Scandinavian Journal of Statistics Theory and Applications

doi: 10.1111/j.1467-9469.2012.00798.x

The Current Duration Approach to Estimating Time to Pregnancy

NIELS KEIDING, OLUF K. HØJBJERG HANSEN and DITTE NØRBO SØRENSEN

Department of Biostatistics, University of Copenhagen

RÉMY SLAMA

Team of Environmental Epidemiology Applied to Reproduction and Respiratory Health, Inserm-Univ. J. Fourier (U823) Joint Research Center, Grenoble, France

ABSTRACT. Time-to-pregnancy (TTP) is the duration from the time a couple starts trying to become pregnant until they succeed. It is considered one of the most direct methods to measure natural fecundity in humans. Statistical tools for designing and analysing time to pregnancy studies belong to survival analysis, but several features require special attention. Prospective designs are difficult to carry out and retrospective (pregnancy-based) designs, being widely used in this area, do not allow efficiently including couples remaining childless. A third possible design starts from a cross-sectional sample of couples currently trying to become pregnant, using *current duration* (backward recurrence time) as basis for the estimation of TTP. Regression analysis is then most conveniently carried out in the accelerated failure time model. This paper surveys some practical and technical-statistical issues in implementing this approach in a large telephone-based survey, the *Epidemiological Observatory of Fecundity in France (Obseff)*.

Key words: accelerated failure time, backward recurrence time, censoring, competing risks, fecundity, fertility treatment, generalized Gamma distribution, isotonic estimator, open birth intervals, Pareto distribution, reproductive epidemiology

1. Introduction

TTP is the duration from the time a couple starts attempting to become pregnant until they succeed. It is considered one of the most direct methods to measure natural (biological) fecundity in humans. The tools for designing and analysing TTP-studies form a special subfield of reproductive epidemiology, relying on careful use of survival analysis techniques, for a review, see Scheike & Keiding (2006).

As is usual in epidemiology, *prospective (follow-up) designs* are the easiest to analyse, but a purely incident cohort study is difficult to organize, primarily because of the fact that some couples may not plan to start a pregnancy attempt long in advance.

There are variants of the prospective design in which recruitment is easier. A particular version is the *historically prospective design*, in which couples (in fact, usually women) are asked to recall how long they attempted to become pregnant for well-defined occasions, and how each of these attempts ended. Such surveys may be influenced by recall biases, but may, in principle, be analysed as prospective follow-up studies.

The most common *retrospective* design is based on asking women who have become pregnant, how long it took them. This is easier to conduct, but hard to analyse, because it conditions not only on success of the attempt (hence excluding infertile and underrepresenting sub-fertile), but also on couples *not having given up*, a somewhat under-emphasized feature which can seriously distort effect estimates for determinants (such as the age of the woman) that may affect both fecundity and the risk of giving up trying (Basso *et al.*, 2000; Juul *et al.*, 2000).

In view of the practical difficulties with the prospective design and the interpretative problems with the pregnancy-based design Keiding *et al.* (2002) described the possibility of basing the design on a cross-sectional survey of couples (women) who are asked whether they are currently attempting to get pregnant, and if so, for how long have they attempted. Based on the distribution of this *current duration* it is in principle possible to derive the distribution of TTP. This idea had earlier been proposed by Weinberg & Gladen (1986).

Epidemiologically the current duration design shares with the prospective design the ability to include infertile and subfertile couples, as well as the inability to register *accidental pregnancies*. These are pregnancies without a TTP, which are sometimes considered to correspond to a TTP of zero, a probably very strong assumption. But like the pregnancy-based design, the survey will only identify those who have not given up trying.

The purpose of this paper is to discuss the statistical aspects of the current duration design. We base the exposition on experiences primarily gathered around the design and analysis of the Observatoire Épidémiologique de la Fertilité en France (Observatory of Fecundity in France, Obseff) (Slama *et al.*, in press).

The idea of inferring the distribution of a duration variable from the distribution of incomplete durations until a given time has been discussed for many years, particularly in social science and demography contexts. A prominent example is the 'open birth intervals' – the duration between the last birth and the time of a survey. In an important paper Allison (1985) summarized these developments and focused on the connection to survival analysis, as we shall also do here.

Other examples of the modelling problem include the study by Ali *et al.* (2001) of current duration of use of contraceptive pills. Yamaguchi (2003), working on residential mobility, proposed using a generalized Gamma model for the 'last episode' duration since the last move before a survey. He then went on to introduce the *accelerated failure time (AFT) regression model* which allows direct inference on the regression coefficients in the underlying mobility model from a fit to the durations. Cristóbal *et al.* (2007) used local polynomial smoothing to develop non-parametric regression estimators based on current duration data.

In sections 2–4 we introduce the two classical designs of TTP studies: prospective and retrospective (pregnancy-based) as well as the proposed current duration design. We focus throughout on concrete details such as adequate statistical modelling of the fact that many couples give up before they succeed in getting pregnant and how to handle couples who initiate fertility treatment. We outline non-parametric and parametric statistical modelling of current duration data in sections 5 and 6, emphasizing the sensitivity of the estimated densities of current duration near zero. Following Yamaguchi (2003), we describe in section 7, AFT models for regression analysis of current duration data; section 8 is devoted to a survey of censoring problems for the current duration approach; section 9 illustrates the developed methods with our experience from the French Observatory of Fecundity, and section 10 contains a brief discussion of the current duration approach in this context, with a few references to other recent applications of the approach.

2. Classical designs for TTP analysis: prospective design

We start the exposition with the *prospective cohort design*, following a group of couples from initiation until conception

$$\text{initiation} \xrightarrow{\pi(t)} \text{conception} \tag{1}$$

with $t =$ duration since initiation and $\pi(t)$ the hazard rate of conception.

This is a simple survival analysis problem where some durations may be 'administratively' censored because conception had not yet happened at the end of study.

However, the attempt to become pregnant may be *discontinued* before conception happens. Discontinuation can have many explanations: the couple may split or one of them may die, they may give up after waiting very long, the woman may become too old. (A special role is played by various types of initiating fertility treatment, to which we return later).

The simple survival analysis (1) is then modified to the competing risks situation

$$\text{initiation} \begin{cases} \xrightarrow{\pi(t)} \text{conception} \\ \xrightarrow{\varphi(t)} \text{giving up} \end{cases} \tag{2}$$

where administrative censoring may again happen.

Most discussions of the prospective design in the literature (e.g. Buck *et al.*, 2004; Tingen *et al.*, 2004) seem to aim for an estimate of the partial survival function $\exp(-\int_0^t \pi(s)\,\mathrm{d}s)$, that is, regarding 'giving up' as independent censoring of the biologically determined TTP distribution in a world where nobody gives up.

One will however expect that both intensities $\pi(t)$ and $\varphi(t)$ will often be affected by the same covariates (e.g. age of the woman), so that a necessary condition for independent censoring will be that these covariates are included in the statistical model in which case censoring may be considered conditionally independent given these covariates. See Andersen & Keiding (2012) for a recent exposition of interpretations in competing risks analysis.

Another possibility is to take a strict competing risks point of view, focusing on the *cumulative incidence*, that is, the probability

$$\int_0^t \exp\left(-\int_0^s (\pi(u)+\varphi(u))\,\mathrm{d}u\right)\pi(s)\,\mathrm{d}s$$

that conception happens before duration t among all who started at duration 0. This quantity is not used much in practice (if at all), perhaps because it has lost the desired property of the 'pure' TTP of depending mainly on biological, as opposed to behavioural, determinants.

The main problem with the prospective design is the difficulty of recruiting participants, particularly if these are intended to be representative of a well-defined population (Joffe *et al.*, 2005). As a result, prospective studies in the pure sense are usually quite small (a few hundred participants), see Bonde *et al.* (1998) and Buck *et al.* (2004).

2.1. Alternative prospective designs

A modification of the prospective design is the *historically prospective design* where couples (in practice usually women) are asked to recall a specific pregnancy attempt (such as the first, or the most recent, possibly still ongoing) and report how long they tried and if and how the attempt ended. For an example see Karmaus & Juul (1999).

Another modification could be termed a *prevalent cohort design* where couples identified at a cross-sectional sample as currently trying (with known initiation time) are followed up. An interesting web-based version of this design is currently running in Denmark (Mikkelsen *et al.*, 2009; Wise *et al.*, 2011).

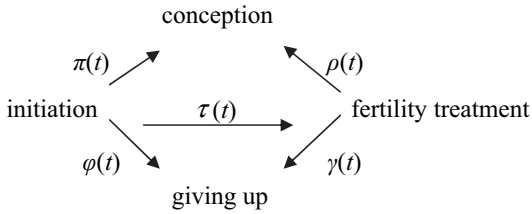
Analytically both of these modifications are analogous to the 'simple' or 'pure' prospective design above, using proper delayed entry techniques.

2.2. Prospective design and fertility treatment

Couples who have waited for some time to conceive will often seek medical help to become pregnant. This can take various forms, from medical consultation leading to prescriptions of hormones or other drugs, to full-fledged assisted reproduction procedures.

In the present study we focus on the formal problems around including decisions to seek medical help by restricting attention to one general action 'fertility treatment', the precise practical interpretation of which will vary according to context.

We then have the multi-state model



(3)

where $\rho(t)$ and $\gamma(t)$ may of course also depend on duration since start of fertility treatment.

Two issues have arisen here: first, the original purpose of TTP studies was to approximate a direct study of couples' natural biological fecundity. One may argue that it is difficult to maintain this interpretation if medical intervention is allowed, even though such intervention may be increasingly common in today's world. On the other hand, it is not without problems either to consider fertility treatment as censoring or as competing risk.

If one regards fertility treatment as censoring, typically a hypothesis of conditionally independent censoring needs to be justified. In this case, the statistical model must contain those relevant risk factors (age of the woman being again a generic example) that may affect both the chance of conception and the inclination to seek fertility treatment.

Another possibility would be to focus on a cumulative incidence-type measure such as probability of (conceiving before t, before giving up and before starting fertility treatment)

$$= \int_0^t \exp\left(- \int_0^s (\pi(u) + \varphi(u) + \tau(u))\,\mathrm{d}u \right) \pi(s)\,\mathrm{d}s$$

but we would then get even further away from the original intention of interpreting TTP as a reasonably pure indicator of biological fecundity.

2.3. Time to fertility treatment

A variant of the above analysis considers time to fertility treatment in the prospective design. In the framework of model (3) Moreau *et al.* (2010) essentially interpreted conception and giving up as censorings, focusing on the partial survival function $\exp(-\int_0^t \tau(s)\,\mathrm{d}s)$. This has the worrisome interpretation of describing waiting time to initiating fertility treatment in a world where conceptions (and giving up) never occur.

Here, an interesting cumulative incidence-type quantity may be the probability of initiating fertility treatment before time t among all couples starting an attempt to become pregnant:

$$\int_0^t \exp\left(- \int_0^s (\pi(u) + \varphi(u) + \tau(u))\,\mathrm{d}u \right) \pi(s)\,\mathrm{d}s.$$

This is being studied by Duron-Martinaud *et al.* (submitted), using a prevalent cohort design.

3. Classical designs for TTP analysis: retrospective design

The typical retrospective design is based on pregnant women, interviewed at a maternity clinic or around birth of the child. Statistically this means that we *condition* on ending in conception in model (2):

$$\text{initiation} \begin{cases} \xrightarrow{\pi(t)} & \text{conception} \\ \xrightarrow{\varphi(t)} & \text{giving up.} \end{cases}$$

Although most sources on retrospective TTP are surprisingly implicit about this, the intended target distribution is presumably again the partial density

$$\pi(t)\exp\left(-\int_0^t \pi(s)\,\mathrm{d}s\right)$$

generated by the hazard $\pi(t)$. This is TTP in a world where nobody gives up.

However, the retrospective design estimates the distribution of conception given that it happens before giving up, with density

$$\frac{\pi(t)\exp\left(-\int_0^t [\pi(s)+\phi(s)]\,\mathrm{d}s\right)}{\int_0^\infty \pi(t)\exp\left(-\int_0^t [\pi(s)+\phi(s)]\,\mathrm{d}s\right)\mathrm{d}t} \tag{4}$$

where there is usually no direct information on the distribution of time to giving up.

The consequences were exemplified by Scheike & Keiding (2006); for a simple example, assume that $\pi_a(t)$ and $\varphi_a(t)$ are both constant in t but depend on the woman's age a at initiation. This corresponds to assuming that the model (2) is generated from independent exponential waiting times to conception (T) and giving up (U) with hazards π_a and φ_a (we omit the subscript a for ease of notation). The joint distribution (T, U) has density

$$\pi\varphi\,\mathrm{e}^{-\pi t-\varphi u}$$

giving the conditional distribution $(T, U \mid T<U)$ the density

$$(\pi+\varphi)\varphi\,\mathrm{e}^{-\pi t-\varphi u}$$

and the density (4) of the observed distribution $T \mid T<U$ reduces to

$$(\pi+\varphi)\,\mathrm{e}^{-(\pi+\varphi)t}.$$

Compared to the partial density

$$\pi\,\mathrm{e}^{-\pi t}$$

we see that we get an exponential distribution with a higher hazard: $\pi+\varphi$ rather than π.

There is a solid biological evidence that the conception rate π_a decreases with age a, particularly for $a>35$ years (Fédératíon CECOS *et al.*, 1982; Dunson *et al.*, 2004). However, if the hazard φ_a of giving up increases sufficiently fast with age a, this may well outweigh the decrease in π_a, yielding an apparent, counter-intuitive *increase with age* of the conception rate, as often empirically observed (e.g. Jensen *et al.*, 2000) and illustrated in a Monte-Carlo simulation assuming a heterogeneous fecundity pattern in the population (Juul *et al.*, 2000). Independently of Jensen *et al.* (2000), Basso *et al.* (2000) gave a mathematical analysis in discrete time similar to the one in continuous time above as well as a Monte-Carlo simulation assuming heterogeneous but age-independent fecundabilities across the population. Basso *et al.* (2000) also analysed

data from a European study, showing that the hazard of giving up was strongly increasing with age in Denmark but age-independent in Northern Italy.

3.1. Censoring

In addition to the special character of the event of giving up trying, there are other, less controversial types of censoring in TTP studies.

Purely administrative censoring (end of study), so well known from prospective follow-up, is in principle absent from the retrospective approach. However, it is quite common to focus on the beginning of the estimated TTP distribution and artificially censor it after about one or two years, motivated partly by lack of confidence in the precision of large retrospectively recalled TTPs, partly by the less interesting nature and validity of the right tail of the TTP distribution. Finally, by focusing on the first year of the TTP distribution, most of the interference of fertility treatment will be avoided.

4. Current duration design

In the current duration design a sample of women (formally: couples) are asked whether they are currently trying to become pregnant, and if so, for how long they have tried. This design was first proposed by Weinberg & Gladen (1986) who used discrete time models. We follow here the continuous time derivations by Keiding *et al.* (2002), see van Es *et al.* (2000).

In the simplest model (2) the survey will reach those who are currently trying because they have not had a conception *and* they have not given up. Using again T for time to conception and U for time to giving up we hence focus on the underlying distribution of $X = T \wedge U$.

(Note that in the simple example of independent exponentials X has density $(\pi + \varphi)\,\mathrm{e}^{-(\pi+\varphi)x}$; exactly the same as the imposed target distribution in the retrospective design. In section 8, we return to the discussion of other censoring problems, such as how to handle fertility treatment).

In the current duration design one observes Y, the time elapsed since initiation, in renewal process contexts termed *backward recurrence time*. If initiations happen in calendar time according to a non-homogeneous Poisson process with intensity $\beta(t)$, then the density $g(y, t_0)$ of the current durations at time t_0 is proportional to $\beta(t_0 - y)S(y)$, where $S(x) = \int_x^\infty f(u)\,\mathrm{d}u$ is the survival function of $X = T \wedge U$ corresponding to the density function $f(x)$. We assume finite mean of $X : E(X) = \int_0^\infty S(x)\,\mathrm{d}x < \infty$.

Over the short time spans usually under study it will be reasonable to assume steady state with $\beta(t) = \beta$ constant, in which case the density

$$g(y, t_0) = g(y) = S(y)/E(X)$$

is decreasing in y and finite at $0 : 0 < g(0) < \infty$. The survival function $S(x)$ may thus be directly obtained from an estimate $\hat{g}$ of g as

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

5. Non-parametric estimation in the current duration design

As surveyed by Keiding *et al.* (2002) non-parametric maximum likelihood estimation (NPMLE) of the decreasing density $g(x)$, finite at 0, has been studied carefully since Grenander (1956). The key problem identified by Woodroofe & Sun (1993) is that the NPMLE $\hat{g}(x)$ is inconsistent at 0: the probability limit for large samples of $\hat{g}(0+)$ is larger than $g(0)$.

Balabdaoui *et al.* (2011) offered the interesting intuitive explanation of this inconsistency that since the NPMLE does not know whether or not $g(0)$ is finite, it overshoots.

Sun & Woodroofe (1996) proposed a penalized NPMLE consistent at 0 with an intricate evaluation of the necessary calibration constants.

Another possibility is to note that since $\hat{g}(\varepsilon)$ is consistent for $\varepsilon > 0$, we always have the consistent estimator $\hat{g}(x)/\hat{g}(\varepsilon)$ of the conditional survival function $S(x)/S(x)S(\varepsilon)$ of X given that $X > \varepsilon$. This was suggested by Hans van Houwelingen (see Keiding *et al.*, 2002) and carried through mathematically by Kulikov & Lopuhaä (2006). Keiding *et al.* (2002) illustrated the consistency problems on simulated data as well as on data from a large European TTP study; see the next section for further graphical illustrations.

Finally, Pal (2009) combined the above approaches with that of Banerjee & Wellner (2001) to obtain a parameter-free limit distribution of the penalized likelihood ratio.

6. Parametric estimation in the current duration design

The non-parametric approach in the previous section is not yet quite ready for routine practical use, and we have explored possible parametric models suitable for this purpose: remember that the density $g(y)$ of current duration needs to be decreasing and finite at 0. This rules out the log-normal and almost all Weibull models.

Keiding *et al.* (2002) noted that if Y is Pareto (λ, μ), that is, with density

$$g(y) = \frac{\lambda\mu}{(1+\mu y)^{\lambda+1}}$$

and with survival function $(1+\mu y)^{-\lambda}$, then the survival function of X is

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

or Pareto $(\lambda+1, \mu)$.

Observe that the Pareto distribution may be interpreted as a Gamma-mixture of exponential distributions: if $Y \mid \theta$ is exponential, $g(y \mid \theta) = \mathrm{e}^{-\theta y}$, and θ is Gamma (λ, μ), then Y is Pareto (λ, μ). Hence, this choice corresponds to Weinberg and Gladen's (1986) discrete-time choice of a beta-mixture of geometric distributions.

Yamaguchi (2003) parameterized the distribution of $\log X$ by the more flexible location-scale family of generalized Gamma distributions earlier studied by Farewell & Prentice (1977). In addition to location and scale, there is a shape parameter λ, giving the underlying density for $W = (\log X - \mu)/\sigma$, as

$$h(w) = \frac{|\lambda|}{\Gamma(\lambda^{-2})}(\lambda^{-2})^{\lambda^{-2}} \exp[\lambda^{-2}(\lambda w - \mathrm{e}^{\lambda w})], \quad \lambda \neq 0$$

$$= \frac{1}{\sqrt{2\pi}} \exp(-w^2/2), \quad \lambda = 0.$$

See Yamaguchi (2003) for explicit expressions for the density g of Y.

Maximum likelihood estimation is directly feasible for these parametric models; because of the complexity of the generalized Gamma likelihood we calculate confidence bounds using a bootstrap approach.

6.1. Illustrations of NPMLE

Figure 1 shows the true density function g of current durations (black curve) distributed according to a generalized Gamma distribution with realistic parameter values ($\lambda = 0.7$,

$\mu = 0.6, \sigma = 2.2$). Ten data sets ($n = 1000$) of current durations were simulated from this distribution and fitted by maximum likelihood (red curves). The blue curves are the (unpenalized) NPMLE. The values of $\hat{g}(0)$ are indicated by tick-marks; the true value $g(0) = 0.17$, the estimated $\hat{g}(0)$:

	Simulation									
	1	2	3	4	5	6	7	8	9	10
Gen. Gamma fit	0.15	0.17	0.19	0.19	0.17	0.20	0.14	0.15	0.12	0.16
NPMLE	0.18	0.45	1.58	0.42	0.14	0.32	0.13	0.19	0.44	5.11

Figure 2 shows the corresponding true survival function of TTP (black curve). The corresponding estimated generalized Gamma-based survival functions for TTP are the red curves, which fit nicely. The blue curves are $\hat{g}(x)/\hat{g}(0)$ with $\hat{g}$ the unpenalized NPMLE. The overshoot of $\hat{g}(0)$ implies that some of the curves are too low.

Figure 3 shows the same curves, conditional on $T \geq 1$ and now fitting well.

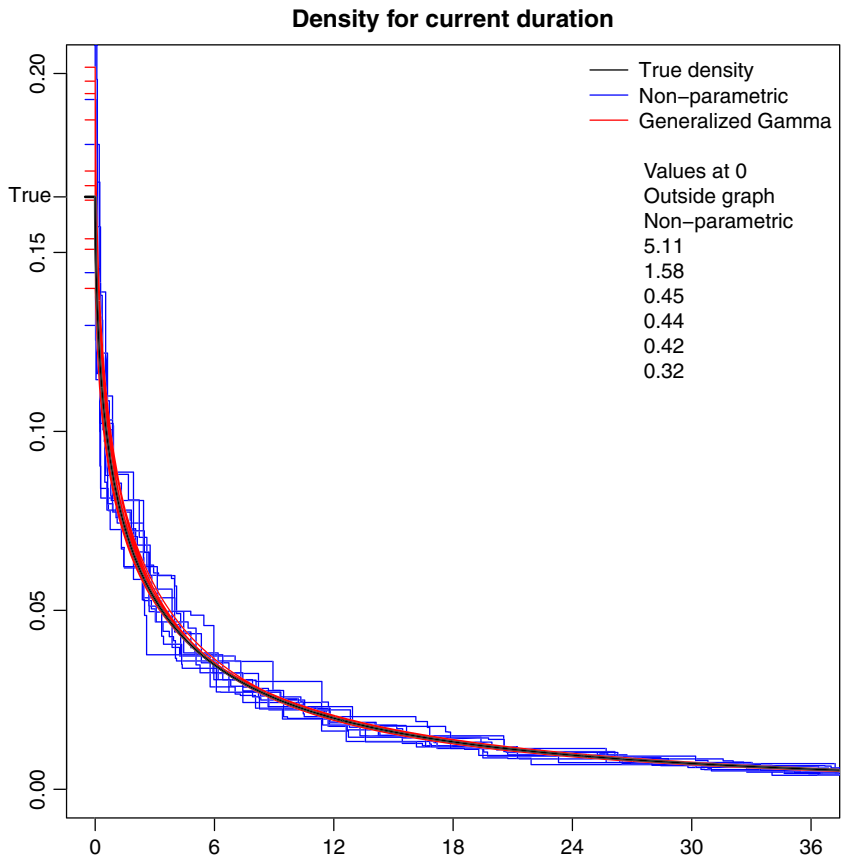


Fig. 1. Generalized Gamma density function ('true', black) from which ten samples ($n = 1000$) were simulated. Generalized Gamma fits (red) and NPMLE fits (blue). True $g(0) = 0.17$, estimated $g(0)$ indicated by tick marks and in the legend if outside the graph.

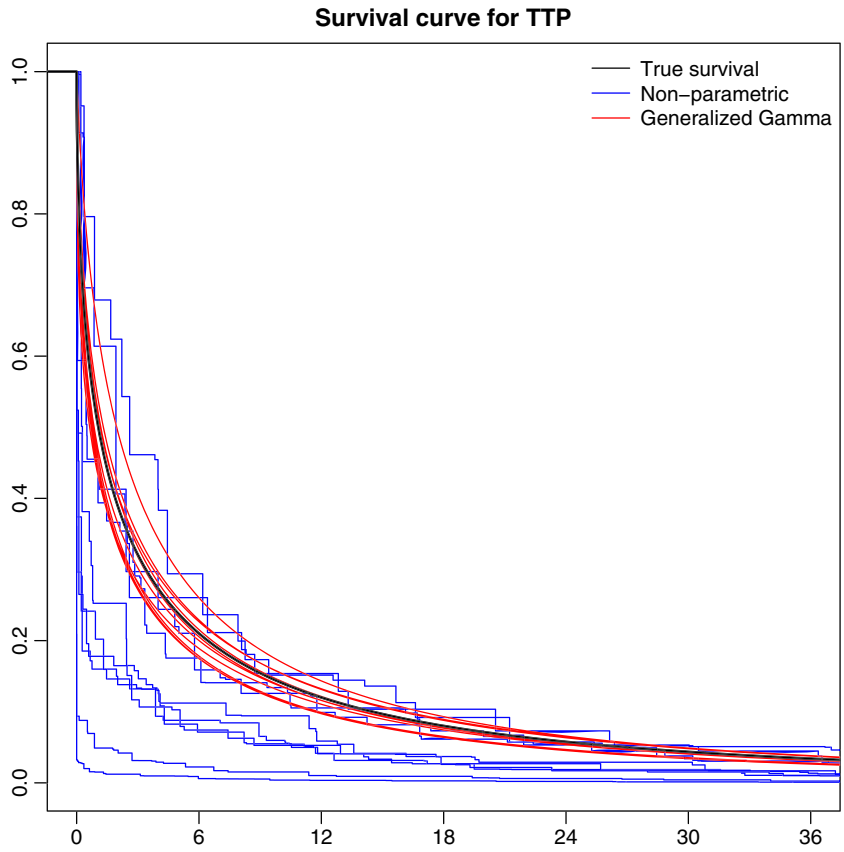


Fig. 2. Generalized Gamma distribution survival function of TTP ('true survival', black curve) with estimated (red, blue) derived from the densities of current duration in Fig. 1.

6.2. Comparison of parametric models and sensitivity to data precision

As we shall see in the practical application, it is difficult to distinguish the fit of the above two classes of parametric models based on usual goodness-of-fit methods. A further problem, illustrated primarily for the non-parametric approach but also relevant for parametric modelling, is the sensitivity of the estimation algorithms to observations close to 0. By the nature of current duration data, it is unrealistic to believe that the survey data are precise down to days, so we found it useful to evaluate how much the parameter estimates changed if observations were grouped into realistic intervals.

We simulated data to imitate realistic distributions (sample size: 1000) of current durations from our practical experience, to be presented below. Besides Pareto and generalized Gamma distributions, we also simulated from a truncated normal distribution, a Gamma distribution and a t distribution. Maximum likelihood estimation was then performed on the directly simulated continuous data as well as on data grouped according to the following schedules, all in months:

(i) [0, 1), [1, 2), [2, 3), ...;
(ii) [0, 1.5), [1.5, 3), [3, 4.5), ...;
(iii) [0, 0.5), [0.5, 1.5), [1.5, 2.5), ...;
(iv) [0, 1.5), [1.5, 2.5), [2.5, 3.5), ...; and
(v) [0, 0.5), [0.5, 1), [1, 1.5), ...

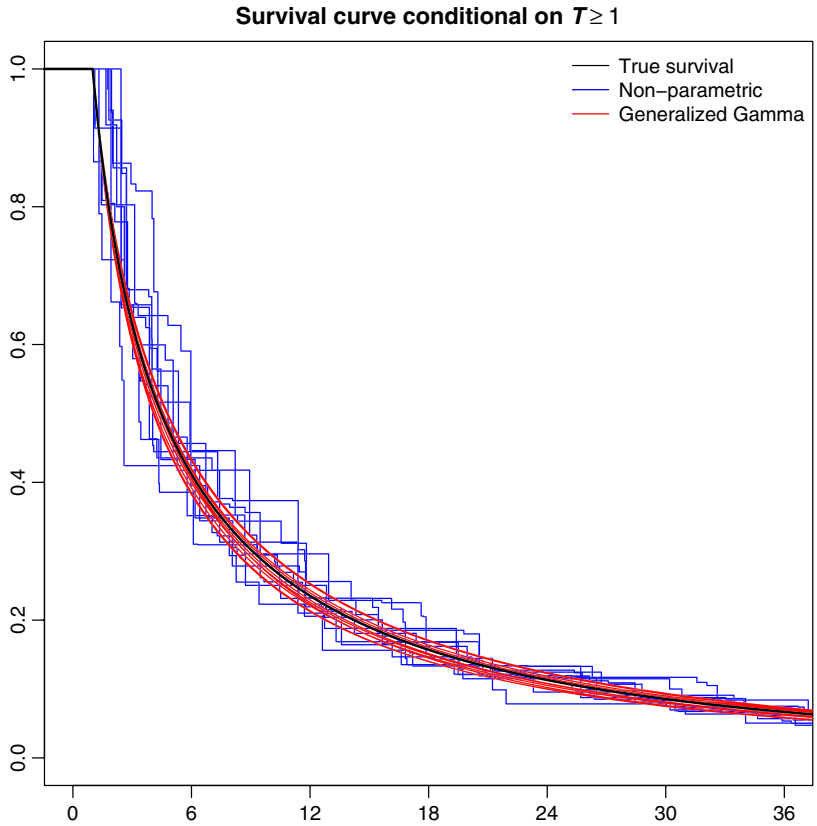


Fig. 3. Generalized Gamma distribution survival function of TTP ('true survival', black curve) with estimated (red, blue) derived from the densities of current duration in Fig. 1 but, conditioned on $T \geq 1$.

The most important results were:

For data simulated from the generalized Gamma distribution, we first used the same distribution for the fit. The closest fit to the true values was obtained with schedule (v), followed by the 'Raw' (i.e. continuous) data, see Fig. 4, upper panel. It was remarkable that the different fits to the TTP distribution were almost coinciding on the 'observed' current duration data, even though they yielded quite different estimated TTP.

Using instead the Pareto distribution provided almost exactly the same fit using raw data and any of the grouping schedules, but the fit was quite different from the true distribution (Fig. 4, middle panel).

If data were simulated from the Pareto distribution, the estimates based on the Pareto distribution fitted very nicely under all schedules. The generalized Gamma distribution fitted well under grouping schedule (v) and reasonably under the other schedules, including raw data (Fig. 4, lower panel).

The generalized Gamma distribution (but not the Pareto distribution) was also flexible enough to fit the three additional distributions listed above.

Our general conclusions, based on these simulations as well as on our experience in fitting actual data from the Obseff, are that

- the Pareto distribution is not sufficiently flexible for fitting realistic distributions, but on the other hand robust against the grouping schedule; and

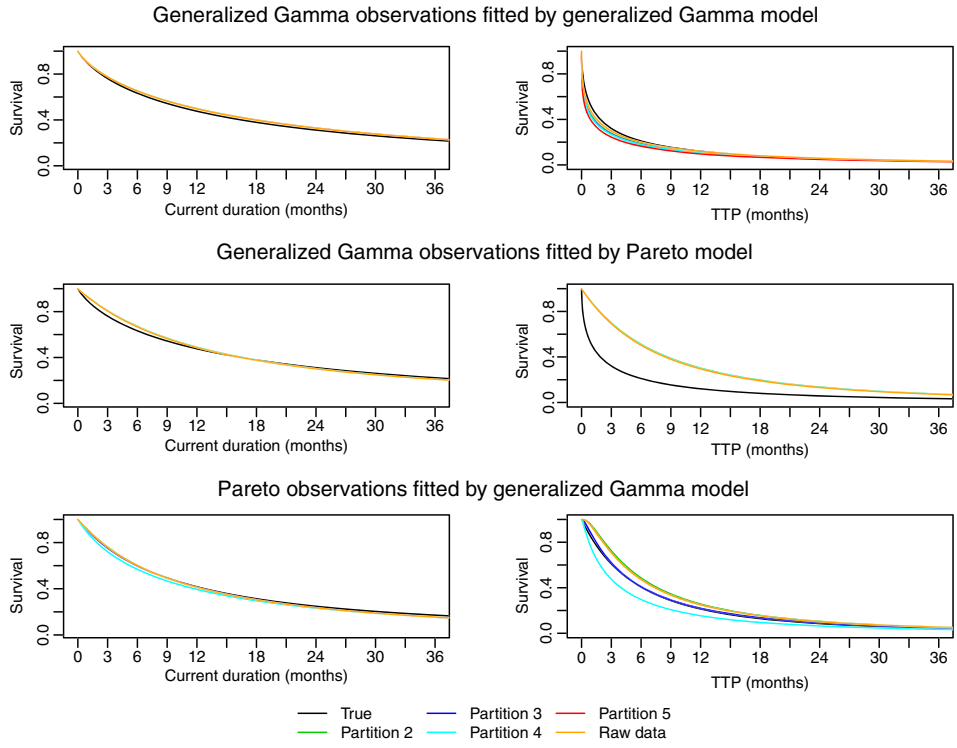


Fig. 4. Simulated current duration ($n=1000$) from parametric models fitted using the same or a different parametric model, using various grouping schedules.

- the generalized Gamma distribution has a suitable balance between robustness and flexibility. Concerning bias versus possible gains in robustness from different grouping schedules, only schedule (v) would be a candidate, but there would be at most a minor gain compared to analysing the 'raw' (continuous time) data.

In practice, it seems to be justified to use the 'raw' data.

7. Regression models for the current duration design

The transformation from current duration to underlying TTP distribution turns out to favour *accelerated failure time models*, see Keiding *et al.* (2011) for a brief historical sketch.

If Y satisfies an accelerated failure time model with survival function

$$P(Y>y\,|\,z)=S_0(y\,\mathrm{e}^{\beta z})$$

where the baseline survival function S_0 can belong to a parametric or more general model, then

$$g(y\,|\,z)=g_0(y\,\mathrm{e}^{\beta z})\,\mathrm{e}^{\beta z}$$

with g_0 the density of S_0, and

$$S(x\,|\,z)=P(X>x\,|\,z)=\frac{g(x\,|\,z)}{g(0\,|\,z)}=\frac{g_0(x\,\mathrm{e}^{\beta z})}{g_0(0)}$$

again satisfying an AFT model, this time with baseline survival function $g_0(\cdot)/g_0(0)$. As indicated in section 1, this seems to have been first observed explicitly by Yamaguchi (2003).

We see that the AFT model for Y needs to satisfy that g_0 is decreasing and that $0 < g_0(0) < \infty$ for all g_0; the restriction for the AFT model for X is that all baseline distributions must have finite expectation. The practical consequence of this simple relationship is that effects of co-variates on time-acceleration in the underlying survival distribution X can be directly estimated from an AFT model for Y. Full parametric specification is possible, and it is directly seen that in the Pareto class of models discussed above, μ is a scale parameter, so that a specification $\mu = e^{\beta z}$ would allow routine maximum likelihood estimation. Yamaguchi (2003) used the location parameter of the generalized Gamma distribution of $\log X$ as specified above.

There have been several approaches to semi-parametric modelling in this context. Thus Mokveld (2007), following up on preliminary work by van Es *et al.* (2000, 2006) derived semi-parametric efficient estimators of the Euclidean regression parameters β with minimal restrictions on the underlying distribution while Mukherjee (2006), in an unpublished Ph.D. dissertation, derived a spline-based semi-parametric estimator in the spirit of Cosslett (2004).

In the present situation we have in principle full observation of all backward recurrence times, hence the AFT model is to be fitted to uncensored data, which is just linear regression of $-\log Y$ on z, the specification of the baseline model family (S_0) essentially only affecting the weights in this regression analysis. Therefore analysis may be performed using standard software, an obvious initial possibility being unweighted least squares. Note that this is equivalent to assuming log-normal baseline for the current durations *but* since the log-normal density is not everywhere decreasing, this choice has no interpretation in the above framework.

In practice, however, we prefer to work on current durations censored at some fixed time (such as 36 months), and AFT algorithms for censored data are required. A convenient possibility is the published lss algorithm (Jin *et al.*, 2006; Huang & Jin, 2007) which generalizes the Buckley & James (1979) approach to AFT time models, thereby furnishing an explicitly implemented semiparametric AFT algorithm. For uncensored data, the lss algorithm yields rather similar results as the unweighted least squares approach.

We have noted an appealing invariance of the AFT structure when going from backward recurrence time or current duration to underlying survival distribution. A similarly simple structure does not in general exist for proportional hazards, for which the survival function of Y given z is

$$P(Y > y \mid z) = S_0(y)^{\exp(\beta z)}$$

with density

$$g(y \mid z) = \exp(\beta z) S_0(y)^{\exp(\beta z)-1} g_0(y)$$

so that the survival function of X given z becomes

$$S(x \mid z) = P(X > x \mid z) = \frac{g(x \mid z)}{g(0 \mid z)} = S_0(x)^{\exp(\beta z)-1} \frac{g_0(x)}{g_0(0)}$$

and this is usually not of proportional hazards form, a notable exception being if S_0 is Weibull.

8. Censoring problems in analyses based on the current duration design

In the beginning of section 4 we emphasized that the current duration design necessarily covers only those couples still trying, that is, they have neither become pregnant *nor given up trying*. We had however already noted in section 3 that this feature also applies to pregnancy-based retrospective TTP studies.

There are two other 'censoring' issues common to TTP studies in general but particularly important for the current duration design: problematic interpretability of very long reported TTP or current durations; and how to handle fertility treatment.

First, most formulations of the central question 'how long have you tried' may give equivocal answers if sexual frequency is low or contraception practices lenient (e.g. because of recognized subfertility). It is common to restrict attention to the first year(s) of attempt, either disregarding (truncating) long reported durations, or at least censoring them. The following lemma is useful in this context.

Let $g(y)$ and $G(y)$ be the density and distribution function of current duration Y. Then the current duration methodology is based on the fact that the survival function of the underlying distribution of $X(= \text{TTP} \wedge U)$ is

$$S(x) = g(x)/g(0).$$

Lemma. *Consider the distribution of Y right-truncated at y_0, with density*

$$g_0(y) = g(y)/G(y_0)$$

Then for $x < y_0$

$$\frac{g_0(x)}{g_0(0)} = \frac{g(x)/G(y_0)}{g(0)G(y_0)} = S(x)$$

so that for $0 < x < y_0$ the survival function of X is correctly recovered from the truncated distribution of Y given $Y < y_0$.

The lemma invites the routine use of the truncation at y_0 (i.e. discarding all individuals with reported current durations larger than y_0). Another possibility is to estimate $S(x)$ from the distribution of Y *censored* at y_0, but here the long current durations will still be playing some role.

Secondly, couples may initiate fertility treatment. As mentioned above, there are in practice various stages: contacting medical expertise, receiving hormonal treatment, various kinds of assisted reproduction techniques; but in this exposition we simplify the situation to one generic action 'starting fertility treatment'. We are then in the multi-state model (3) and our survey may hit the eligible couple either before or after starting fertility treatment, or before conception or before giving up.

The methodology for analysis of current duration data will be directly relevant for either

(i) couples who have not become pregnant, not given up and not yet started fertility treatment, corresponding to targeting the time $T \wedge U \wedge F$, $F =$ time from initiation to fertility treatment, in effect treating fertility treatment initiation in the same way as giving up; or

(ii) couples who have not become pregnant and not given up, disregarding the information on fertility treatment. The interpretation here is difficult in the light of the original ambition of studying *natural* biological fecundity, unless one takes the radical view that fertility treatment is useless. However, this approach can make sense if one aims to characterize the waiting time of couples in a 'real life' situation taking into account the fact that persistency in trying is not infinite and that treatments exist, with a certain degree of efficiency.

There is of course the third possibility: for those who have started fertility treatment but not yet become pregnant and not yet given up, one will in principle have retrospective information on the complete waiting time from initiation to fertility treatment, as well as the duration on fertility treatment. We have not attempted to derive the necessary methodology for analysis of this information, and we suspect that it may not be defendable to proceed along this course in practical cases.

9. The Observatory of Fecundity in France

The current duration approach was used in a large French telephone survey on TTP. See Slama *et al.* (2006) for a detailed feasibility study and Slama *et al.* (to appear) for the full

study. Women were eligible if they were aged 18–44 years, were currently living with a male partner, and neither woman nor man used any method to avoid pregnancy. Women who did not have a sexual intercourse in the last two months before interview were excluded, as were women who had delivered in the last three months before the interview. For the purpose of the present illustration, we exclude women who had initiated fertility treatment. The principal response variable, current duration of unprotected intercourse (CDUI) was defined as the time elapsed between the start of the period of unprotected intercourse and the interview, see Slama *et al.* (2006) for details. It follows from the exposition above that the underlying estimand 'TTP' then is to be interpreted as time *from* initiation *to* the minimum of (discovered) pregnancy, end of attempt for other reasons, and initiation of fertility treatment. For full details of the main study, see Slama *et al.* (to appear).

We illustrate here the above methodology on the 268 eligible women with no previous children and no missing values for the covariates of interest.

There were some very large 'current durations of unprotected intercourse'. Since the interpretation of these responses could be doubtful (did the couples really intend to become pregnant?) we decided to artificially censor all reported current durations at 36 months. This also reduced the dependence of the fit on the tail of the generalized Gamma distribution.

We first fit Pareto and generalized Gamma distributions to the observed distribution of CDUI, see Fig. 5 which indicates very similar fit for these two classes of distribution. However, as seen from Fig. 6, and Table 1, the estimated distributions of TTP differ, even in view of the estimation uncertainty expressed by the pointwise bootstrap confidence limits. Our informal simulation studies and general experience with the models make us believe that the greater flexibility of the generalized Gamma distribution family makes that fit the more credible one; but this cannot be judged from the comparison of fits in Fig. 5.

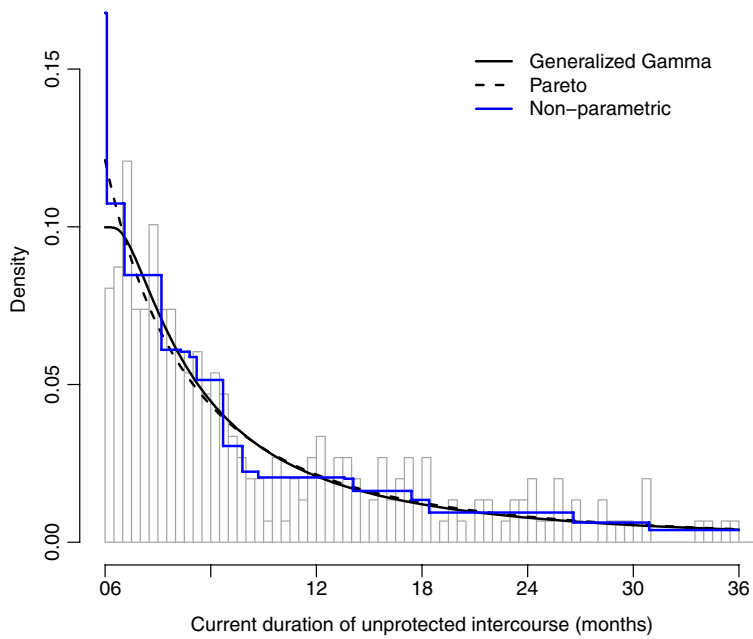


Fig. 5. Women with no previous children in the Obseff study. Observed current durations and fitted generalized Gamma and Pareto densities (based on data censored at 36 months). The NPMLE (based on all data) is also fitted.

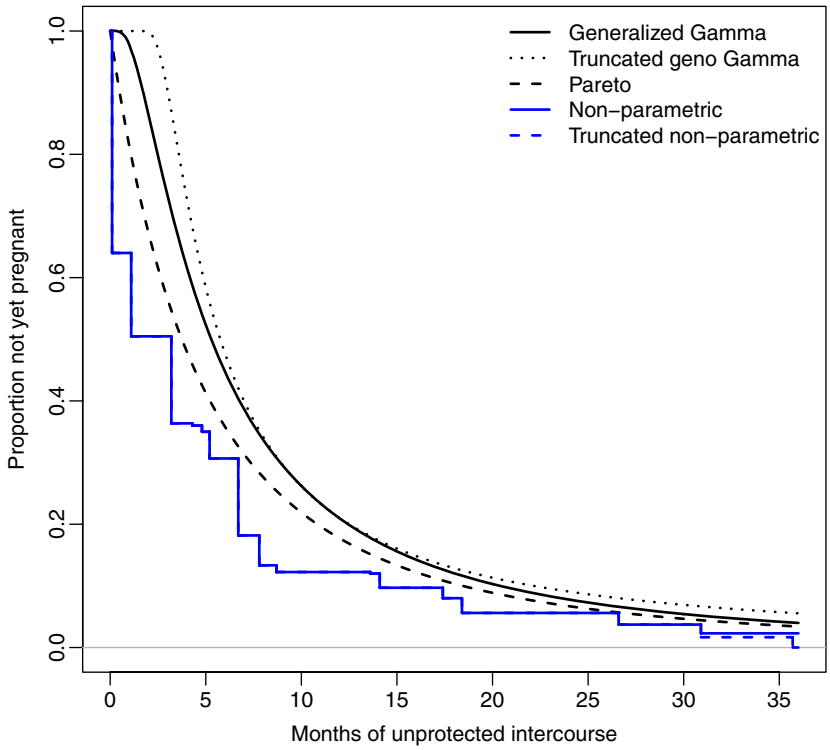


Fig. 6. Women with no previous children in the Obseff study. Estimated survival functions for time to pregnancy based on the fits in Fig. 5. The fits based on truncating the observed data at 36 months are also included.

Table 1. *Women with no previous children in the Obseff study. Estimation characteristics (with* 95 *per cent bootstrap confidence limits) for distribution of TTP using generalized Gamma and Pareto distributions, censored or truncated at* 36 *months*

TTP in months	Generalized Gamma (based on censored data)	Generalized Gamma (based on truncated data)	Pareto (based on censored data)
Lower quartile	2.87 (0.79, 5.2)	3.89 (1.74, 5.3)	1.42 (1.03, 1.98)
Median	5.29 (2.36, 7.7)	5.77 (3.42, 7.4)	3.75 (2.76, 5.1)
Upper quartile	10.4 (6.74, 13.6)	10.4 (7.3, 13.3)	8.84 (6.7, 11.6)

The NPMLE is seen (Fig. 5) to overshoot at 0, leading to a lower estimated survival curve in Fig. 6 than the parametric models.

Though the inconsistency of the non-parametric maximum likelihood estimator at 0 does not directly generalize to the parametric fits studied here, the difficulty remains, and we therefore add information on the result of conditioning on CDUI ≥ 1 month. Figure 7 shows that the conditional densities given TTP ≥ 1 month are now very similar for Pareto and generalized Gamma as well as for the NPMLE. Note that the estimated distributions of TTP using current durations *censored* at 36 months and *truncated* at 36 months are *not* the same, since observations larger than 36 do count in the fit for censored but not for truncated data. Table 1 shows that the difference is modest.

Table 2 illustrates the multivariate accelerated failure time regression of the observed current durations of unprotected intercourse on four important covariates using generalized Gamma distribution, Pareto distribution, the lss algorithm and ordinary least squares (OLS).

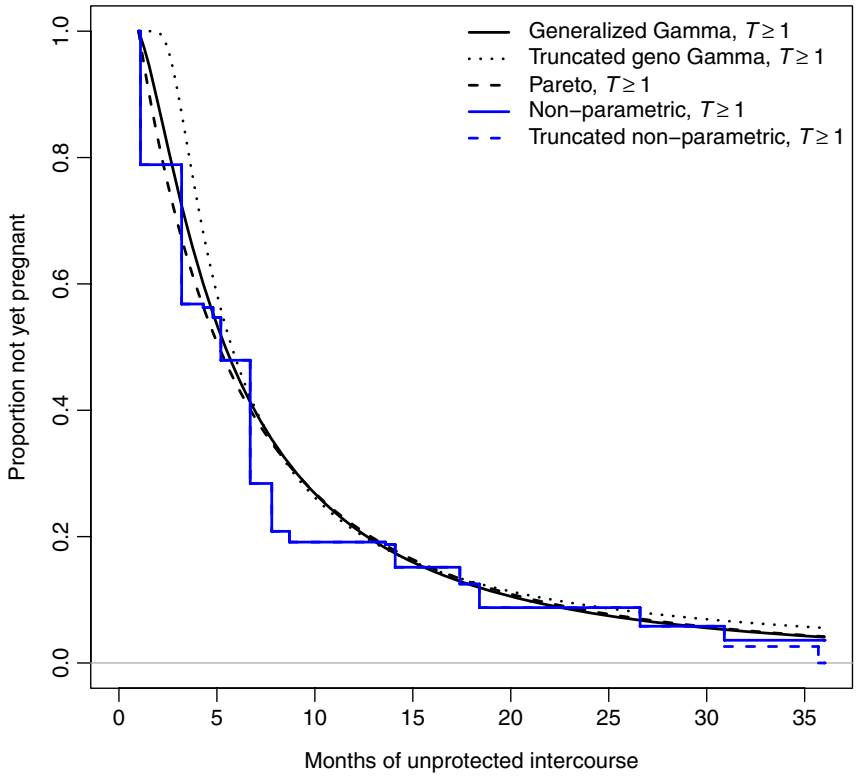


Fig. 7. Women with no previous children in the Obseff study. Estimated conditional survival functions for TTP, given TTP > 1 month, based on the fits in Fig. 5 as well as those based on truncating the data at 36 months.

Rather similar estimates are obtained for all four choices despite the fact, explained above, that lss and OLS do not have a direct interpretation in this framework. In particular, lss and OLS give very similar estimates.

10. Discussion

10.1. Heterogeneity of fecundability

The early literature on TTP was focused on heterogeneity in ability to conceive (fecundability) within and between couples, see Ecochard (2006) for a comprehensive survey. In the present study we ignore this important issue, essentially focusing on describing a snapshot of a population's fecundity from one or a sequence of cross-sectional surveys. We should of course acknowledge the mixture interpretation of the beta-geometric and Pareto models, but our focus here has been on estimating fecundity and identifying observable determinants.

Similarly, we have disregarded the interesting intra-couple (or intra-woman) correlation deriving from couples having several children, see e.g. Scheike *et al.* (1999) for models for the pregnancy-based design.

10.2. Discrete versus continuous time

The early literature on TTP, including Weinberg & Gladen (1986), was mostly formulated in discrete time, one argument being that conception can only occur once per menstrual cycle,

Table 2. *Women with no previous children in the Obseff study. AFT regression of CDUI using generalized Gamma distribution, Pareto distribution, ordinary least squares and LSS regression. Time ratios and* 95 *per cent bootstrap based confidence intervals*

		Generalized Gamma	Pareto	OLS	LSS
Covariate	No.	Time ratio	Time ratio	Time ratio	Time ratio
Tobacco consumption at recruitment					
Non-smokers	167	1	1	1	1
Smokers	101	1.19 (0.79, 1.84)	1.20 (0.83, 1.80)	1.08 (0.76, 1.64)	1.08 (0.71, 1.63)
Age at recruitment					
0–17	4	11.54 (0.61, *)	12.05 (0.83, *)	9.35 (1.06, *)	8.52 (0.50, *)
18–24	53	1.95 (1.20, 3.51)	1.94 (1.23, 3.27)	2.14 (1.24, 3.52)	2.21 (1.34, 3.37)
25–29	100	1	1	1	1
30–34	64	1.11 (0.61, 2.01)	1.13 (0.69, 1.93)	1.12 (0.66, 1.89)	1.08 (0.67, 1.76)
35–39	44	1.06 (0.58, 1.95)	1.03 (0.58, 1.69)	0.95 (0.46, 1.63)	0.95 (0.53, 1.73)
40–44	3	0.21 (0.02, 0.98)	0.20 (0.02, 0.65)	0.19 (0.03, 1.00)	0.19 (0.05, 0.76)
Frequency of sexual intercourse					
<1 per month	11	0.98 (0.27, 3.06)	1.03 (0.34, 2.61)	0.94 (0.33, 2.68)	0.95 (0.39, 2.34)
1–3 per month	45	2.08 (1.17, 3.81)	2.09 (1.26, 3.55)	1.85 (1.05, 3.36)	1.81 (1.01, 3.25)
1–2 per week	98	1.17 (0.71, 1.81)	1.20 (0.82, 1.81)	1.18 (0.78, 1.76)	1.16 (0.78, 1.72)
≥ 3 per week	114	1	1	1	1
Menstrual cycle length					
<27 days	61	1	1	1	1
27–29 days	114	0.87 (0.49, 1.47)	0.91 (0.54, 1.48)	0.85 (0.50, 1.38)	0.86 (0.52, 1.40)
≥ 30 days	93	1.08 (0.64, 1.90)	1.08 (0.65, 1.75)	1.04 (0.55, 1.71)	1.05 (0.65, 1.71)

*large value, >100.

which is therefore the natural discrete time unit. However, our impression is that most practical applications approximated a menstrual cycle to equal a month, getting back to chronological time. Moreover, there is within-woman (between two successive cycles) and between-woman variability in the cycle day of ovulation, which is unobserved, in addition to recall and declaration errors, so that actual declared durations often correspond to non-integer number of months. Our choice of continuous time is mostly for practical and mathematical convenience.

10.3. Applicability of the current duration approach

Our experience so far with the current duration approach to TTP studies is that it is in fact possible to obtain estimates of the distribution of the target variable

time to conception *or* giving up

where initiation of fertility treatment in our view should preferably be viewed analogously to giving up.

The composite nature of this target variable may worry practitioners, but it should be remembered that the ubiquitous pregnancy-based design suffers the same restrictions, in addition to its biased sampling of infertile and subfertile, usually excluded. However, accidental pregnancies escape attention from the current duration design, but can be identified in pregnancy-based TTP studies (with the open question of deciding which value of TTP might be attributed to these pregnancies).

Technically, parametric inference using the generalized Gamma distribution seems at the moment to be the primary recommendation. It would be desirable to develop portable software for the calculations using the parametric approach as well as for the non-parametric maximum likelihood approach.

For identification of observed determinants, the AFT models form an easily applicable tool, and our experience is that the form of the underlying distribution has little influence on the results.

10.4. The Observatory of Fecundity in France

For this particular study the two follow-up surveys will present unique opportunities for additional prospective and prevalent cohort-type follow-up analyses which will allow concrete empirical validation of the estimates obtained from the current duration approach.

10.5. Other applications of the current duration approach

We mentioned in section 1 that there is a long history in the analysis of 'open birth intervals' in demography, and 'last episode data' in survey analysis. In particular, Ali *et al.* (2001) studied the current duration of use of contraceptive pills. Their study used two parametric classes of distribution in parallel: the Weibull and the log-logistic, and fitted linear models to the log (true 'survival time') using maximum quasi-likelihood. Though both of these distributions are AFT, leading to models

$$S(x \mid z) = \exp\left[-(x\, \mathrm{e}^{\beta z})^{\alpha}\right]$$

and

$$S(x \mid z) = (1 + (x\, \mathrm{e}^{\beta z})^{\alpha})^{-1}$$

for the underlying time X, these authors did not take advantage of the AFT interpretation of their models.

A recent application directly inspired by the present work was the study by McLaughlin *et al.* (2010) on the influence of childhood adversities on adult psychiatric disorders, based on a cross-sectional survey with information on age at onset and (if relevant) offset of the most recent episode of the disorder.

10.6. Prevalent cohort studies

We conclude by noting that because of the equivalence between backward and forward recurrence time distributions, accelerated failure time models would be similarly useful in situations where forward recurrence times are modelled under stationarity. An important class of such problems are follow-up data from prevalent cohort studies with unknown initiation date (Brookmeyer & Gail, 1987; Keiding, 1992).

Acknowledgements

Jean Bouyer and Béatrice Ducot are acknowledged for their role in the Observatory of Fecundity. The Observatory of Fecundity was funded by grants from Institut National de Veille Sanitaire (InVS), from the French National Research Agency (ANR) and from the French Agency for Food, Environmental and Occupational Health and Safety (ANSES).

References

Ali, M. M., Marshall, T. & Babiker, A. G. (2001). Analysis of incomplete durations with application to contraceptive use. *J. Roy. Statist. Soc. A* **164**, 549–563.

Allison, P. D. (1985). Survival analysis of backward recurrence times. *J. Amer. Statist. Assoc.* **80**, 315–322.

Andersen, P. K. & Keiding, N. (2012). Interpretability and importance of functionals in competing risks and multistate models. *Statist. Med.* to appear

Balabdaoui, F., Jankowski, H., Pavlides, M., Seregin, A. & Wellner, J. (2011). On the Grenander estimator at zero. *Statist. Sinica* **21**, 873–899.

Banerjee, M. & Wellner, J. A. (2001). Likelihood ratio tests for monotone funtions. *Ann. Statist*. **29**, 1699–1731.

Basso, O., Juul, S. & Olsen, J. (2000). Time to pregnancy as a correlate of fecundity: differential persistence in trying to become pregnant as a source of bias. *Int. J. Epidem.* **29**, 856–861.

Bonde, J. P., Ernst, E., Jensen, T. K., Hjollund, N. H., Kolstad, H., Henriksen, T. B., Scheike, T., Giwercman, A., Olsen, J. & Skakkebæk, N. E. (1998). Relation between semen quality and fertility: a population-based study of 430 first-pregnancy planners. *Lancet* **352**, 1172–1177.

Brookmeyer, R. & Gail, M. H. (1987). Biases in prevalent cohorts. *Biometrics* **43**, 739–749.

Buck, G. M., Lynch, C. D., Stanford, J. B., Sweeney, A. M., Schieve, L. A., Rockett, J. C., Selevan, S. G. & Schrader, S. M. (2004). Prospective pregnancy study designs for assessing reproductive and developmental toxicants. *Env. Health Persp*. **112**, 79–86.

Buckley, J. & James, I. (1979). Linear regression with censored data. *Biometrika* **66**, 429–436.

Cosslett, S. R. (2004). Efficient semiparametric estimation of censored and truncated regression via a smoothed self-consistency equation. *Econometrica* **72**, 1277–1293.

Cristóbal, J. A., Alcalá, J. T. & Ojeda, J. L. (2007). Nonparametric estimation of a regression function from backward recurrence times in a cross-section sampling. *Lifetime Data Anal*. **13**, 273–293.

Dunson, D. B., Baird, D. D. & Colombo, B. (2004). Increased infertility with age in men and women. *Obst. Gynecol.* **103**, 51–56.

Duron-Martinaud, S., Slama, R., Ducot, B., Bohet, A., Keiding, N., Moreau, C. & Bouyer, J. *Cumulative incidence rate of medical consultation for fertility problems – results from a prospective random probability survey in France* (submitted).

Ecochard, R. (2006). Heterogeneity in fecundability studies: issues and modelling. *Statist. Meth. Med. Res.* **15**, 141–160.

Farewell, V. T. & Prentice, R. L. (1977). A study of distributional shape in life testing. *Technometrics* **19**, 69–75.

Fédération CECOS, Schwartz, D. & Mayaux, M. J. (1982). Female fecundity as a function of age: results of artificial insemination in 2193 nulliparous women with azoospermic husbands. *N. Engl. J. Med.* **306**, 404–406.

Grenander, U. (1956). On the theory of mortality measurement, part II. *Skand. Akt.* **39**, 125–153.

Huang, L. & Jin, Z. (2007). LSS: An S-Plus/R program for the accelerated failure time model to right censored data based on least-squares principle. *Comp. Meth. Progr. Biomed.* **86**, 45–50.

Jensen, T. K., Scheike, T., Keiding, N., Schaumburg, I. & Grandjean, P. (2000). Selection bias in determining the age dependence of waiting time to pregnancy. *Amer. J. Epidem.* **152**, 565–572.

Jin, Z., Lin, Y. & Ying, Z. (2006). On least-squares regression with censored data. *Biometrika* **93**, 147–161.

Joffe, M., Key, J., Best, N., Keiding, N., Scheike, T. & Jensen, T. K. (2005). Studying time to pregnancy by use of a retrospective design. *Amer. J. Epidem.* **162**, 115–124.

Juul, S., Keiding, N. & Tvede, M. (2000). Retrospectively sampled time-to-pregnancy data may make age-decreasing fecundity look increasing. *Epidemiology* **11**, 717–719.

Karmaus, W. & Juul, S. (1999). Infertility and subfecundity in population-based samples from Denmark, Germany, Poland and Spain. *Europ. J. Pub. Health* **9**, 229–235.

Keiding, N. (1992). Independent delayed entry (with discussion). In *Survival analysis: State of the art*. (eds J. P. Klein & P. K. Goel), 309–326. Kluwer, Dordrecht.

Keiding, N., Kvist, K., Hartvig, H., Tvede, M. & Juul, S. (2002). Estimating time to pregnancy from current durations in a cross-sectional sample. *Biostatistics* **3**, 565–578.

Keiding, N., Fine, J. P., Hansen, O. H. & Slama, R. (2011). Accelerated failure time regression for backward recurrence times and current durations. *Statist. Probab. Letters* **81**, 724–729.

Kulikov, V. N. & Lopuhaä, H. P. (2006). The behavior of the NPMLE of a decreasing density near the boundaries of the support. *Ann. Statist.* **34**, 742–768.

McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M. & Kessler, R. C. (2010). Childhood adversities and adult psychiatric disorders in the national comorbidity survey replication II. Associations with persistence of DSM-IV disorders. *Arch. Gen. Psych.* **67**, 124–132.

Mikkelsen, E. M., Hatch, E. E., Wise, L. A., Rothman, K. J., Riis, A. & Sørensen, H. T. (2009). Cohort profile: The Danish web-based pregnancy planning study – 'Snart-gravid'. *Int. J. Epidem* **38**, 938–943.

Mokveld, P. J. (2007). *The accelerated failure time model under cross sectional sampling schemes*. Ph.D. thesis, University of Amsterdam.

Moreau, C., Bouyer, J., Ducot, B., Spira, A. & Slama, R. (2010). When do involuntarily infertile couples choose to seek medical help? *Fertil. Steril.* **93**, 737–744.

Mukherjee, R. (2006). *On accelerated failure time models for forward and backward recurrence times*. Ph.D. thesis, University of Wisconsin – Madison.

Pal, J. K. (2009). End-point estimation for decreasing densities: asymptotic behavior of the penalized likelihood ratio. *Scand. J. Statist.* **36**, 764–781.

Scheike, T. & Keiding, N. (2006). Design and analysis of time to pregnancy. *Stat. Meth. Med. Res.* **15**, 127–140.

Scheike, T., Petersen, J. H. & Martinussen, T. (1999). Retrospective ascertainment of recurrent events: an application to time to pregnancy. *J. Amer. Statist. Assoc.* **94**, 713–725.

Slama, R., Ducot, B., Carstensen, L., Lorente, C., de La Rochebrochard, E., Leridon, H., Keiding, N. & Bouyer, J. (2006). Feasibility of the current duration approach to study human fecundity. *Epidemiol.* **17**, 440–449.

Slama, R., Hansen, O., Ducot, B., Bohet, A., Sorensen, D., Giorgis Allemand, L., Eijkemans, M. J. C., Rosetta, L., Thalabard, J. C., Keiding, N. & Bouyer, J. Estimation of the frequency of involuntary infertility on a nation-wide basis. To appear in *Hum.Reprod*.

Slama, R., Keiding, N. & Jensen, T. K. A prospective study of the association between sperm concentration and long-term involuntary infertility taking into account informative censoring (in progress).

Sun, J. & Woodroofe, M. (1996). Adaptive smoothing for a penalized NPMLE of a non-increasing density. *J. Statist. Plann. Inf.* **52**, 143–159.

Tingen, C., Stanford, J. B. & Dunson, D. B. (2004). Methodologic and statistical approaches to studying human fertility and environmental exposure. *Env. Health Persp*. **112**, 87–93.

van Es, B., Klaassen, C. A. J. & Mokveld, Ph. J. (2006). Current duration versus length biased sampling: does it pay to be patient? In *Prague Stochastics 2006* (eds M. Hušková & M. Janžura), 466–474. Matfyzpress, Prague. arXiv: math/0603497v1.

van Es, B., Klaassen, C. A. J. & Oudshoorn, K. (2000). Survival analysis under cross sectional sampling: length bias and multiplicative censoring. *J. Statist. Plann. Inf.* **91**, 295–312.

Weinberg, C. S. & Gladen, B. C. (1986). The beta-geometric distribution applied to comparative fecundability studies. *Biometrics* **42**, 547–560.

Wise, L. A., Mikkelsen, E. M., Rothman, K. J., Riis, A. H., Sørensen, H. T., Huybrechts, K. F. & Hatch, E. E. (2011). A prospective cohort study of menstrual characteristics and time to pregnancy. *Amer. J. Epidem.* **174**, 701–709.

Woodroofe, M. & Sun, J. (1993). A penalized maximum likelihood estimate of $f(0+)$ when f is non-increasing. *Statist. Sinica* **3**, 501–515.

Yamaguchi, K. (2003). Accelerated failure-time mover-stayer regression models for the analysis of last episode data. *Sociol. Methodol.* **33**, 81–110.

Received December 2011, in final form February 2012

Niels Keiding, Department of Biostatistics, University of Copenhagen, P.O. Box 2099, DK-1014 Copenhagen K, Denmark.
E-mail: nike@sund.ku.dk