HEALTH ECONOMICS
Health Econ. **11**: 551–566 (2002)
Published online 2 August 2002 in Wiley InterScience (www.interscience.wiley.com). **DOI:**10.1002/hec.739

Bayesian estimation, simulation and uncertainty analysis: the cost-effectiveness of ganciclovir prophylaxis in liver transplantation

David J. Vanness[a,*] and W. Ray Kim[b]
[a] *Division of Health Care Policy & Research, Mayo Clinic, Rochester, Minnesota, USA*
[b] *Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, Minnesota, USA*

Summary

This paper demonstrates the usefulness of combining simulation with Bayesian estimation methods in analysis of cost-effectiveness data collected alongside a clinical trial. Specifically, we use Markov Chain Monte Carlo (MCMC) to estimate a system of generalized linear models relating costs and outcomes to a disease process affected by treatment under alternative therapies. The MCMC draws are used as parameters in simulations which yield inference about the relative cost-effectiveness of the novel therapy under a variety of scenarios. Total parametric uncertainty is assessed directly by examining the joint distribution of simulated average incremental cost and effectiveness. The approach allows flexibility in assessing treatment in various counterfactual premises and quantifies the global effect of parametric uncertainty on a decision-maker's confidence in adopting one therapy over the other. Copyright © 2002 John Wiley & Sons, Ltd.

Keywords cost-effectiveness; Bayesian analysis; Markov Chain Monte Carlo; uncertainty

Introduction

Quantification of uncertainty has been attracting increasing attention in the cost-effectiveness literature (see, e.g., Manning *et al.* [1], Briggs and Sculpher [2] and the Consensus Statement from the Conference on Economic Modelling [3]). Uncertainty arising from population variation, parametric imprecision, and even choice of model – all should be addressed [4]. Attention to uncertainty is not just an affirmation of the need to conduct statistically rigorous research. It should also reflect that uncertainty is important to decision-makers.

Bayesian methods may be particularly well suited to the assessment of uncertainty from a decision-maker's perspective [5]. Bayes' Law is, after all, one mathematical representation of mechanisms by which statistically rational decision-makers might assess their knowledge about uncertain propositions [6]. Through the ubiquity of expected utility [7] and games of incomplete information [8], Bayesian principles have become intrinsic to microeconomic theory.

Despite being second-nature in theoretical work [9], Bayesian empirical analysis has been limited by the scarcity of tractable models based on conjugate priors (whose posteriors are simple functions of observed data). However, the advent of cheap, fast

*Correspondence to: Division of Health Care Policy and Research, Mayo Clinic, 200 First Street SW, Rochester, Minnesota 55905, USA. Tel.: +1 507 538 1240; fax: +1 507 284 1731; e-mail: vanness.david@mayo.edu

Received 16 October 2001
Accepted 15 March 2002

computation and efficient Markov Chain Monte Carlo (MCMC) techniques has resulted in a surge of published Bayesian analyses [10]. MCMC generates samples from the simulated posterior distribution of parameters, even under quite complicated models. It therefore provides useful glimpses into a Bayesian decision-maker's posterior knowledge about relationships between treatments, covariates and outcomes.

In this paper, we analyze cost-effectiveness data collected alongside a clinical trial by estimating a structural model of resource use and outcomes using MCMC. We then conduct a number of simulations using the MCMC parameter samples to predict the joint distribution of costs and outcomes under a variety of counterfactual premises. This approach can quantify the impact of global parametric uncertainty in terms of the probability that one technology dominates (both lower cost and improved outcomes) another in a variety of settings.

Background

We examine two competing therapies for prevention of cytomegalovirus (CMV) infection in liver transplant patients. CMV infection generally remains asymptomatic in the general population. However, in immunosuppressed patients (such as those taking anti-rejection medication), CMV infection can progress to full-blown CMV disease characterized by tissue invasion [11]. CMV disease significantly increases morbidity and mortality in solid organ transplantation [11–15].

CMV seronegative recipients (R−) of organs from CMV seropositive donors (D+) are at highest risk of CMV disease [16]. Two recent trials report that prolonged administration of intravenous or oral ganciclovir reduces CMV disease in R−/D+ cases [17,18].

We analyze direct medical cost (proxied by billed charges) and CMV outcomes (infection and disease) from a randomized trial in which sequential use of intravenous ganciclovir and high-dose oral acyclovir (GCV group) resulted in a significant reduction in CMV disease in R−/D+ liver transplant recipients, when compared to high-dose oral acyclovir alone (ACV group) [19]. A previous analysis concluded that patients with CMV disease had 49% higher charges and that R−/D+ patients in the GCV group had significantly lower charges than R−/D+ patients in the ACV group [20].

Data

In the trial, 167 patients were enrolled from two transplant centers over a period of 3.5 years. Patients were randomly assigned to receive either 800 mg oral acyclovir four times daily for 120 days (ACV group), or 5 mg/kg ganciclovir intravenously every 12 h for the first 14 days, followed by oral acyclovir (800 mg four times daily) for the remaining 106 days (GCV group).

Recipient and donor CMV serologies were obtained before transplant. CMV infection was defined to be the isolation of CMV from any body fluid or tissue other than urine. CMV disease was defined to be CMV infection together with CMV syndrome or tissue invasion with CMV such as pneumonia, enteritis and hepatitis [11].

Of the enrollees, 147 patients (88%) were transplanted at our institution. Data collection continued for 120 days, starting from the day of transplantation. Inpatient and outpatient charges (in constant 1990 dollars), including professional fees and charges for acyclovir and ganciclovir were analyzed. Analysis was restricted to 134 patients with complete follow-up data.

Outcomes of interest are any CMV infection (including those who progress to CMV disease), $y_i^{\mathrm{C}} \in \{0,1\}$, CMV disease, $y_i^{\mathrm{D}} \in \{0,1\}$, ICU days, $y_i^{\mathrm{F}} \in \mathbb{Z}^+$, non-ICU hospital days, $y_i^{\mathrm{H}} \in \mathbb{Z}^+$ and total charges, $y_i^{\mathrm{G}} \in \mathbb{R}^+$. Covariates are age, sex, surgical blood loss (in units), surgical time (in seconds) and donor/recipient seropositivity. Continuous covariates were centered around zero and rescaled by their standard deviations. Untransformed summary statistics are presented in Table 1.

Methods

Statistical model

CMV status, ICU days, non-ICU hospital days and total charges are modeled as a set of independent generalized linear models, presented as a directed graph in Figure 1. We model CMV status using a sequential probit specification [21].[a]

Table 1. Means of study variables

	ACV group	GCV group
Independent variables		
Age (yrs.)	50.89	49.4
Female (=1)	0.43	0.43
R − /D − (= 1)	0.09	0.09
R − /D + (= 1)	0.17	0.18
R + /D − (= 1)	0.25	0.26
R + /D + (= 1)	0.49	0.46
Surgery time (s)	24 013.91	22 986.46
Surgical blood use (units)	7.53	6.05
Dependent Variables		
Any CMV infection (=1)	0.45	0.25
CMV disease (=1)	0.2	0.11
ICU days	5.54	3.97
Non-ICU days in hospital	12.94	12.45
Total charges (1990 US$)	138 849.58	125 810.36
n	69	65

Let x_i^C be a row vector of covariates for patient i that are associated with propensity for becoming infected with CMV, and c be the associated parameter column vector. Then the probability of becoming infected with CMV is $P(Y^C = 1|x_i^C) = \Phi(x_i^C c)$, where Φ is the standard normal CDF.

Once infected, some patients will progress to full-blown CMV disease, while others will have infection only. Let x_i^D be a row vector of covariates for patient i that are associated with propensity for progressing to CMV disease, and d be the associated parameter column vector. Probability of progression is then $P(Y^D = 1|x_i^D, Y_i^C = 1) = Y_i^C \Phi(x_i^D d)$.

A key covariate in both infection and progression is donor/recipient serostatus, which is classified into three categories: recipients of seronegative grafts (D−), seronegative recipients of seropositive grafts (R−/D+) and seropositive recipients of seropositive grafts (R+/D+). Because no patient

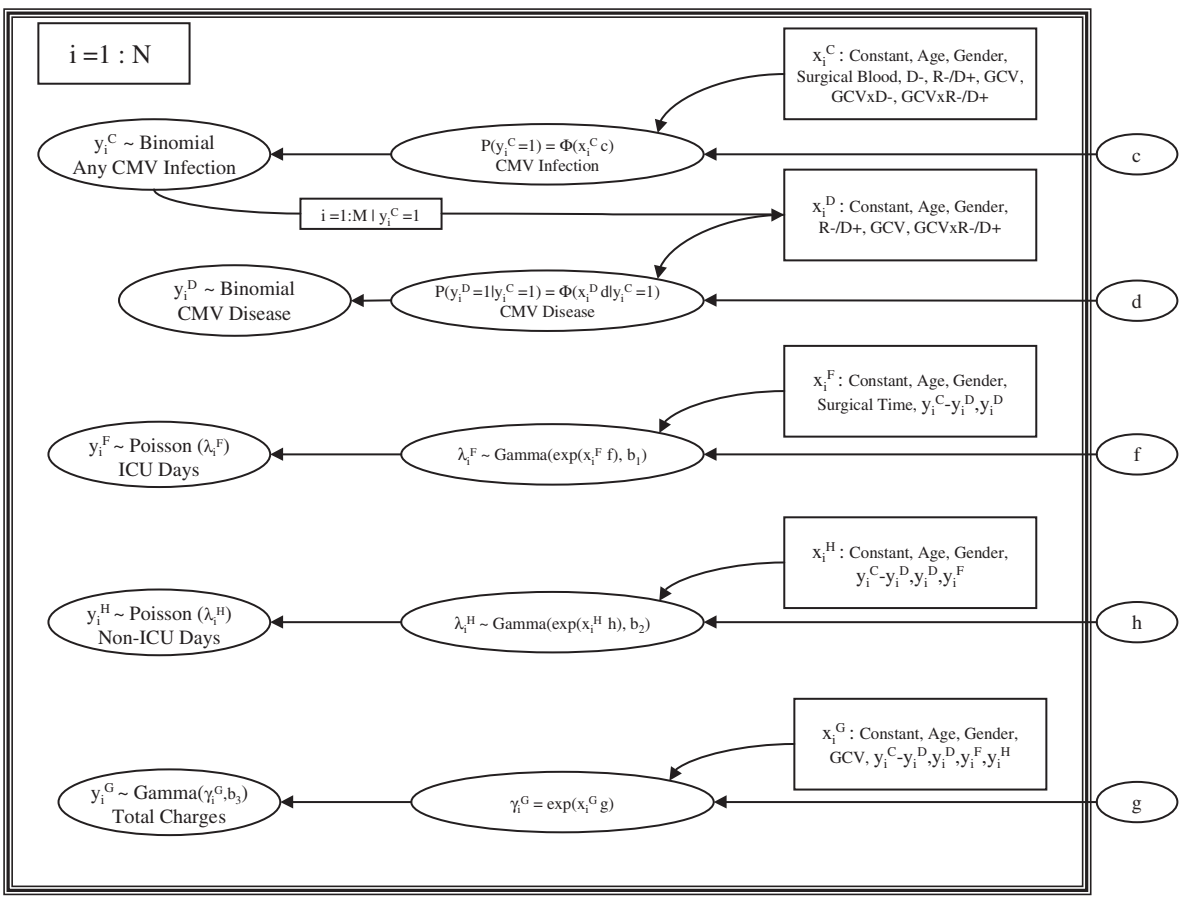


Figure 1. Directed graphical representation of the model

who received a seronegative graft (D−) went on to develop full-blown CMV disease, D− would be a perfect predictor of disease progression. We therefore assume that such patients are at zero risk for CMV disease, and exclude them from the progression model.[b]

Treatment with GCV (relative to ACV control) enters the model as a dichotomous variable, which equals one when a patient receives GCV. Treatment is also interacted with the serostatus risk categories, reflecting the hypothesis that ganciclovir may have a different effect for different serostatus risk categories. Because patients may also be exposed to CMV through blood, we include the quantity of surgical blood used as a predictor of infection. Infection and progression may also differ by recipient age and sex.

We hypothesize that CMV infection[c] and disease increase the overall length of stay in the hospital as well as requiring a higher number of days spent in intensive care. Length of stay in the ICU is assumed to be a function of CMV status and covariates and is modeled using a negative binomial specification [23]. Let x_i^{F} be a row vector of covariates associated with days in the ICU (including a constant, age, sex, total surgical time and the patient's CMV outcome), and f the associated column parameter vector. Assume ICU stay, y_i^{F}, is distributed:

$$y_i^{\mathrm{F}} \sim \mathrm{Poisson}(\lambda_i^{\mathrm{F}}),$$

$$\text{where } \lambda_i^{\mathrm{F}} \sim \mathrm{Gamma}(\exp(x_i^{\mathrm{F}} f), b_1).$$

We assume that a longer stay in the ICU increases the length of time spent convalescing in non-ICU areas of the hospital. We therefore model non-ICU days (with variables superscripted by H) as a function of CMV status, ICU days and covariates in a negative binomial framework identical to the previous equation (excluding surgical time, which we assume impacts non-ICU days only indirectly through requiring longer initial ICU stays).

We hypothesize that CMV status affects direct medical cost in the following ways: (1) requiring more resources necessary to treat the infection; (2) extending the amount of time the patient spends in an intensive care environment as the infection is treated; (3) extending the overall hospital stay as the recipient recovers; and (4) requiring more intense medical followup after the initial hospitalization. The direct cost of GCV treatment may be offset if GCV improves outcomes. Let x_i^{G} be a row vector of covariates associated with total charges (age, sex, ICU days, non-ICU days, CMV infection only, CMV disease and GCV treatment), and g the associated column vector of parameters. We use a gamma model to estimate total charges [24]:

$$D_i^{\mathrm{G}} \sim \mathrm{Gamma}(\exp(x_i^{\mathrm{G}} g), b_3).$$

Simulation

The simulation mirrors the empirical model, except that the levels of the model are nested (i.e., each level of the simulation depends on the simulated result(s) of one or more of the previous levels). First, all trial participants, regardless of their actual trial arm, are 'assigned' to ACV therapy by setting the value of the GCV column in covariate matrices X^{C}, X^{D} and X^{G} to a 134×1 vector of zeros. Outcomes are then simulated assuming this assignment to ACV.

Let j index the draws from the MCMC simulated posterior. Note that there are 35 parameters in the model. Therefore, each MCMC draw j from the posterior generates a 35×1 vector of parameters θ_j. Given the hypothetical assignment to GCV or ACV, and a draw j, outcomes are simulated using the procedure outlined below.

Within θ_j, the 9×1 CMV parameter sub-vector, c_j is multiplied into covariates X^{C} for all participants i using the standard normal distribution link function to obtain CMV infection probabilities for each patient assuming they have received ACV (control) therapy. Uniform random numbers are then drawn to simulate CMV infections $\hat{y}_i^{\mathrm{C}}(x_i^{\mathrm{C}}|\mathrm{GCV}=0; c_j)$ in the group. Next, d_j is multiplied into covariates X^{D} for all non-D− participants who were simulated to become infected. Similarly, uniform random numbers are then drawn to simulate whether or not CMV disease occurs $\hat{y}_i^{\mathrm{D}}(x_i^{\mathrm{D}}|\mathrm{GCV}=0; \hat{y}_i^{\mathrm{C}}=1; d_j)$ for each patient with simulated infection.

Simulated CMV status for all individuals becomes a covariate column in X^{F}, which is multiplied into the draw of ICU day parameters f_j using an exponential link to yield the shape parameter of the gamma distribution for the mean number of ICU days. Gamma-distributed random numbers are then generated for each individual, which in turn become parameters for the Poisson distribution for ICU days. Poisson random numbers are generated to yield a simulated outcome of ICU days, $\hat{y}_i^{\mathrm{F}}(x_i^{\mathrm{F}}|\mathrm{GCV}=0; \hat{y}_i^{\mathrm{C}}, \hat{y}_i^{\mathrm{D}}; f_j)$,

for each individual. Simulated ICU days and simulated CMV status then become covariates in X^{H}. Given the draw for parameters for h_j, simulated non-ICU hospital days, $\hat{y}_i^{H}(x_i^{H}|\text{GCV} = 0; \hat{y}_i^{C}, \hat{y}_i^{D}, \hat{y}_i^{F}; h_j)$, are generated in a similar manner.

Finally, simulated CMV status, ICU and non-ICU hospital days become covariates in X^{G} and are multiplied into the draw for parameters g_j using an exponential link function to yield shape parameters for each individual. Gamma-distributed random numbers are drawn to simulate total charges, $\hat{y}_i^{G}(x_i^{G}|\text{GCV} = 0; \hat{y}_i^{C}, \hat{y}_i^{D}, \hat{y}_i^{F}, \hat{y}_i^{H}; g_j)$ for each individual.

The exercise is then repeated, assigning all participants to GCV therapy by setting the GCV column equal to a vector of ones. Incremental outcomes are then calculated at the individual level by taking the difference (ganciclovir minus acyclovir control). Given simulated charges and outcomes for each individual under both treatment and control, for each draw j we calculate average incremental charges $\Delta\hat{Y}_j^{G}$, average attributable CMV infection risk reduction $\Delta\hat{Y}_j^{C}$ and average CMV disease risk reduction $\Delta\hat{Y}_j^{D}$ over all participants.[d]

$$\Delta\hat{Y}_j^{G} = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{y}_i^{G}(x_i|\text{GCV} = 1; \theta_j) - \hat{y}_i^{G}(x_i|\text{GCV} = 0; \theta_j)\right)$$

$$\Delta\hat{Y}_j^{C} = \frac{1}{N}\sum_{i=1}^{N} -\left(\hat{y}_i^{C}(x_i|\text{GCV} = 1; \theta_j) - \hat{y}_i^{C}(x_i|\text{GCV} = 0; \theta_j)\right)$$

$$\Delta\hat{Y}_j^{D} = \frac{1}{N}\sum_{i=1}^{N} -\left(\hat{y}_i^{D}(x_i|\text{GCV} = 1; \theta_j) - \hat{y}_i^{D}(x_i|\text{GCV} = 0; \theta_j)\right)$$

Repeating the process for each draw of θ_j yields a cloud of simulated average incremental charges and average attributable risk reductions which captures the remaining uncertainty about model parameters after the data from the trial have been incorporated into the posterior via Bayes' law. Posterior joint distributions of $(\Delta\hat{Y}^{G}, \Delta\hat{Y}^{C})$ and $(\Delta\hat{Y}^{G}, \Delta\hat{Y}^{D})$ were then analyzed using a bivariate kernel density estimator (normal kernel, bandwidth set by Silverman's rule of thumb) [25].

Results

Estimation

To reflect naive beliefs, diffuse, independent, normal priors (mean = 0, precision = 1.0E-06) were chosen. Scale parameters $b_1 = 0.211651$, $b_2 = 0.194171$, and $b_3 = 6.86703E - 5$ were estimated separately from the data as the inverse of the observed ratio of the variance to the mean of ICU days, non-ICU days and charges, respectively.[e] MCMC (using the Metropolis algorithm of WinBUGS v.1.3) was used to sample from the simulated posterior distributions of the parameters using three parallel Markov chains started from different random initial values [26].[f]

Moderate autocorrelation and cross-parameter correlation was detected during the estimation procedure, indicating possibly slow mixing for some parameters. We thinned the Markov chains (keeping every 20th draw) until 30 000 draws were collected for each chain, of which, we retained the second half (15 000 draws). In order to assess convergence, we monitored the Metropolis acceptance rate during estimation, and performed tests by Geweke [27], Brooks, *et al.* [28,29], Raftery and Lewis [30] and Heidelberger and Welch [31] as implemented in BOA v.1.0.0 [32]. All tests were consistent with convergence of all three chains, although the Heidelberger and Welch test suggested that a small number of additional draws be discarded for a small number of parameters. The Raftery and Lewis and Heidelberger and Welch tests both indicated that we had sufficient draws to estimate the 0.025 and 0.075 percentiles of the posteriors with acceptable accuracy (Type I error rate of 5%). After assessing convergence, we kept only the last 7500 draws from each chain and combined them to form a collective sample of 22 500 draws from the simulated posterior distributions. Posterior means, medians, 95% credible regions, and $P(\theta > 0)$ are presented in Table 2.

We consider the first stage of the sequential probit, which estimates the probability of infection conditional on treatment, serostatus risk category, blood use, age and gender. Note that the excluded risk category is the medium risk (R+/D+) group. The results suggest that ganciclovir significantly reduces infection risk among the medium risk group (mean = −0.8562; $P(> 0) = 0.0064$). GCV effectiveness was reduced somewhat in low-risk

Table 2. Results from the empirical model

Model parameter	Mean	Std Dev.	MC error	95% credible interval		
				Lower	Upper	$P(>0)$
	Any CMV Infection					
$c0$ Constant	0.0860	0.2558	0.0023	−0.4089	0.5886	0.6337
$c1$ Female	−0.1962	0.2531	0.0018	−0.6878	0.3007	0.2199
$c2$ Age (standardized)	−0.0285	0.1277	0.0008	−0.2795	0.2231	0.4079
$c3$ Blood use (standardized)	0.2031**	0.1247	0.0008	−0.0384	0.4507	0.9502
$c4$ Donor negative	−0.8910**	0.3748	0.0030	−1.6450	−0.1741	0.0077
$c5$ Recipient negative/Donor positive	0.7331*	0.4574	0.0037	−0.1682	1.6250	0.9481
$c6$ Treated with ganciclovir	−0.8562**	0.3492	0.0030	−1.5380	−0.1689	0.0064
$c7$ Ganciclovir×D−	0.3149	0.6074	0.0047	−0.8567	1.5130	0.7036
$c8$ Ganciclovir×R−/D+	0.6246	0.6576	0.0055	−0.7166	1.8520	0.8273
CMV disease conditional on infection						
$d0$ Constant	−0.7747**	0.4175	0.0037	−1.5950	0.0364	0.0272
$d1$ Female	1.0257**	0.5065	0.0039	0.0363	2.0220	0.9802
$d2$ Age (standardized)	0.2014	0.2466	0.0017	−0.2824	0.6832	0.7924
$d3$ Recipient negative/Donor positive	1.3620**	0.6431	0.0053	0.1357	2.6430	0.9870
$d4$ Treated with ganciclovir	1.7972**	0.7862	0.0070	0.2691	3.3340	0.9933
$d5$ Ganciclovir×R−/D+	−3.4469**	1.1156	0.0104	−5.6290	−1.2520	0.0002
ICU days						
$f0$ Constant	0.1436*	0.1113	0.0009	−0.0785	0.3570	0.9003
$f1$ Female	0.0624	0.1434	0.0010	−0.2160	0.3447	0.6694
$f2$ Age (standardized)	0.0499	0.0716	0.0004	−0.0890	0.1908	0.7568
$f3$ Surgical time (standardized)	0.1673**	0.0695	0.0005	0.0319	0.3046	0.9887
$f4$ CMV infection only	0.0363	0.1856	0.0012	−0.3152	0.4104	0.5840
$f5$ CMV disease	0.2743*	0.1878	0.0013	−0.0835	0.6491	0.9252
Non-ICU hospital days						
$h0$ Constant	0.7605**	0.0885	0.0007	0.5886	0.9345	1.0000
$h1$ Female	0.0418	0.1089	0.0008	−0.1790	0.2481	0.6536
$h2$ Age (standardized)	−0.0085	0.0517	0.0003	−0.1105	0.0908	0.4318
$h3$ CMV infection only	−0.0791	0.1432	0.0009	−0.3621	0.1976	0.2935
$h4$ CMV disease	0.1364	0.1468	0.0009	−0.1500	0.4213	0.8242
$h5$ Days in ICU	0.0389**	0.0081	0.0001	0.0231	0.0547	1.0000
Total Charges						
$g0$ Constant	2.0329**	0.0640	0.0006	1.9070	2.1590	1.0000
$g1$ Female	−0.0922	0.0571	0.0004	−0.2011	0.0219	0.0532
$g2$ Age (standardized)	−0.0185	0.0281	0.0002	−0.0733	0.0359	0.2558
$g3$ CMV infection only	0.0325	0.0733	0.0005	−0.1147	0.1720	0.6744
$g4$ CMV disease	0.1176*	0.0823	0.0005	−0.0463	0.2775	0.9223
$g5$ Days in ICU	0.0236**	0.0068	0.0001	0.0102	0.0370	0.9996
$g6$ Non-ICU hospital days	0.0075**	0.0042	0.0000	−0.0010	0.0155	0.9608
$g7$ Treated with ganciclovir	−0.0149	0.0583	0.0004	−0.1261	0.1026	0.4021

*Denotes Posterior $P(>0)$ or $P(<0) > 0.90$.
**Denotes Posterior $P(>0)$ or $P(<0) > 0.95$.

and high-risk groups, as reflected by positive (but conventionally non-significant) coefficients on the interaction terms: (GCV × D− mean = 0.3149; $P(>0) = 0.7036$) and (GCV × R−/D+ mean = 0.6246; $P(>0) = 0.8273$). Other covariates were as expected with higher surgical blood use and R−/D+ serostatus increasing infection risk, and D− serostatus decreasing infection risk.

The second stage of the sequential probit estimates the probability that those infected with CMV will progress to full-blown CMV disease, conditional on treatment, serostatus risk, age and gender. Recall that since no D− patients developed CMV disease, such patients were excluded. The excluded risk category remains the medium risk (R+/D+) group. The results suggest that GCV increases the risk of progression of CMV infection to disease for R+/D+ patients (mean = 1.7972; $P(>0) = 0.9933$). However, we must remember from the first stage that GCV significantly reduced the overall rate of infection for such medium risk patients. The results for the interaction of GCV and R−/D+ (mean = −3.4469; $P(>0) = 0.0002$) suggest that GCV strongly reduces progression rates for high-risk patients. As expected, patients in the high-risk serostatus group were more likely to progress to CMV disease (mean = 1.3620; $P(>0) = 0.9870$). Interestingly, female patients were also significantly more likely to progress (mean = 1.0257; $P(>0) = 0.9802$).

The associations of CMV infection with ICU days (mean = 0.0363; $P(>0) = 0.5840$) and non-ICU days (mean = −0.0791; $P(>0) = 0.2935$) are relatively weak. CMV disease, on the other hand, has a strong positive association with days in the ICU (mean = 0.2743; $P(>0) = 0.9252$) and a less-strong positive association with non-ICU days in the hospital (mean = 0.1364; $P(>0) = 0.8242$). As expected, longer surgical time is associated with a longer stay in the ICU (mean = 0.1673; $P(>0) = 0.9887$). As expected, a longer ICU stay is strongly associated with a longer non-ICU stay (mean = 0.0389; $P(>0) = 1$).

Both ICU (mean = 0.0236; $P(>0) = 0.9996$) and non-ICU (mean = 0.0075; $P(>0) = 0.9608$) hospital days are strongly positively associated with higher total charges. As expected, ICU days appear to be more 'expensive' than non-ICU days.

CMV infection demonstrates only a weak positive association with higher charges (mean = 0.0325; $P(>0) = 0.6744$). However, CMV disease (independent from its association with longer ICU and non-ICU stays) appears positively associated with higher charges (mean = 0.1176; $P(>0) = 0.9223$). Independent of its association through reducing CMV infection and disease risk (and resultant changes in resource use), GCV shows no strong association with charges (mean = −0.0149; $P(>0) = 0.4021$).

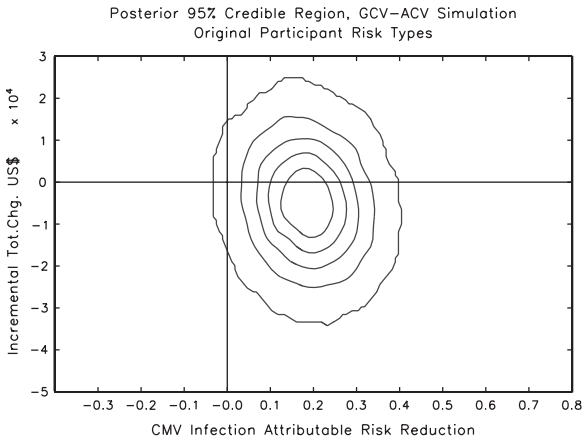
In summary, ganciclovir appears to reduce CMV infection rates for medium risk R+/D+ cases and reduce progression of infection to full-blown disease for R−/D+ cases. Lower rates of CMV disease appear likely to reduce resource use both directly and through shortened ICU times, and overall length of stay. The apparent differences by serostatus group in CMV infection and disease outcomes under treatment underscores the desirability of building risk type into our simulations.

Simulation

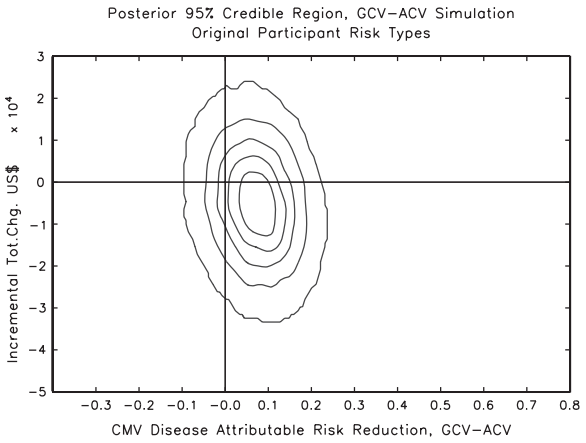
Simulations were conducted with Gauss (v.3.6) given the 22,500 MCMC draws. The first set of simulations examines the original group of trial participants in their original risk categories. The left graph of Figure 2 shows a contour plot of the 95% posterior credible region for the joint distribution of incremental total charges and attributable CMV infection risk reduction $(\Delta\hat{Y}^{G}, \Delta\hat{Y}^{C})$. There appears to be a negative slope of the credible region, indicating that preventing CMV infection reduces total charges. The bivariate posterior mean implies that after observing the trial, a decision-maker can expect a savings of $4,840 and 18.6% reduction in CMV infection risk when using GCV prophylaxis relative to ACV. The decision-maker can have reasonable confidence in the desirability of GCV in this case, given that 98.1% of the posterior density indicates improved CMV infection rates and 67.5% indicates lower charges. About 66.5% of the posterior density lies in the lower-right (dominant) quadrant indicating both lower expenses and better outcomes.

The right graph of Figure 2 presents the CMV disease case. The credible region spans all four quadrants of the plane, with 85.7% of the posterior density indicating improved CMV disease rates, 67.5% indicating lower total charges, and 59.4% indicating GCV dominance. The posterior mean is an incremental charge reduction of $4,840 and 7.2% reduction in CMV disease risk.[g]

The next set of simulations considers outcomes if all enrolled patients had actually been high risk (R−/D+) cases. We can repeat the above simulations, also imposing that R − /D+ = 1 and D− = 0 for all trial participants.[h] The left graph in Figure 3 presents the CMV infection case. Relative

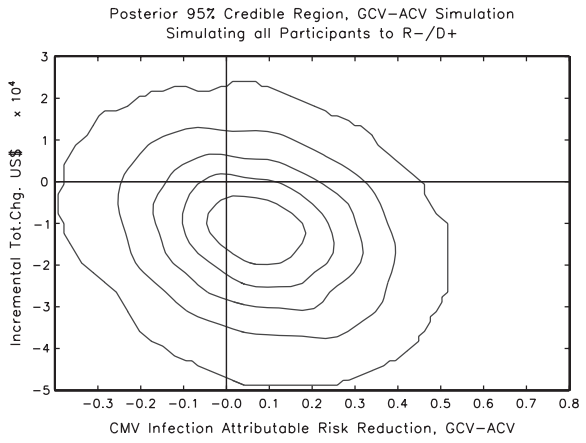


Any CMV infection: $(\Delta \hat{Y}^G, \Delta \hat{Y}^C)$.

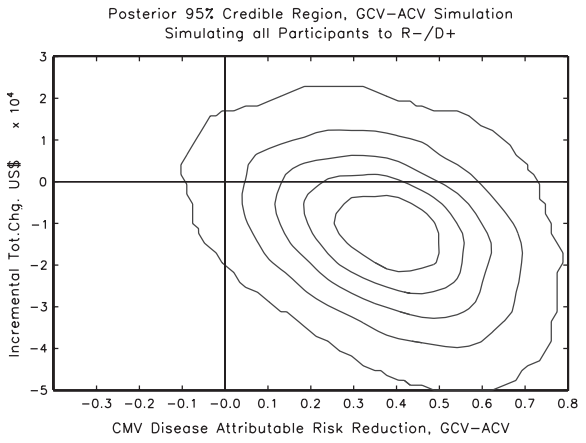


CMV disease: $(\Delta \hat{Y}^G, \Delta \hat{Y}^D)$.

Figure 2. Posterior 95% credible region: original risk types



Any CMV infection: $(\Delta \hat{Y}^G, \Delta \hat{Y}^C)$.

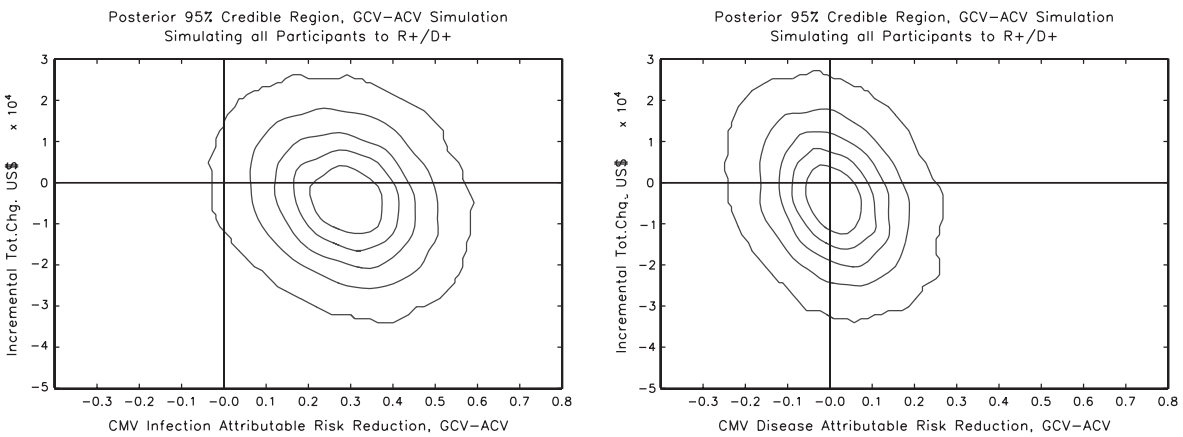


CMV disease: $(\Delta \hat{Y}^G, \Delta \hat{Y}^D)$.

Figure 3. Posterior 95% credible region: simulated R−/D+

to the simulations of participants in their actual risk categories, the R−/D+ CMV infection credible region is larger (indicating greater global uncertainty), lower (indicating greater charge reductions) and more to the left (indicating lower effectiveness). However, for CMV *disease* (the right graph of Figure 3) in the R−/D+ case, the credible region is lower and more to the right, indicating larger reductions in charges and CMV disease rates than in the original risk group simulation. For CMV disease, 97.6% of the posterior density indicates that GCV reduces CMV disease rates, 80.1% indicates lower charges and 79.0% of the posterior density is in the dominant quadrant. The posterior mean is a reduction in charges of $12 249 and 37.3% reduction in CMV disease risk. This simulation suggests that for R−/D+ cases, the benefit of GCV as a protective agent may be more so through lower rates of progression of infection to full-blown disease than through lower rates of infection.

The final simulations suppose that all trial participants were R+/D+. The left graph of



Any CMV infection: $(\Delta \hat{Y}^G, \Delta \hat{Y}^C)$. CMV disease: $(\Delta \hat{Y}^G, \Delta \hat{Y}^D)$.

Figure 4. Posterior 95% credible region: simulated R+/D+

Figure 4 presents CMV infection risk reduction. In the R+/D+ simulation, as in the original risk category simulation, nearly the entire region lies within the upper and lower right quadrants for CMV infection. The credible region is larger than in the original case, reflecting greater global uncertainty, with 98.5% of the posterior density indicating lower CMV infection rates and 60.1% of the posterior density in the dominant quadrant. The posterior mean is an incremental charge reduction of $3,594 and a 28.5% reduction in CMV infection risk. However, for CMV disease in the R+/D+ simulation (the right graph of Figure 4), a decision-maker would likely be unconvinced of a risk reduction from GCV use. Only 54.6% of the posterior density indicates lower CMV disease rates, 60.8% indicates lower charges, and only 37.2% of the posterior is in the dominant quadrant indicating both lower disease rates and charges. The simulations suggest that for R+/D+ cases, the protective effect of GCV is not compelling and seems to come through lower rates of infection rather than lower rates of progression.

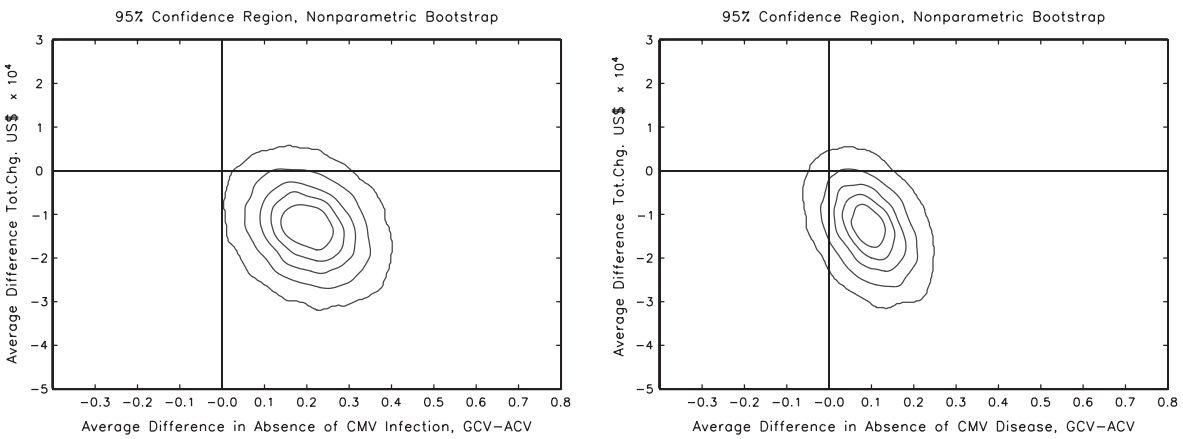
Bootstrap

For comparison with the parametric simulations, a nonparametric (ideal) bootstrap analysis was also conducted [33]. Specifically, we resampled with replacement from our original dataset to create 22,500 replicate datasets of the same length in each treatment group ($n = 65$ in GCV group; $n = 69$ in ACV group). We used the percentile-t method and kernel density estimates to identify the 95% confidence ellipsoid for the joint distribution of average incremental cost and outcomes (separately for infection and disease) [34] as shown in Figure 5. The bootstrap case for GCV is quite strong. The base case (original sample) mean is a 20.3% lower CMV infection risk, 9.5% lower CMV disease risk and $13 039 lower total charges. The bootstrap confidence intervals show 99.3% of the bootstrap density associated with improved infection risk, 93.1% associated with improved disease risk, and approximately 94.2% associated with lower total charges for both infection and disease. For infection and disease, 93.7% and 88.6% of the bootstrap densities lie in the dominant quadrant, respectively.

Discussion

Limitations

There are a number of limitations to our data. Restricting analysis to complete responses may bias estimates if assignment to treatment is correlated with being a non-completer and if non-completers have systematically different outcomes. Our perspective is also limited by reliance upon billed charges rather than true direct medical cost. The reasons to prefer standardized measures of cost over billed charges are well known [35]. The



Any CMV infection: $(\Delta \hat{Y}^G, \Delta \hat{Y}^C)$. CMV disease: $(\Delta \hat{Y}^G, \Delta \hat{Y}^D)$.

Figure 5. Bootstrap 95% confidence ellipsoid

difference between costs and charges may be heterogeneous among the various service and facility line-items of a patient's bill. If the difference varies systematically between treatment and control groups, then the results of this analysis may be biased. However, in a single institution, charges are likely to be proportional to direct medical resource use. It is unlikely that the sign of the charge difference differs from the sign of a true cost difference, but magnitudes may differ. However, patterns of patient care and resource use may be different at other institutions. Thus, generalizability of results may be limited. A related limitation is that we are unable to model the potential for transplant center-specific effects, since all data came from one center.

Furthermore, our outcomes do not consider other important endpoints, such as enhanced health-related quality of life or productivity gains associated with avoidance of CMV disease, in either the short or long-run. This study therefore is clearly not from the societal perspective and is not an example of the US Panel on Cost Effectiveness 'reference case' [36]. As such, the study cannot be used for prioritization of resources in an entire health economy. However, the limited provider perspective is still valuable for informing clinical decision-making.

Additionally, a key source of uncertainty is not reflected in the analysis. We assume that we have correctly chosen functional forms and other elements of model specification in our example. There is always the potential, for example, that an omitted variable is correlated with selection (either into the treatment group, or even into the trial itself) and outcome.

Furthermore, we have estimated all of the equations separately and have thus made the strong assumption that error terms are not correlated across equations. If an unobservable variable affects both the probability of developing CMV infection or disease and also separately affected length of stay or cost, then our estimates could potentially be biased. For example, suppose that there is an unmeasured overall 'strength' variable that improves the body's ability to fight off CMV and also independently shortens recuperative times. Our estimate of the effect of CMV on length of stay and cost would be biased upward.[i]

Our choice of flat, independent priors, may also be challenged. By assigning such priors, we assume that the decision-maker has no previous knowledge about the effectiveness of GCV relative to ACV. A fully Bayesian analysis would have synthesized the existing literature on prophylactic treatment for CMV and incorporated that information in the priors.

Simulation and uncertainty in technology assessment

Econometric simulation provides a useful tool for analysis of cost-effectiveness data collected alongside clinical trials. It allows for counterfactual ('what-if') questions of potential interest to both

clinical decision-makers and broad policy-makers. For example, one might ask: How would outcomes in a group of trial participants have differed had they all been assigned to treatment (or control)? or Would outcomes be more favorable if treatment is restricted to high risk patients? Simulation affords greater flexibility in controlling for individual characteristics than simple comparisons of average outcomes between treatment and control groups. Even in a clinical trial context, where randomization and stratification are used to control for selection bias and groupwise comparability, a structural model may be desirable to parse out the effect of confounding covariates so that results may be extrapolated to other populations see (Table 3).

One must be careful not to treat the model as if its parameters are known. Uncertainty about parameters is often assessed in the current literature through one or two (or occasionally three) way deterministic sensitivity analysis. Such sensitivity analysis is a good check of a model's robustness and helps to answer the question: Are outcomes sensitive to a small set of parameters?

From a decision-maker's perspective, however, it might be even more interesting to ask: 'How much don't we know about the relationship between treatment, covariates and outcomes, and how much does that lack of knowledge limit our ability to infer desirability of one technology relative to another?' Analysis of the joint posterior distribution cost-effectiveness given the uncertainty of all parameters is necessary to answer that question.

Assessing global uncertainty deterministically requires covering all permutations of the upper and lower bounds of the confidence intervals for all parameters. Such an analysis for a simulation with J parameters requires 2^J calculations – which can quickly get out of hand. 'Probabilistic sensitivity analysis' overcomes this limitation by randomly drawing vectors from a specified joint distribution of parameters (see Doubilet *et al.* [37]). The distribution can be chosen *ad hoc*, for example, as a uniform distribution with upper and lower limits equal to $\pm$ 20% of the mean.

Ideally, probabilistic sensitivity analysis ought to reflect a decision-maker's state of knowledge in a systematic, data-driven way. For example, Manning and Mullahy [38] suggest that using observed data, one could estimate parameters and their variance-covariance matrices using maximum likelihood or other standard techniques and then draw randomly from the normal distributions implied by the calculated confidence intervals. In two recent analyses in the gastroenterology field, Pasta *et al.* [39] and Lord and Asante [40] used the nonparametric bootstrap, rather than relying on consistent parametric estimates, to generate the draws for probabilistic sensitivity analysis. Briggs *et al.* [41] discuss use of the nonparametric bootstrap to identify the confidence interval around the incremental cost effectiveness ratio.

One might reason that the nonparametric bootstrap, by placing no restriction on the functional form of the relationships between treatment, covariates and outcomes, would be more conservative than a parametric model [33]. However,

Table 3. Bayesian simulation expected incremental total charges and attributable risk reduction

Serostatus risk group	Posterior mean risk reduction (%)	Posterior probability of improved outcome (%)	Posterior mean incremental charges	Posterior probability of lower charges (%)	Posterior probability of GCV Dominance (%)
Outcome: any CMV infection					
Original risk groups	18.6	98.1	−$4 840	67.5	66.5
Assuming all R−/D+	7.1	66.2	−$12 252	80.1	55.6
Assuming all R+/D+	28.5	98.5	−$3 594	60.8	60.1
Assuming all D−	10.6	83.6	−$2 616	58.4	37.2
Outcome: CMV disease					
Original risk groups	7.2	85.7	−$4 840	67.5	59.4
Assuming all R−/D+	37.3	97.6	−$12 249	80.1	79.0
Assuming all R+/D+	0.8	54.6	−$3 597	60.8	37.2

through its reliance upon the 'plug-in' principle, the simple non-parametric bootstrap requires that the empirical distribution sufficiently describes the population distribution. The non-parametric bootstrap demands the very strong assumption that the sample being studied is both representative and independent and identically distributed [42].

However, we cannot conclude that structural modeling is superior because it does not require representativeness.[j] As mentioned above, one should be aware that the structural model instead requires the equally (or even more) strong assumption of correct specification. In other words, we assume that the only variables that systematically affect outcomes are contained within the model. Hence, so long as we believe this assumption, we can use the model to extrapolate predictions for any population (collection of covariates) of interest.

In this study, the non-parametric bootstrap results seemed less conservative than the structural model results (see Table 4). Though, strictly speaking, credible regions and confidence ellipsoids are not directly comparable [1], the area of the bootstrapped confidence ellipsoid is substantially smaller than the posterior credible region. The strong assumption of representativeness may have contributed to the greater certainty. By drawing from a relatively small pool of possible individual outcomes, the variability of differences in outcomes between groups is also likely to be small. The statistical model and simulation, on the other hand, allow for greater variability of potential individual outcomes.

The bootstrap analysis also shows a larger incremental reduction in charges associated with GCV than does the Bayesian simulation. The nonparametric bootstrap relies upon the randomization mechanism to equalize all other differences between the GCV and ACV groups that may affect outcomes, while a structural model (whether estimated by maximum likelihood, Bayesian or other techniques) relies upon randomization only to assure that treatment selection is exogenous. Of course, omitted variable bias may always present problems. However, if the model is correctly specified, and the parameter estimates are unbiased, then simulation based on that structural model can parse out just the differences in outcomes associated with treatment.

Why Bayesian analysis?

The debate on the nature of Bayesian versus classical analysis has gone on for decades. Much of this debate has centered on the notion that while classical statistics may be intuitive in the context of often-repeated experiments, interpretation may be difficult in less controlled or 'one-shot' analyses.

In the classical paradigm, the $q\%$ confidence interval represents the range which, if an experiment were repeated an arbitrary number of times, would contain the 'true' parameter $q\%$ of the time. In 1964, Aitchison [43] noted:

> While this statement is comforting to a statistician making repeated use of R [the tolerance region] to provide his customers with tolerance regions, its meaning for the individual customer is not at all easy to specify. The author has recently experienced the difficulty of trying to sell the frequentist approach to engineers whose work is crucially concerned with the problems of tolerance regions . . . The frequentist might then argue that the probabilistic statement, while not of direct application to the engineer's needs, is intended to give him comfort through the hope that his particular experiment is one of the lucky ones belonging to this proportion q in a population of hypothetical experiments.

In medical technology assessment, we often are faced with less-than-perfectly controlled experiments. Furthermore, one can make the case that technology adopters are less like experimentalists in a lab who can repeat experiments, and more like engineers who have to 'get it right the first time'

Table 4. Nonparametric bootstrap simulation expected incremental total charges and attributable risk reduction

	Base case risk reduction (%)	Percent of bootstrap density showing improved outcome (%)	Base case incremental total charges	Percent of bootstrap density showing lower charges (%)	Percent of bootstrap density showing GCV dominance (%)
Any CMV infection	20.3	99.3	−$13 039	94.2	93.7
CMV disease	9.5	93.1	−$13 039	94.3	88.6

when they choose which technology to adopt. The engineer's 'intolerance for ambiguity' discussed above may indeed be very similar to the clinician's fear of choosing the wrong treatment for an individual patient.

The Bayesian approach may therefore hold promise in making the statistical methods of health technology assessment more believable and accessible to clinicians. Eddy *et al.* [44] made early inroads in this area by using a Bayesian approach to meta-analysis. Parmigiani *et al.* [45] introduced the concept of using Bayesian resampling methods together with a (non-empirical) Markov simulation to help assess the global uncertainty over the cost-effectiveness ratio. Briggs [4] has also drawn a Bayesian linkage between Monte Carlo simulation and prior knowledge about population characteristics and chance node probabilities in decision-analytic models.

Recent work by Craig *et al.* [46] demonstrates the usefulness of MCMC in counterfactual simulation based on an observational study. By using MCMC techniques, the authors were able to sample from the joint posterior distribution of an empirical Markov model of the progression of diabetic retinopathy, accounting for potentially non-random assignment to treatment. The authors then used the MCMC draws in a counterfactual simulation of the disease's natural progression, i.e., what would have occurred had observed patients not received treatment. Our study builds upon their work, extending it into the realm of cost-effectiveness alongside clinical trials.

Heitjan *et al.* [47] applied Bayesian methods to sample from posterior joint distributions using three previously reported clinical trials as examples. Their approach provides visual representations of global uncertainty in the incremental cost-effectiveness plane using contour plots. Al and Van Hout [48] used conjugate priors to perform a Bayesian cost-effectiveness analysis of previously reported data from a clinical trial comparing stenting to balloon angioplasty. O'Hagan and Stevens developed a framework for use of MCMC in cost-effectiveness analysis alongside clinical trials [49]. While their framework is based on the same principles as ours, their models for cost and effectiveness are not conditional. Rather, they identify a set of parameters related only to group average cost and effectiveness and do not examine counterfactual, individual simulations. Our study extends the concepts of these works into structural empirical models which are more amenable to counterfactual simulation.

In future analysis, one might consider the decision to adopt a new technology as a complex, empirical, Bayesian hypothesis test, with a fully specified loss function. Such a loss function might even value the distribution of outcomes over individuals (rather than simply the average outcome) as a factor important to decision-makers [Mullahy J. Which cost-effectiveness ratio? Evaluating health policies and medical technologies in stochastic contexts, University of Wisconsin, Madison; unpublished manuscript]. The MCMC estimation and simulation approach can be combined with such loss functions to yield very useful and powerful statements of the form: 'Given prior beliefs $p(\theta)$, observed outcomes Y and covariates X under technologies A and B, and model $Y = F(X, \theta)$, a decision-maker with loss function $L(Y)$ would have $Z\%$ confidence that technology A (or B) is preferred.'

Conclusion

Our study suggests that a decision-maker observing information from a recent clinical trial on GCV relative to ACV prophylaxis for CMV post liver transplant would be confident that GCV is associated with lower total charges and improved CMV disease outcomes for high risk R−/D+ cases. The decision-maker might have somewhat less confidence of the relative desirability of GCV in medium risk R+/D+ cases.

This analysis demonstrates the feasibility of using MCMC-generated draws from the simulated posterior distribution of parameters from an empirical model in analyzing the joint distribution of incremental cost and effectiveness in a simulation. The impact of parametric uncertainty remaining after observing the data is reflected in the simulations conducted using these MCMC draws. The combination of Bayesian estimation of structural models linking treatment, covariates and outcomes with counterfactual simulation, may be a useful tool for assessing the impact of parametric uncertainty in cost-effectiveness analysis.

Acknowledgements

This work was funded in part by a grant (R01-DK-34238) from the National Institute of Diabetes &

Digestive & Kidney Diseases, National Institutes of Health. The authors thank an anonymous reviewer, John Forbes, Kirsten Long, John Mullahy, Jim Naessens, Niels Neymark, Dan Sargent, Chris Seplaki and participants in the 10th European Workshop on Econometrics and Health Economics for helpful comments. All errors are our own.

Notes

a. An earlier version of the model used a generalized ordered probit framework [22] with separate hurdles for infection and disease. That model required the hurdle for disease to be greater than the hurdle for infection for all individuals in all MCMC iterations. Because this condition was violated frequently, we chose to use the sequential probit specification. Results from an ordered probit specification in which we forced the ordering condition to hold by truncation did not differ substantially from the sequential probit.
b. Because an earlier version of the model used a generalized ordinal probit framework, we were able to include D− and interactions. However, as expected, the model indicated that D− serostatus raised the hurdle for CMV disease by an arbitrarily large amount. An anonymous reviewer noted that estimating an arbitrarily large coefficient using MCMC may negatively affect chain convergence.
c. Note that we re-parameterize CMV infection when it is used as a covariate so that it equals zero when CMV disease equals one. Therefore, as a covariate, CMV infection should be interpreted as 'CMV infection only,' i.e., infection that did not progress to full-blown CMV disease. As a dependent variable, CMV infection should be interpreted as 'Any CMV infection,' i.e., *infection whether or not it progresses to CMV disease*.
d. Note that the negative sign in the summed terms for $\Delta\hat{Y}_j^{\mathrm{C}}$ and $\Delta\hat{Y}_j^{\mathrm{D}}$ are necessary to express the outcome in terms of a benefit (infection and disease avoidance) rather than a harm (infection and disease rate).
e. An earlier version of the model treated the scale parameters as unknown. Though the model successfully estimated the parameters, we observed a high level of correlation between the scale parameter and the associated constant term in the shape parameter. Because this correlation may negatively affect convergence, we chose to estimate the scale parameters separately.
f. Code for the MCMC estimation and simulations are available from the author upon request.
g. As noted by an anonymous reviewer, expected incremental total charges should be the same regardless of whether we plot against infection or disease. Because results reported in the table are the expected value of the bivariate distribution determined by numerical integration, they may differ by a few dollars due to discretization error.
h. If there were a strong correlation between serostatus and other covariates, such as age and gender, then changing only the serostatus variable might be a strong assumption. An alternative, more conservative, approach suggested by an anonymous reviewer would be to use only those cases that actually were high risk. In our study, there is little correlation between serostatus and other explanatory variables. Results from the more conservative simulation (available from the authors on request) were very similar to results as reported. Researchers using our method may wish to pay close attention to whether such correlations exist.
i. We thank an anonymous reviewer for providing this example.
j. We thank an anonymous reviewer for drawing attention to this point.

References

1. Manning WG, Fryback DG, Weinstein MC. Reflecting uncertainty in cost-effectiveness analysis. In *Cost-effectiveness in Health and Medicine*, Gold MR, Siegel JE, Russell LB, Weinstein MC (eds). Oxford University Press: Oxford, 1996.
2. Briggs AH, Sculpher M. Sensitivity analysis in economic evaluation: a review of published studies. *Health Econ* 1995; **4**: 359–361.
3. Milne R, Nuitjen M, Weinstein M, Zhou X. Decision analytic modelling in the economic evaluation of health technologies: a consensus statement. (Consensus statement and background papers from the Consensus Conference on Guidelines on Economic Modelling in Health Technology Assessment held in Sheffield, England, 22–23 April, 1999). *Pharmacoeconomics* 2000; **17**: 443–444.
4. Briggs AH. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 2000; **17**: 479–500.
5. Spiegelhalter DJ, Myles JP, Jones DR, Abrans KR. An introduction to Bayesian methods in health technology assessment. *Br Med J* 1999; **319**: 508–512.
6. El Gamal MA, Grether DM. Are people Bayesian? Uncovering behavioral strategies. *J Am Stat Assoc*, 1995; **90**: 1137–1145.
7. von Neumann J, Morgenstern O. *Theory of Games and Economic Behavior*. Princeton University Press: Princeton, NJ, 1944.
8. Harsanyi J. Games with incomplete information played by Bayesian players. *Manage Sci* 1967; **14**: 159–182.
9. Geweke J. Bayesian econometrics and forecasting. *J Econom* 2001; **100**: 11–15.
10. Malakoff D. Bayes offers a 'new' way to make sense of numbers. *Science* 1999; **286**: 1460–1464.

11. Wiesner RH, Marin E, Porayko MK, Steers JL, Krom RA, Paya CV. Advanced in the diagnosis and treatment and and prevention of cytomegalovirus infections after liver transplantation. *Gastroenterol Clin North Am* 1993; **22**(2): 351–366.
12. Rook AH. Interactions of cytomegalovirus with the human immune system. *Rev Infect Dis* 1988; **10**(Suppl 3): S460–S467.
13. Paya CV, Wiesner RH, Hermans PE, *et al.* Risk factors for cytomegalovirus and severe bacterial infections following liver transplantation: a prospective multivariate time-dependent analysis. *J Hepatol* 1993; **18**(2): 185–195.
14. Paya CV. Opportunistic infections: risk factor or virus-associated complication? *Transplant Proc* 1994; **12**.
15. Patel R, Paya CV. Infection in solid-organ transplant recipients. *Clin Microbiol Rev* 1997; **10**(1): 86–124.
16. Patel R, Snydman DR, Rubin RH, Ho M, *et al.* Cytomegalovirus prophylaxis in solid organ transplant recipients. *Transplantation* 1996; **61**(9): 1279–1289.
17. Winston DJ, Wirin D, Shaked A, Busuttil RW. Randomised comparison of ganciclovir and high-dose acyclovir for long-term cytomegalovirus prophylaxis in liver-transplant recipients. *Lancet*, 1995; **346**(8967): 69–74.
18. Gane E, Saliba F, Valdecasas GJ, O'Grady J, Pescovitz MD, Lyman S, Robinson CA. Randomised trial of efficacy and safety of oral ganciclovir in the prevention of cytomegalovirus disease in liver-transplant recipients. *Lancet* 1997; **350**(9093): 1729–1733.
19. Badley AD, Seaberg EC, Porayko MK, *et al.* Prophylaxis of cytomegalovirus infection in liver transplantation: A randomized trial comparing a combination of ganciclovir and acyclovir to acyclovir. *Transplantation* 1997; **64**(1): 66–73.
20. Kim WR, Badley AD, Wiesner RH, *et al.* The economic impact of cytomegalovirus infection after liver transplantation. *Transplantation* 2000; **69**(3): 357–361.
21. Maddala GS. *Limited-dependent and Qualitative Variables in Econometrics*, Chapter 2. Econometric Society Monographs. Cambridge University Press: Cambridge, 1983.
22. Terza JV. Ordinal probit: a generalization. *Commun Stat – Theory Methods* 1985; **14**: 1–11.
23. Cameron AC, Trivedi PK. Econometric models based on count data: comparisons and applications of some estimators and tests. *J Appl Econom* 1986; **1**: 29–53.
24. Blough DK, Ramsey SD. Using generalized linear models to assess medical care costs. *Health Services Outcomes Res Methodol* 2000; **1**: 185–202.
25. Silverman BW. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall: London, 1986.
26. Spiegelhalter D, Thomas A, Best N. *WinBUGS Version 1.3 User Manual*, April 2000.
27. Geweke J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In *Bayesian Statistics 4*, Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds). Oxford University Press: Oxford, 1992.
28. Brooks S, Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graphical Stat* 1998; **7**: 434–455.
29. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Stat Sci* 1992; **7**: 457–511.
30. Raftery AL, Lewis S. How many iterations in the Gibbs sampler? In *Bayesian Statistics 4*, Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds). Oxford University Press, Oxford.
31. Heidelberger P, Welch P. Simulation run length control in the presence of an initial transient. *Oper Res* 1983; **31**: 1109–1144.
32. Smith BJ. *Bayesian Output Analysis Program (BOA) Version 1.0.0 User's Manual*, June 2001.
33. Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. Chapman & Hall: London, 1993.
34. Hall P. On the bootstrap and likelihood-based confidence regions. *Biometrika* 1987; **74**: 481–493.
35. Finkler SA. The distinction between costs and charges. *Ann Internal Med* 1982; **96**: 102–109.
36. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-effectiveness in Health and Medicine*. Oxford University Press: Oxford, 1996.
37. Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation. A practical approach. *Med Decision Making* 1985; **5**: 157–177.
38. Mullahy J, Manning WG. Statistical issues in cost-effectiveness analysis. In *Valuing Health Care: Costs, Benefits and Effectiveness of Pharmaceuticals and Other Medical Technologies*, Sloan F (ed). Cambridge University Press: New York, 1994.
39. Pasta DJ, Taylor JL, Henning JM. Probabilistic sensitivity analysis incorporating the bootstrap: an example comparing treatments for the eradication of *helicobacter pylori*. *Med Decision Making* 1999; **19**: 353–363.
40. Lord J, Asante MA. Estimating uncertainty ranges for costs by the bootstrap procedure combined with probabilistic sensitivity analysis. *Health Econ* 1999; **8**: 323–333.
41. Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: A non-parametric approach to confidence interval estimation. *Health Econ* 1997; **6**(4): 327–340.
42. Young GA. Bootstrap: more than a stab in the dark? *Stat Sci* 1994; **9**: 382–395.
43. Aitchison J. Two papers on the comparison of Bayesian and frequentist approaches to statistical problems of prediction: Bayesian tolerance regions.

J Roy Stat Soc Ser B (Methodological) 1964; **26**: 161–175.
44. Eddy DM, Hasselblad V, Schachter RD. *Meta-analysis by the Confidence Profile Method: The Statistical Synthesis of Evidence*. Academic Press: Boston, 1992.
45. Parmigiani G, Samsa GP, Ancukiewicz M, Lipscomb J, Hasselblad V, Matchar DB. Assessing uncertainty in cost-effectiveness analyses: application to a complex decision model. *Med Decision Making* 1997; **17**: 390–401.
46. Craig BA, Fryback DG, Klein R, Klein BEK. A Bayesian approach to modelling the natural history of a chronic condition from observations with intervention. *Stat Med* 1999; **18**: 1355–1371.
47. Heitjan DF, Moskowitz AJ, Whang W. Bayesian estimation of cost-effectiveness ratios from clinical trials. *Health Econ* 1999; **8**: 191–201.
48. Al MJ, Van Hout BA. A Bayesian approach to economic analyses of clinical trials: the case of stenting versus balloon angioplasty. *Health Econ* 2000; **9**: 599–609.
49. O'Hagan A, Stevens JW. A framework for cost-effectiveness analysis from clinical trial data. *Health Econ* 2001; **10**: 303–315.

