International Journal of Technology Assessment in Health Care, **16:4** (2000), 1013–1023.

MEASURING THE EFFECT OF CLINICAL GUIDELINES ON PATIENT OUTCOMES

Deborah A. Marshall
Kit N. Simpson
Edward C. Norton
Andrea K. Biddle
University of North Carolina at Chapel Hill

Mike Youle
Chelsea and Westminister Hospital, London

Abstract

Objectives: To identify and examine the methodologic issues related to evaluating the effectiveness of treatment adherence to clinical guidelines. The example of antiretroviral therapy guidelines for human immunodeficiency virus (HIV) disease is used to illustrate the points.
Methods: Regression analysis was applied to observational HIV clinic data for patients with CD4+ cell counts less than 500 per μL and greater than 50 per μL at baseline (n = 704), using Cox proportional hazards time-varying covariates models controlling for baseline risk. The results are compared with simpler models (Cox model [without time-varying covariates] and logistic regression). In addition, the effect of including a measure of exposure to antiretroviral guidelines in the model is explored.
Results: This study has three implications for modeling clinical guideline effectiveness. To capture events that are time-sensitive, a duration model should be used, and covariates that are time-varying should be modeled as time-varying. Thirdly, incorporating a threshold measure of exposure to reflect the minimum period of time for guideline adherence required for a measurable effect on patient outcome should be considered.
Conclusions: The methods proposed in this paper are important to consider if guidelines are to evolve from being a tool for summarizing and transferring the results of research from the literature to clinicians into a practical tool that influences clinical practice patterns. However, the methodology tested in this study needs to be validated using additional data on similar patients and using data on patients with other diseases.

Keywords: Practice guidelines, Proportional hazards models, Acquired immunodeficiency syndrome, Longitudinal studies, Outcome assessment (health care)

The recent surge of interest in clinical guidelines is evidenced by the explosion of publications in the literature produced by various professional organizations, research institutes, and governments prescribing guidelines for a wide variety of diseases and conditions

The authors gratefully acknowledge and thank Dr. Brian Gazzard, director of the Kobler HIV Clinic at the Chelsea and Westminister Hospital in the United Kingdom, for providing the data and clinic-specific information for this work. This work was completed as part of a doctoral dissertation undertaken on a Studentship awarded by the Alberta Heritage Foundation for Medical Research.

(2;4;6;17;25;26;27). Practice guidelines have been defined by the Institute of Medicine Committee to Advise the Public Health Service on Clinical Practice Guidelines as "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances" (6,38). The underlying hypothesis is that "scientific evidence and clinical judgement can be systematically combined to produce clinically valid, operational recommendations for appropriate care that can and will be used to persuade clinicians, patients, and others to change their practices in ways that lead to better health outcomes and lower health care costs" (17,4).

Although it seems intuitively correct that the result of applying guidelines in practice will lead to a positive impact on patient outcomes, the assumption underlying the use of evidence-based clinical guidelines, that patients in regular practice (and not in a study environment) will do better if they are treated according to guidelines developed from clinical trials, is not yet proven. For that reason, we must systematically evaluate the impact of using guidelines to inform practice. This evaluation must be performed using a design that is able to capture both the effect that may be expected from using a better therapy and the effect expected simply from the standardization of practice. This evaluation will often require analysis of observational data, study of differences in treatment over time between patients treated in general practice and patients treated according to recommendations in clinical practice guidelines, and evaluation of the effect of these differences on patient outcomes. Unfortunately, the literature does not provide clear guidance on how to measure guideline adherence with observational data, nor is it very helpful with regard to choosing how best to operationalize the multidimensional concept of a good outcome for a population. This problem becomes especially difficult for populations with complex chronic conditions in whom patient stage and severity may affect the choice of guideline as well as the outcomes expected.

This paper begins by discussing the issues that need to be considered when evaluating treatment adherence to clinical guidelines, and then examines and compares the results of applying regression models that differ in their approach to changes over time for both dependent and independent variables. The example of human immunodeficiency virus (HIV) disease is used to illustrate these issues because it is an area of medicine for which there exist well-defined guidelines regarding treatment with antiretroviral therapy and prophylaxis for opportunistic infections. These guidelines are based on a combination of results from randomized controlled trials (RCTs) that are generally of short duration and expert clinical experience, and it is not clear how the recommendations translate into practice. HIV disease is also a complex chronic disease in which patient status and treatment change frequently, and as a consequence patients' treatment status changes over time between being adherent to guidelines and not being adherent to guidelines. The methods proposed in this paper are important to consider if guidelines are to evolve from being a tool for summarizing and transferring the results of research from the literature to clinicians into a practical tool that influences clinical practice patterns.

GUIDELINES FOR ANTIRETROVIRAL THERAPY IN HIV DISEASE

In this study, patient treatment as noted in the clinical database was compared to the recommended treatment based on the guidelines for antiretroviral therapy developed by the International AIDS Society Panel (5), the U.S. Department of Health and Human Services (12), and the British HIV Association (14). Four main considerations were addressed in the 1996 guidelines from the International AIDS Society Panel: a) when to initiate therapy; b) which types of drugs to use; c) when to change therapy; and d) which types of drugs to use when a change in therapy is indicated. It was recommended that therapy be initiated for all HIV-infected patients with symptomatic disease and for asymptomatic patients with fewer than 500 CD4+ cells per μL or with a rapidly declining CD4+ cell count. Preferred initial

drug regimens with the most demonstrated clinical benefits included specific combinations of two nucleoside analogs, and possibly a protease inhibitor. The guidelines for treatment of HIV disease change rapidly compared to other specialties, and these guidelines incorporated data from both clinical trials and expert opinion about the best approaches to treatment.

The data for this analysis were from 1994, and adherence to the guidelines was considered in this context, since treatment practice patterns were different from today's and adherence to guidelines by clinicians was if anything lower than at present because of less positive views about the utility of therapy. Protease inhibitors were available only through a clinical trial, so treatment was limited to nucleoside reverse transcriptase inhibitors (didanosine [ddI], zalcitabine [ddC], zidovudine [ZDV or AZT], stavudine [d4T], lamivudine [3TC]) and non-nucleoside reverse transcriptase inhibitors (nevirapine or delavirdine). It was not common practice to treat patients with combination antiretroviral therapy. Consequently, in this study guideline adherent treatment was defined as treatment with one or more antiretroviral drugs in patients with CD4+ cell counts less than 500 per μL. For the purposes of this paper, prescribed treatment, as noted in the clinical database, that conformed to the recommended treatment with antiretroviral therapy was called guideline adherence. However, guideline adherence by clinicians did not take into account patient compliance to a prescribed treatment regimen.

FRAMEWORK FOR ANALYSIS: ISSUES TO CONSIDER

Several important considerations need to be taken into account when approaching the evaluation of adherence to clinical guidelines using observational data. These are discussed below in the context of this analysis.

Selection of Appropriate Patients for Analysis

The first basic issue is to determine the appropriate patients to whom the guidelines apply, and the corresponding appropriate outcome measures for these patients. This determination consists of two steps—defining the analytical data set at baseline, and creating a group that is suitable for the outcome of interest.

The analytical data set for the evaluation of adherence to antiretroviral guidelines included patients with CD4+ cell counts below 500 per μL at baseline who were never enrolled in a clinical trial during the observation period. Once the analytical data set was defined, the group of patients that should be included in the analysis was considered in the context of the outcome measure. Death and progression to acquired immunodeficiency syndrome (AIDS) were included as a combined outcome measure because of the small number of deaths during the observation period. As a consequence, patients whose status was AIDS at baseline had to be deleted in this analysis because they were not eligible to progress to AIDS during the course of the observation window.

Degree of Control for Baseline Risk

Observational data are considered less rigorous than data from controlled clinical trials because of the lack of randomization into treatment groups and the lack of a comparison group. Thus, selection bias is one of the greatest threats to validity in observational studies. Consequently, it is important to control for baseline risk and to test for interactions among the variable of interest and covariates.

Although the importance of controlling for baseline risk when analyzing observational data is well recognized, this control is often limited by the availability of data and the technical requirements of the regression model (8;21). The number of variables needed to be limited in this study because of the small number of clinical events during the observation

window. Therefore, the variables were selected carefully based on previous results from the literature about which risk factors are the most important predictors of outcome. In addition, the number of levels used to define the categorical variables was minimized. In evaluating the effect of antiretroviral therapy on the risk of death or progression to AIDS during the observation period, baseline CD4+ cell count was stratified initially into five categories to reflect the varying level of risk at baseline according to the HIV literature (19;20), but it was subsequently collapsed into a dichotomous categorical variable to allow testing for interactions (24).

Duration Models for the Dependent Variable

Choosing the appropriate statistical model is important for obtaining the most precise parameter estimates. The outcome of interest in a clinical guideline study is often time-sensitive. For example, both death and progression to AIDS are time-sensitive because either outcome could happen immediately after the start of the study period, just before the end, or not at all during the study. Analyses that ignore time when modeling the dependent variable have larger standard errors than models that incorporate time. Although one common way to model outcomes is with a logit model, in which the dependent variable equals one if the patient has the event during the study, a logit model is less precise than a duration model. Duration models use information about the timing of the event to get more precise estimates than a logit model, which ignores that information (3;7).

Of the many kinds of duration models, the Cox proportional hazard model (hereafter referred to as the Cox model) is appropriate when the primary purpose is to estimate the relative hazard of the outcome. The results can be used to test whether a person who adheres to clinical guidelines is less likely to experience the outcome of interest. The Cox model has the advantage of not imposing assumptions about the distribution of the underlying hazard, but the disadvantage of not being able to estimate the baseline hazard rate. The Cox model is widely used, and we will compare it to a logit model to show how the results improve when taking into account the information on when the outcome happens (3;10;18). The Cox model can be estimated in most commercially available statistical software packages (1).

Time-varying Models for the Independent Variables

The conventional approach to the analysis of the effect of clinical practice guidelines on patient outcomes does not capture how adherence can change over time. Patients' treatment may start the study in adherence but later lapse. Using a simple dummy variable to indicate adherence may lead to considerable bias if adherence changes much over time. The proper analysis makes the treatment adherence to clinical guidelines a time-varying independent variable. Time-varying independent variables are not possible in linear regression or logit models, but are possible in duration models like the Cox model (1;7) (Table 1).

Although the Cox model in its basic form does not account for changes in the value of the independent variables included in the model, it can be modified to control for time-varying covariates (the Cox proportional hazards time-varying covariate model, hereafter referred to as the Cox time-varying model) (1). This modification is another reason to choose the Cox model over nonduration models or other types of duration model—the capability of including time-varying covariates to get unbiased estimates of adherence. The Cox time-varying model then controls for whether the treatment adhered to clinical guidelines early in the study, or late, or switched back and forth. The ability to account for changes in covariates over time is particularly important for patients with complex chronic diseases such as HIV disease. Clinical indicators for antiretroviral treatment, such as CD4+ cell counts, change frequently, and accordingly patient status with regard to guideline recommendations also changes frequently.

Table 1. Comparison of Analytical Regression Methods

Regression model	Dependent variable: Death or progression to AIDS	Independent variable: Adherence to antiretroviral guideline
Cox proportional hazards time-varying covariate model	Time to death or progression to AIDS	Time-varying Dichotomous 1 = adherent to guideline at that point in time 0 = not adherent to guideline at that point in time
Cox proportional hazards model	Time to death or progression to AIDS	Constant over time Dichotomous 1 = ever adherent to guideline during observation period 0 = never adherent to guideline during observation period
Logistic regression model	Dichotomous 1 = death or progression to AIDS 0 = alive with no progression to AIDS	Constant over time Dichotomous 1 = ever adherent to guideline during observation period 0 = never adherent to guideline during observation period

Measure of Exposure for Guideline Adherence

As discussed in the previous two sections, when analyzing the effect of adherence to guidelines, it is important to use methods that allow for the effect of changes over time in both the dependent and the independent variables. The application of the Cox time-varying model allows for changes over time in both the dependent and the independent variable. However, the interpretation of the Cox time-varying model, as the change in risk of the outcome for each change by 1 day in the adherence to guidelines, is practically cumbersome. One further consideration is how to incorporate the effect of guideline adherence into the model so that it can be interpreted in a clinically meaningful way. Intuitively, it makes sense that the effect of adhering to a guideline would not be observed in a measurable way as a patient outcome (death, death or progression to AIDS, decline in CD4+ cell count) until some threshold amount of exposure through adherence to the guidelines has been exceeded. In other words, the benefits of adhering to the guidelines become apparent only after some minimum period of time.

In this paper, we have included regression models in which a separate dichotomous variable, representing an exposure threshold, was added. The measure of exposure for each individual was calculated as the proportion of time that the treatment was adherent to the guideline over the course of that individual's observation time. Initial regression models considered exposure as a continuous variable, and models were estimated for all threshold levels from a series of models measuring exposure in 10% increments. The model with the optimal threshold point was identified as that for which the parameter estimates most closely approximated the estimates for the model where exposure was continuous. Patients then were categorized as exceeding or not exceeding the threshold. This classification allows the results of the regression analysis to be interpreted as the minimum exposure period for which a treatment must remain adherent to the guidelines in order to have an effect.

The value of this threshold varied depending on the specific relationship being examined. For example, the effect of exposure to antiretroviral guidelines was incorporated into the analysis of death or progression to AIDS. The optimal threshold point for exposure to antiretroviral guidelines was determined to be 10%. This finding made it possible to estimate the effect on risk of death or progression to AIDS if a patient's treatment had been

adherent to guidelines for a minimum of about 1 month (10% of the observation period of 365 days). In this way, the negligible effect of adherence to guidelines for short periods of time, for which it is unreasonable to expect an effect to manifest itself, does not dilute the effect of a more substantial exposure period. This procedure also provides an approach that can be translated into clinical practice more easily.

METHODS

The specific example described in this paper considers only the relationship of adherence to guidelines for antiretroviral therapy and the combined outcome of death or progression to AIDS. The methods applied to the analysis of these clinical data reflect the five issues detailed above.

The analytical data set included only patients with CD4+ cell counts less than 500 per μL at baseline who were candidates for antiretroviral therapy according to the guidelines. From this data set, only patients who had AIDS at baseline were excluded, because the outcome measures included both death and progression to AIDS. Regression models included variables to control for baseline AIDS status, baseline CD4+ cell count as a dichotomous variable (greater than 200 cells per μL), and significant interaction terms.

As described previously, the Cox time-varying model that takes into account changes in both the dependent and the independent variables over time is the preferred methodologic approach for this analysis (1;7;10;18). For each day in the 1-year observation window, the CD4+ cell count, AIDS status, and opportunistic infection status were determined in order to evaluate whether the prescribed treatment adhered to the antiretroviral guideline. Model goodness-of-fit was assessed using the likelihood ratio test and deviance and Pearson chi-square test statistics (24). The results of this approach were then compared with those from simpler models (Cox model without time-varying covariates and logistic regression) that might commonly be applied (Table 1) (16). The direction and magnitude of parameter estimates and the size of standard error estimates were compared.

In addition, the effect of including in the model a measure of exposure to antiretroviral guidelines was explored. A separate continuous variable was created to measure exposure and was included in the Cox time-varying model. To make the model more easily interpretable and clinically meaningful, a threshold point was selected by comparing the coefficient estimates with the model in which the measure of exposure was defined as a continuous variable. The optimal threshold point was identified as that for which the parameter estimates for adherence to guideline most closely approximated the estimates for the model in which exposure was continuous. The value of this threshold varied depending on the parameters being examined.

DATA SOURCE AND DEFINITIONS

The data for this study came from the Kobler Center HIV clinical patient database in London. All patients attending the clinic were HIV-positive, and may have been referred to the clinic from another hospital or physician or may have come directly to the clinic without a referral. All HIV patients (n = 1,894) who attended the clinic between January 1 and December 31, 1994 were included and followed for at least 1 year from the first visit date, unless the patient died. The index date for each patient was the date the patient first visited the clinic during this time period, at which time a baseline CD4+ cell count was obtained. The observation period of 1 year represents a time period over which data might typically be available in a clinic setting and over which one might expect to be able to observe an effect of differences in guideline adherence.

Patients included for analysis in this study must have had a valid CD4+ cell count measurement between 51 and 500 per μL on the index date. Patients with CD4+ cell

counts greater than 500 per μL (n = 399) were excluded because these patients are not generally recommended for antiretroviral treatment according to the guideline recommendations being examined in this study. Patients with CD4+ cell counts less than 50 per μL (n = 397) were not included. The data are from an era prior to the introduction of highly active antiretroviral therapy, and in that era it was likely that patients with such a low CD4+ cell count would be receiving rescue therapy. A patient with a CD4+ cell count less than 500 per μL who was concurrently on a prescription for one or more antiretroviral drugs was considered to be receiving treatment that adhered to the guidelines.

The outcome variable examined as an example in this paper was death or progression to AIDS. For the Cox models (both time-varying and not time-varying), it was measured as the time to the first event that occurred within the observation period. For the logistic regression model, it was measured as a dichotomous variable that was coded as one if the patient ever experienced either of these outcomes at any point, and zero otherwise.

The independent variable of interest was adherence to antiretroviral guidelines. For the Cox time-varying model, it was measured as a time-varying dichotomous variable, coded as one when the treatment was adherent to antiretroviral guidelines at that point, and zero otherwise. For the Cox model and the logistic regression model, death or progression to AIDS was measured as a time-constant dichotomous variable coded as one if the treatment ever adhered to antiretroviral guidelines during the observation period, and zero otherwise.

RESULTS

The results of the analysis for adherence to antiretroviral therapy guidelines using the outcome of death or progression to AIDS are described below as an example to illustrate the issues discussed earlier.

The first issue was to select the appropriate patients to whom the antiretroviral guideline applies and who should be included in the analysis. Of the 1,894 patients identified in the cohort, 704 patients were included in the analysis. All of these patients were eligible for antiretroviral therapy according to the guidelines, and the outcome of death or progression to AIDS appears to be a reasonable choice as an outcome measure. There were 126 events over the 1-year observation window for this group of 704 patients.

The control for severity of disease at baseline was a dichotomous measure of CD4+ cell count at baseline. More detailed categorical analysis of CD4+ cell count could not be supported by these data. No significant interaction terms were found in this analysis.

The expected result of this analysis was that patients whose treatment adhered to antiretroviral guidelines would be less likely (have a lower relative hazard) to die or progress to AIDS during the observation period. The results with the preferred approach, using a Cox time-varying model, showed a negative but statistically nonsignificant effect ($p > .05$) of antiretroviral guideline adherence on death or progression to AIDS (Table 2). The direction of this effect was consistent with the expected result.

This result contrasted with results from the simpler models, using Cox and logistic regression analyses that measured adherence as a constant over time. Both of these models showed a positive and statistically significant effect ($p < .05$) of adherence to antiretroviral guidelines on death or progression to AIDS (Table 2). The difference may be bias in the way of measuring adherence to guideline as compared to that of the Cox time-varying model. These results suggest that patients whose treatment adheres to antiretroviral guidelines are more likely to die or progress to AIDS. The magnitude of the effect is greater (0.70 versus 0.48) for the logistic model than for the Cox model. It should be noted that the standard error estimates in the Cox model were smaller than those for the logistic model, reflecting the higher precision of the Cox model.

Table 2. Results of Regression Models for Death or Progression to AIDS by Adherence to Antiretroviral Therapy Guidelines

Variable	Cox proportional hazards time-varying covariate model	Cox proportional hazards model	Logistic regression model
Intercept	—	—	−0.62[b] (0.17)
Adherence to antiretroviral therapy guideline	−0.51 (0.42)	0.48[a] (0.19)	0.70[b] (0.23)
Baseline CD4+ cell count >200 per μL	−1.79[b] (0.19)	−1.63[b] (0.19)	−1.86[b] (0.22)

Results are expressed as model parameter estimates with standard errors in parentheses.
[a] Statistical significance at the $p = .05$ level.
[b] Statistical significance at the $p = .01$ level.

When exposure to antiretroviral guidelines was incorporated as a separate variable in the regression analyses, the effect of the treatment's ever being adherent to antiretroviral guidelines and the effect of a minimum exposure could be differentiated. The Cox time-varying model results in Table 3 suggest that patients whose treatment adhered to antiretroviral guidelines for more than 10% of the observation period (about 1 month) had a lower risk of dying or progression to AIDS (relative hazard = 0.73, $p < .05$) compared with individuals whose treatment never adhered or adhered for less than 10% of the observation period. The parameter estimate capturing the effect of ever being on antiretroviral therapy was negative, suggesting a reduced risk, and the parameter estimate capturing the effect of exposure for 10% of the observation period was positive but smaller in magnitude. Initially, this result may seem surprising, but individuals who are exposed for longer will also tend to be sicker and have a longer opportunity to experience the event. Treatment with antiretroviral therapy may delay the progression of disease, but the event may still occur within the observation window.

The Cox model and the logistic regression model again showed effects different from the Cox time-varying model (Table 3). The parameter estimates of the Cox model and the logistic model were again the same in direction and similar in magnitude, and neither was statistically significant.

Table 3. Results of Regression Models for Death or Progression to AIDS by Adherence to Antiretroviral Therapy Guidelines Including Measure of Exposure

Variable	Cox proportional hazards time-varying covariate model	Cox proportional hazards model	Logistic regression model
Intercept	—	—	−0.61[b] (0.18)
Adherence to antiretroviral therapy guideline	−0.96[a] (0.44)	0.43 (0.30)	0.58 (0.38)
Baseline CD4+ cell count >200 per μL	−1.67[b] (0.19)	−1.63[b] (0.19)	−1.85[b] (0.22)
Antiretroviral therapy guideline adherence for >10% of patient observation time	0.65[b] (0.21)	−0.08 (0.32)	−0.17 (0.42)

Results are expressed as model parameter estimates with standard errors in parentheses.
[a] Statistical significance at the $p = .05$ level.
[b] Statistical significance at the $p = .01$ level.

DISCUSSION

Despite the widespread and increasing interest in clinical guidelines, there is still much work to be done in evaluating the effect on patient outcomes of adhering to these guidelines. This paper illustrates some of the issues and focuses on the critical importance of the statistical modeling methods used in the analysis of these data. As demonstrated using the example of adherence to antiretroviral guidelines in HIV disease using an observational clinic database, the specification of guideline adherence to include time dependency can dramatically alter the results of the analysis. Researchers are encouraged to explore ways of evaluating these types of data that include time-dependent covariate analysis and some measure of exposure time reflecting guideline adherence.

There is an inherent paradox in the recommended approach to develop guidelines based on the most rigorous study designs (RCTs), on one hand, and their intended application in clinical practice on the other (4;9;25;28). Basing guideline development on RCTs has been criticized by some because such studies include, by design, only a carefully selected group of patients. In addition, they do not represent the conditions in general practice under which the intervention would be applied, both because RCTs are more likely to include university and teaching centers where clinical expertise is higher than the average, and because the study protocol requires more intensive monitoring. On the other hand, observational studies are criticized for their susceptibility to bias because of the lack of an equivalent comparison group. Although it is possible to control statistically for some of the baseline differences between groups, this control is generally not considered sufficient and may be limited by the availability of control variables as well as the stability of the model, depending on the number of events that occur. A recent review of studies comparing the results of RCTs and prospective nonrandomized studies found that the estimate of effect was different depending on the study design, but the relationship was not consistent (22).

In contrast to RCTs, which have shown that antiretroviral therapy improves outcomes for patients with HIV disease in the short run (11;13;15;23), our study examined whether such an improvement is also found when clinical guidelines are used in clinical practice. In other words, the intent of the study was to examine the effectiveness of guidelines under routine clinical conditions instead of the efficacy under idealized conditions. Our findings showed that, for the analysis of complex chronic diseases such as HIV, the best approach models the effects of changes in independent variables over time. Furthermore, a great amount of attention should be paid to how patient stage, severity of disease, and outcomes are measured.

Tradeoffs must be made between the need for detailed control of severity at baseline and model parsimony requirements to assure power to detect outcome differences. Our study illustrated the complex interactions between patient health status measures, such as CD4+ cell count and viral load, and guideline adherence, which was based on these health status indicators. We managed this complexity by partitioning the population by guideline relevance and by specifying outcomes based on their relevance to the 1-year time frame selected for the study. This procedure meant that we had to rely on surrogate marker changes for some patient group guideline combinations and on clinical events for others. The alternative would be to use a longer time frame in the analysis—trading off timeliness of results to inform practice improvement in favor of better evidence on clinically relevant outcomes. Given the rapid evolution of knowledge in HIV treatment, such a trade-off may not be desirable. It is clear from our findings that the use of simple logistic regression analysis to examine the impact of guidelines is inadequate. Exposure models do better but are complex to fit and demanding to interpret. Thus, there is a need for the development of new statistical methods to capture the outcome effects associated with guideline adherence in observational data.

This study has three implications for modeling clinical guideline effectiveness. First, the basic model should be a duration model, to capture events that are time-sensitive. This study used a Cox proportional hazards model. Second, covariates that are time-varying should be modeled as time-varying. To treat time-varying covariates as fixed if they in fact vary will lead to biased estimates. Modern statistical software makes modeling time-varying covariates much easier. Third, the level of exposure should reflect the minimum period of time for guideline adherence in order to capture an effect on the outcome. The results from this study suggest that the effect of time, both for the dependent variable using duration models and for the independent variable of interest (adherence to antiretroviral guidelines), must be captured in these analyses. However, the methodology tested in this study needs to be validated using additional data on similar patients and using data on patients with other diseases.

The results of this study imply that analyses done without reflecting time dependencies in the dependent and independent variables may inadequately capture the effect of guidelines on patient outcomes. The results using these different methodologies vary so substantially that inappropriate conclusions about the direction, magnitude, and statistical significance of adherence to guidelines could be made. As the number of clinical practice guidelines expands and access to them through electronic means is facilitated, it is ever more critical that we understand the consequences of their widespread application. As noted previously, the rigor of the clinical studies supporting the intervention is generally the key criterion used to judge the quality of a clinical guideline (9). The most rigorous study design is considered to be the RCT, which reflects the efficacy of an intervention in a highly selected patient population. Ironically, guidelines are generally intended for application in routine clinical practice, and it is not certain that best practice recommendations will always translate as expected from the results of RCTs. Appropriate tools need to be developed to measure the effect of the guidelines despite the limitations of observational data.

Furthermore, much development needs to take place in how we treat the concept of guideline adherence. Medical practice is passing through a paradigm shift. Choice of therapy used to be determined by a physician's personal experience and judgment of patient preferences, sometimes informed by consultation with experts. The focus was on tailoring the treatment to best meet the individual needs and preferences of the patient. Research in the last 20 years has shown that this approach to the practice of medicine resulted in slow diffusion of scientific knowledge and large practice variations. The use of clinical guidelines may improve the scientific basis on which treatments are selected and decrease practice variations. However, the price of this may be a poorer "fit" with the needs and preferences of some patients, resulting in worse outcomes for this subgroup. This situation has implications for both medical education and outcomes research. Medical educators must focus more on issues of how best to tailor guidelines to be compatible with local resource availability and preferences and how to identify patients who should be treated outside a guideline. Outcome researchers need to develop measures that capture appropriate divergence from a guideline as well as methods to empirically identify treatment processes that are more parsimonious than current guidelines with similar outcomes and care enrichments that fall outside guidelines but have important health effects at low marginal costs. To accomplish the latter we need improved statistical methods and a better understanding of how baseline stage and severity and health outcomes are best measured for specific diseases.

REFERENCES

1. Allison PD. *Survival analysis using the SAS system: A practical guide*. Cary, NC: SAS Institute Inc.; 1995.
2. American Medical Association. *Diagnostic and therapeutic technology assessment. DATTA procedures*. Chicago, Ill: American Medical Association; 1992.

3. Annesi I, Moreau T, Lellouch J. Efficiency of the logistic regression and Cox proportional hazards models in longitudinal studies. *Stat Med*. 1989;8:1515-1521.
4. Canadian Task Force on the Periodic Health Examination. *The Canadian guide to clinical preventive health care*. Ottawa, Ontario: Canada Communication Group; 1994.
5. Carpenter CCJ, Fischl MA, Hammer SM, et al. Antiretroviral therapy for HIV infection in 1996. Recommendations of an international panel. *JAMA*. 1996;276:146-154.
6. *Clinical practice guidelines. Directions for a new program. Committee to Advise the Public Health Service on Clinical Practice Guidelines*. Washington, DC: National Academy Press; 1990.
7. Collett D. *Modelling survival data in medical research*. London: Chapman & Hall; 1994.
8. Concato J, Peduzzi P, Holford TR, et al. Importance of events per independent variable in proportional hazards analysis, I: Background, goals, and general strategy. *J Clin Epidemiol*. 1995;48:1495-1501.
9. Cook DJ, Greengold NL, Ellrodt G, et al. The relation between systematic reviews and practice guidelines. *Ann Intern Med*. 1997;127:210-216.
10. Cox DR, Oakes E. *Analysis of survival data*. London: Chapman & Hall; 1984.
11. Eron JJ, Beniot SL, Jemsek J, et al. Treatment with lamivudine, zidovudine or both in HIV-positive patients with 200 to 500 CD4+ cells per cubic millimeter. *N Engl J Med*. 1995;333:1534-1539.
12. Feinberg MB, Carpenter C, Fauci AS, et al. Guidelines for the use of antiretroviral agents in HIV-infected adults and adolescents. *Ann Intern Med*. 1998;128:1079-1100.
13. Fischl MA, Richman DD, Grieco MH, et al. The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. A double-blind, placebo-controlled trial. *N Engl J Med*. 1987;317:185-191.
14. Gazzard BJ, Moyle GJ, Weber J, et al. British HIV Association guidelines for antiretroviral treatment of HIV seropositive individuals. *Lancet*. 1997;349:1086-1092.
15. Hammer SM, Katzenstein DA, Hughes MD, et al. A trial comparing nucleoside monotherapy with combination therapy in HIV-infected adults with CD4 cell counts from 200–500 per cubic millimeter. *N Engl J Med*. 1996;335:1081-1090.
16. Hosmer DW, Lemeshow S. *Applied logistic regression*. New York: John Wiley & Sons Inc.; 1989.
17. Institute of Medicine Committee on Clinical Practice Guidelines. *Guidelines for clinical practice: From development to use*. Washington, DC: National Academy Press; 1992.
18. Kleinbaum DG. *Survival analysis*. New York: Springer-Verlag; 1995.
19. Mellors JW, Kingsley LA, Rinaldo CR, et al. Quantitation of HIV-1 RNA in plasma predicts outcome after seroconversion. *Ann Intern Med*. 1995;122:573-579.
20. Mellors JW, Rinaldo CR, Phalguni G, et al. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. *Science*. 1996;272:1167-1170.
21. Peduzzi P, Concato J, Feinstein AR, et al. Importance of events per independent variable in proportional hazards regression analysis, II: Accuracy and precision of regression estimates. *J Clin Epidemiol*. 1995;48:1503-1510.
22. Reeves B, Maclehose R, Harvey I, et al. Comparison of estimates of effectiveness from randomised and non-randomised studies. Paper presented at the Annual Meeting of the International Society of Technology Assessment in Health Care, Edinburgh, U.K., June 20–23, 1999.
23. Saravolatz LD, Winslow DK, Collins G, et al. Zidovudine alone or in combination with didanosine or zalcitabine in HIV-infected patients with the acquired immunodeficiency syndrome or fewer than 200 CD4 cells per cubic millimeter. *N Engl J Med*. 1996;335:1099-1106.
24. Stokes ME, Davis CS, Koch GG, *Categorical data analysis using the SAS system*. Cary, NC: SAS Institute Inc; 1995.
25. U.S. Preventive Services Task Force. *Guide to clinical preventive services*. 2nd ed. Baltimore, MD: Williams and Wilkins; 1996.
26. Woolf SH. Practice guidelines: A new reality in medicine, I: Recent developments. *Arch Intern Med*. 1990;150:1811-1818.
27. Woolf SH. *AHCPR interim manual for clinical practice guideline development*. Rockville, MD: Agency for Health Care Policy and Research; 1992. AHCPR Publication No 91-0018.
28. Woolf SH, Battista RN, Anderson GM, et al. Assessing the clinical effectiveness of preventive maneuvers: Analytic principles and systematic methods in reviewing evidence and developing clinical practice recommendations. *J Clin Epidemiol*. 1990;43:891-905.