However, the INVEST study population differs significantly from a general hypertensive population by including only patients with documented CAD. In patients with CAD low SBP (and more importantly DBP, which was not included in the analyses of our study) may compromise coronary perfusion and cause cardiac ischemia.[31] While it is important to note that neither the previous INVEST report nor our study, both post hoc observational analyses, can establish that low blood pressures causes cardiovascular outcomes, they do show that CAD patients with low SBP are at an increased risk to experience cardiovascular outcomes and thus suggest caution in lowering SBP in hypertensive patients with CAD.

Modeling Assumptions

The 7 models presented in Table 4-2 make implicit assumptions about the mechanism by which SBP affects the risk for cardiovascular events. The following section will summarize and contrast these assumptions of how SBP over time affects cardiovascular outcomes. At one end of the spectrum is the baseline SBP model. By ignoring any changes in SBP after the beginning of follow-up, the baseline SBP model assumes that SBP at a single historic point in time (baseline) acts as a proxy for a patient's cardiovascular risk, while short term changes in SBP after patient enrollment do not significantly change the cardiovascular risk that has been defined by numerous years of SBP history. In contrast, the average SBP model assumes that both historic and actual SBP values affect cardiovascular risk. The respective contributions of historic and actual SBP

values, however, are difficult to quantify since they are dependent on the length of follow-up. Specifically, the contribution of recent SBP values diminishes as follow up time increases. In order to adjust for the fact that the average SBP model would apply the same weight to patients who were followed over multiple years as to patients who were followed for only a short period of time, the time-weighted average SBP model was introduced by weighing each subject's observation by its individual follow-up time.

As opposed to the previous models that rely either completely (baseline model) or partially (both average SBP models) on historic SBP values, the two short term time-dependent models lie at the other end of the spectrum of possible assumptions. These models assume that there is no effect of SBP history but rather assume that SBP at each specific point in time determines the risk for a cardiovascular event at this moment in time. The short term time-dependent SBP models do not differ in these modeling assumptions but rather in the choice of the best available SBP measurement at a given point of follow-up. The prior visit model assumes that SBP at each time point is best approximated by the last measured SBP observation prior to this time point while the following visit model assumes that the best approximation of SBP for the same time point is the first observed measurement after this point in time (Figure 3-2). The rationale for the prior model is straightforward: for each time point it simply uses the last measured SBP value. The rationale for the next visit model is slightly more complex: since antihypertensive treatment is potentially changed at each visit (after response to treatment has been assessed by measuring BP), SBP may change shortly after the visit and thus, true SBP at a given point after this visit may actually be better reflected by the SBP observed at the next visit. However, this approach is potentially biased because for patients who experience an event, no SBP values are available after the event. As a consequence the next visit model treats time periods directly before events

systematically different than all other time periods between visits and therefore may produce a biased estimate.

The updated average model combines aspects of the previous models in that it incorporates SBP history by using an average SBP but also allows change over time by calculating a new (thus time-dependent) average at each observed follow-up time. Last, the Poisson model makes assumptions that closely resemble those of the time-dependent prior visit model as it assigns patient time to the last observed SBP category.

The remainder of this section discusses the SBP models presented in Figures 4-14 to 4-19 in more detail. Several comparisons between SBP models are of specific interest.

Baseline SBP Model

Compared to all other models, the baseline SBP model (Figure 4-14) estimates a substantially smaller risk associated with the two highest SBP categories, while its risk estimates for low SBP categories are of a similar magnitude as the other models. This observation can be explained by considering what happened to the average SBP during the INVEST. Since the INVEST protocol was aimed at controlling patients' SBP, most patients who had high SBP at baseline experienced a reduction in SBP in the first months follow up as shown in Figures 4-2 and 4-3. Thus, for patients with high SBP at baseline, the baseline SBP value poorly represents the true SBP over follow-up (which is likely to be lower than baseline) and as a consequence models that use the baseline value to predict the risk of an outcome event will underestimate the risk for patients with high baseline SBP. In contrast, this systematic misclassification affects patients with a low baseline SBP to a much lesser extend since these patients were likely to remain in a low SBP category. It is therefore not surprising that the baseline SBP model is comparable in its risk estimates for low SBP categories to the other presented models but it results in much lower estimates for the highest SBP categories.

Average SBP Models

On the surface, the use of an average of a surrogate over follow-up to predict the risk of an outcome of interest has many appealing features. Averages over follow-up allow simple modeling without the need to explicitly incorporate time-dependent changes in the measure. They incorporate implicitly and to some extend intuitively such time-dependent changes and create a single value for every subject that allows straightforward incorporation in standard regression models. As a result, averages over follow-up have been used to estimate the effects of surrogates on clinical outcomes in various disease states such as diabetes, hyperlipidemia, and hypertension.[31, 34, 35]

However, the simulations presented in Chapter 4 demonstrate that the apparent simplicity of this approach comes at a price. Considerable bias may be introduced when modeling an outcome with an average over follow-up (Table 4-3). Specifically, for SBP categories that (1) include a large proportion of patients at baseline and (2) include fewer and fewer patients over the course of follow up, overestimation of the true risk associated with these SBP categories is likely. In the INVEST, the above conditions are met for the highest three SBP categories. Table 4-3 shows that these three SBP categories at baseline include the largest proportions of patients (with SBP >160 mm Hg including the single largest proportion) and include continuously fewer patients over the subsequent six months of follow-up (with the most significant reduction in patients observed in the category of SBP >160 mm Hg). As a result, the risk for the highest three SBP categories is systematically overestimated by the average SBP model. Specifically, compared to the updated-mean model, the fixed average model overestimates the risk for a cardiovascular event by 56%, 34%, and 14% for SBP >160 mm Hg, SBP 150 to 160 mm Hg, and SBP 140 to 150 mm Hg, respectively. The updated mean model is selected as the comparator because like the average model, it uses an average of SBP to predict cardiovascular outcomes,

but as opposed to the average model does not introduce bias as shown by the second set of simulations in Chapter 4 (Table 4-4).

The weighted average model (Figure 4-16) estimates results that lie between the average model and the updated mean model. This is expected since the weighted average model is subject to the same bias as the average model (it uses the identical averages) but reduces the impact of the bias by weighting observations according to the length of their follow up time. Since observations with events in the first follow-up period (and thus rather short follow-up time) are most responsible for the bias as discussed above, such weighting will, while not eliminating the bias, reduce its magnitude.

Short Term SBP Models (Time-Dependent)

The two short term time-dependent SBP models differ substantially in their results. Compared to the model that for each period between visits carries forward the SBP value from the first visit (updated previous model), the model that utilizes the SBP from the latter visit for the same interval (updated next model) estimates higher risks associated with SBP categories above 140 mm Hg and lower risks for SBP categories below 130 mm Hg. Considering that mean SBP in the INVEST was lowered substantially over follow-up, this pattern is expected. The time-dependent Cox model used in our analyses compares at each time an event occurs, the SBP of a patient who experiences this event with all other patients' SBP at the same time. Thus, a shift from the SBP measured at the visit prior to the event time to the SBP measured at the visit following the event will on average reduce the SBP of the patients at risk (i.e., those who have not experienced an event and are uncensored). However, the SBP of patients with events are unaffected by the shift because (see Figure 3-2) patients are censored at the time of the event and therefore no post-event SBP measurements exist. Consequently, compared to the updated previous model, the updated next model compares identical SBP values for patients with events,

to on average lower SBP values for patients without events and therefore overestimates the risk for high SBP categories and underestimates the risk for low SBP categories.

When comparing the two short term time-dependent models to the updated mean model it becomes apparent, that the prior model's results are very similar to the updated mean model. The similarity of results is somewhat surprising considering the significant differences in SBP operationalization. While the updated prior model incorporates only the most recently observed SBP values, the updated mean model does incorporate both historic and current SBP values. Both models would therefore be expected to provide differing results. The similarity of both approaches in the INVEST may result from the limited follow-up time and the relative stability of mean SBP after the initial six months of follow-up, but may not generalize to different follow-up periods or disease states. Even so, in the INVEST it is reassuring that both unbiased time-dependent models provide similar estimates for the association of SBP with the cardiovascular risk.

More complex modeling approaches that combine short term time-dependent SBP measures with one or more lagged historic SBP values (i.e., BP at each time point would be represented by the most recent SBP as well as SBP measures from fixed time intervals prior to the most recent visit) are possible, however, such models are more difficult to interpret since SBP values with different lag times may be associated with different estimates of risk, the and thus limited in their clinical utility.

Model Selection

The question arising from the presented models is which of the presented modeling approaches should ultimately be used to estimate the effects of SBP on the risk of cardiovascular outcomes, or more generally the effects of a surrogate on clinical outcomes? While no single correct answer to this question exists—any specific decision will always be influenced by a

multitude of factors—a number of points deserve consideration and should guide the selection of modeling approaches. First and foremost, the assumptions underlying the modeling approach should be closely aligned with the hypothesized biological mechanism. If the hypothesized biological mechanism suggests that a surrogate exerts its effect on clinical outcomes in an immediate fashion, a short term time-dependent model may be preferable. However, the absence of established biological models as well as limitations in data availability and quality may complicate this decision. If no established biological models exist, multiple models with varying assumptions should be produced and compared. If data availability or quality does not allow the selection of a preferred model, the consequences and limitations of choosing a different model should be explicitly discussed.

When data availability permits, our study suggests that an updated mean model should be considered over more extreme (baseline or short term time-dependent) models since it incorporates both historic and current values of the surrogate. More complex models that use both current and lagged surrogate values may also be appropriate (especially if overall follow-up is extremely long and as a consequence recent SBP values would contribute less and less to the updated mean as time progresses), but results obtained from such models are commonly difficult to interpret and communicate.

A second general consideration should be the avoidance of models that likely introduce bias. As discussed above this will generally be the case for baseline models if an intervention or natural progression of the disease leads to a directed change of the surrogate over follow-up, which in turn leads to a systematic over- or underestimation of the true value of the surrogate over follow-up. Even without such directed change, the use of a single measurement at baseline

may introduce an estimate that, due to the effect of measurement error, is biased towards the null (regression dilution bias).[36, 37]

The use of a fixed average over follow-up will under most conditions introduce bias and should be generally avoided. If the data allow calculation of an average (i.e., surrogate values are not limited to baseline), a time-dependent updated mean model should be used instead of a fixed average model. Lastly, the choice between the two short term time-dependent models is difficult. While the updated next model may provide more precise approximation of a surrogate than the updated previous model, especially when changes in drug therapy are common at visits and time-intervals between visits are long, the differential treatment of time periods before events may introduce bias. Whether the potential bias introduced in the updated next model through this differential treatment of time periods preceding events or the potential bias resulting from systematic measurement error introduced by the updated next model is a greater threat to the validity of the estimate has to be evaluated in the context of each specific study.

Unfortunately, the comparison of model diagnostics such as generalized R^2 (a proxy for the fit of the model), is not helpful in the process of model selection. Unlike the R^2 in linear regression models, the generalized R^2 cannot be interpreted as the proportion of variation in the dependent variable that is explained by the covariates included in the model but is only interpretable as a number between 0 and 1 that gets larger when the covariates are associated more strongly with the outcome.[30] More importantly, the generalized R^2, as all measures of model fit, does not distinguish between true associations and associations resulting from bias. Consequently, a biased model will generally produce a larger generalized R^2 than an equivalent unbiased model. In our study, for instance, the biased average model produces a larger generalized R^2 than the unbiased updated mean model.

These considerations regarding the selection of models to estimate the effect of a surrogate on a clinical outcome are not mere academic but rather have profound consequences for the establishment of treatment guidelines and clinical practice.

Time-dependent Confounding

The previous section suggests that the selection of a modeling approach for a surrogate measure has substantial consequences for the resulting estimates of risk and shows that time-dependent modeling approaches are preferable under most circumstances. However, the use of a time-dependent surrogate may introduce time-dependent confounding by any type of treatment that has nonsurrogate mediated effects on the clinical outcome, affects the surrogate, and whose initiation is not independent of the surrogate. Such time-dependent confounding by treatment was not considered in the previously presented modeling approaches. Specifically, the previously presented time-dependent modeling approaches do not include adjustments for the concurrent use of antihypertensive medication and thus, may produce biased estimates of the effects of SBP on cardiovascular outcomes. Time dependent confounding would occur if antihypertensive drug use has effects on cardiovascular outcomes which are not mediated by SBP control since initiation and change of antihypertensive treatments is dependent on SBP control. For example, an increase in the number of antihypertensive drugs will at the same time increase the likelihood of SBP control and (assuming non-SBP mediated beneficial effects of antihypertensive drugs on outcome) reduce the patients risk to experience an outcome event. Thus, a model that does not control for time-dependent confounding by antihypertensive drug use will likely overestimate the beneficial effects of SBP control by attributing both the effects of SBP control and the beneficial effects of an increased number of antihypertensive drugs to SBP control. Controlling for time-dependent confounding by antihypertensive drug use would therefore reduce the estimated beneficial effect of SBP control compared to the estimate obtained from a standard model.

However, our results do not support the presence of a time-dependent confounding effect of antihypertensive treatment. Both the standard time-dependent Cox model and the marginal structural Cox model estimate an almost identical, strong beneficial effect of SBP control with HRs of 0.54 (95% CI 0.48-0.60) and 0.55 (95% CI 0.50-0.61), respectively. If time-dependent confounding by treatment (i.e., a non-SBP mediated beneficial or detrimental effect of antihypertensive drugs) had been present, the HR estimated by the marginal structural model would have been expected to be closer to 1.0, (i.e., show a weaker association between SBP control and outcome). In other words, the two complementary analyses suggest that the type of antihypertensive treatment does not affect cardiovascular outcomes, as long as SBP is adequately controlled. However, it is important to note, that the drug effect was only operationalized as the number of antihypertensive total, and study drugs with no regard for drug class or dose, and that trandolapril, an antihypertensive with demonstrated outcome benefits for patients with specific comorbidities, was per INVEST protocol prescribed to all patients with an indication. To our knowledge this is the first time that inverse probability of treatment weighting was used to obtain an estimate of the effects of a surrogate controlling for time-dependent confounding by treatment.

Additional marginal structural models were created to estimate the effects of aggressive versus standard antihypertensive treatment (more than two versus less two or less concurrent total antihypertensive drugs), controlling for time-dependent confounding by SBP. In this scenario, the fact that the aggressive treatment is more likely to be initiated in patients with uncontrolled SBP (confounding by indication), would lead to an underestimation of the beneficial effect of antihypertensive therapy (or, if the effects of such negative selection outweigh the beneficial effects of aggressive treatment, in the estimation of a harmful effect of

aggressive treatment) because uncontrolled SBP would at the same time increase the likelihood of treatment with aggressive antihypertensive therapy and increase the risk for a cardiovascular event.

The estimate for aggressive antihypertensive treatment obtained from the marginal structural model is lower than the estimate from the standard Cox model but the 95% CIs of both estimates overlap (HR 0.81, 95% CI 0.71-0.92 and HR 0.96, 95% CI 0.87-1.07 for marginal structural and the standard Cox models, respectively). However, while the models are not significantly different from each other, it is noteworthy that the marginal structural Cox model shows a significant benefit of aggressive antihypertensive therapy, while the standard Cox model fails to show such benefit. Sensitivity analyses were conducted to assess the influence of the definition of aggressive antihypertensive therapy on the results. The marginal structural Cox models consistently estimate lower point estimates for the HR associated with aggressive antihypertensive therapy than the standard Cox models for all four additional definitions of aggressive antihypertensive therapy. Of the four definitions the differences between the two models were significant for one definition (aggressive antihypertensive therapy defined as more than three concurrent total antihypertensive drugs), borderline significant for another (aggressive antihypertensive therapy defined as more than one concurrent study antihypertensive drug) and not significant for the remaining two definitions.

Although the differences between the standard Cox models and the marginal structural Cox models did not reach statistical significance in the main analysis and three of the four additional analyses, the trend for a larger beneficial effect of aggressive antihypertensive treatment versus standard antihypertensive treatment was consistent across all analyses. Assuming correct model specification and no violation of the assumptions necessary for the

marginal structural Cox model, the estimate for the beneficial effect of aggressive versus standard antihypertensive therapy can be interpreted as the effect that would have been observed in a RCT that compared aggressive to standard antihypertensive treatment over the course of follow-up. Thus the estimate is causally interpretable and includes all pathways (including SBP mediated pathways) through which aggressive antihypertensive therapy tends to improve the clinical outcome, regardless of SBP values preceding the initiation of aggressive antihypertensive therapy.

These results are in both magnitude and direction comparable to prior studies that have used inverse probability of treatment weighted estimates to estimate unbiased effects of treatments in observational studies in the presence of time-dependent confounding. Specifically, an observational study that assessed the effectiveness of methotrexate in patients with rheumatoid arthritis estimated a mortality HR of 0.6 (95% CI 0.4-0.8) using a standard time-dependent Cox model and an mortality HR of 0.4 (95% CI 0.2-0.8) using a marginal structural Cox model that controlled for time-dependent confounding by several prognostic factors. Another observational study evaluated the aspirin component of the Physicians' Health Study, a randomized controlled trial that evaluated the effectiveness of aspirin in the prevention of cardiovascular disease. While the trial established a strong reduction in first myocardial infarction in patients treated with aspirin, and was stopped early because of it, it failed to detect a significant beneficial effect of aspirin on cardiovascular mortality. At the time of the Physicians' Health Study, the effectiveness of aspirin in the prevention of secondary cardiovascular mortality already had been established and lead to increased use of aspirin in patients that had experienced a nonfatal cardiovascular event. The observational study reanalyzed the clinical trial data (similar to our study) disregarding the original randomization and controlling for time-dependent

confounding by nonfatal cardiovascular events using a marginal structural Cox model. The standard as treated model controlling for predictors of aspirin exposure at baseline as well as cardiovascular risk factors estimated a mortality HR of 0.81 (95% CI 0.57-1.15), while the marginal structural Cox model that additionally controlled for time-dependent confounding by nonfatal cardiovascular events reduced the estimated HR to 0.74 (95% CI 0.48-1.15). Lastly, a study that assessed the effectiveness of zidovudine in reducing mortality in patients with HIV found a significant detrimental effect of zidovudine using standard methods that did not control for time-dependent confounding by CD4 cell count (HR 2.3, 95% CI 1.9-2.8), but was able to show a beneficial effect of zidovudine using a marginal structural Cox model (HR 0.7, 95% CI 0.6-1.0). Of these three studies, only the study that investigated the effectiveness of zidovudine found a statistically significant difference between the HR estimates obtained from a standard time dependent Cox model and a marginal structural Cox model, but as in our study, all three studies estimated consistently higher effectiveness of the treatment under study when a using marginal structural Cox model that controlled for time-dependent confounding by a risk factor for the clinical outcome that was at the same time a predictor of treatment initiation and subsequently affected by the treatment.

Limitations

The present study has several noteworthy limitations. All analyses were conducted in a retrospective, observational design. While our study utilized the dataset of INVEST, a large international randomized controlled clinical trial, the presented analyses are post hoc and are independent of the initial randomization. Therefore our analysis of SBP cannot establish a causal association between SBP and cardiovascular risk. Unmeasured confounding factors may exist that affect SBP as well as cardiovascular risk. In addition, the generalizability of our study is limited to elderly hypertensive patients with CAD. Specifically, the increase in risk associated

with lower SBP categories is likely not representative of a healthier population of patients with uncomplicated hypertension. However, the INVEST population represents a large and important high risk population that has not been at the center of antihypertensive research efforts and therefore warrants comprehensive investigation.

Another limitation of our study is the measurement error inherent in the study of hypertension. Measurement error has been widely reported as a problem in the analysis of hypertension studies.[38-40] While the INVEST protocol tried to minimize measurement error by providing standardized BP measurement instructions following JNC-VI to all providers, a certain extend of measurement error is unavoidable. However, it is unlikely that the remaining measurement error is systematic. Random measurement error may affect our results in ways other than reducing the precision of the provided results. If the absolute magnitude of SBP variation increases with higher SBP values then the categorization of SBP into 10 mm Hg categories may affect our results in that the observed variability in the continuous SBP measure would lead to a larger extend of variation in SBP categories at larger values of SBP.

The use of marginal structural Cox models to adjust for time-dependent confounding through inverse probability of treatment weighting introduces another set of limitations to the causal analyses in our study. Inverse probability of treatment weighting as to date established in statistical theory requires a number of assumptions and a rather simple data structure. First, the independent variable of interest has to be binary. As clearly shown in Figures 4-12 to 4-19 the relationship between SBP and clinical outcomes is more complex than our categorization of SBP into controlled versus uncontrolled. Although we excluded patients with very low SBP to account for the observed J-shape, our operationalization of SBP is a rather poor representation of the data. The same applies to the complementary analysis of antihypertensive drug use. Our

study categorized antihypertensive drug use in aggressive versus standard antihypertensive therapy defined as the use of three or more concurrent total antihypertensive drugs. This definition does not distinguish between specific drugs and drug classes and does not take dosing of each specific drug into account. If treatment effects vary between specific drugs and doses, our model will not show these differences and rather produce an estimate reflective of the average treatment combinations used in the aggressive and standard antihypertensive therapy groups. Thus, our results are of limited value in aiding in the selection of specific antihypertensive drugs. However, we conducted four additional analyses varying the definition of aggressive antihypertensive therapy to assess how sensitive our results are to the treatment definition and results were consistent across all analyses. The fact that the marginal structural Cox model estimates a larger benefit for aggressive antihypertensive therapy than the standard Cox model regardless of the drug number cutoff used to define aggressive antihypertensive therapy, shows that more antihypertensive drugs tend to produce more beneficial outcomes than less antihypertensive drugs, when preceding SBP values that may influence the addition of antihypertensive drugs to an individual's treatment regimen are controlled for. This beneficial effect is likely mediated through a more pronounced effect on SBP, since the results presented in Table 4-5 suggest the absence of non-SBP mediated treatment effects.

Second the method requires the assumption, that once initiated, exposure (in our analyses either SBP control, or aggressive antihypertensive therapy) is not discontinued until the end of follow-up or censoring. This assumption is not met in the INVEST (Figures 4-4, 4-9, and 4-10). We therefore artificially altered the data in order to meet the assumption of the model by keeping BP controlled after BP control was first achieved or keeping patients on aggressive antihypertensive therapy once the respective cut-off was reached. This was necessary, because

the marginal structural model, at present, is not able to incorporate complex modeling approaches that allow reversion of exposure. Lastly, the pooled logistic regression model that was used to estimate the marginal structural Cox model assumes that observations are equally spaced, which was not the case in the INVEST where the initial five visits occurred in six week intervals while the remaining visits occurred in six month intervals. Thus, the pooled logistic regression model may not be fully equivalent to the Cox model it replaces.

The estimation of a surrogate (SBP control in our study) in the presence of time-dependent confounding by treatment raises concerns about its interpretability. While controlling for time dependent confounding by a surrogate is, if all assumptions are fully met and all models are correctly specified, essentially equivalent to a RCT where treatment initiation is independent of the surrogate, the same thought experiment does not hold for the analysis of time-dependent confounding of a surrogate by treatment. Controlling for time-dependent confounding by treatment in the estimation of the effect of a surrogate on a clinical outcome would—as a thought experiment—assume that values of the surrogate are randomized to subjects regardless of their treatment pattern (in our study, SBP control status would be randomly assigned to patients regardless of the number of antihypertensive drugs they are taking). While this obviously would not be possible in reality, and the estimate is therefore more difficult to communicate, we do believe in the validity of the presented approach.

Future Research

The marginal structural models presented in our study demonstrate, in principle, the importance of considering time-dependent confounding in the analysis of chronic disease drug studies involving surrogates. However, the limitations described above presently limit the clinical utility of marginal structural models in the analysis of disease states with complex treatment patterns such as hypertension. In order to make more clinically useful inferences,

future research will have to extend causal methods such as marginal structural models, to allow a more realistic representation of the observed data. Specifically, the method should be extended to allow multi-category independent variables and both initiation as well as discontinuation of exposure. Another important area of research is the evaluation of dynamic treatment regimens from observational data. While our analysis compared nondynamic regimens (e.g., aggressive antihypertensive treatment throughout follow-up versus standard treatment throughout follow-up), questions involving dynamic treatment regimens (e.g., initiation of antihypertensive therapy at SBP of greater than 135 mm Hg versus initiation of antihypertensive therapy at 145 mm Hg) may be equally important for clinical practice. Such analyses require artificial censoring of patients once they deviate from the defined treatment regimens followed by weighting by the inverse probability of censoring to adjust for the potential biased introduced by the artificial censoring.[41]

Summary and Conclusions

The complex interplay of surrogate measures, pharmacological treatments, and clinical outcomes makes the analysis of observational studies of chronic diseases challenging. Our study shows that the estimation of surrogate effects on a clinical outcome is highly dependent on the operationalization of the surrogate and that several commonly used approaches may introduce severe bias to the analysis.

While it is conceptually clear that time-dependent confounding by a surrogate is problematic for the estimation of treatment effects in observational studies, our study is the first to empirically support the presence of such time-dependent confounding in the context of hypertension. Our results suggest that time-dependent confounding by SBP, leads to an underestimation of the effectiveness of antihypertensive treatment. No evidence for time-dependent confounding of the effect of SBP control by antihypertensive treatment was found,

suggesting, that antihypertensive treatment as modeled in our analysis does not affect cardiovascular outcomes in pathways other than SBP. However, limitations in the methods that were used to allow the inclusion of time-dependent confounders in the analysis required considerable simplification of the data structure and thus, these methods at present have only limited practical utility. As such causal methods are further developed, they may become more useful in the analysis of large observational hypertension studies.

LIST OF REFERENCES

1. World Health Organization. Preventing chronic diseases: a vital investment: WHO global report. World Health Organization, Geneva, Switzerland. 2005.

2. Centers for Disease Control and Prevention. Chronic disease overview. Centers for Disease Control and Prevention, Atlanta, GA. Revised 2005. http://www.cdc.gov/nccdphp/overview.htm. Accessed January 2007.

3. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology.* Sep 2000;11(5):550-560.

4. Pepine CJ, Handberg EM, Cooper-DeHoff RM, et al. A calcium antagonist vs a non-calcium antagonist hypertension treatment strategy for patients with coronary artery disease. The International Verapamil-Trandolapril Study (INVEST): a randomized controlled trial. *Jama.* Dec 3 2003;290(21):2805-2816.

5. Pepine CJ, Handberg-Thurmond E, Marks RG, et al. Rationale and design of the International Verapamil SR/Trandolapril Study (INVEST): an Internet-based randomized trial in coronary artery disease patients with hypertension. *J Am Coll Cardiol.* Nov 1998;32(5):1228-1237.

6. Temple R. A regulatory authority's opinion about surrogate endpoints. In: Nimmo WS TG, ed. *Clinical Measurement in Drug Evaluation*. New York: Wiley; 1995.

7. Fleming TR, DeGruttola V, DeMets DL. Surrogate endpoints. *AIDS Clin Rev.* 1997:129-143.

8. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med.* Apr 1989;8(4):431-440.

9. Psaty BM, Weiss NS, Furberg CD, et al. Surrogate end points, health outcomes, and the drug-approval process for the treatment of risk factors for cardiovascular disease. *Jama.* Aug 25 1999;282(8):786-790.

10. Code of Federal Regulations. Approval based on a surrogate endpoint or on an effect on a clinical endpoint other than survival or irreversible morbidity.: Code of Federal Regulations, Title 21, Volume 5, (21CFR314.410), Revised 2006.

11. D'Agostino RB, Jr. Debate: The slippery slope of surrogate outcomes. *Curr Control Trials Cardiovasc med.* 2000;1(2):76-78.

12. Fleming TR. Surrogate endpoints and FDA's accelerated approval process. *Health Aff (Millwood).* Jan-Feb 2005;24(1):67-78.

13. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med.* Oct 1 1996;125(7):605-613.

14. Sobel BE, Furberg CD. Surrogates, semantics, and sensible public policy. *Circulation.* Mar 18 1997;95(6):1661-1663.

15. Lin DY, Fleming TR, De Gruttola V. Estimating the proportion of treatment effect explained by a surrogate marker. *Stat Med.* Jul 15 1997;16(13):1515-1527.

16. Chobanian AV, Bakris GL, Black HR, et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *Jama.* May 21 2003;289(19):2560-2572.

17. Lewington S, Clarke R, Qizilbash N, Peto R, Collins R. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet.* Dec 14 2002;360(9349):1903-1913.

18. Burt VL, Whelton P, Roccella EJ, et al. Prevalence of hypertension in the US adult population. Results from the Third National Health and Nutrition Examination Survey, 1988-1991. *Hypertension.* Mar 1995;25(3):305-313.

19. Franklin SS, Larson MG, Khan SA, et al. Does the relation of blood pressure to coronary heart disease risk change with aging? The Framingham Heart Study. *Circulation.* Mar 6 2001;103(9):1245-1249.

20. Turnbull F. Effects of different blood-pressure-lowering regimens on major cardiovascular events: results of prospectively-designed overviews of randomised trials. *Lancet.* Nov 8 2003;362(9395):1527-1535.

21. Psaty BM, Lumley T, Furberg CD, et al. Health outcomes associated with various antihypertensive therapies used as first-line agents: a network meta-analysis. *Jama.* May 21 2003;289(19):2534-2544.

22. LeLorier J. The value of lowering blood pressure. *Can J Cardiol.* Jan 2006;22(1):63-64.

23. Mitka M. Experts ponder treating prehypertension. *Jama.* May 10 2006;295(18):2125-2126.

24. Julius S, Nesbitt SD, Egan BM, et al. Feasibility of treating prehypertension with an angiotensin-receptor blocker. *N Engl J Med.* Apr 20 2006;354(16):1685-1697.

25. Hartzema AG, Tilson HH, Chan KA, ed. *Pharmacoepidemiology: An Introduction.* 4th ed. Cincinnati, OH: Harvey Whitney Books; 2007; in press.

26. Hernan MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology.* Sep 2000;11(5):561-570.

27. Choi HK, Hernan MA, Seeger JD, Robins JM, Wolfe F. Methotrexate and mortality in patients with rheumatoid arthritis: a prospective study. *Lancet.* Apr 6 2002;359(9313):1173-1177.

28. Cook NR, Cole SR, Hennekens CH. Use of a marginal structural model to determine the effect of aspirin on cardiovascular mortality in the Physicians' Health Study. *Am J Epidemiol.* Jun 1 2002;155(11):1045-1053.

29. Pepine CJ, Kowey PR, Kupfer S, et al. Predictors of adverse outcome among patients with hypertension and coronary artery disease. *J Am Coll Cardiol.* Feb 7 2006;47(3):547-551.

30. Allison PD, SAS Institute. *Survival analysis using the SAS system: a practical guide*. Cary, NC: SAS Institute; 1995.

31. Messerli FH, Mancia G, Conti CR, et al. Dogma disputed: can aggressively lowering blood pressure in hypertensive patients with coronary artery disease be dangerous? *Ann Intern Med.* Jun 20 2006;144(12):884-893.

32. D'Agostino RB, Lee ML, Belanger AJ, Cupples LA, Anderson K, Kannel WB. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. *Stat Med.* Dec 1990;9(12):1501-1515.

33. Elliott WJ, Hewkin AC, Kupfer S, Cooper-DeHoff R, Pepine CJ. A drug dose model for predicting clinical outcomes in hypertensive coronary disease patients. *J Clin Hypertens (Greenwich).* Nov 2005;7(11):654-663.

34. Henricsson M, Nilsson A, Groop L, Heijl A, Janzon L. Prevalence of diabetic retinopathy in relation to age at onset of the diabetes, treatment, duration and glycemic control. *Acta Ophthalmol Scand.* Dec 1996;74(6):523-527.

35. Matsuzaki M, Kita T, Mabuchi H, et al. Large scale cohort study of the relationship between serum cholesterol concentration and coronary events with low-dose simvastatin therapy in Japanese patients with hypercholesterolemia. *Circ J.* Dec 2002;66(12):1087-1095.

36. MacMahon S, Peto R, Cutler J, et al. Blood pressure, stroke, and coronary heart disease. Part 1, Prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet.* Mar 31 1990;335(8692):765-774.

37. Rosner B, Spiegelman D, Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. *Am J Epidemiol.* Oct 1990;132(4):734-745.

38. Hypertension Detection and Follow-up Program Co-operative Group. Variability of blood pressure and the results of screening in the hypertension detection and follow-up program. *J Chronic Dis.* 1978;31(11):651-667.

39. Bennett S. Blood pressure measurement error: its effect on cross-sectional and trend analyses. *J Clin Epidemiol.* Mar 1994;47(3):293-301.

40. Marshall T. Misleading measurements: modeling the effects of blood pressure misclassification in a United States population. *Med Decis Making.* Nov-Dec 2006;26(6):624-632.

41. Hernan MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic Clin Pharmacol Toxicol.* Mar 2006;98(3):237-242.

BIOGRAPHICAL SKETCH

Tobias Gerhard was born and raised in Hammelburg, Germany. He holds a Bachelor of Pharmacy degree from Albert-Ludwigs University in Freiburg. In 2002, he joined the Department of Pharmacy Health Care Administration at the University of Florida as an Alumni Fellow to pursue a PhD in pharmacoepidemiology. Tobias has authored and coauthored several peer-reviewed publications and presented at national and international conferences. He is interested in the pharmacoepidemiology and pharmacogenetics of cardiovascular and mental health drugs.