mode estimates of the relevant parameters. Would that not be a simpler approach than Gibbs sampling in many cases?

(b) Including HIV test reports is important, but, as the authors say, it is not easy. With AIDS diagnoses being postponed by new treatments, will not the use of HIV test data become increasingly important in assessing the further development of HIV incidence and prevalence?

(c) What do the authors think about the future of back-calculation? With new antiretroviral treatments producing a dramatic decline in the incidence of AIDS, what role is left for statistical methods of prediction? Could the great efforts that have gone into developing back-calculation be applied in other fields? In other words, what is the generalizability of this methodology?

Brian Dangerfield (*University of Salford*)

It is useful to be presented with a full description of developments in back-calculation (BC) techniques in so far as they are employed as part of the preparation of official projections of new AIDS cases in England and Wales, the most recent of which were published in 1996 (Day, 1996). Since the first official projections were published in 1988 (Cox, 1988) the range of methodologies underpinning the various projections has been reduced somewhat such that now BC has a pre-eminent role.

However, it is not immediately apparent that BC is a superior methodology for making short-term projections of AIDS diagnoses than, say, a carefully constructed deterministic transmission model which is fitted to existing AIDS data. There are uncertainties in using either approach (admittedly reduced in the case of BC by the developments described in the paper presented) but a return to the pluralism of the Cox report would be welcome. Our 1993 projections, made using data to the end of 1992, and for the UK (not England and Wales) together with four other European countries (Dangerfield and Roberts, 1996) highlighted even then an incipient peak in the incidence of new AIDS cases, although an *ex post* appraisal (Roberts and Dangerfield, 1998) of these projections has revealed a systematic overestimation almost certainly caused by the lack of inclusion of treatment effects in the 1993 version of the model.

Methodologically speaking, the proponents of BC, employed for projecting new cases of disease with a long incubation time, are on stronger ground when they can reasonably assert, *a priori*, that the temporal profile of infection has run its course. Thus Anderson *et al.* (1996) employed BC techniques in projecting the spread of bovine spongiform encephalopathy in cattle. In this case, the incidence of new infections via contaminated feedstuffs could be reasonably thought to be 'negligible' by mid-1994, directly resulting from the feed ban introduced in mid-1988. This adjective cannot be used to describe the current incidence of HIV.

The authors model pre-AIDS treatment in a manner analogous to extending the mean sojourn time in the asymptomatic phase of the incubation period in a transmission model. Their work does not extend to including the recent extremely encouraging effects arising through the introduction of triple-combination antiretroviral therapy from mid-1996. One clinician has described this as producing the greatest turnaround in prognosis that he has seen in 25 years in medicine (Carpenter, 1997). The unanswered question at present for the Government, public health officials and modellers alike is if and when the dam of cases on treatment effectively bursts.

Andrew Gelman (*Columbia University, New York*) **and John Carlin** (*Royal Children's Hospital, Melbourne*)

The authors have succeeded in using complex hierarchical modelling to combine information from many sources to estimate HIV and AIDS incidence rates. Their thoroughness and perseverance with such a complex modelling exercise is very impressive. As the complexity of a modelling exercise increases, it does, however, raise more critically the question of judging the credibility of the final results. The authors have performed fairly extensive sensitivity analyses by varying what they believe are the key elements of the model and exploring the resulting changes in conclusions. They have not, however, examined the related but distinct issue of model fit. It is possible, at least in principle, that none of the models used actually provides a very good fit to the data (in which case the implications of the models in terms of projections might not be trustworthy).

In particular, perhaps the authors can simulate replicated data sets from the posterior predictive distribution under their model (as in Gelman *et al.* (1996) or by using related methods, e.g. as described in Gelfand *et al.* (1992) and Weiss (1994)) and compare these with various data summaries, including but not limited to Figs 1 and 4 and graphical summaries of Tables 1 and 2. More specific realized

discrepancies of interest might summarize individual information estimated from the model such as times between infection and diagnosis.

The question of model fit has concerned us because it appeared to be a key difficulty in our own much more limited attempt at implementing Bayesian back-projection (Carlin and Gelman, 1993). One difference in the models that we were using is that they were based on estimating the HIV infection intensity in monthly or quarterly, rather than annual, units. The authors' choice of a stepwise constant form based on annual units seems a rather coarse representation of the process. In other respects, however, the model presented here is much more sophisticated than our attempt, and we have every reason to expect that it fits the data adequately. But we believe that the authors' substantive recommendations would be more defensible if the model fit were examined systematically.

Another question is the extent to which the results are influenced by the two seroprevalence estimates. The authors rely on estimated standard errors associated with the surveys but other possible sources of error from these surveys, such as sampling biases, do not appear to have been considered.

J. F. Lawless (*University of Waterloo*)

The authors are to be commended for bringing such a comprehensive Bayesian analysis to bear on the important problems of HIV incidence and AIDS forecasting. The comparison of analyses under various assumptions about the key ingredients of back-calculation (i.e. the HIV incidence function, the incubation or progression model and the reporting-delay model) provide insight concerning their effects on the estimation of the incidence of HIV and AIDS and give a realistic picture of uncertainty. However, the combination of survey seroprevalence data and AIDS surveillance data without an initial assessment of bias and compatibility is potentially worrisome.

The overall picture of the epidemic and estimated HIV incidence curves are similar to those from other recent studies. Major improvements in estimation must of course rely on improved data. It is interesting that the incubation distribution models used in back-calculation have been virtually the same for the past 9 or 10 years. Studies involving the long-term follow-up of HIV-infected individuals and information on the effects of treatment and on the extent to which different groups of people have access to health care have not led to significant improvements in the population incubation model; in particular, there is still considerable uncertainty about the progression to AIDS for a substantial proportion of people infected as far back as 10–20 years.

Reporting delays also continue to pose problems. It is difficult to provide satisfactory adjustments when the delay process is not stable over time. Hesselager and Witting (1988) and Lawless (1994) have discussed ways of minimizing the dependence on long-term stability assumptions, but the real solution lies in improving the reporting process.

Ian Marschner (*Harvard School of Public Health, Boston*)

This paper provides a nice presentation of a comprehensive back-calculation framework. The inclusion of seroprevalence surveys is a welcome addition to back-calculation analyses. The authors note the potential to include information on HIV tests for individuals with a diagnosis of AIDS. An important extension is the use of such information for individuals without diagnoses of AIDS. In many surveillance systems new HIV positive test results are reported. Such data potentially provide a large amount of additional information about the incidence of HIV because the number of AIDS-free HIV positive individuals is typically a large proportion of the total number infected. The authors correctly note that it is invalid to include HIV test data in the form of a simple inequality for the infection time. Although the beta model for the HIV test times was clearly intended only to illustrate a point, it is important to note that in practice models of the testing process will need to be more complex. Such models will need to account for a dependence on both calendar time and time since infection, for example by using a model such as

$$\Pr(\text{test at time } x | \text{infection at time } t \leqslant x \text{ and no previous positive test}) = 1 - \exp(-a_x b_{t-x}).$$

Although the framework presented is flexible and comprehensive, I question the claim by the authors that the continuous time Bayesian framework has a special advantage over previous formulations, with respect to the inclusion of individual-specific information such as HIV testing history. Previous methods using finely partitioned discrete time models with smoothing have allowed a convenient incorporation of information such as age at diagnosis and risk group. Furthermore, methods under current development in an EM framework with smoothing (building on the work of Marschner (1994)) allow a