An Economic Cost Analysis of Oral Ganciclovir Prophylaxis for the Prevention of CMV Disease

Gordon G. Liu[1] and Joel Hay[2]

Received March 8, 2000; accepted May 4, 2000

Purpose. The study conducted an economic cost analysis of oral ganciclovir prophylaxis in preventing cytomegalovirus (CMV) disease for AIDS patients in a randomized clinical trial setting.

Methods. Data were generated from patient interviews, medical records, and case reports from a multi-center, randomized, double-blind, and placebo-controlled pharmacoeconomic study appended to a clinical trial. The outcomes were measured in monthly cost per patient. Various cost functions were tested in the context of sample-selection model (SSM) and two-part model (TPM), and were estimated using both the ordinary least squares (OLS) and the bounded influence estimation (BIE) methods.

Results. The use of informal caregiver services did not differ significantly between patients in the treatment group and those in the placebo group. The OLS estimates for the ganciclovir prophylaxis arm showed a reduced, but statistically insignificant use of formal care in both outpatient and inpatient settings. The BIE results for the ganciclovir prophylaxis arm, in contrast, showed a significant reduction of 27% in hospital cost among hospital users, and 44% among the total sample of AIDS patients. The monthly total cost function also identified a decreasing but insignificant trend due to the treatment effect.

Conclusions. At the methodological level, this study demonstrated the value of employing more rigorous econometric techniques in identifying subtle treatment effects on cost outcomes from clinical trial data in the economic assessment of medical technologies. At the empirical level, the study concluded that beyond its demonstrated efficacy of preventing CVM disease among AIDS patients, ganciclovir prophylaxis did not lead to additional health care costs, other than the cost of the drug therapy.

KEY WORDS: cost; AIDS; cytomegalovirus; ganciclovir; pharmacoeconomics; sample selection bias.

INTRODUCTION

Study Objectives

As the health care system becomes increasingly focused on cost containment, the pharmaceutical industry has begun to routinely incorporate cost and cost effectiveness data collection during Phase Three clinical trials for FDA marketing approval. These clinical trials are not usually specifically designed for this type of economic data collection. Not all clinical trial sites are able to collect usable economic data, patients may drop out of the economic data collection protocol even if they stay in the clinical trial, and clinical trials are generally powered to have enough people to evaluate clinical safety and efficacy, rather than cost. Thus, advanced econometric modeling techniques are often required to evaluate clinical trial-based economic data. This analysis considers these issues in the context of a specific clinical trial.

Ganciclovir is a well-established treatment for CMV disease in AIDS patients. Until the availability of an oral ganciclovir regimen, use of this medication to prevent rather than treat CMV disease was not clinically practical. Since ganciclovir prophylaxis has been shown to be a safe and efficacious regimen for reducing the incidence of CMV disease among advanced AIDS patients (1), it is now reasonable to evaluate whether the medication is economically justified. The prophylaxis regimen is not inexpensive (costing about $50 per day, based on average wholesale price). It is therefore important to determine the net impact of ganciclovir on medical and non-medical service utilization and treatment costs when used in CMV prophylaxis.

This study conducts an economic cost analysis of oral ganciclovir prophylaxis, for the prevention of cytomegalovirus (CMV) disease in persons with AIDS. The analysis is a prospective pharmacoeconomic study based on a Phase III clinical trial (ICM/GAN$_c$1654/USA, Roche 1995) designed, conducted, and monitored as a multi-center, randomized, double-blind, and placebo-controlled project by Roche Pharmacoeconomic Research (1). The patients' maximum study enrollment period was 18 months.

Study Design and Data Collection

In this double-blinded study, patients with AIDS were randomly assigned to two trial arms: a treatment group receiving oral ganciclovir prophylaxis, and a control group on placebo. In the clinical trial, eligible individuals included HIV infected subjects, whose CD4 lymphocyte measures were <50 cells/mm^3, or <100 cells/mm^3 with an AIDS defining illness. All subjects had CMV infection, but were free from CMV disease at enrollment, confirmed by antibody test or urine culture. Exclusion criteria included history of past or present CMV disease, or previous treatment for CMV. The primary clinical endpoint of this study was the development of CMV disease. Clinical diagnosis of CMV retinitis was determined by fundoscopic examination by a board certified ophthalmologist. Other CMV diseases assessed included, visceral disease, involving the gastrointestinal tract, lung, or liver. CMV disease was diagnosed, for the purposes of the pharmacoeconomic analysis, on physician-reported events rather than biopsies. The rationale behind use of physician-reported rather than protocol-defined CMV disease as an endpoint was that treatment for CMV disease would be given in cases of presumptive as well as confirmed CMV.

The economic cost analysis was conducted on a subset of patients randomized to the clinical trial determined by 11 centers (out of the 19 centers participated in the clinical trial) who were willing to participate in the pharmacoeconomic protocol. For the 11 selected centers, medical care utilization was collected in the clinical trial for the period up to the point of diagnosis of CMV disease. It should be noted that the termination of clinical trial observation for patient health care utilization at the diagnosis of CMV disease may bias this cost analysis against ganciclovir prophylaxis. This is because ganciclovir

[1] University of North Carolina at Chapel Hill, Department of Pharmaceutical Policy and Evaluative Sciences, Beard Hall, Room 205Q, Chapel Hill, North Carolina 27599-7360. (e-mail: ggliu@unc.edu)

[2] Department of Pharmaceutical Economics & Policy, University of Southern California, CHP 140, 1540 E. Alcazar Street, Los Angeles, California 90089. (e-mail: jhay@usc.usc.edu)

0724-8741/00/0800-0911$18.00/0 © 2000 Plenum Publishing Corporation

prophylaxis has been shown to significantly reduce the incidence of CMV disease (1), and the average use of resources would have been greater for the control group if those developing CMV disease had been followed for longer.

Health care utilization data were taken from five major sources. The sections below outline the pharmacoeconomic data collection forms that were used during the clinical trial:

A) Health care utilization I (HCU1). HCU1 was for patient-reported social services and home-based care provided by paid professionals.

B) Health care utilization II (HCU2). HCU2 was for patient-reported informal care services given by friends or relatives.

C) Health care utilization III (HCU3). HCU3 was for all outpatient care services and procedures. The data were collected in service units from patient medical records at the sites of health care delivery. For consistency across sites, we estimated medical costs by weighting units of services at all clinical sites with prices taken from the 1993 Medi-Cal 5% sample claims database. This Medi-Cal database included both charges and allowed payments, and can be taken as reasonably representative of the range of health care prices reflecting both the amounts charged to third parties and the resource cost for delivering these services (2).

D) Hospital care utilization (HS1). Hospital care data were drawn from the clinical trial case report form. Form HS1 collected duration of stay in days and primary reason for admission. Unit cost of hospital care was derived from the 1993 Medi-Cal average hospital contracted payments.

E) Concomitant medication (CM). Data on use of medications were drawn from the clinical trial concomitant medication records. For concomitant medications, two strategies were performed for pricing depending on whether the drug was a brand name or generic product. In particular, for a brand name drug a unit price from Medi-Cal was attached by the National Drug Codes (NDC). For generic products the daily costs were estimated based upon the most frequently used drugs and their cost figures provided by Roche Pharmacoeconomic Research.

ANALYTICAL METHODS

Model Specifications

Similar to the typical distribution of annual medical expenditures in other general populations (3), HCU1, HCU2 and hospital care data in this study also showed a large portion of "0" observations (N_0), with the remaining non-zero observations being highly skewed to the left (N_1). These "0" observations may have been due to missing data, or could reflect actual zero expenditures. There are several major types of missing data: missing forms, patients without pharmacoeconomic data, and forms where some resource use variables were blank. Although both missing data and zero data cannot be distinguished in the analysis, their impact on the analysis was taken as a data-censoring problem.

Broadly speaking, there are two different methodological approaches to dealing with health care costs with censored data: econometric analysis and survival analysis. In general, economists employ the former, while epidemiologists routinely use the latter (3–7). In this study, we take the econometric modeling approach to formulate our analytical framework. Along this line, two major econometric methods were considered in our analytical modeling framework: the two-part model (TPM) and sample-selection model (SSM).

Following the TPM method (3), an individual utilization of health care can be measured in two steps: whether to use care at all, and how much to use if the individual decides to use care. In essence, the first step describes the incidence or probability of utilization, while the second looks at the volume or quantity of utilization. The two steps convey quite different information on one's utilization behavior (8–9). The TPM utilization behavior was modeled as:

$$I_i = Z_i\,\alpha + u_i,\ u_i \sim N(0, \sigma_u^2)\ \textit{for}\ N = N_1 + N_2$$

$$H_i | I_i > 0 = X_i\beta + v_i,\ v_i \sim N(0, \sigma_v^2)\ \textit{for}\ N_1$$

where I_i is a binary choice variable indicating whether individual i ever used any health care services in the study period; H_i stands for the consumed quantify of health care expenditures, conditional on $I_i > 0$. Z_i and X_i are two vectors of explanatory variables; α (or β) is the corresponding parameter vector; u_i and v_i are two stochastic terms which generally are assumed to follow a normal distribution with zero mean and constant variance. A fundamental assumption for the TPM method is that the stochastic term v_i follows its own normal distribution, independent from the distribution of u_i.

Alternatively, the SSM method (4) maintains that the two stochastic terms may be correlated, leading to a statistical sample selection bias in the usual OLS estimates of the expenditure function H_i. Following Heckman (4), a statistical correction term can be employed to control for such selection bias. In sum, the crucial difference between the TPM and SSM centers on the difference in the fundamental assumption on the distribution relationship between the two stochastic terms u_i and v_i. Unfortunately, existing theories and previous empirical work to date are yet not conclusive as to which method performs better overall in terms of model fitting (3,5,10).[3] However, one can test the null hypothesis of no correlation between the two stochastic terms on the basis of SSM estimates. Following this strategy, we first estimate equation (1) using SSM method. Depending upon the SSM testing results, we would choose the TPM if the null hypothesis of no sample selection was not rejected. Alternatively, we would rely on the SSM results if the null hypothesis was rejected.

Estimation Approaches

Average Costs vs. Total Costs

Total cost is simply the sum of all utilization units, each multiplied by respective unit cost. However, as stated earlier,

[3] Using bootstrap sampling approach, Lipscomb *et al.* (7) conducted a comprehensive comparison of single-equation model, two-part model, and the Cox proportional-hazards model. According to three alternative criteria (root mean square error; mean absolute error, and log score), the log-transformed two-part model edged out all other models including the proportional-hazards model based on a national sample of 21,546 Medicare-covered ischemic stroke patients. For deriving the predictive distribution of cost, the log-transformed two-part and proportional-hazards models are superior. In terms of predicted mean or median cost, the two part models and the log-transformed single-equation model are statistically indifferent.

data were only collected while subjects stayed in the study. That is, there was no follow-up monitoring of the subjects after their termination due to CMV disease, adverse events or subjects' wishes to withdraw. As a result, patient's time spent in the study is dependent upon the effect of the study drug intervention. That is, the duration of stay may be an endogenous variable, and the associated per patient costs are a function of the duration. To control for the problem of endogeneity, monthly average cost models are chosen over total cost models for analyzing each type of care.[4] Total average cost model is also carried out using the sum of all average costs across all subgroups.

OLS vs. BIE Approach

Very often the estimates of a model may be highly influenced by a few extreme observations which may distort the real relationship for the whole population. There are several approaches to outlier problems based upon the information generated from residuals. For instance, a simple approach is to drop off the outliers (if any), and re-estimate the model again using the data without the outliers. Obviously, the major problem of this drop-off approach is the loss of information associated with the deleted outliers. In this study, rather than simply dropping the outliers, we employed the Bounded Influence Estimation (BIE) method suggested by Welsch (11). The basic principle of the BIE method is to minimize $\Sigma w_i (\ln H_i - X_i\beta)$ which puts a lower weight on outliers than on non-outlier observations.[5] For a study with substantial outliers, like this one, the BIE results sharply contrast with OLS estimates, which would assume an equal weight contribution of every observation to the model estimation.

EMPIRICAL RESULTS

Data

There were five subgroups of health care utilization: social services and home-based care provided by paid professionals (HCU1), informal care given by non-paid helpers such as friends and relatives (HCU2); office visits, procedures, and ancillary services obtained at outpatient health care settings (HCU3); hospital care in terms of days in hospital (HS1), and concomitant medications (CM) which included both brand name drugs and generics. All data but HCU3 were captured at regularly scheduled visits within the clinical trial.

For both paid and unpaid caregiver services, there were no cost measures directly captured on the case report forms. The basic unit of service utilized was the reported number of caregiver hours received by patients for each activity. The utilization data were obtained from patients' surveys at baseline, month 4, month 8, month 12, and month 18. To price these patient self-reported data services, we assumed a $12/hour wage rate for all the labor cost involved in these home-based care services such as housekeeping, day care, counseling and transportation.

Medical services and procedures were classified and coded by Current Procedural Terminology Codes 4th Edition (CPT4). We then estimated a price from the 1993 Medi-Cal short-form claim file and attached it to each of the service CPT4 codes. The service-specific mean value paid by Medi-Cal (meanpaid) was employed to price all the medical services in the study.

Hospital care data were drawn from the clinical trial report file and the Medi-Cal data. In particular, the health status file, HS1, was used to estimate patients' inpatient care utilization. We then priced daily hospital room cost using $780/day, a Medi-Cal statewide contract average rate in 1993.

For all the concomitant medications taken by subjects, we took two different approaches to assigning costs to brand name drugs vs. generics. First, we identified all the brand name drugs and their NDC codes (these were those drugs for which generics were not available). We then priced each medication using Medi-Cal paid mean drug cost index. For generics, since we were not able to associate them with the Medi-Cal drug cost files by National Drug Codes (NDC), we used the actual daily costs provided by Roche Pharmacoeconomic Research, based on review of IMS reports on the most frequently prescribed entity in each generic drug category. Because it could not be determined which drug was actually taken by the patient due to multiple manufacturers as well as brand names available, a weighted average estimate was calculated, using the top three selling products listed in the IMS data.

Conceptually, the sum of all sub-group average costs gives the average total cost. As shown above, however, since various sources had to be used to produce the different sub-category cost measures, not many subjects had complete data for all across-category cost variables. For this reason, the total cost variable may be underestimated overall due to those with missing data being treated as "0". Such an underestimation concern, however, should not be taken to be critical with respect to the purpose of this analysis because there was no evidence that the missing data occurred in any non-random fashion across the study trial group and control group. No attempt was made to reconstruct missing cost variable observations. For variables used as explanatory variables in the cost function estimates, however, we set all missing values to the sample mean value of that variable.

Mean Difference Analysis

Before performing econometric analysis, we first conducted a simple two-tailed *T-test* to examine whether there was a significant difference in mean value between treatment group and control group for key variables including all the cost variables, some risk factors, and other explanatory variables. The risk factors to be considered included the duration of hospital stay, mortality, and adverse events. Since almost identical results

[4] Technically, assume the original conditional utilization model was measured in units of total cost per patient, and X_i contained a logged time variable, i.e., log of the months stayed in the study. After the log month variable was moved over to the right hand side with some rearrangements, it is straightforward that we would end up with an equation with the log of the average health care utilization as the dependent variable.

[5] $$w_i = \begin{cases} 1 & if\ |DFFITS_i| \le 0.34 \\ 0.34/|DFFITS_i| & if\ |DFFITS_i| > 0.34 \end{cases}$$

where $$DFFITS_i = \left[\frac{h_{ii}}{1 - h_{ii}}\right]^{1/2} \frac{\hat{e}_i}{\sqrt{(1 - h_{ii})s^2}}$$

$$h_{ii} = X_i(X'X)^{-1} X_i'$$

$$s^2 = MSE = \sum_{N1} \hat{e}_i/(N_1 - k_1)$$

Table 1. Definitions of Variables

Variables	Definitions
Paid caregiver services (HCU1)	Average monthly cost of social and home-based care provided by paid professionals
Unpaid caregiver services (HCU2)	Average monthly cost of informal care given by unpaid care givers
Outpatient care (HCU3)	Average monthly cost of outpatient care services and procedures provided by health care professionals, priced out using Medi-Cal paid mean values
Hospital care	Average monthly cost of hospital care, priced at Medi-Cal rate of $780 per day at an acute hospital
Concomitant medications	Average monthly cost of concomitant medications (not including ganciclovir), priced at Medi-Cal paid mean values by NDC codes
All care	Average monthly cost of all services and medical care: (AC1 + AC2 + AC3MP + ACHP + ACRXMP)
Medical care	Average monthly cost of direct medical care (AC3MP + ACHP + ACRXMP)
Treatment	Dummy variable 1 for the drug intervention, and 0 for placebo
CMV	Dummy variable 1 for CMV disease, and 0 for no CMV disease
Death	Dummy variable 1 for subjects who died during the study period
Adverse event	Dummy variable 1 for subjects with adverse events
LOS in study	Number of months stayed in the clinical study from beginning till any dropout
White	Dummy variable 1 for white, and 0 for others
Age	Years of age
Education	A categorical variable, being 1 for no formal school through 12 for post-graduate
Income	A categorical variable, being 1 for annual household income of $4,999 or less through 10 for 100,000 and over
Weeks FTW	Number of weeks worked full time during the past 12 months
Δ Health status	A categorical variable for changes in health status compared to 1 year ago at the baseline level, being 1 for much better; 2 for somewhat better; 3 for about the same; 4 for somewhat worse; and 5 for much worse
QOL	A categorical variable for baseline self-reported quality of life, being scaled from 0 for the worst possible through 10 for the best possible
Health status	A categorical variable for self-rated health status, being 1 for excellent, 2 for very good, 3 for good, 4 for fair and 5 for poor
Candidiasis	Dummy variable 1 for a disease history of oral/pharyngeal candidiasis
Leukoplakia	Dummy variable 1 for a disease history of oral/hairy leukoplakia
Herpes zoster	Dummy variable 1 for a disease history of herpes zoster
CD4	Average CD4 counts at baseline

were obtained using Medi-Cal paid median and mean values, we report only the results of the cost analysis based on paid mean values. Definitions of variables used in the study are given in Table 1.

According to the univariate *T-test* results for all cost differences (see Table 2) only the monthly cost of hospital care differed between the treatment and control groups, with ganciclovir prophylaxis patients reporting lower average hospital costs at the 13% significance level. The cross-group differences in terms of all other costs were statistically insignificant. Among risk factors, the mean difference T-test results in Table 3 suggest that the ganciclovir subjects show a two-week longer average stay in the clinical trial before developing CMV disease (P < 0.12). We also investigated the relationship between the ganciclovir prophylaxis intervention and two other variables: death and adverse drug event. There was no evidence that the study drug intervention altered the likelihood of these two events, suggesting these two variables could be treated as exogenous.

In Table 3, three explanatory variables were also found to be significantly different between treatment group and control group before the treatment intervention. The three variables are number of weeks worked full time (Weeks FTW) during the past 12 months, change in health status and quality of life scores (QOL). This finding suggests that even though subjects were randomly assigned at the beginning of the randomized study,

Table 2. T-tests of Differences in Average Monthly Cost Variables Across Treatment Groups

Variables	Sample size control	Treatment	Mean control	Treatment	Prob > T
Paid home services	51	112	$26	$62	0.41
Unpaid home services	50	112	$491	$518	0.88
Outpatient care	56	107	$84	$97	0.70
Hospital care	239	486	$508	$374	0.13
Concomitant medications	166	319	$120	$152	0.16
Monthly total costs	29	50	$969	$1027	0.81
Monthly medical costs	45	72	$614	$604	0.95

Table 3. T-tests of Differences in Explanatory Variables Across Treatment Groups*

Variables	Sample size control	Treatment	Mean control	Treatment	Prob > T
Death	239	486	4%	6%	0.20
Adverse event	239	486	16%	19%	0.35
LOS in study	239	486	10.4	11	0.12
White	239	486	84%	81%	0.29
Age	239	486	40	40	0.24
Education	239	486	9.11	8.97	0.38
Income	239	486	5.05	5.09	0.78
Weeks FTW	239	486	21	24	0.07**
Δ Health status	239	486	3.38	3.23	0.07**
QOL	239	486	6.87	7.07	0.11
Health status	239	486	3.18	3.11	0.28
Candidiasis	239	486	71%	74%	0.37
Leukoplakia	239	486	33%	30%	0.41
Herpes zoster	239	486	30%	29%	0.76
CD4	239	486	27.05	25.84	0.44

* The table includes baseline demographic variables and study findings.
** Significant at the 10% level.

the dropouts over time may have led to a non-random truncated sample with significant cross-group differences for the three variables. If these variables have a strong relationship with the cost functions, univariate analysis of the mean difference in costs could be biased. We also recognized that these cost data did not follow a normal distribution, therefore the T-test results were not reliable. As a result, further econometric analysis was warranted.

Econometric Analysis

PE-Test

To model the cost function appropriately, we compared linear with log-linear model specifications using the PE-test procedures (12). A common problem may arise with the test when both model specifications are either accepted or rejected at the same time. Although the PE-test may not be as powerful as the Wald, Lagrange multiplier, and likelihood ratio tests, the authors suggest that it has sufficient power for most empirical research and it is much easier to carry out. According to the PE-test results, all the semi-log model specifications appeared to perform better than the linear counterparts.

BIE Estimation

For paid and unpaid caregiver support, and hospital care models, since there were a large number of non-spenders involved, these models were estimated using both SSM and TPM methods. According to the SSM estimates, sample selection bias was not statistically significant among all the cost functions except for the unpaid caregiver cost function (a small proportion of the total average costs). As a result, there was no need to correct for selection bias in our data and therefore, we discuss results given by the TPM method.[6]

For outpatient care, medications, and total average cost models, since there were few non-spenders, unconditional cost models were estimated using the total number of observations. Concerning the possible outlier effect on the conditional cost functions, we conducted both the OLS and the BIE estimations. Based upon the overall model fitting performance, BIE estimates appeared to be more reliable and robust than the OLS estimates. As a result, we focus on discussing the empirical results given by the BIE model estimates, while providing the OLS model estimates as reference.

Paid and Unpaid Caregiver Cost

We first discuss results for the social services and home-based care provided by paid professionals. There were total of 163 subjects with returned caregiver support surveys, and only 24 of them reported positive use of paid caregiver support. The choice model suggested no treatment impact on the probability of using the social and home-based care services. Some non-treatment variables, however, seemed to affect the probability. For instance, the use of the services was found to be significantly more likely among the subjects with higher household income, higher CD4 counts, and a history of oral hairy leukoplakia or herpes zoster. Among those who ever used any of the services, the amount of care was also indifferent between the treatment group and the control group.[7]

Some patients also received informal care services that were offered by unpaid caregivers such as friends, relatives, or family members. Of the 162 subjects contained in original data,

[6] The Sample selection model estimates are available upon request.

[7] A potential problem, however, may exist with this model due to the sample size of the non-zero spenders. This was because over 85% of subjects (139) in this group were non-spenders, leaving only 15% (24) of the observations for estimation of the conditional cost model. As a result, the fitted residuals were substantial, leading to an upward bias of the estimated paid caregiver average costs. Therefore, we considered this estimation equation to be imprecise. However, because the magnitude of spending on paid caregiver support is quite low, the lack of precision in estimating this equation should have little impact on the overall cost estimates followed.

Table 4. Costs of Outpatient Care Services (AC3MP)

Variable	OLS model β coefficients	BIE model β coefficients	BIE model $e^{b\Delta X}-1$**
Intercept	0.864 (0.476)	0.509 (0.638)	—
Treatment	0.038 (0.883)	−0.037 (0.871)	−0.037
White	0.910 (0.003)*	1.097 (0.0001)*	1.995
Age	−0.022 (0.148)	−0.017 (0.191)	−0.017
Education	0.095 (0.120)	0.098 (0.073)*	0.103
Income (H)	0.035 (0.588)	0.045 (0.437)	0.046
Weeks FTW	−0.003 (0.646)	−0.003 (0.632)	−0.003
Δ Health status	−0.079 (0.528)	−0.075 (0.489)	−0.073
QOL	0.096 (0.260)	0.102 (0.171)	0.108
Health status	0.290 (0.095)*	0.299 (0.054)*	0.349
Candidiasis	0.481 (0.115)	0.397 (0.141)	0.487
Leukoplakia	−0.122 (0.673)	−0.081 (0.747)	−0.078
Herpes zoster	0.099 (0.730)	0.026 (0.919)	0.027
CD4	0.002 (0.776)	0.002 (0.650)	0.002
Death	−0.457 (0.307)	−0.298 (0.452)	−0.258
Adverse event	0.561 (0.113)	0.469 (0.133)	0.598
F-Statistic	2.145	3.104	3.104
R^2	0.184	0.246	0.246
Sample mean (log)	3.379	3.395	3.395
Observations	158	158	158

Notes: 1. The P-value (T-test) for each coefficient is given in parentheses; 2. *Significant at the 10% level; 3. **To better interpret the β coefficients, we derive the % change in the cost function in response to a unit change in each explanatory variable X. Suppose the cost function C is a semi-log function of X: $LnC = b_0 + bX$, then the % change in C can be expressed as $\Delta C/C = e^{b\Delta X}-1$, where Δx stands for a one unit change in X.

94 reported unpaid caregiver service use. According to our TPM estimates, the probability of using informal care services did not differ between the treatment group and the control group. The conditional cost model seemed to show a substantial reduction (about 40%) in caregiver support for the drug intervention group, but the effect was not statistically significant. Household income, again, showed a significant role in increasing unpaid caregiver support by about 18%.[8]

Outpatient Care Cost

Of the total 163 subjects, only 4 were non-spenders. As a result, a log cost model was estimated for all spenders. The results are reported in Table 4, where column 1 shows the OLS estimates, column 2 shows the BIE estimates, and column 3 shows the marginal percentage change in the BIE cost function in response to a unit change in each explanatory variable. The results found no evidence that outpatient care service costs differed between the ganciclovir prophylaxis group and the control group. However, utilization of medical services differed substantially by race, education, and health status. In particular, whites on average spent on outpatient care almost twice as much as non-white groups. Subjects with a higher education level and higher self-rated health status also tended to use the medical care services more than their counterparts. That whites and better-educated groups consumed more outpatient care services may be attributed in part to their more generous health insurance, but we did not have data on subjects' insurance status. We also looked at the treatment effect of ganciclovir on the use of concomitant medications, which include both brand name and generic medications. We found no significant impact of the study drug on the use of concomitant medications. A detailed description of the estimates on the use of concomitant medications is available upon request.

Hospital Care Cost

The primary hospital care data were taken from the clinical data files, and therefore available for most subjects. There were 725 subjects in the hospital data file, the largest sample size of all. Of the 725 subjects, 240 of them had used hospital care during the study period. Table 5 presents the two-part model results. The logit model in column 1 shows that the drug intervention tended to reduce the probability of seeking hospital care, but this was only significant at the 15% level. People who died or experienced adverse drug events were much more likely to be hospitalized than others. Another determinant of hospital cost was patients' baseline CD4 counts, which showed a highly significant negative association with the hospital costs, indicating that lower CD4 count was associated with increased use of hospital care.

The conditional BIE cost model (column 3) indicates, however, that among hospital care users, patients on ganciclovir

[8] Due to the relatively small sample size and insignificant impact, detailed estimates on the paid and unpaid informal care services are not presented in the Tables, but they are available upon request.

Table 5. Costs of Hospital Care (ACHP)

Variable	Logit model β coefficients	OLS model β coefficients	BIE model β coefficients	Expected BIE model r = b + a(1 − P)**
Intercept	−0.788 (0.408)	6.210 (0.0001)*	6.448 (0.0001)*	—
Treatment	−0.255 (0.148)	−0.256 (0.098)*	−0.268 (0.059)*	−0.439
White	−0.364 (0.102)*	−0.014 (0.941)	−0.030 (0.865)	−0.273
Age	−0.010 (0.430)	−0.009 (0.369)	−0.011 (0.238)	−0.018
Education	0.082 (0.083)*	0.038 (0.396)	0.036 (0.385)	0.090
Income (H)	−0.039 (0.372)	−0.057 (0.155)	−0.056 (0.134)	−0.082
Weeks FTW	−0.003 (0.550)	−0.005 (0.279)	−0.004 (0.314)	−0.006
Δ Health status	−0.037 (0.695)	0.055 (0.513)	0.037 (0.633)	0.012
QOL	−0.003 (0.967)	0.072 (0.217)	0.058 (0.278)	0.056
Health status	0.133 (0.310)	0.016 (0.896)	0.017 (0.884)	0.106
Candidiasis	0.335 (0.089)*	0.250 (0.166)	0.272 (0.111)	0.497
Leukoplakia	0.302 (0.093)*	0.098 (0.542)	0.093 (0.528)	0.295
Herpes zoster	−0.031 (0.868)	−0.120 (0.460)	−0.092 (0.536)	−0.113
CD4	−0.016 (0.001)*	−0.004 (0.334)	−0.004 (0.289)	−0.05
Death	1.904 (0.0001)*	0.504 (0.032)*	0.575 (0.009)*	1.848
Adverse event	0.595 (0.005)*	0.302 (0.097)*	0.306 (0.067)*	0.704
F-Statistic		1.424	1.677	
R^2		0.087	0.101	
Sample mean (log)		6.589	6.614	
Observations	725	239	239	239

Notes: 1. The P-value for each coefficient is given in parentheses; 2. *Significant at the 10% level; 3. **Assume the logit probability function of using hospital care is $P(C > 0) = (1 + e^{-aX})^{-1}$, and the log transform of conditional cost function is $LnC(C > 0) = bX + u$, the expected marginal effect of C with respect to a change in X can be written as $r = \Delta lnC/\Delta X = b + a(1 - P)$ (15). The expected model is derived from combining the probability of seeking care (logit) and the conditional cost of care (BIE). Following the expected model, the % change in hospital cost in response to a unit change in each explanatory variable can be estimated by $r = b + a(1 - P)$, where b and a represent the estimated coefficients of the logit and BIE models respectively, and P is the average probability of using hospital care by anyone in the total sample of AIDS patients.

prophylaxis had lower hospital costs than control group patients by about 27% at the 6% significance level. Furthermore, when taking into account the probability of being a hospital user in the conditional BIE cost function (column 4), the drug intervention would be expected to result in over 44% reduction in monthly hospital cost among the total sample of AIDS patients. This finding implies that the study drug not only can lower AIDS patients' likelihood of using hospital care, but also would reduce the length of stay or intensity of care required in hospitalization.

Total Medical Care Cost

Total monthly medical cost was defined as the sum of the three direct medical care costs: outpatient services, medications and hospital care. Since various cost data were obtained from different sources, only 117 subjects were fully captured (including non-spenders) across all three cost categories. For others without complete information on any one of the three costs, we were unable to tell if the incomplete case was due to missing data or due to real non-use of the services. As a result, we decided to include in the total cost analysis only those patients who has complete information for all three cost component utilization files.

Based upon our estimates in Table 6, both OLS and BIE models show that ganciclovir prophylaxis lowered total average medical costs per patient. However, the effects were not statistically significant. There was probably not enough power to detect a difference due to the small sample size. Death events during the clinical trial remained the most statistically significant determinant of the total medical costs. We found that patients who died during the clinical trial consumed more than three times as much medical care as those who remained alive throughout the study.

DISCUSSION

Spector *et al.* (1996) found the 12-month cumulative rates of CMV disease to be 26% in the placebo group and 14% in the ganciclovir group, indicating a significant relative risk reduction of 49% ($P < 0.001$). The incidence of CMV retinitis after 12 months was 12% in the ganciclovir group, while it was 24% in the placebo group. Using the same clinical trial, the present paper has presented an economic cost analysis of how oral ganciclovir prophylaxis altered the probability of medical care utilization and the cost of various health care resources consumed during the study period (up to 18 months) for advanced AIDS patients prior to CMV diagnosis. Clinical trial patients who did not receive ganciclovir were more likely to progress to CMV disease, to leave the clinical trial, and thus be eliminated from the monthly cost calculations, while patients on ganciclovir prophylaxis who remained CMV disease-free still reported treatment costs for other AIDS opportunistic infections.

Since this analysis only considered the cost of care prior

Table 6. Total Monthly Medical Costs (TMACMP)

Variable	OLS model β coefficients	BIE model β coefficients	BIE model $e^{b\Delta x}$-1
Intercept	2.803 (0.059)*	2.818 (0.032)*	—
Treatment	−0.149 (0.629)	−0.126 (0.635)	−0.118
White	0.419 (0.239)	0.504 (0.112)	0.655
Age	0.018 (0.300)	0.021 (0.143)	0.021
Education	0.123 (0.105)	0.112 (0.100)*	0.118
Income (H)	−0.144 (0.074)*	−0.122 (0.080)*	−0.115
Weeks FTW	−0.002 (0.824)	−0.003 (0.704)	−0.003
Δ Health status	0.040 (0.796)	0.075 (0.590)	0.077
QOL	0.059 (0.549)	0.025 (0.768)	0.025
Health status	0.247 (0.247)	0.197 (0.287)	0.217
Candidiasis	0.095 (0.795)	0.182 (0.560)	0.199
Leukoplakia	−0.207 (0.538)	−0.173 (0.557)	−0.159
Herpes zoster	−0.646 (0.062)*	−0.723 (0.018)*	−0.515
CD4	−0.002 (0.827)	−0.0004 (0.946)	−0.0004
Death	1.322 (0.006)*	1.417 (0.001)*	3.126
Adverse event	0.589 (0.178)	0.620 (0.114)	0.858
F-Statistic	1.003	1.590	1.590
R^2	0.193	0.275	0.275
Sample mean (log)	6.215	6.284	6.284
Observations	78	78	78

Notes: 1. The P-value (T-test) for each coefficient is given in parentheses; 2. *Significant at the 10% level; 3. **To better interpret the β coefficients, we derive the % change in the cost function in response to a unit change in each explanatory variable X. Suppose the cost function C is a semi-log function of X: $LnC = b_0 + bX$, then the % change in C can be expressed as $\Delta C/C = e^{b\Delta X}$-1, where Δx stands for a one unit change in X.

to an actual diagnosis of CMV, it may have biased the cost comparison against ganciclovir prophylaxis. Recent estimates suggest that the average lifetime cost of AIDS patients with CMV is $30,000–$50,000 higher than for AIDS patients without CMV (13–14). Based on these estimates, assuming CMV would add about $37,700 in lifetime patient care costs, this implies that ganciclovir prophylaxis costs would be offset by an expected $8,440 for the avoided costs of CMV disease.

Putting aside the cost of the prophylactic regimen itself, all cost component models showed a consistent downward impact of ganciclovir prophylaxis on the utilization of health care resources. That is, in all the cost component analyses we found no evidence that the treatment group patients consumed more medical or social support services that the control group patients. While this effect was statistically significant only for hospital costs, hospital care was the most costly component of all the medical services involved. In addition, the sample size associated with hospital care was also the largest and would therefore provide the most robust results of all the parameters analyzed.

In assessing the study results, there are two remaining areas that deserve some discussion. First, some of the equations may lack statistical power due to the small sample size. It should be noted, however, that the cost variables with small sample size, such as paid and unpaid caregiver services (HCU1 and HCU2), were also small in magnitude relative to other resources use, especially to hospital care, the most important driver of cost.

Second, there might be some differences among the AIDS patients with respect to receiving concomitant drug therapies. In particular, many patients were on highly active antiretroviral therapy (HAART). Thus, to better identify the study drug treatment effect, it would be prudent to stratify patient samples by whether or not patients are receiving HAART, and by other therapeutic sub-categories. Clinical trial randomization ensured a valid and unbiased assessment of ganciclovir prophylaxis costs, but therapeutic stratification would enhance estimation efficiency. Unfortunately, the current study sample was too limited to undertake the stratification. Future studies should take into account such a stratification strategy.

Nevertheless, this study concluded with two major observations shedding light on cost analysis in general and the assessment of ganciclovir prophylaxis in particular. First, to analyze clinical trial cost data, it is important to test different model specifications and estimate the models using alternative approaches in consideration of the effects of relatively small sample and outliners. For example, our BIE estimates were able to identify the treatment effect on cost savings in hospital care, which was not picked up with the OLS estimates. In our view, the types of econometric evaluations of the cost data undertaken in this analysis are often necessary for rigorous assessment of health care costs and utilization within the context of clinical trials. We hope that these techniques become more routinely used by health service researchers in the economic assessment of medical technologies.

Second, at the empirical level, we identified that despite of its significant efficacy of preventing CVM disease among AIDS patients (1), ganciclovir prophylaxis did not lead to additional significant treatment costs, other than the cost of daily administration of ganciclovir. Moreover, our refined results indeed found the drug treatment to be cost saving in hospital care utilization.

ACKNOWLEDGMENTS

Roche Pharmaceuticals, Inc. supported this research. We thank Amy Jones, Carol Jamieson, Lou Garrison, and Yong Yuan for their suggestions and assistance in conducting this research.

REFERENCES

1. S. A. Spector, G. F. McKinley, J. P. Lalezari, T. Samo, R. Andruczk, S. Follansbee, P. D. Sparti, D. V. Havlir, G. Simpson, W. Buhles, R. Wong, and M. Stempien. Oral Ganciclovir for the Prevention of Cytomegalovirus Disease in Persons with AIDS. *N. Engl. J. Med.* **334**:1491–1497 (1996).
2. S. Bozzette, R. Parker, and J. Hay. A Cost Analysis of Approved Antiretroviral Strategies in Advanced Human Immunodeficiency Virus Disease. *AIDS.* **7**:355–362 (1994).
3. N. Duan, W. G. Manning, and J. P. Newhouse. A Comparison of Alternative Models for the Demand for Medical Care. *J. Bus. Econ. Stat.* **2**:279–282 (1983).
4. J. Heckman. Sample Selection Bias as a Specification Error. *Econometrica.* **47**:153–61 (1979).
5. W. G. Manning, N Duan, and W. H. Rogers. Monte Carlo Evidence on the Choice Between Sample Selection and Two-part Models. *J. Econometrics*, **35**:59–82 (1987).
6. P. Fenn, A. McGuire, V. Phillips, M. Backhouse, and D. Jones. The Analysis of Censored Treatment Cost Data in Economic Evaluation. *Med. Care.* **33**:851–863 (1995).
7. J. Lipscomb, M. Ancukiewicz, G. Parmigiani, V. Hasselblad, G. Samsa, and D. B. Matchar. Predicting the Cost of Illness: A Comparison of Alternative Models Applied to Stroke. *Med. Decis. Making.* **18** suppl:S39–S56 (1998).
8. R. Andersen and L. A. Aday. Access to Medical Care in the US: Realized and Potential. *Med. Care.* **16**:533–546 (1973).
9. J. F. Alberts, R. Sanderman, J. M. Eimers, and W. J. van den Heuvel. Socioeconomic Inequity in Health Care: A Study of Services Utilization in Curacao. *Soc Sci Med*, **45**:213–220 (1997).
10. J. W. Hay and R. Olsen. Let Them Eat Cake: A Note on Comparing Alternative Models for the Demand of Medical Care. *J. Bus. Econ. Stat.* **1**:115–126 (1984).
11. R. Welsch. Regression Sensitivity Analysis and Bounded Influence Estimation. In: J. Kmenta and J. B. Ramsay (eds), *Eval Econometric Models*, New York. Academic Press (1980).
12. J. Mackinnon, H. White, and R. Davidson. Tests for Model Specification in the Presence of Alternative Hypothesis: Some Further Results. *J. Econometrics.* **21**:53–70 (1983).
13. F. J. Hellinger. The Lifetime Cost of Treating a Person with HIV Infection. *JAMA.* **270**:474–478 (1993).
14. D. Rose and H. S. Sacks. Cost-Effectiveness of Cytomegalovirus (CMV) Diseases Prevention in Patients with AIDS: Oral Ganciclovir and CMV Polymerase Chain Reaction Testing. *AIDS.* **11**:883–887 (1997).
15. W. G. Manning. The Demand for Alcohol: The Differential Response to Price. unpublished manuscript. University of Minnesota (1993).