2003

Estimation of a failure time distribution based on imperfect diagnostic tests

Raji Balasubramanian, *University of Massachusetts - Amherst*
Stephen W Lagakos, *Harvard School of Public Health*

Biometrika (2003), **90,** 1, *pp.* 171–182

Printed in Great Britain

Estimation of a failure time distribution based on imperfect diagnostic tests

By R. BALASUBRAMANIAN and S. W. LAGAKOS

Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts 02115, U.S.A.

rbalasub@biostat.harvard.edu lagakos@biostat.harvard.edu

Summary

Sequentially-administered diagnostic tests used to determine the occurrence of a silent event are sometimes subject to error, leading to false positive and false negative test results. In such cases, standard methods for interval censored data do not give valid estimates of the distribution of the time to the event. We present methods for estimating the distribution of the time to the event that account for multiple types of imperfect diagnostic test, as well as differing periods at risk. We illustrate the methods with simulated data and results from a clinical trial for the prevention of mother-to-infant transmission of HIV in Tanzania.

Some key words: Diagnostic test; Interval censored data; Panel data; Time-to-event methods.

1. Introduction

The onset of some chronic conditions or diseases are asymptomatic and can only be detected by diagnostic tests which indicate whether or not the event has occurred. Observations of this type are sometimes referred to as panel data. When the diagnostic test is perfect and given once, the time to the event is either right- or left-censored, depending on whether the test is negative or positive. When a perfect diagnostic test is given sequentially at different time points in the same individual, the time until the event is interval censored, and statistical methods have been developed to estimate this distribution or to assess factors that may be associated with it (Turnbull, 1976).

In some settings, the diagnostic tests used to assess the occurrence of the event are subject to error, yielding either false positive or false negative results. Examples include liver function tests and biopsies to detect liver disorders (Martin & Friedman, 1998) and DNA PCR tests to detect HIV infection. For the latter, infection is the unobserved biological event of the HIV integrating into the host's RNA/DNA, and test sensitivity is a function of how soon the test is administered after infection (Dunn et al., 2000). Here standard methods for interval censored data are not in general valid. An added complication in some applications is that an individual may only be at risk for the event at certain times; an example is an infant's risk for HIV infection while being breast-fed by an infected mother.

Previous work related to this setting includes accounting for imperfect diagnostic tests in the estimation of the probability that an event has occurred (Dunn & Ades, 1996; Tsai et al., 1994), as opposed to estimation of the distribution of time until the event, or has not allowed imperfect diagnostic tests. Hughes & Richardson (2000) estimate the distribution of time of vertical transmission of hiv and allow the diagnostic test given at the time of birth to have imperfect specificity. Balasubramanian & Lagakos (2001) allow imperfect test specificity but consider a specific type of test sensitivity and consider only settings where the diagnostic test can be given following the event of interest. Neither method allows differing exposure periods.

In this paper, we present statistical methods for estimating the distribution of the time until an event whose occurrence is assessed by sequentially-administered and imperfect diagnostic tests, in settings where individuals can also have differing periods of exposure to the forces that place them at risk for the event. In § 2, we present notation used in the paper and discuss certain assumptions that are made in order to develop the likelihood. In § 3, we develop the likelihood for special forms of the sensitivity function of the diagnostic tests and discuss identifiability and parameter estimation. In § 4, we illustrate the methods using data from a study of vitamin supplements for preventing mother-to-infant hiv-1 transmission in Tanzania (Fawzi et al., 1998).

2. Notation and assumptions

Let T refer to a random variable denoting the time until an event of interest for an individual and suppose that the distribution of T depends upon some binary exposure process, $E(t)=1[t \leqslant \tau_E]$, where $E(t)=1$ or $E(t)=0$, indicates that the individual is, or is not, exposed at time t. For example, in the setting of transmission of hiv from a mother to her foetus or newborn, τ_E might denote the age at which an infant is weaned, in which case $E(t)$ would equal one during pregnancy and while the infant is breast-fed. We assume that the joint distribution of (T, τ_E), as specified by their cause-specific hazard functions, is

$$\begin{aligned}\lambda_T\{t \mid E(u), 0 \leqslant u \leqslant t\} &= \lim_{h \to 0} \frac{1}{h} \operatorname{pr}\{T < t+h \mid T \geqslant t, E(u), 0 \leq u \leqslant t\} \\ &= \lambda(t)E(t)\end{aligned}$$

and

$$\lambda_E(t) = \lim_{h \to 0} \frac{1}{h} \operatorname{pr}(\tau_E < t+h \mid \tau_E \geqslant t, 1[T \leqslant u], 0 \leqslant u \leqslant t),$$

for some functions $\lambda(.)$ and $\lambda_E(.)$; that is, the risk of the event occurring at time t equals zero when $E(t)=0$ and is $\lambda(t)$ when $E(t)=1$. Note that T does not influence the exposure process in the sense that the hazard function for τ_E at time t does not depend on the history of the failure time process prior to t. The exposure process thus acts to ‘turn off’ the underlying hazard function $\lambda(.)$ governing the risk of the event, but otherwise does not modify this risk nor is it affected by the occurrence of the event. For convenience, we set $T=\infty$ when the event of interest does not occur.

With this model, $\lambda_E(.)$ also represents the marginal hazard function of τ_E, and the conditional distribution of T, given τ_E, has density function

$$f(t \mid \tau_E) = \lambda(t)e^{-\Lambda(t)}1[t \leqslant \tau_E]$$

for $0 \leqslant t < \infty$ and mass $e^{-\Lambda(\tau_E)}$ at infinity, where $\Lambda(t)=\int_0^t \lambda(u)\,du$. The marginal distribution of T has density $\lambda(t)e^{-\Lambda(t)}e^{-\Lambda_E(t)}$ for $0 \leqslant t < \infty$ and mass $\int_0^\infty e^{-\Lambda(t)}\lambda_E(t)e^{-\Lambda_E(t)}\,dt$ at infinity, where $\Lambda_E(t)=\int_0^t \lambda_E(u)\,du$. The usual failure time setting in which exposure is not explicitly considered corresponds to the limiting case where $\tau_E = \infty$, in which case $\lambda(.)$ represents the marginal hazard function of T, and the marginal density and cumulative distribution function of T are given by

$$f(.) = \lambda(.)e^{-\Lambda(.)}, \quad F(.) = 1 - \exp\{-\Lambda(.)\}.$$

The goal is to estimate the hazard function $\lambda(.)$, or equivalently the corresponding density function $f(.)$ or distribution function $F(.)$.

We assume throughout that T cannot be directly observed but that information about T is available from possibly imperfect diagnostic tests of the occurrence of the event that determines T. For example, a DNA PCR test might be administered to an individual to assess whether or not the event of HIV infection may have occurred. The observed data for the individual are denoted by $(J, \tau, v, r, E(.))$, where $\tau = (\tau_1, \ldots, \tau_J)$ is a $J \times 1$ vector denoting the times at which diagnostic tests are given, $v = (v_1, \ldots, v_J)$, where v_j is an indicator denoting the type of diagnostic test given at time τ_j, and $r = (r_1, \ldots, r_J)$ is the vector of binary test results, where $r_j = 1$ indicates that the test given at time τ_j was positive and $r_j = 0$ indicates that the test was negative, for $j = 1, \ldots, J$. We note that a 'test' can actually represent a battery of several different diagnostic tests given in a predetermined way. We allow the number, times and types of tests to be determined adaptively, with J and (τ_j, v_j) being some known deterministic or probabilistic function of $\{(\tau_i, v_i, r_i), i = 1, \ldots, j-1\}$ and of $\{E(u), 0 \leqslant u \leqslant \tau_{j-1}\}$.

If $g\{r | \tau, v, E(.), t\}$ denotes the conditional probability mass function of r, given τ, v, $E(.)$ and $T = t$, we assume that the individual's J test results are conditionally independent given T; that is,

$$g\{r | \tau, v, E(.), t\} = \prod_j g(r_j | \tau_j, v_j, t). \qquad (2{\cdot}1)$$

This assumption is analogous to the common assumption in measurement error regression models that the conditional distribution of the response variable, given the covariate and its proxy, is the same as the conditional distribution given only the covariate. In our setting, it means that the observed values of other diagnostic tests do not provide additional information about the distribution of a particular diagnostic test from that provided by the actual time of the event. Then, for an admissible set of values for J, τ and v, it is shown in the Appendix that

$$g\{J, \tau, v, r | t, E(.)\} = k_0 \prod_{j=1}^{J} g(r_j | \tau_j, v_j, t),$$

where the proportionality constant, k_0, is a consequence of the decision rule used to schedule visits. Thus, for an admissible set of values for J, τ and v, we can write

$$g\{J, \tau, v, r | E(.)\} = k_0 \int_0^\infty \prod_{j=1}^{J} \phi(t, \tau_j, v_j)^{r_j} \{1 - \phi(t, \tau_j, v_j)\}^{1-r_j}\, dF\{t | E(.)\}, \qquad (2{\cdot}2)$$

where $\phi(t, \tau_j, v_j) = \mathrm{pr}(r_j = 1 | T = t, \tau_j, v_j)$ and where 0^0 is taken to be zero. Note that, for $t \leqslant \tau$, $\phi(t, \tau, v)$ denotes the sensitivity of test v, whereas, for $t > \tau$, $\phi(t, \tau, v)$ denotes the complement of the test specificity. When $\tau_E = \infty$ and the diagnostic test is perfect, that is $\phi(t, \tau, v) = 1[t \leqslant \tau]$, the setting reduces to the special case of interval censored data.

3. Likelihood, identifiability and estimation

3·1. *Time-independent sensitivity and specificity*

Suppose that $\phi(t, \tau, v) = \phi_{\delta,v}$, where $\delta = 1[t \leqslant \tau]$; that is, the probability of a positive test of type v given after the event has occurred, or test sensitivity, is given by ϕ_{1v}, and the probability of a positive test given prior to the event, or the complement of the test specificity, is given by ϕ_{0v}. Then equation (2·2) reduces to

$$g\{J, \tau, v, r \mid E(.)\} = k_0 \int_0^\infty \prod_{j:\tau_j < t} \phi_{0v_j}^{r_j}(1-\phi_{0v_j})^{1-r_j} \prod_{j:\tau_j \geqslant t} \phi_{1v_j}^{r_j}(1-\phi_{1v_j})^{1-r_j}\, dF\{t \mid E(.)\},$$

where we use $T = \infty$ to denote the fact that the event does not occur prior to τ_E.

Let $0 < \tau_1^* < \ldots < \tau_K^*$ denote the distinct values of $\{\tau_1, \ldots, \tau_J, \tau_E\}$ that do not exceed τ_E and assume that no more than one diagnostic test is given at each test time, although situations in which more than one test is given to an individual at a single time point are easily accommodated. Note that $K = J + 1$ when $\tau_E > \tau_J$ and that $K \leqslant J$ when $\tau_E \leqslant \tau_J$. Since the terms in the products do not vary for $t \in (\tau_{k-1}^*, \tau_k^*)$, for $k = 1, \ldots, K$, $f\{t \mid E(.)\} = 0$, for $t > \tau_E$, and $dF\{\infty \mid E(.)\} = 1 - F(\tau_E)$, this becomes

$$g\{J, \tau, v, r \mid E(.)\} = k_0 \sum_{k=1}^{K+1} c_k \int_{\tau_{k-1}^*}^{\tau_k^*} f(t)\, dt = \sum_{k=1}^{K+1} c_k \{F(\tau_k^*) - F(\tau_{k-1}^*)\}, \qquad (3·1)$$

where $\tau_0^* = 0$, $\tau_{K+1}^* = \infty$, $c_{K+1} = \prod_{j=1}^J \phi_{0v_j}^{r_j}(1-\phi_{0v_j})^{1-r_j}$ and

$$c_k = \prod_{j:\tau_j \leqslant \tau_{k-1}^*} \phi_{0v_j}^{r_j}(1-\phi_{0v_j})^{1-r_j} \prod_{j:\tau_j \geqslant \tau_k^*} \phi_{1v_j}^{r_j}(1-\phi_{1v_j})^{1-r_j} \quad (k \leqslant K).$$

The likelihood function for a random sample of N individuals is a product of terms, each having the general form in equation (3·1). Let $0 < \omega_1 < \ldots < \omega_M$ denote the distinct values of τ_k^* for the N individuals. Then, if we use (3·1), the likelihood function is proportional to

$$L(\theta) = \prod_{i=1}^N \sum_{m=1}^{M+1} c_{im} \theta_m, \qquad (3·2)$$

where $\omega_0 = 0$, $\omega_{M+1} = \infty$, $\theta_m = F(\omega_m) - F(\omega_{m-1})$ and c_{im} is the value of

$$c_m = \prod_{j:\tau_j \leqslant \omega_{m-1}} \phi_{0v_j}^{r_j}(1-\phi_{0v_j})^{1-r_j} \prod_{j:\tau_j \geqslant \omega_m} \phi_{1v_j}^{r_j}(1-\phi_{1v_j})^{1-r_j}$$

or $\prod_{j=1}^J \phi_{0v_j}^{r_j}(1-\phi_{0v_j})^{1-r_j}$ for the ith individual, depending on whether $\omega_m \leqslant \tau_E$ or $\omega_m > \tau_E$, respectively.

It follows directly from the form of (3·2) that, without further assumptions, $\lambda(.)$ is estimable at most up to the interval probabilities $\{\theta_m, m = 1, \ldots, M\}$. When $\omega_M = \infty$, the sum of the θ_m equals one, but otherwise is less than one. In the former case, this expression can be rewritten as a linear combination of $\theta_1, \ldots, \theta_{M-1}$. In the degenerate case where $\phi_{0v} = \phi_{1v}$, r is independent of T, so that the likelihood function is noninformative about $\lambda(.)$. When $\phi_{0v} \neq \phi_{1v}$, it follows from Gentleman & Geyer (1994) that the logarithm of $L(\theta)$ is a strictly concave function of θ with a unique maximum provided that $\text{rank}(C) = M$, where C denotes the $N \times M$ matrix with elements c_{im}. Thus, when the sensitivity and specificity are known, the maximum likelihood estimator of θ can be obtained from (3·2) by joint maximisation of the loglikelihood using numerical techniques, as is illustrated in § 4. When the number of parameters is small, approximate confidence intervals and standard errors for components of θ, or functions of these components, can be obtained from the Hessian of the loglikelihood function. More generally, bootstrap estimates, based on bootstrap samples yielding the same ω as in the original dataset, are more stable.

When either ϕ_{0v} or ϕ_{1v} is unknown, the likelihood viewed as a function of the unknown values of θ and ϕ_{0v} or ϕ_{1v} does not have the structural form assumed by Gentleman & Geyer (1994), and thus their result does not apply. In this case, identifiability arguments can be established by other means that may require further assumptions that reduce the dimensionality of the vector of unknown parameters. To see one way of proceeding in this situation, suppose there is a single type of test and let $p_j = \text{pr}(r_j = 1)$ for $j = 1, \ldots, J_N$, where J_N is the number of unique test times among the N observations. Then, when $\tau_E = \infty$ for every individual, it follows that $M = J_N + 1$, where $\sum_{j=1}^{J_N+1} \theta_j = 1$, and

$$p_j = \phi_1 F(\tau_j) + \phi_0\{1 - F(\tau_j)\} \quad (j = 1, \ldots, J_N). \tag{3·3}$$

Since the p_j are directly estimable from the observed test results, θ becomes estimable if its dimensionality is reduced by one or two, depending on whether one or both of ϕ_0 and ϕ_1 are unknown. In § 3·2 we achieve this by setting some of the components of $\theta_1, \ldots, \theta_M$ to be equal. Alternatively, as we illustrate below, a parametric form for $\lambda(.)$ can be assumed to reduce the dimensionality of θ.

When $E(.)$ is not identically one, $M > J_N + 1$, so that more about $\lambda(.)$ is estimable than when $\tau_E = \infty$ for all individuals. For example, suppose that each individual is exposed either in the first half or the second half of the visit interval $(\tau_1, \tau_2]$. Then the likelihood contribution for those exposed in the first half of the interval involves $F(\tau_{12}) - F(\tau_1)$, where τ_{12} is the midpoint, while, for those exposed only in the second half, the likelihood contribution involves $F(\tau_2) - F(\tau_{12})$. Hence, $F(.)$ becomes identified at τ_{12} in addition to τ_1 and τ_2.

Asymptotically, standard methods can be used to demonstrate that the maximum likelihood estimator of θ is identifiable, consistent and asymptotically normal provided that the number of distinct visit times is finite.

To illustrate these points, suppose that T has an exponential distribution, $\tau_E = \infty$ for all individuals, and a diagnostic test is given at times $\{5, 15, 25, 35, 45\}$, with each test time having a 0·4 probability of being missing, so that different subjects will have tests done at different subsets of $\{5, 15, 25, 35, 45\}$. This yields $J_N = 5$ and $M = 6$, with

$$\theta_1 = F(5), \quad \theta_j = F(10j - 5) - F(10j - 15) \quad (j = 2, \ldots, 5), \quad \theta_6 = 1 - \theta_1 - \ldots - \theta_5.$$

Consider first the case where ϕ_0 and ϕ_1 are known. Figure 1 gives $F(t)$ and the average values of maximum likelihood estimate of $F(t)$ for $t = 5$, 15, 25, 35 and 45 obtained from (3·2) based on 25 simulations, where $N = 100$, λ is selected to yield $\text{pr}(T > 45)$ equal to 0·3 and 0·7, and the sensitivity, ϕ_1, and specificity, $1 - \phi_0$, equal to various combinations of 0·33, 0·70 and 0·99. Also shown is the average of the maximum likelihood estimates based on the incorrect assumption that T is interval censored between the first positive diagnostic test and the preceding negative diagnostic test. To facilitate viewing the results, the values of the estimated distribution functions, which are only estimable at the test times, are connected with lines. When test sensitivity or specificity is low, i.e. 0·33, the interval censored estimator can be severely biased. The bias is somewhat less when the sensitivity and specificity are higher, i.e. 0·7. In both cases the direction of the bias depends on the expected value of T. The proposed method, as would be expected, performs well.

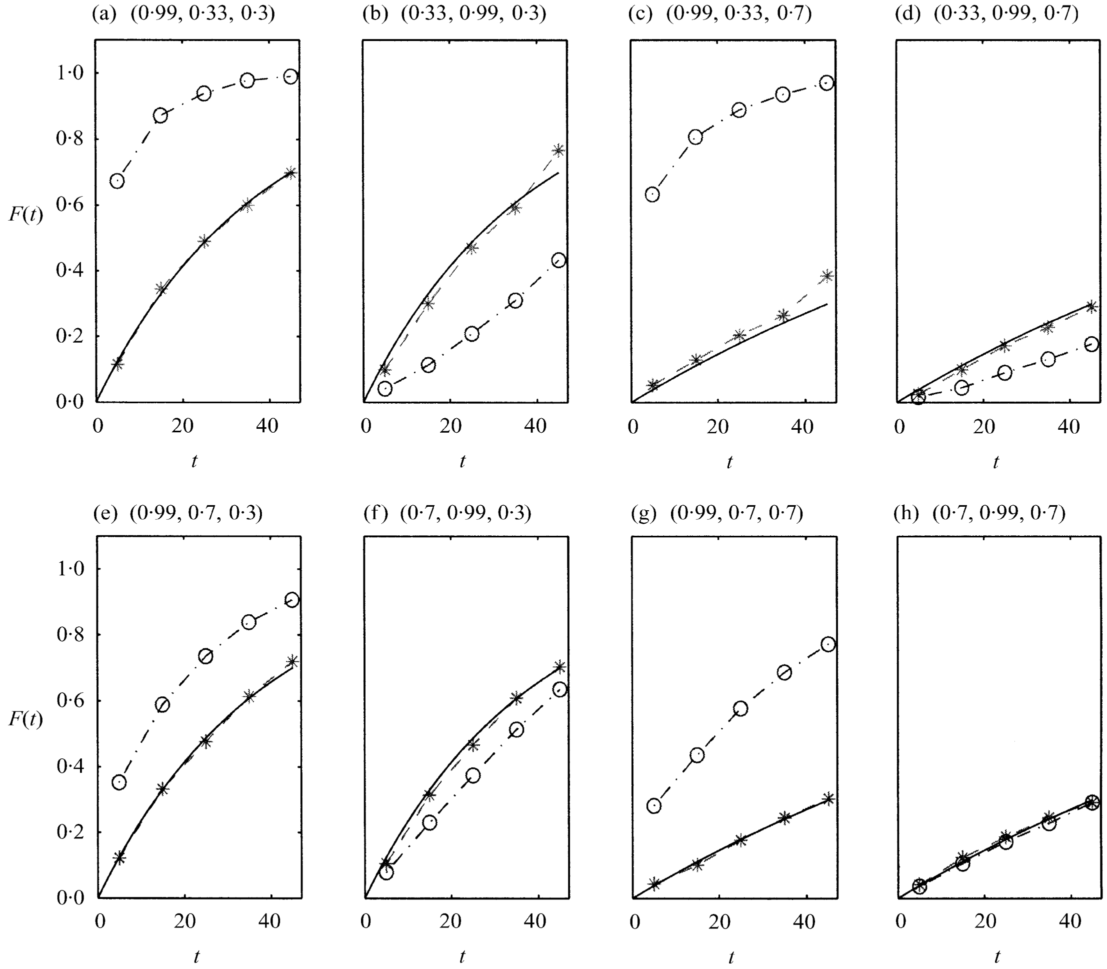

Fig. 1. Estimated cumulative distribution functions, with titles referring to (sensitivity, specificity, $\text{pr}(T > 45)$). Dash-dotted lines with circles correspond to interval censoring, dashed lines with stars correspond to the proposed method and solid lines correspond to $F(t)$.

For this same example, suppose now that both ϕ_0 and ϕ_1 are unknown, but that T is assumed to have the Weibull distribution; that is, $\lambda(t) = \lambda\gamma t^{\gamma-1}$, where $\lambda > 0$ and $\gamma > 0$ are unknown parameters. Evaluation of (3·3) gives

$$\phi_1 - p_j = (\phi_1 - \phi_0)e^{-\lambda\tau_j^{\gamma}} \quad (j = 1, \ldots, 5).$$

Since this system of equations is over-specified, it follows that the unknown parameters $(\lambda, \gamma, \phi_0, \phi_1)$ are estimable from (3·2) with θ_m expressed in terms of λ and γ.

3·2. *Time-dependent sensitivity*

As seen in § 2, $g\{J, \tau, v, r \mid E(.)\}$ can be expressed as

$$g\{J, \tau, v, r \mid E(.)\} = k_0 \int c(t)\, dF\{t \mid E(.)\}, \tag{3·4}$$

where $c(t)$ is a function of the sensitivity and specificity of the diagnostic tests and the individual's test results. In general, for sensitivity functions that depend on τ and t, this integral equation does not simplify easily. In this section, we consider a special class of diagnostic tests in which specificity is time-independent, denoted by $1-\phi_{0v}$, and sensitivity depends on t and τ through their difference, $\tau-t$, and at some stage reaches a constant level. This model is motivated by the behaviour of diagnostic tests administered to detect the presence of HIV in infants, such as DNA PCR and HIV culture tests, in which the sensitivity can be low for approximately 2 weeks following infection and is thereafter high and constant (Dunn et al., 1995, 2000). This is based on rates of test positivity in non-breast-fed infants known to have been HIV infected in utero or at birth, where these rates are observed to rise steadily until 2 weeks of age and thereafter remain constant. Suppose that $\mathrm{pr}(r=1\,|\,t,\tau,v,t\leqslant\tau)=\phi_{1v}(\tau-t)$, where $\phi_{1v}(u)$ is nondecreasing in u for $0\leqslant u\leqslant l_v$ and is constant in u, $\phi_{1v}(u)=\phi_{2v}$ say, for $u>l_v$ and $v=1,\ldots,V$. Let $\mathscr{T}_v$ represent the times of diagnostic test v and let $\mathscr{T}$ represent the distinct times of all diagnostic tests. Define $(\tau_1^*,\ldots,\tau_K^*)$ to be the unique elements of $\{\mathscr{T}_1-l_1\cup\ldots\cup\mathscr{T}_V-l_V\cup\mathscr{T}\cup\tau_E\}$ that satisfy $0<\tau_1^*<\ldots<\tau_K^*=\tau_E$. Equation (3·4) can then be expressed as

$$g\{J,\tau,v,r\,|\,E(.)\}=k_0\sum_{k=1}^{K+1}c_k\int_{\tau_{k-1}^*}^{\tau_k^*}G_k(t)f(t)\,dt, \qquad (3{\cdot}5)$$

where $\tau_0^*=0$, $\tau_{K+1}^*=\infty$, $G_{K+1}(t)=1$, $c_{K+1}=\prod_{j=1}^{J}\phi_{0v_j}^{r_j}(1-\phi_{0v_j})^{1-r_j}$,

$$G_k(t)=\prod_{j:\tau_{k-1}^*<\tau_j<\tau_k^*+l_{v_j}}\{\phi_{1v_j}(\tau_j-t)\}^{r_j}\{1-\phi_{1v_j}(\tau_j-t)\}^{1-r_j}\quad(k=1,\ldots,K),$$

$$c_k=\prod_{j:\tau_j\leqslant\tau_{k-1}^*}\phi_{0v_j}^{r_j}(1-\phi_{0v_j})^{1-r_j}\prod_{j:\tau_j\geqslant\tau_k^*+l_{v_j}}\phi_{2v_j}^{r_j}(1-\phi_{2v_j})^{1-r_j}\quad(k=1,\ldots,K).$$

Further simplification of equation (3·5) in general requires additional assumptions about the form of $\phi_{1v}(\tau-t)$. For the special case when

$$\phi_{1v}(\tau-t)=\begin{cases}\phi_{1v}, & \text{for } 0\leqslant\tau-t<l_v,\\ \phi_{2v}, & \text{for } l_v\leqslant\tau-t,\end{cases}$$

where $\phi_{1v}<\phi_{2v}$ for all v, we have

$$g\{J,\tau,v,r\,|\,E(.)\}=k_0\sum_{k=1}^{K+1}c_k\{F(\tau_k^*)-F(\tau_{k-1}^*)\}, \qquad (3{\cdot}6)$$

where

$$\begin{aligned}c_k=&\prod_{j:\tau_j\leqslant\tau_{k-1}^*}\phi_{0v_j}^{r_j}(1-\phi_{0v_j})^{1-r_j}\prod_{j:\tau_j\geqslant\tau_k^*+l_{v_j}}\phi_{2v_j}^{r_j}(1-\phi_{2v_j})^{1-r_j}\\ &\times\prod_{j:\tau_{k-1}^*<\tau_j<\tau_k^*+l_{v_j}}\phi_{1v_j}^{r_j}(1-\phi_{1v_j})^{1-r_j}\end{aligned}$$

for $k=1,\ldots,K$ and $c_{K+1}=\prod_{j=1}^{J}\phi_{0v_j}^{r_j}(1-\phi_{0v_j})^{1-r_j}$.

The likelihood function for a random sample of N individuals is a product of terms, each having the general form in (3·6). Let $0<\omega_1<\ldots<\omega_M$ denote the distinct times τ_k^* for the N individuals. Then, from (3·6), the likelihood function is proportional to

$$L(\theta)=\prod_{i=1}^{N}\sum_{m=1}^{M+1}c_{im}\theta_m, \qquad (3{\cdot}7)$$

where $\omega_0 = 0$, $\omega_{M+1} = \infty$, $\theta_m = F(\omega_m) - F(\omega_{m-1})$, and c_{im} is the value of

$$c_m = \prod_{j:\tau_j \leqslant \omega_{m-1}} \phi_{0v_j}^{r_j}(1-\phi_{0v_j})^{1-r_j} \prod_{j:\tau_j \geqslant \omega_m + l_{v_j}} \phi_{2v_j}^{r_j}(1-\phi_{2v_j})^{1-r_j} \times \prod_{j:\omega_{m-1} < \tau_j < \omega_m + l_{v_j}} \phi_{1v_j}^{r_j}(1-\phi_{1v_j})^{1-r_j}$$

or $\prod_{j=1}^{J} \phi_{0v_j}^{r_j}(1-\phi_{0v_j})^{1-r_j}$ for the ith individual, depending on whether ω_m is no greater than or greater than τ_E, respectively.

Without further assumptions, $\lambda(.)$ is at most identifiable up to the vector of parameters θ. As in the preceding section, $L(\theta)$ is a strictly concave function of θ with unique maximum if $\operatorname{rank}(C) = M$, where C denotes the $N \times M$ matrix with elements c_{im}. When the sensitivity function and specificity are known, the maximum likelihood estimate of θ can thus be obtained from (3·7) using standard numerical techniques. When the sensitivity and specificity functions are not fully known, further assumptions are needed to ensure the estimability of θ. As with time-independent sensitivities and specificities, these can be guided by the set of J_N equations relating the probabilities p_j to $\lambda(.)$ and the unknown sensitivity and specificity parameters. We illustrate this in § 4 by constraining some of the components of θ in adjacent time intervals to be equal.

As with the likelihood function for time independent sensitivity and specificity, the dimension of θ is greater for a non-constant exposure process than when $\tau_E = \infty$ for all individuals. In addition, the non-constant form of $\phi_{1v}(\tau - t)$ allows more aspects of $\lambda(.)$ to be estimated than when the sensitivity is time-independent. When there is a single type of test in these settings, it can be shown that the number of components of θ is equal to $2J_N - M_J + 1$, where M_J denotes the number of inter-visit intervals equal in length to l. This varies from $J_N + 2$, when the consecutive visit times are spaced by l, to $2J_N + 1$, when no inter-visit time equals l. This is illustrated in the example to follow where, for example, information about the risk of hiv during pregnancy is estimable in settings where diagnostic tests are given only following birth.

4. Example

We illustrate the proposed methods with data from a Tanzanian trial of the prevention of mother-to-child transmission of hiv during pregnancy or while breast-feeding; see Fawzi et al. (1998) for details of the trial design and results. Here T denotes the time to hiv infection of the infant, and the exposure process $E(t)$ equals 1 during pregnancy and while the infant is breast-fed, and zero thereafter. Two types of diagnostic test were used, dna pcr and the definitive elisa antibody test. The former were scheduled at birth, 6 weeks of age and every three months thereafter up to 2 years of age, with the elisa antibody test being given at or after 18 months of age. In practice, most infants had multiple missed visits and actual visit times were not always as scheduled.

We analyse the subset of dna pcr tests done prior to 2 years of age and elisa tests done between 18 months and 2 years. A total of 786 infants had at least one dna pcr test during the first 2 years, with a median of 2 tests, the range being from 1 to 7 and with most being done prior to 3 months. A total of 144 infants were also given an elisa antibody test between 18 months and 2 years. Of the 144 infants, 17 were found to be hiv-infected and 127 were found to be uninfected at the time of the elisa test. The median duration of breast-feeding was 18 months, with a range of 1 to 45 months. Thus, the dataset includes two types of diagnostic test and differential exposure patterns among the infants. In order to reduce the number of parameters estimated, test times after 2 weeks of age were grouped to the nearest month, as were the times at which breast-feeding was stopped. This results in 67 disjoint time intervals between −14 days and 2 years of age, where 0 is taken to be the time of birth. To incorporate the overwhelming clinical evidence that a majority of transmissions occur at the time of birth, we imposed the constraint that the probability of HIV transmission in the interval including the time of birth is the largest among all intervals in the time period between −14 days and birth.

We fitted the step function sensitivity model described in § 3·2 for the DNA PCR tests with $l = 14$ days, $\phi_0 = 0{\cdot}01$ and $\phi_2 = 0{\cdot}93$. The values for l, ϕ_0 and ϕ_1 were based on a review of the literature on diagnostic tests; see Dunn et al. (1995, 2000), Kalish et al. (1997) and Owens et al. (1996). We estimate the parameter ϕ_1 as this is not well characterised in the literature. The ELISA antibody test is taken to have perfect sensitivity and specificity. This leads to estimable parameters provided the dimensionality of θ is reduced to equal the number of unique test times in the dataset, namely 34. We achieve this by assuming that the daily probability of HIV transmission is the same from −14 to −7 days of age, from −7 to −3 days of age, during week 1, during week 2, during weeks 2–4, monthly thereafter until month 6 and then once every 2 months until 2 years of age, where we have relabelled the time axis so that time 0 corresponds to birth. This results in 20 independent components of θ, corresponding to the values of $F(t)$ at −14, −7, −3, 0, 7, 14, 30, 61, 91, 122, 152, 183, 243, 304, 381, 442, 503, 564, 655 and 730 days.

Figure 2(a) displays the resulting estimate of $F(.)$ when $E(t) = 1$ for all t, that is when breast-feeding continues until 2 years of age. The estimates are obtained by numerical

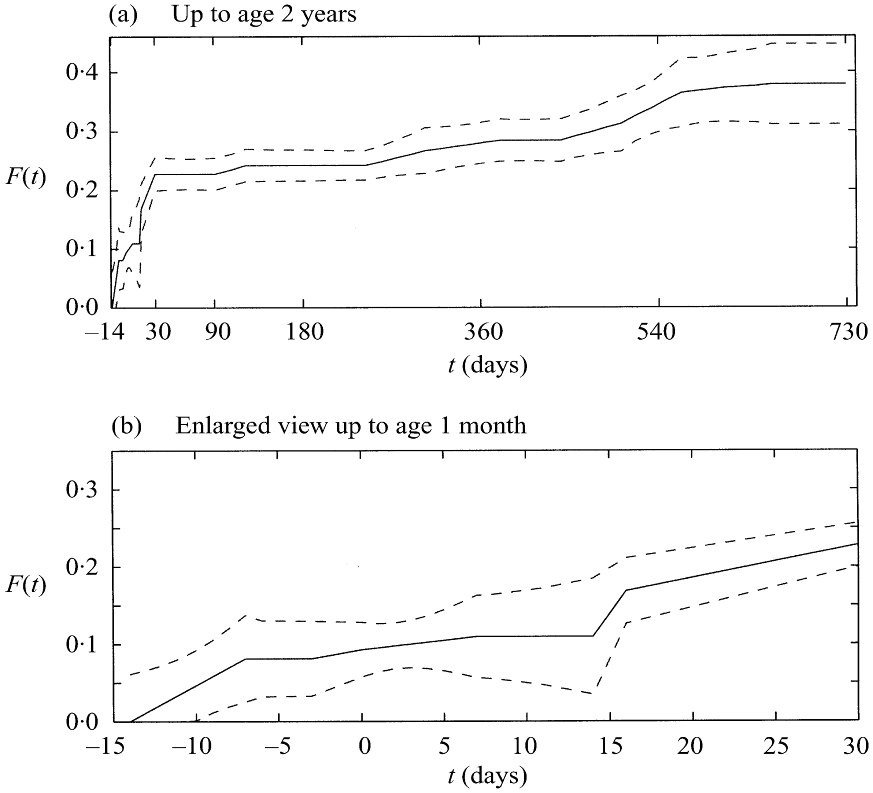

Fig. 2. Example. Estimated cumulative distribution functions of the timing of vertical transmission of HIV and approximate 95% pointwise confidence intervals.

maximisation of the loglikelihood using Matlab, which uses sequential quadratic programming methods based on obtaining solutions to the Kuhn–Tucker equations. Also shown are approximate 95% pointwise confidence intervals based on a normal approximation using a bootstrap variance estimator. The corresponding estimate of ϕ_1 is 0·70. For ease of viewing, the estimable values of F at distinct visit times are connected by straight lines. Figure 2(b) gives an enlarged view of the estimate of $F(t)$ until one month after birth. The estimated probability of vertical transmission by 2 years is 0·379, with 95% confidence interval (0·311, 0·446), and the estimated probability of transmission prior to and during birth is 0·093, with 95% confidence interval (0·057, 0·128). The estimated cumulative distribution function for T corresponding to a finite value of τ_E, 360 days, say, is the same as the estimate in Fig. 2 up to this value, and is thereafter horizontal. Overall, the cumulative risk of transmission from breast-feeding increases quickly during the first month of life and then more gradually.

To assess the dependency of the estimated parameter values on the assumed values of l, ϕ_2 and ϕ_0, we reanalysed the data using different plausible values of these quantities. For values of $l=12$, 14 and 16 days respectively, the estimated probability of infection prior to $-l$ and the cumulative distribution function during the period between one month and 2 years of age were generally similar. For times between -10 and 30 days, the estimated cumulative distribution function based on $l=12$ days placed a somewhat higher probability of HIV transmission at the time of birth when compared to the corresponding estimates for $l=14$ and 16 days. For values of ϕ_2 equal to 0·89, 0·91, 0·93 and 0·95 respectively, and for values of $1-\phi_0$ equal to 0·95, 0·97 and 0·99 respectively, the estimated cumulative distribution functions were relatively stable, especially beyond 1 month of age. Further details regarding these analyses as well as the data are available from the authors upon request.

5. Discussion

It was seen in § 3 that standard methods for interval censored data can be biased when the diagnostic tests used to screen for the event of interest are imperfect. The bias will be small when the test sensitivities and specificities are high but otherwise depends in a complicated way on the distribution of T, the test times and the sensitivity and specificity functions, and can be substantial. The proposed approach is flexible in terms of allowing different types of imperfect diagnostic test and different exposure histories among individuals.

The methods developed in this paper can be extended in several ways. Covariates can be incorporated by assuming a parametric or semiparametric regression model for $\lambda(.)$, such as a proportional hazards or accelerated failure time model. Another extension is to generalise the exposure process. The simple binary process considered here could be extended to an alternating process, allowing several periods of exposure for the same individual. This might occur, for example, for assessing pregnancy risk in women who use contraceptives on an intermittent basis, and leads to a likelihood contribution similar to (3·1), but with $F(.)$ replaced by the conditional distribution of T given the individual's exposure history. Another extension would allow several types of exposure. For example, in a study of vertical transmission of HIV, suppose that all mothers receive the same treatment during pregnancy but that at 2 weeks of age infants are randomly assigned to breast-feeding in combination with anti-HIV therapy or formula feeding. In such settings, one would like to use data from all mother-infant pairs to estimate the risk of transmission during pregnancy and up to 2 weeks of age. This can be accommodated by taking $E(t)$ to be a categorical variable with values in $\{0, \ldots, M\}$, and assuming that $\lambda\{t \mid E(.)\} = \lambda_m(t)$ when $E(t) = m$ and zero otherwise. When each individual is subject to only one type of risk, this reduces to just M one-sample problems. However, when an individual can be exposed to different risks, methods analogous to those developed in § 3 apply directly.

For applications such as vertical transmission of HIV, it would also be useful to extend the methods to allow for competing risks. For example, in the Tanzania data, 18·7% of the infants died by 2 years of age. To illustrate how ignoring information on infant mortality can affect results, we reanalysed the data assuming that all infants who died, and whose last visit date was before 2 years of age, became HIV positive by the time of their last visit. We accommodated this additional information in our model by assuming that the test on this last visit was perfect and positive for HIV infection. Under the model assumptions made in the example, we obtained a cumulative probability of infection of 0·45 for an infant breast-fed up to 2 years of age. Note that one could assume numerous such 'worst case' scenarios to obtain a plausible bound on the cumulative distribution function of time to HIV infection. In this setting, it would be worth exploring extensions based on the models considered by Hughes & Richardson (2000).

Finally, methods for time-dependent sensitivity or specificity functions of other forms could be explored. The methods developed in § 3·2 are easily extended to more general step-functions, but extensions for complicated functions are more challenging because the integral in (3·4) may not be solvable for $f(.)$ without additional assumptions.

Acknowledgement

This work was supported by a grant from the National Institutes of Health and a predoctoral fellowship from the Howard Hughes Medical Institute. We are grateful to Drs Wafaie Fawzi and Donna Spiegelman for their insights and for providing the data used in § 4.

Appendix

Derivation of (2·2)

Let $\eta_j = (\tau_j, v_j, r_j)$ and $\eta = (\eta_1, \ldots, \eta_J)$. The conditional probability mass function of (J, τ, v, r), given $T = t$ and $E(.)$, can be written

$$\begin{aligned} g\{J, \tau, v, r, \mid E(.), t\} &= \left[\prod_{j=1}^{J} g\{\eta_j \mid \eta_1, \ldots, \eta_{j-1}, t, E(.)\}\right] \times g\{J \mid \eta, t, E(.)\} \\ &= \prod_{j=1}^{J} g\{\tau_j, v_j \mid \eta_1, \ldots, \eta_{j-1}, t, E(.)\} \\ &\quad \times g\{J \mid \eta, t, E(.)\} \prod_{j=1}^{J} g\{r_j \mid \eta_1, \ldots, \eta_{j-1}, \tau_j, v_j, t, E(.)\}. \qquad (\text{A·1}) \end{aligned}$$

where $g\{\tau_j, v_j \mid \eta_1, \ldots, \eta_{j-1}, t, E(.)\}$ denotes the rule for determining the time and type of the jth test and $g\{J \mid \eta, t, E(.)\}$ denotes the probability that the adaptive procedure terminates after the Jth test. Applying the conditional independence assumption in (2·1) gives

$$g\{J, \tau, v, r, \mid E(.), t\} = \left[g\{J \mid \eta, t, E(.)\} \prod_{j=1}^{J} g\{\tau_j, v_j \mid \eta_1, \ldots, \eta_{j-1}, t, E(.)\}\right] \prod_{j=1}^{J} g(r_j \mid \tau_j, v_j, t). \quad (\text{A·2})$$

We assume that the rules for determining (J, τ_j, v_j) are known probabilistic or deterministic functions of only $\eta_1, \ldots, \eta_{j-1}$ and $E(u)$ for $0 \leqslant u \leqslant \tau_{j-1}$, that place all the mass for τ_j at values larger than τ_{j-1}. The special case of a prespecified visit/test schedule, (J^*, τ^*, v^*), say, is obtained by taking $g\{\tau_j, v_j | \eta_1, \ldots, \eta_{j-1}, t, E(.)\}$ and $g\{J | \eta, t, E(.)\}$ to be 1 for $(\tau_j, v_j, J) = (\tau_j^*, v_j^*, J^*)$ and zero otherwise. For a deterministic adaptive procedure, each of the probability mass functions is also degenerate, and in both cases the term in square brackets in (A·2) equals 1. For probabilistic rules, for example, when a subset of individuals are tested more intensively, the term in square brackets in (A·2) is a known constant, $k_0 = k_0\{\eta, E(.)\}$, say.

References

Balasubramanian, R. & Lagakos, S. W. (2001). Estimation of the timing of perinatal transmission of HIV. *Biometrics* **57**, 1048–58.

Dunn, D. T. & Ades, A. E. (1996). Estimating the HIV vertical transmission rate and the pediatric AIDS incubation period from prospective data. *J. Am. Statist. Assoc.* **91**, 935–43.

Dunn, D. T., Brandt, C., Krivine, A., Cassol, S., Roques, P., Borkowsky, W., De Rossi, A., Denamur, E., Ehrnst, A., Loveday, C., Harris, J., McIntosh, K., Comeau, A., Rakusan, T., Newell, M. & Peckham, C. (1995). The sensitivity of HIV-1 DNA polymerase chain reaction in the neonatal period and the relative contributions of intra-uterine and intra-partum transmission. *AIDS* **9**, F7–F11.

Dunn, D. T., Simonds, R. J., Bultery, M., Kalish, L. A., Moye, J., deMaria, A., Kind, C., Rudin, C., Denamur, E., Krivine, A., Loveday, C. & Newell, M. L. (2000). Interventions to prevent vertical transmission of HIV-1: effect on viral detection rate in early infant samples. *AIDS* **14,** 1421–8.

Fawzi, W., Msamanga, G., Spiegelman, D., Urassa, E., McGrath, N., Mwakagile, D., Antelman, G., Mbise, R., Herrera, G., Kapiga, S., Willet, W. & Hunter, D. (1998). Randomized trial of effects of vitamin supplements on pregnancy outcomes and T cell counts in HIV-1 infected women in Tanzania. *Lancet* **351**, 1477–82.

Gentleman, R. & Geyer, C. (1994). Maximum likelihood for interval censored data: Consistency and computation. *Biometrika* **81**, 618–23.

Hughes, J. & Richardson, B. (2000). Analysis of a randomized trial to prevent vertical transmission of HIV-1. *J. Am. Statist. Assoc.* **95**, 1032–43.

Kalish, L. A., Pitt, J., Lew, J., Landesman, S., Diaz, C., Hershow, R., Hollinger, F., Pagano, M., Smeriglio, V. & Moye, J. (1997). Defining the time of fetal or perinatal acquisition of human immunodeficiency virus type 1 infection on the basis of age at first positive culture. *J. Inf. Dis.* **175**, 712–5.

Martin, P. & Friedman, L. (1998). Assessment of liver function and diagnostic studies. In *Handbook of Liver Disease*, Ed. L. S. Friedman and E. B. Keeffe, pp. 1–14. Philadelphia, PA: Churchill Livingstone.

Owens, D., Holodniy, M., McDonald, T., Scott, J. & Sonnad, S. (1996). A meta-analytic evaluation of the polymerase chain reaction for the diagnosis of HIV infection in infants. *J. Am. Med. Assoc.* **275**, 1342–61.

Tsai, W., Goedert, J., Orazem, J., Landesman, S., Rubinstein, A., Willoughby, A. & Gail, M. (1994). A nonparametric analysis of the transmission rate of human immunodeficiency virus from mother to infant. *Biometrics* **50**, 1015–28.

Turnbull, B. W. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *J. R. Statist. Soc.* B **38**, 290–5.

[*Received August* 2001. *Revised March* 2002]