where $\alpha = \log[\hat{\theta}_K/\hat{\theta}_{K-1}]$ and K is sufficiently large. Thus, the most recent infection rates are estimated by extrapolating an exponential function.

Another choice of the penalty function is

$$J = \sum[\theta_i - 2\theta_{i+1} + \theta_{i+2}]^2.$$

Then, estimates of recent infection rates are approximately

$$\hat{\theta}_{K+j} \approx \hat{\theta}_K + j\delta, \tag{4}$$

where $\delta = [\hat{\theta}_K - \hat{\theta}_{K-1}]$. Thus the most recent infection rates are estimated by extrapolating a linear function.

The piecewise constant step function model for $I(s)$ that was used in the early work on backcalculation assumes that infection rates are constant over intervals. Simulation studies of Rosenberg, Gail and Pee (1991) suggest choosing a last step of 4 to 4.5 years in length. Recent infection rates under this model are estimated by

$$\hat{\theta}_{K+j} = \hat{\theta}_K. \tag{5}$$

Estimates of recent infection rates obtained by backcalculation are essentially extrapolations of trends in $I(s)$. Equations (3) through (5) are different examples of mathematical functions that have been used for such extrapolations and result from different choices of the roughness penalties or parametric assumptions on $I(s)$. Estimates of recent infection rates based on backcalculation are highly dependent on the degree of smoothing λ, the penalty J and the parametric model for $I(s)$.

Appreciable improvements in our ability to reconstruct infection rates may come, not from alternative smoothing procedures or parametric models but rather from obtaining empirical data on recent infection rates.

4. FUTURE PROSPECTS FOR FORECASTING AND RECONSTRUCTING THE AIDS EPIDEMIC

Early in the AIDS epidemic, the only reliable data for monitoring the epidemic was AIDS-incidence data. Since the development of the HIV antibody test in the mid-1980s, numerous surveys of HIV seroprevalence have been conducted. Infection rates have also been directly estimated in several cohorts. Our ability to reconstruct infection rates may drastically improve by incorporating external information about recent infection rates and HIV seroprevalence derived from cohort studies and cross-sectional surveys.

There is considerable underreporting of AIDS cases to national and regional AIDS surveillance registries in developing countries, especially in Africa. Projections of the course of the epidemic in developing countries must rely more on HIV seroprevalence and seroincidence surveys than on AIDS-incidence data. While U.S. AIDS-incidence data are relatively complete, more reliable assessments of the scope of the epidemic may be obtained by considering HIV-seroprevalence and HIV-seroincidence data as well. For example, extensive HIV-seroprevalence surveys among childbearing women are extraordinarily useful for forecasting the future numbers of pediatric AIDS cases. Statistical approaches that combine data from multiple sources (e.g., AIDS-incidence data, HIV-seroprevalence and -seroincidence surveys, incubation distributions) are promising and may considerably improve the accuracy of assessments of the scope of the epidemic.

Comment: Assessing Uncertainty in Backprojection

John B. Carlin and Andrew Gelman

Bacchetti, Segal and Jewell are to be congratulated for providing not only a comprehensive review of an important problem in applied statistics but also for introducing a number of new ideas that should have a practical impact on understanding the course of the HIV epidemic. On a semantic detail, we wonder why the authors (and others) have adopted the term "backcalculation," rather than "backprojection," which seems to carry a more appropriate connotation of uncertain inference (as well as being shorter!).

The authors rightly emphasize the sensitivity of backprojection estimates to assumptions about the incubation distribution, but they seem strangely reluc-

John B. Carlin is Deputy Head, Clinical Epidemiology and Biostatistics Unit, Royal Children's Hospital, Melbourne, Victoria 3052, Australia. Andrew Gelman is Assistant Professor, Department of Statistics, University of California, Berkeley, California 94720.

tant to exercise any judgement about the choices available here. Their comment that attempts to model secular changes in the incubation distribution due to the effects of treatment and so on "run counter to the original spirit of backcalculation" seems a little peculiar in the sense that backcalculation has only avoided such efforts by making very strong and unverifiable assumptions of simple mathematical forms for this distribution. We feel that sensitivity analyses (or more formal Bayesian analyses) should be guided as far as possible by informed judgements involving a synthesis of as much of the available evidence as possible. For example, some forms for the incubation distribution might be judged to be more appropriate than others for each of the different subgroups analysed in the paper. It would be helpful to see in the rejoinder a plot of the four distribution functions they have used (especially as three of the four come from yet-to-be published work). In analyses of Australian data (Becker et al., 1993), similar sensitivity to assumptions about incubation is observed, but the range of results is interpreted in the framework of a general model that allows for nonstationarity in explicitly defined ways, related to assumptions about the effect of treatment and trends in treatment practice (Becker and Motika, 1992). As well as allowing for treatment effects, one of the underlying models used in these analyses is a log-logistic survival function, which has a bounded hazard function that rises rapidly and then declines slowly over a long period.

The authors make an important new contribution in allowing the incubation distribution and reporting delay distribution to change over time in a nonparametric, data-determined fashion. Their method appears to be a useful exploratory tool in searching for patterns over time, but it would be attractive to be able to relate the results back to an interpretable model that might have other support, external to the data. Of course, a major difficulty is the lack of identifiability of patterns of secular change as distinct from aspects of "genuine" (i.e., not treatment-influenced) natural history.

The authors suggest that Bayesian approaches may be useful in quantifying uncertainty in backprojections of HIV incidence. In principle such an approach is very attractive because of the conceptually simple framework within which all sources of uncertainty can be jointly accommodated. For example, a simple approach to the problem of alternative incubation models would be to use a discrete prior distribution over a small set of reasonable alternatives. Bayesian analysis would then allow the data to provide an update to the assumed distribution of models (although information in the data might be rather limited). Similarly, a Bayesian approach automatically averages uncertainty with respect to all unknown parameters including, for example, the authors' λ_θ, which determines the degree of smoothing in the point estimates of θ. Another advantage, in principle, is the possibility of modeling subgroups within a consistent overall framework.

Bayesian analysis requires, however, that all model-based input to the analysis be expressed in the form of (prior) probability distributions, rather than, for example, nonparametric assumptions accompanied by more or less ad hoc estimation procedures. Although model specifications may end up being almost as difficult to justify as the assumptions behind the authors' approach, the Bayesian framework would seem to provide a more satisfying avenue for exploring "what-if" sensitivity analyses.

Putting aside the question of uncertainty due to model specifications, we are puzzled by the relatively small role apparently played by sampling variability in the reported estimates. In particular, estimates of θ_i for i near n should have large uncertainty (even conditional on the model, including the incubation distribution), since under all reasonable models, progression to AIDS in the first year or two of infection is extremely unlikely (in the notation of the article, $D_{i\,i+d}$ is very small for small d), implying that the data y contain almost no information about these recent infection rates. Put very simply, infection rates could jump substantially in the last year or two, but this would make little difference to the observed y_i's. The simulation-based estimates of uncertainty do not seem to reflect adequately this source of variability, perhaps because when the data are uninformative, smoothing takes over. This provides another motivation for a Bayesian approach, where the probability calculus guarantees that posterior intervals reflect *all* uncertainty, both sampling-based and model-dependent, given the data.

We take this opportunity to report briefly on some preliminary work on implementing a Bayesian approach. The model we have currently programmed assumes that the incubation distribution is known, which is clearly unrealistic, as Bacchetti, Segal and Jewell emphasize, but nevertheless provides a useful starting point. It also ignores the problems of reporting delay: that is, we suppose that $n^* - n$ is sufficiently great that $R_j = 1$ for all j. Our approach is closely related to the authors' penalized-likelihood approach, except that the penalty function is interpreted as a prior distribution and we endeavor to compute full posterior distributions rather than modes with approximate standard errors. It is certainly a point of major agreement that the Bayesian *or* penalized-likelihood approaches represent an improvement over the use of parametrically specified forms for θ, including the convenient but implausible step-function model (e.g., Rosenberg and Gail, 1991).

We have experimented with Bayesian analysis based

on a simple random walk prior for the unknown infection rates in the square-root scale. Adapting the authors' notation (except that we find it more convenient to index the observed data from $i = 1$ rather than from $i = 0$), let $\gamma_i = \sqrt{\theta_i}$ and assume a prior distribution such that

$$\gamma_i | \{\gamma_j : j < i\}; \sigma^2 \sim N(\gamma_{i-1}, \sigma_i^2),$$

where $\gamma_0 = 0$, and a simple but fairly general form is allowed for the variances, σ_i^2, by putting $\sigma_i^2 = \sigma^2 s_i$, where s_i is given a fixed form and σ^2 is to be estimated. In practice, so far, we have let s_i be constant. The prior distribution is then completed by specifying a prior density for σ^2, a convenient and nonrestrictive choice being the conjugate inverse-gamma density. The random walk prior is somewhat arbitrary but provides a minimal structure that is nonstationary and hence noninformative with respect to the level of γ_i. The square-root scale has some appeal within a Poisson sampling framework, as we see below.

Conditioning on the variance hyperparameter σ^2, and assuming a Poisson sampling model for the observed y_i, one obtains the following joint log-posterior density for $\Gamma = \{\gamma_i : i = 1, \ldots, n\}$ (up to a constant):

$$\log p(\Gamma | y, D, \sigma^2) = \sum_{i=1}^{n} \left[y_i \log\left(\sum_{k=1}^{i} \gamma_k^2 D_{ki} \right) - \sum_{k=1}^{i} \gamma_k^2 D_{ki} \right] - \sum_{i=1}^{n} \left(\log \sigma_i^2 + \frac{1}{2\sigma_i^2} (\gamma_i - \gamma_{i-1})^2 \right).$$

The first term in this expression is the same as the authors' $\log(L_m)$ in Expression (3), and the contribution from the prior is clearly analogous to their penalty function, except that they use the log scale instead of the square root and have taken second differences rather than first. The first choice (the log scale) seems eminently reasonable, especially because it ensures nonnegativity of the θ_i, but it would be interesting to know why the authors choose the second difference for their penalty function. Including in the above expression a term for the prior density of σ^2 would give an unconditional joint posterior density for Γ and σ^2, but direct computation based on this expression does not seem possible.

Instead, we turn to the very flexible computational framework for Bayesian analysis that is provided by the Gibbs sampler (e.g., see Gelman, 1992). We have implemented a version that uses the same data augmentation as Bacchetti, Segal and Jewell use for their EM algorithm. Let $\mathbf{x}$ represent the array $\{x_{ij} = \Sigma_k x_{ijk}\}$ (since we are ignoring reporting delays). Then to implement the Gibbs sampler we need to draw in turn from the conditional distributions of *each* of the unknown quantities, $\mathbf{x}$, Γ, σ^2, given currently drawn values of each of the other unknowns, as well as the data y (which of course remain fixed through the whole process). If a discrete prior distribution over alternative incubation models were introduced, this could also be readily incorporated into the analysis. Under relatively mild conditions this procedure converges to produce values from the joint posterior distribution, from which the marginal distribution of Γ (or θ) is of primary interest. Convergence may be conveniently monitored using the methods described by Gelman and Rubin (1992).

Of the three conditional distributions required, two are relatively straightforward to simulate. First, the conditional distribution of $\mathbf{x}$, given everything else, breaks into a product of n multinomial distributions each based on an observed y_i. Second, given the normal prior density for Γ, the update for σ^2 has a conjugate form under the inverse-gamma prior.

The only difficulty arises in simulating from the conditional distribution of Γ or θ given $\mathbf{x}$ and σ^2 (note that, given $\mathbf{x}$, we have conditional independence with respect to the data y). For these simulations we have employed a two-step procedure, first approximating the Poisson (complete data) likelihood given as the authors' Expression (4) with a normal distribution for $z_i = \sqrt{x_i}$, and then correcting this approximation using the generalized Metropolis algorithm (Metropolis et al., 1953; Gelman, 1992). With the normal approximation and normal prior in the square root, a conjugate update of Γ is possible, either performing the sampling one time point at a time (i.e., progressively updating $\gamma_1, \gamma_2, \gamma_3, \ldots$, etc.) or using time-series methods related to the Kalman filter to update the entire Γ sequence at once. The Metropolis correction method involves obtaining trial values using the approximate distribution and comparing importance ratios between trial value and current value to determine whether the update should be accepted.

A simple modification of the Gibbs sampler produces estimates of posterior modes, which should be essentially the same as the authors' maximum penalized-likelihood estimates, if σ^2 is held fixed. Our program is currently very successful at mode finding, despite the Poisson-normal approximation, but there are difficulties in the sampling of Γ, apparently because the normal approximation is sometimes so poor that Metropolis updating takes place with very low probability. Different computational strategies may be required, perhaps working with the authors' (implicit) log-normal prior in place of our square-root specification.

ACKNOWLEDGMENT

This work was supported by an Australian Commonwealth AIDS Research Grant.