

Estimating time to pregnancy from current durations in a cross-sectional sample

NIELS KEIDING[†], KAJSA KVIST, HELLE HARTVIG, MADTS TVEDE

Department of Biostatistics, University of Copenhagen, Blegdamsvej 3, DK-2200, Copenhagen N, Denmark

N.Keiding@biostat.ku.dk

SVEND JUUL

Department of Epidemiology and Social Medicine, University of Aarhus, Vennelyst Boulevard 6, DK-8000, Aarhus C, Denmark

SUMMARY

A new design for estimating the distribution of time to pregnancy is proposed and investigated. The design is based on recording *current durations* in a cross-sectional sample of women, leading to statistical problems similar to estimating renewal time distributions from backward recurrence times. Non-parametric estimation is studied in some detail and a parametric approach is indicated. The results are illustrated on Monte Carlo simulations and on data from a recent European collaborative study. The role and applicability of this approach is discussed.

Keywords: Cross-sectional study; Historically prospective study; Nonparametric maximum likelihood; Renewal process; Retrospective study; Time-to-pregnancy.

1. INTRODUCTION

The time it takes for a couple from initiating attempts to become pregnant until conception occurs (*time-to-pregnancy* or TTP) is gaining importance as a measure of natural fecundity (Weinberg and Dunson, 2000). The two most common and obvious designs are a *cohort* (follow-up) study where couples are followed forward in time from when they start attempting to become pregnant, or a *retrospective* study of pregnant women where couples are interviewed about when they started their attempt to become pregnant. A variation of the cohort study is the *historically prospective* design where a general sample (usually of women) from the population are asked to recall their reproductive history.

In principle the cohort approach allows direct estimation and should therefore be desirable, but to recruit a representative cohort is very difficult, as in practice severe self-selection effects are to be expected. Also, even if recruitment were successful it is very resource-demanding and time-consuming to conduct an accurate follow-up (see Bonde *et al.* (1998) for an example of a prospective study). The historically prospective study suffers from recall bias and also mixes experience over a long calendar time period. Most of the practical experience on distribution of time to pregnancy therefore comes from retrospective studies of pregnant women. The main difficulty with this design is that it intrinsically conditions on becoming pregnant, which is often ignored in the analysis, distorting the evaluation of

[†]To whom correspondence should be addressed

the effects of certain covariates (in particular, age of the woman), which are related to both biological and behavioural aspects of fecundity (Juul *et al.*, 2000).

It would therefore seem worthwhile (Keiding, 1999) to explore a third possible design: collect from a cross-sectional sample of women (couples) those that are currently attempting to become pregnant in order to obtain from each of these the time they have been attempting. We shall show that the distribution of these times identifies the distribution of actually realized waiting times (to pregnancy or to unsuccessful end of attempt) in the population.

Section 2 of the paper derives the distribution of current duration and compares it to the distribution of the observations in the other designs, with regard also to possible observed or unobserved heterogeneity. Section 3 surveys nonparametric estimation of a decreasing density g , including how to perform the vital correction for inconsistency of $\hat{g}(0)$. Monte Carlo studies in Section 5 evaluate these estimators, while Section 4 reports a first attempt at parametric modelling. In Section 6 data from a recent European collaborative study are used to illustrate the ideas, while Section 7 is a concluding discussion.

2. THE DISTRIBUTION OF CURRENT DURATION

For each attempt at becoming pregnant let T be the time to pregnancy, U the time to discontinuation without pregnancy (for reasons such as death of the woman, disappearance of partner, couple give up trying; in some cases start of fertility treatment should perhaps be included) and V the time to discontinuation of follow-up since the start of the attempt. We are interested in the distribution of T . In a *prospective* design the problem reduces to standard survival analysis with T as the time to endpoint and $\min(U, V) = U \wedge V$ the time to censoring. In the *retrospective* design (based on pregnant women) we have a complete sample from the conditional distribution of $T|T < U$. Note that this situation is different from right truncation of T by U , which corresponds to observing the conditional distribution of $(T, U)|T < U$, see e.g. Keiding and Gill (1990).

In the proposed *current-duration* design let $X = T \wedge U$ be the waiting time until termination for whatever reason, successful or not, with probability density $f(x)$, survival function $S(x) = \int_x^\infty f(a) da$ and expectation $\mu_X = \int_0^\infty xf(x) dx = \int_0^\infty S(x) dx$, which we shall assume finite. To derive the distribution of current durations we follow Keiding (1991) and Lund (2000). In a (calendar time, duration) Lexis diagram each pregnancy attempt is represented by a line segment from $(t, 0)$ to $(t + X, X)$ with t the calendar time at initiation. Cross-sectional sampling takes place at some fixed time t_0 , and the sampled attempts are those crossing the vertical half-line $(t_0, 0)$ to (t_0, ∞) . Assuming that initiations happen according to a Poisson process in calendar time t with intensity $\beta(t)$, the observed experienced waiting time at t_0 , $Y = X \wedge V = T \wedge U \wedge V$, has density proportional to $\beta(t_0 - y)S(y)$.

Our focus will be on the time-homogeneous situation $\beta(t) = \beta$, which should suffice in most situations where only short calendar intervals are considered for each 'cross-section'. Here, as shown by Cox (1969), Y will be distributed as a backward recurrence time in a renewal process in equilibrium with renewal distribution $f(x)$, that is, the *density* of Y is

$$g(y) = S(y)/\mu_X.$$

Note in particular that $0 < g(0) < \infty$. Thus, Y has a *decreasing density* proportional to the *survival function* of X .

To estimate the distribution of Y thus amounts to estimating a distribution in some class of distributions on $[0, \infty)$ with decreasing density g , taking a finite value at the origin. An estimate of the survival function $S(x)$ of X is then obtained by setting

$$\hat{S}(x) = \hat{g}(x)/\hat{g}(0).$$

A main difficulty with the retrospective design is that careful account of the sampling procedure is required to avoid distortion of the effect of covariates, even in the often unrealistically simple case where the time T to pregnancy may be assumed conditionally independent of time U to discontinuation given a covariate Z : let $f(t|z)$ and $k(u|z)$ be the densities of $(T|Z)$ and $(U|Z)$, then the joint density of $(T, U|Z)$ is $f(t|z)k(u|z)$. We observe $(T|Z, \{T < U\})$ which has density

$$c f(t|z) \int_t^\infty k(u|z) du. \tag{2.1}$$

Since there is usually no reasonable independent information on $k(u|z)$, f becomes unidentifiable. As an aside, note that in an ordinary right truncation situation, i.e. observation of $(T, U|Z, \{T < U\})$, $f(t|z)$ would be identifiable via the retro-hazard $f(t|z) / \int_0^t f(s|z) ds$, see Keiding (1992).

A further consequence of (2.1) is that in case of completely or partially unobserved heterogeneity Z the mixture across Z is different for the retrospective data than for the prospective cohort data on which we want information: as is intuitively obvious, the subfertile are under-represented, and the infertile not represented at all.

On this background the situation for the current duration design is more encouraging. It is here natural to assume independence of (X, Z) and the censoring time V ; let $f(x|z)$ be the density of $(X|Z)$. As remarked by van Es *et al.* (2000), the same derivation of the distribution of the current duration $Y|Z$ as above will be valid:

$$g(y|z) = \frac{S(y|z)}{E(X|z)} \tag{2.2}$$

with $S(x|z) = \int_x^\infty f(s|z) ds$. From an estimate $\hat{g}(y|z)$ we may therefore estimate the conditional survival function $S(x|z)$ by $\hat{g}(x|z) / \hat{g}(0|z)$.

For unobserved heterogeneity denote by $h(z)$ the distribution of Z in the population, so that $f(x, z) = f(x|z)h(z)$ is the joint distribution of X and Z . As above, the distribution of (X, Z) in the sample has density $xf(x, z) / E(X)$, yielding the density of Z in the sample $h(z)E(X|z) / E(X)$, so that, by 2.2, the density of the observed current durations is

$$g(y) = \int g(y|z) \frac{h(z)E(X|z)}{E(X)} dz = \frac{1}{E(X)} \int S(y|z)h(z) dz.$$

This means that $\hat{g}(x) / \hat{g}(0)$ will be an estimator of

$$\int S(x|z)h(z) dz = S(x)$$

even under unobserved heterogeneity.

3. NONPARAMETRIC ESTIMATION

We return to the situation without covariates. The non-parametric maximum likelihood estimator (NPMLE) of a decreasing density was derived by Grenander (1956) and generalized to right-censored data by Denby and Vardi (1986). Woodroffe and Sun (1993) pointed out (in the uncensored case) that $\hat{g}(0)$ is inconsistent and suggested a consistent penalized NPMLE of $g(0)$; this work was elaborated by Sun and Woodroffe (1996) who developed an adaptive penalized NPMLE.

In our application we will usually observe *uncensored* current durations Y_1, \dots, Y_n . However, we may not always want to place too much reliance on durations larger than $\xi = 3$ years, say, and therefore may

prefer to artificially censor all durations at 3 years, that is, rather than Y_i consider $(Y_i \wedge 3, I\{Y_i \leq 3\})$. This is only a small problem because of the following reproducibility property of the Denby–Vardi estimator: if $\xi_1 < \xi_2$ and the estimators based on censoring at ξ_i are denoted \hat{g}_{ξ_i} , $i = 1, 2$, then $\hat{g}_{\xi_1}(y) = \hat{g}_{\xi_2}(y)$ for $0 < y < \xi_1$. In particular $\hat{g}_{\xi}(y) = \hat{g}(y)$ for $0 < y < \xi$ if $\hat{g} = \hat{g}_{\infty}$, the Grenander estimator based on the uncensored data. We should note here that our practical experience has revealed that the identity of the Denby–Vardi iteration with the closed form formula quoted below requires a quite stringent convergence criterion in the algorithm.

In the uncensored case we follow Woodroffe and Sun (1993) in assuming all observations distinct and denote by $0 < y_1 < \dots < y_n < \infty$ the ordered values of 0 and the independent, identically distributed observations Y_1, \dots, Y_n from the decreasing density g . The NPMLE is then given as the left-continuous step function

$$\hat{g}_n(y) = \min_{0 \leq r \leq k-1} \max_{k \leq s \leq n} \frac{s-r}{n(y_s - y_r)}, \quad y_{k-1} < y \leq y_k,$$

$$\hat{g}_n(y) = 0 \text{ for } y > y_n.$$

Woodroffe and Sun explained that $\hat{g}_n(0+)$ does converge in probability, but that the limit is greater than $g(0+)$ with probability one. Because our real interest is in $\hat{S}(y) = \hat{g}(y)/\hat{g}(0)$, this issue requires further study.

Woodroffe and Sun proposed to use the penalized log-likelihood function

$$l_{\alpha}(g) = \sum_{i=1}^n \log g(y_i) - n\alpha g(0+),$$

where g varies across all decreasing left-continuous densities and $\alpha > 0$ is a penalty, or smoothing parameter, $\alpha = 0$ corresponding to the NPMLE. The penalized NPMLE has the surprisingly simple explicit representation as the NPMLE on the perturbed data points $\alpha + \gamma y_k$, $k = 1, \dots, n$, where γ is defined implicitly by

$$\gamma = \min_{k=1, \dots, n} \left(1 - \frac{\alpha k/n}{\alpha + \gamma y_k} \right);$$

Woodroffe and Sun provided an explicit formula for γ solving this equation. If $\alpha = \alpha_n \rightarrow 0$ such that $n\alpha_n \rightarrow \infty$ for $n \rightarrow \infty$, Woodroffe and Sun (1993) proved consistency of the PNPML \hat{g}_{α} and discussed its asymptotic distribution. For finite n , the follow-up paper by Sun and Woodroffe (1996) proposed using

$$\alpha = 0.649 \beta^{-1/3} n^{-2/3}$$

with

$$\beta = -\frac{1}{2}g(0)g'(0).$$

This leads to an adaptive procedure, switching between estimating β from the current estimate of g and estimating g from the current estimate of β . For the exponential distribution g with intensity λ we get $\beta = \frac{1}{2}\lambda^3$, where we have used the ad hoc estimate $\lambda \approx (\log 2)/(\text{empirical median})$. We have used this calculation in general, thus avoiding the adaptive iteration, after preliminary Monte Carlo studies which indicated that this choice works satisfactorily in a reasonably broad class of mixed exponential (Pareto) distributions g .

An alternative way out of the inconsistency for $\hat{g}(0+)$ (suggested to us by Hans van Houwelingen) is to study the conditional distribution of X given $X > x_0$, with x_0 small; since $\hat{g}(x_0)$ is consistent we can directly estimate $S_{x_0}(x) = S(x)/S(x_0)$ by $\hat{g}(x)/\hat{g}(x_0)$. In principle, asymptotic theory should be available for $x_{0n} \rightarrow 0$ at a suitably slow rate as $n \rightarrow \infty$.

4. PARAMETRIC MODELLING

For choosing a parametric model it is most convenient to start from a class $\{g_\theta(y)\}$ of decreasing densities with $g_\theta(0) < \infty$, such as mixtures of exponential densities. One such example, for which both the density $f(x)$ and the survival function $S(x)$ have explicit representations, is the *Pareto* model with density

$$g(y) = \lambda\mu(1 + \mu y)^{-\lambda-1}, \quad y > 0 \tag{4.3}$$

for $\lambda > 0$, $\mu > 0$, and survival function

$$S_g(y) = (1 + \mu y)^{-\lambda}.$$

If the current durations follow this distribution, then the prospectively observed times to pregnancy will have survival function

$$S(x) = g(x)/g(0) = (1 + \mu x)^{-\lambda-1}$$

which is a Pareto $(\lambda + 1, \mu)$ -distribution. This means that parameters of the TTP distribution may be fitted directly from the current durations, typically by maximum likelihood, and likelihood ratio tests may be used. In practice the heavy-tailed Pareto distribution can only be expected to serve satisfactorily if censored artificially at some maximum time of interest (typically $\frac{1}{2}$, 1, 2 or 3 years). However, censoring is easily accommodated. A current duration censored at y_0 contributes a factor $S_g(y_0) = (1 + \mu y_0)^{-\lambda}$ to the likelihood.

5. SIMULATION STUDIES OF THE NPMLE

For a brief numerical illustration, we simulated 999 replications of 100 Pareto-distributed times from the density $g(y)$ in 4.3 with $\lambda = 1.25$, $\mu = 0.2$. As seen in Figure 1(a), the NPMLE is too large at 0+, with $\hat{g}(0+)$ having 0.1, 0.5, 0.9 quantiles of 0.23, 0.49, 2.33, to be compared with the true $g(0) = 0.25$. Using $\hat{g}(x)/\hat{g}(0+)$ as estimate of $S(x) = g(x)/g(0)$ yields the unsatisfactory result of Figure 1(b), while Figure 1(c) shows that $\hat{g}(x)/\hat{g}(1.5)$ works well as an estimate of $S(x)/S(1.5) = g(x)/g(1.5)$. Figure 2 shows that the penalized NPMLE $\hat{g}_p(y)$ works well, with quantiles at 0+ of 0.157, 0.202, 0.269, although perhaps with a slight tendency of underestimation; accordingly $\hat{S}_p(x) = \hat{g}_p(x)/\hat{g}_p(0+)$ generally works well as an estimate of $S(x)$, although with a tendency to overestimate at 0+.

6. APPLICATION

For a practical illustration of the idea we consider a subset of data collected by The European Infertility and Subfecundity Study Group (EISSG), see Karmaus and Juul (1999). This study comprised detailed interviews with 6630 women, aged 25–44 years, at 14 locations in five European countries, sampled with the intention of being representative for the general populations in this age span. We restrict attention to 4473 interviewed women from Denmark, Germany (West and East) and Northern Italy. Of these, 3104 had ever attempted to become pregnant, and for each woman the ‘most recent time to pregnancy’ was identified. This waiting period was not necessarily concluded at the time of interview and was interpreted as a *historically prospective* follow-up study as defined above, defined as ending in a success if pregnancy had occurred, censored if the couple was still trying or had given up trying. We shall interpret the couples still trying at interview to be a cross-sectional sample of current durations. A study of this kind allows us to compare estimates from the current duration approach with the ‘truth’ based on the historically prospective study, disregarding possible temporal trends in the latter. The data contain information on

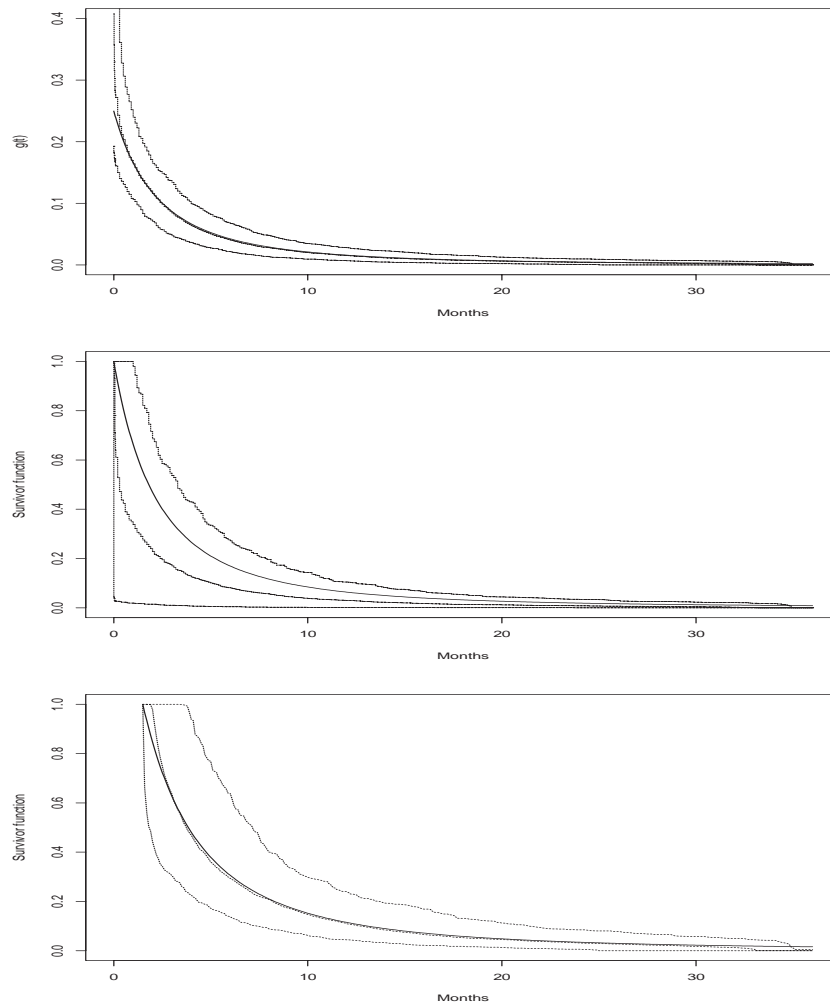


Fig. 1. (a) Observed current durations y simulated from Pareto density g with $\lambda = 1.25$, $\mu = 0.2$ (—). Sample size 100, 999 replications. Medians and pointwise 2.5% and 97.5% confidence limits (\cdots) of the NPMLE $\hat{g}(y)$. (b) Same observations: $S(x) = g(x)/g(0)$ and $\hat{S}(x) = \hat{g}(x)/\hat{g}(0)$ as estimate of survival function of time x to pregnancy. (c) Same observations: $S(x)/S(1.5) = g(x)/g(1.5)$ and $\hat{S}(x) = \hat{g}(x)/\hat{g}(1.5)$ as estimate of survival function of time x to pregnancy, given $x > 1.5$.

whether infertility treatment had been started and about pregnancy outcome, but this information is disregarded in this first illustration.

Certain conventions were necessary to obtain comparable data for the various estimation methods. As indicated earlier, the current duration approach requires stationarity and independence assumptions, in particular the survey plan has to be uninformative with respect to the TTP. Since the surveys were intended to be simple random samples of the female population aged 25–44 years, we restrict attention to attempts started at age 25 or later. Further, while EISSG took the pragmatic approach of including all periods of unprotected intercourse as time to pregnancy, we needed to focus on women actively attempting to become pregnant. Attention was therefore restricted to women answering ‘Yes’ to the question ‘Were

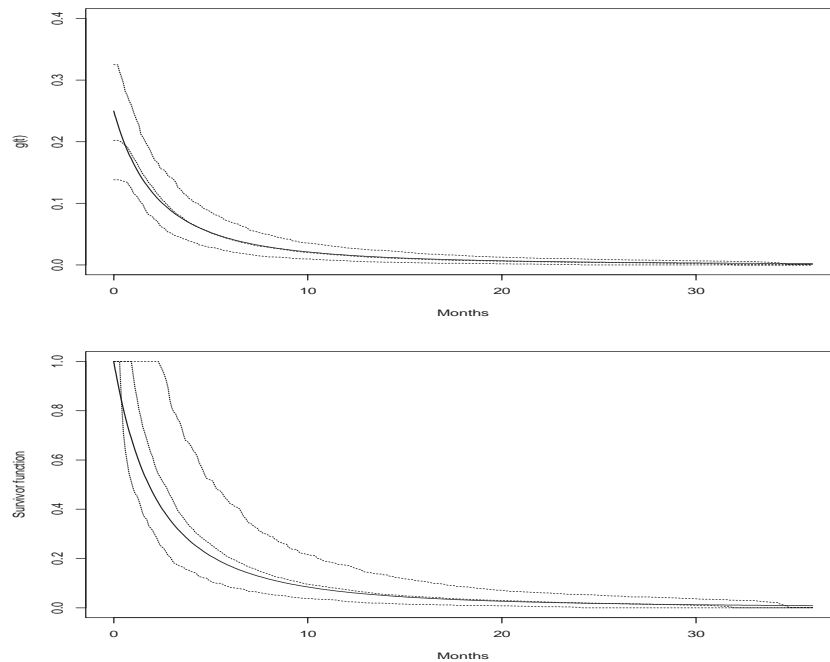


Fig. 2. (a) Same observations as in Figure 1. Penalized NPMLE $\hat{g}_p(x)$. (b) Same observations: estimated survival function $\hat{S}_p(x) = \hat{g}_p(x)/\hat{g}_p(0)$.

you planning to become pregnant?’ Finally we fixed the lag time in knowing whether or not you are pregnant at 0.7 months and reduced current durations for women ‘still trying’ at interview by that amount. In particular, current durations less than that limit were deleted.

The estimate based on current duration data is relevant for the distribution of waiting time to pregnancy T or end of unsuccessful attempt U , whichever comes first, that is $T \wedge U$. Also, since the current duration data carry no information on TTP of length zero, all such TTP were omitted.

6.1 Point estimates

Figure 3 shows the NPMLE and penalized NPMLE based on the 204 women who were still trying at interview, compared with the Kaplan–Meier estimate based on the most recent TTP (as qualified above) for the relevant 1456 women. As with the simulation results, the penalization correction is vital. The result after correction is reasonable but not fully satisfactory during the first few months. It should here of course be remembered that these historically prospective data possibly contain time trends that would not interfere in an actual cross-section.

Note that the focus would usually be on the time to successfully obtained pregnancy T , for which (still conditioning on $T > 0$) the Kaplan–Meier estimator of Figure 4 is obtained by censoring at $U \wedge V$, denoting unsuccessful end of attempt or end of follow-up. By definition, the estimate of Figure 4 is larger than that of Figure 3, but the difference is moderate.

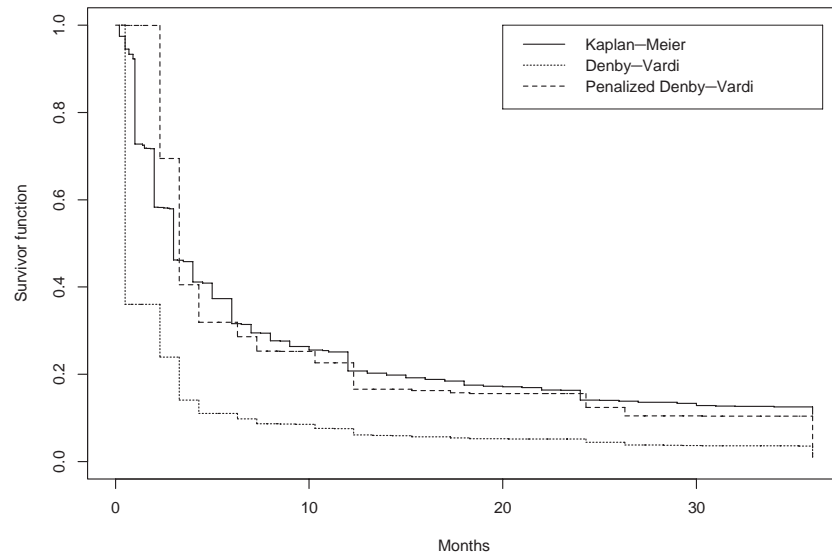


Fig. 3. EISSG data. Time to pregnancy, or end of unsuccessful attempt, conditional on it being positive. Kaplan–Meier estimator based on historically prospective data compared to Denby–Vardi and penalized Denby–Vardi estimators based on current durations.

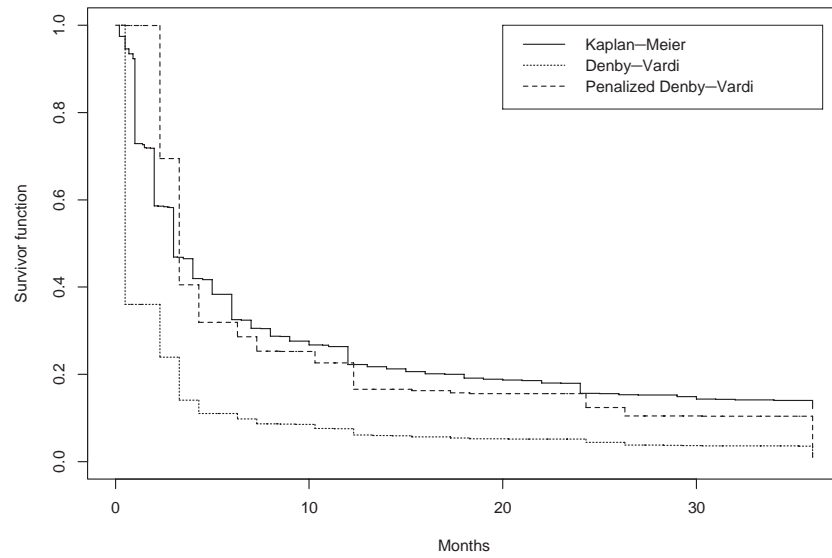


Fig. 4. EISSG data. As Figure 3, except that the Kaplan–Meier estimator now represents time to pregnancy, *censored* at end of unsuccessful attempt or end of follow-up.

6.2 Confidence limits

Pointwise confidence limits were obtained by drawing 999 bootstrap samples from the 204 women still trying, yielding the rather wide intervals in Figure 5.

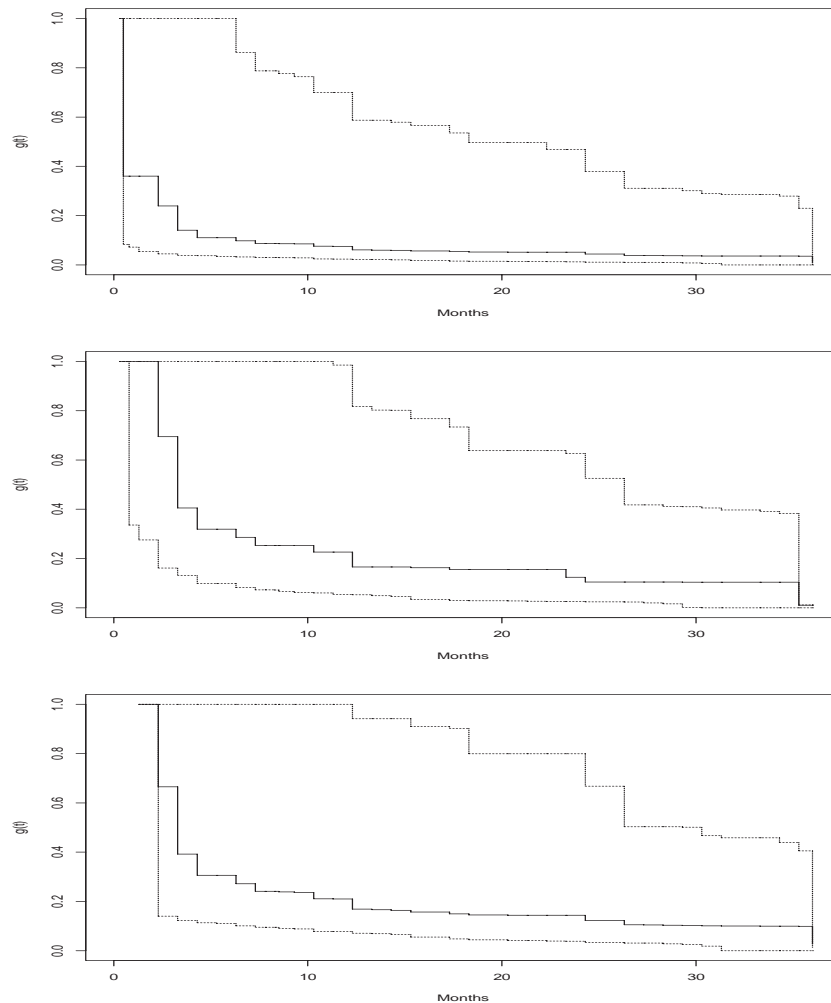


Fig. 5. (a) EISSG data: Denby–Vardi estimate with pointwise 95% bootstrap confidence intervals based on 999 replications. (b) EISSG data: Penalised Denby–Vardi estimate with pointwise 95% bootstrap confidence intervals based on 999 replications. (c) EISSG data: Conditional Denby–Vardi estimate with pointwise 95% bootstrap confidence intervals based on 999 replications.

6.3 Covariates

For an example of the effect of covariates, we group the women into smokers/nonsmokers (ever/never smoked) during the waiting time. Figures 6 and 7 show that the Kaplan–Meier and penalised current-duration estimates yield similar qualitative results: that smoking delays pregnancy. We take the opportunity also to compare with a (synthetic) retrospective approach based on the empirical distributions of those achieving pregnancy (Figure 8) with very similar results, so that for this covariate the problems about covariates for retrospective data (see Section 2) are minor.

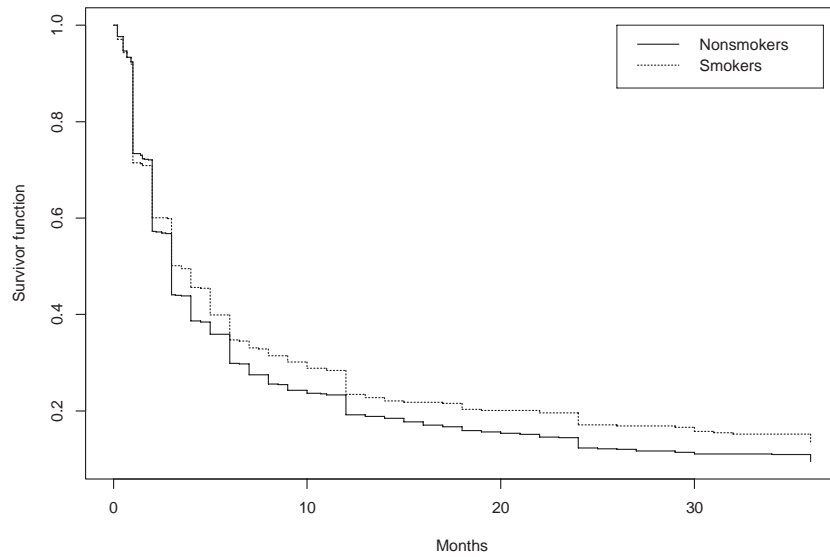


Fig. 6. EISSG data: effect of ever smoked during pregnancy. Time to pregnancy or end of unsuccessful attempt, conditional on it being positive. Kaplan–Meier (historically prospective) estimators.

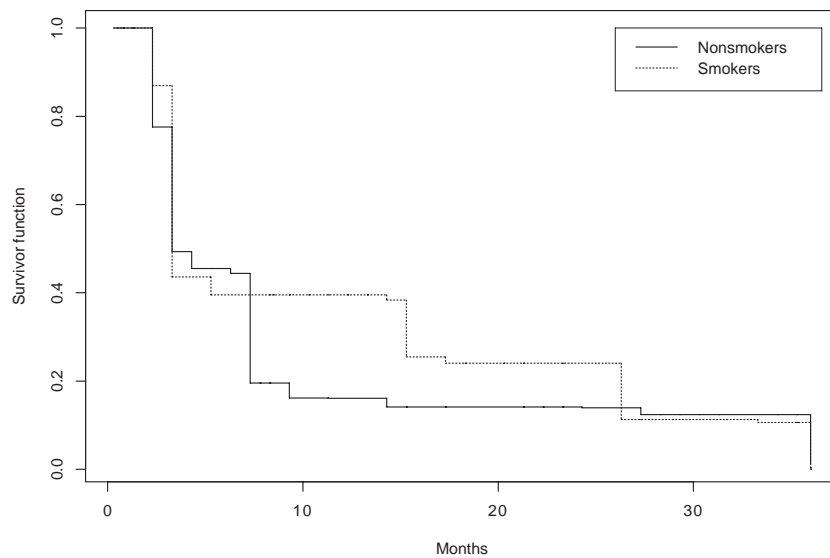


Fig. 7. EISSG data: effect of ever smoked during pregnancy. Penalized Denby–Vardi estimators based on current durations.

6.4 Parametric estimation

Finally, we fitted the Pareto model described above, experimenting with censoring at 15 months and 36 months. Table 1 contains the estimates based on censoring at 15 months, for which the fit was tolerable. The relatively few current durations cannot support particularly precise statements. The parameters λ and

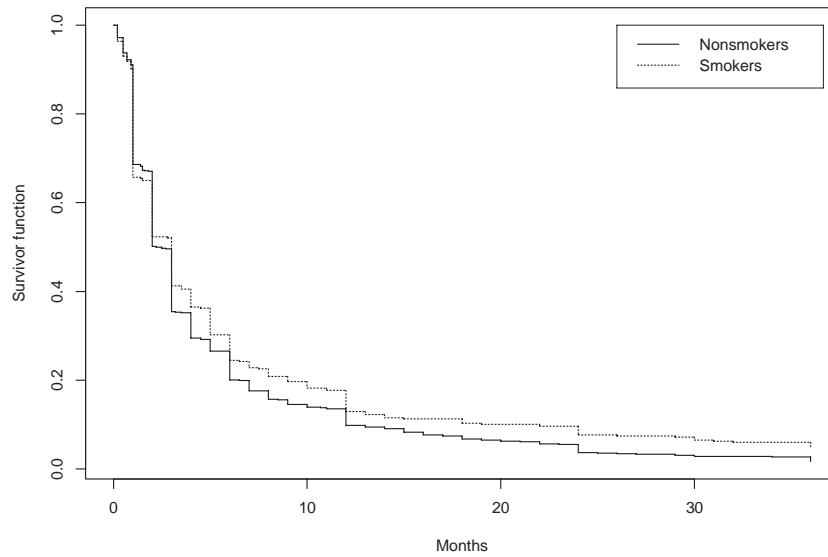


Fig. 8. EISSG data: effect of ever smoked during pregnancy. Kaplan–Meier (retrospective) estimators based on successful pregnancies.

Table 1. Pareto distributed times to pregnancy: parameter estimates based on prospective data and on current-duration data, both censored at 15 months

| | Current duration | | | | Prospective | | | |
|-------------------------|------------------------------------|------------------------|--------------------|-----------|---------------------------------------|------------------------|--------------------|-----------|
| | <i>n</i> | $\hat{\lambda}$ (s.e.) | $\hat{\mu}$ (s.e.) | $-\log L$ | <i>n</i> | $\hat{\lambda}$ (s.e.) | $\hat{\mu}$ (s.e.) | $-\log L$ |
| Smokers | 76 | 1.58(0.22) | 0.148(0.088) | 152.61 | 512 | 1.05(0.15) | 0.230(0.048) | 1130.5 |
| Non-smokers | 130 | 1.29(0.42) | 0.569(0.110) | 235.31 | 937 | 1.38(0.15) | 0.184(0.029) | 2063.1 |
| All | 206 | 1.32(0.37) | 0.429(0.065) | 389.32 | 1449 | 1.24(0.11) | 0.200(0.025) | 3211.4 |
| Test for smoking effect | $\chi^2 = 2.8, d.f. = 2, P = 0.25$ | | | | $\chi^2 = 35.6, d.f. = 2, P < 0.0005$ | | | |

Table 2. Estimated medians and upper quartiles (months)

| | Current duration | | Prospective | |
|-------------|------------------|--------------|-------------|--------------|
| | Median | Upper quart. | Median | Upper quart. |
| Smokers | 3.7 | 9.5 | 4.1 | 11.9 |
| Non-smokers | 1.3 | 3.4 | 3.5 | 9.4 |
| All | 1.6 | 4.3 | 3.7 | 10.3 |

μ are not so descriptive, so Table 2 contains estimates of medians and upper quartiles. These are rather different in the two approaches, although reasonably consistent with the results of the nonparametric analyses, particularly the concept that smoking delays pregnancy.

7. DISCUSSION

The obvious difficulties in currently used designs for time-to-pregnancy studies (prospective, historically prospective, retrospective) were outlined in the introduction and have been extensively discussed in the literature (Weinberg *et al.*, 1993, 1994; Scheike and Jensen, 1997; Scheike and Ekstrøm, 1998; Scheike *et al.*, 1999; Olsen *et al.*, 1998; Olsen, 1999). These difficulties motivate the investigation of new designs, or combination of designs. We have here given a first statistical and empirical evaluation of what we call the *current duration design*, which should be very simple to carry out on a standard population survey basis.

Fecundability studies based on TTP data have well-known and well-documented problems in data validity, hidden selection and confounding, etc. We will however here focus on the *statistical* aspects.

We have studied nonparametric inference, which turned out to require implementation of recent asymptotic theory of penalized nonparametric maximum likelihood. The results were technically encouraging, both in Monte Carlo simulations and in practice, although much larger sample sizes than in our illustration would be required to get sufficiently precise estimates. We have also considered parametric modelling in this problem through the mathematically convenient Pareto model, allowing direct likelihood inference from (possibly censored) current durations as well as prospective data. These results are preliminary, and other classes of distributions might fit better.

There are still several issues open for additional work. First, the recent breakthrough by Banerjee and Wellner (2001) in asymptotic nonparametric likelihood theory for current-status data might well be adapted to the nonparametric likelihood for current-duration data. Secondly, van Es *et al.* (2000) outlined a research programme based on powerful general semiparametric technology for models derived from proportional hazards or accelerated failure time models for $f(x|z)$. This programme has yet to be implemented, and one probably cannot hope to find solutions as elegant as those of Lin *et al.* (1998) and Martinussen and Scheike (2002) for the somewhat similar current status data problem.

We see the main strength of the current duration design in its ease of application, in principle allowing enormous sample sizes. Such studies could provide sentinels for effects of certain risk factors on fecundity, suggesting more intensive and costly, focused studies.

However, the current duration design will usually be insufficient as the only source of information on fecundability. It cannot catch those who never waited, nor can it distinguish between a TTP concluded with a pregnancy and one concluded without. It might often be natural to combine information from the current duration design with evidence from the relatively easily obtained retrospective pregnancy-based data.

As another possibility, Olsen and Andersen (1999) recently suggested a *combination design* based on case-cohort ideas. Their idea is, within a well-defined cohort, to combine a survey at t_0 of those currently waiting with a monitoring of all subsequent pregnancies in a six month to one year period after t_0 .

ACKNOWLEDGEMENT

This research was supported by Public Health Services Grant R01-CA54706-07 from the National Cancer Institute.

REFERENCES

- BANERJEE, M. AND WELLNER, J. A. (2001). Likelihood ratio tests for monotone functions. *Technical Report. Annals of Statistics* **29**, 1699–1731.
- BONDE, J. P. E., ERNST, E., JENSEN, T. K., HJOLLUND, N. H. I., KOLSTAD, H., HENRIKSEN, T. B., SCHEIKE, T., GIWERCMAN, A., OLSEN, J. AND SKAKKEBÆK, N. E. (1998). Relation between semen quality and fertility: a population-based study of 430 first-pregnancy planners. *Lancet* **352**, 1172–1177.

- COX, D. R. (1969). Some sampling problems in technology. In Johnson, N. L. and Smith, H. (eds), *New Developments in Survey Sampling*, New York: Wiley, pp. 506–527.
- DENBY, L. AND VARDI, Y. (1986). The survival curve with decreasing density. *Technometrics* **28**, 359–367.
- GRENANDER, U. (1956). On the theory of mortality measurement, part II. *Skandinavisk Aktuarietidskrift* **39**, 125–153.
- JUUL, S., KEIDING, N. AND TVEDE, M. (2000). Retrospectively sampled time-to-pregnancy data may make age-decreasing fecundity look increasing. *Epidemiology* **11**, 717–719.
- KARMAUS, W. AND JUUL, S. (1999). Infertility and subfecundity in population-based samples from Denmark, Germany, Italy, Poland and Spain. *European Journal of Public Health* **9**, 229–235.
- KEIDING, N. (1991). Age-specific incidence and prevalence: a statistical perspective (with discussion). *Journal of the Royal Statistical Society Series A* **154**, 371–412.
- KEIDING, N. (1992). Independent delayed entry (with discussion). In Klein, J. P. and Goel, P. K. (eds), *Survival Analysis: State of the Art.*, Dordrecht: Kluwer, pp. 309–326.
- KEIDING, N. (1999). Analysis of time-to-pregnancy data. *Scandinavian Journal of Work, Environment & Health* **25** Suppl. 1, 10–11.
- KEIDING, N. AND GILL, R. D. (1990). Random truncation models and Markov processes. *Annals of Statistics* **18**, 582–602.
- LIN, D. Y., OAKES, D. AND YING, Z. (1998). Additive hazards regression with current status data. *Biometrika* **85**, 289–298.
- LUND, J. (2000). Sampling bias in population studies—How to use the Lexis diagram. *Scandinavian Journal of Statistics* **27**, 589–604.
- MARTINUSSEN, T. AND SCHEIKE, T. H. (2002). Efficient estimation in additive hazards regression with current status data. *Biometrika* **89**, (to appear).
- OLSEN, J. (1999). Design options and sources of bias in time-to-pregnancy studies. *Scandinavian Journal of Work, Environment & Health* **25** Suppl 1, 5–7.
- OLSEN, J. AND ANDERSEN, P. K. (1999). We should monitor human fecundity, but how? A suggestion for a new method that may also be used to identify determinants of low fecundity. *Epidemiology* **10**, 419–421.
- OLSEN, J., JUUL, S. AND BASSO, O. (1998). Measuring time to pregnancy. Methodological issues to consider. *Human Reproduction* **13**, 1751–1756.
- SCHEIKE, T. H. AND EKSTRØM, C. T. (1998). A discrete survival model with random effects and covariate-dependent selection. *Applied Stochastic Models and Data Analysis* **4**, 153–163.
- SCHEIKE, T. H. AND JENSEN, T. K. (1997). A discrete survival model with random effects: An application to time to pregnancy. *Biometrics* **53**, 318–329.
- SCHEIKE, T. H., PETERSEN, J. H. AND MARTINUSSEN, T. (1999). Retrospective ascertainment of recurrent events: An application to time to pregnancy. *Journal of the American Statistical Association* **94**, 713–725.
- SUN, J. AND WOODROOFE, M. (1996). Adaptive smoothing for a penalized NPMLE of a non-increasing density. *Journal of Statistical Planning and Inference* **52**, 143–159.
- VAN ES, B., KLAASSEN, C. A. J. AND OUDSHOORN, K. (2000). Survival analysis under cross sectional sampling: length bias and multiplicative censoring. *Journal of Statistical Planning and Inference* **91**, 295–312.
- WEINBERG, C. R., BAIRD, D. D. AND ROWLAND, A. S. (1993). Pitfalls inherent in retrospective time-to-event studies: The example of time to pregnancy. *Statistics in Medicine* **12**, 867–879.

- WEINBERG, C. R., BAIRD, D. D. AND WILCOX, A. J. (1994). Sources of bias in studies of time to pregnancy. *Statistics in Medicine* **13**, 671–681.
- WEINBERG, C. R. AND DUNSON, D. B. (2000). Some issues in assessing human fertility. *Journal of the American Statistical Association* **95**, 300–303.
- WOODROOFE, M. AND SUN, J. (1993). A penalized maximum likelihood estimate of $f(0+)$ when f is non-increasing. *Statistica Sinica* **3**, 501–515.

[Received June 21, 2000; first revision November 26, 2001; second revision February 19, 2002;
accepted for publication February 25, 2002]