Biostatistics (2006), **0**, 0, *pp.* 1–19
doi:10.1093/biostatistics/kxl030

Estimating and modeling the cure fraction in population-based cancer survival analysis

PAUL C. LAMBERT*, JOHN R. THOMPSON

Department of Health Sciences, Centre for Biostatistics and Genetic Epidemiology, University of Leicester, 22-28 Princess Road West, Leicester LE1 6TP, UK
pl4@le.ac.uk

CLAIRE L. WESTON

United Kingdom Children's Cancer Study Group, University of Leicester, Leicester, UK

PAUL W. DICKMAN

Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

SUMMARY

In population-based cancer studies, cure is said to occur when the mortality (hazard) rate in the diseased group of individuals returns to the same level as that expected in the general population. The cure fraction (the proportion of patients cured of disease) is of interest to patients and is a useful measure to monitor trends in survival of curable disease. There are 2 main types of cure fraction model, the mixture cure fraction model and the non-mixture cure fraction model, with most previous work concentrating on the mixture cure fraction model. In this paper, we extend the parametric non-mixture cure fraction model to incorporate background mortality, thus providing estimates of the cure fraction in population-based cancer studies. We compare the estimates of relative survival and the cure fraction between the 2 types of model and also investigate the importance of modeling the ancillary parameters in the selected parametric distribution for both types of model.

Keywords: Cure models; Relative survival; Survival analysis.

1. INTRODUCTION

The treatment of cancer has progressed dramatically over the last few decades with an increasing proportion of patients being cured for many types of cancer. The cure fraction (the proportion of patients cured of disease) is of interest to patients and is a useful measure to monitor trends in survival of curable disease. Standard survival analysis techniques, for example the Cox proportional hazards model, provide no direct estimate of the cure fraction. If it is believed that a proportion of individuals will not experience the event of interest, then it may be appropriate to fit models that explicitly allow for the cure fraction to be estimated and directly modeled.

*To whom correspondence should be addressed.

Much of the previous work on cure models has either analyzed data from children, where other causes of death can effectively be ignored as they are so rare, or analyzed nonfatal outcomes such as disease recurrence. However, if it is of interest to estimate the cure fraction in adults with time to death as the outcome, then mortality due to other diseases clearly needs to be considered. If reliable information on cause of death is available, then a cause-specific analysis can be performed where deaths not due to the disease of interest can be treated as censored observations. However, in population-based cancer studies, cause of death may either not be recorded or be obtained from death certificates, which are often inaccurately recorded (Begg and Schrag, 2002). In addition, the decision of whether a death is due or not due to a particular disease is difficult, opinions may vary and it may be hard to reach an unequivocal conclusion. Therefore, when analyzing survival data in population-based cancer studies, mortality due to other causes is generally incorporated using "relative survival methods" (Cutler and Aztell, 1969) and therefore when modeling the cure fraction, it will also be important to take account of mortality due to other causes.

Relative survival is the ratio of observed (all-cause) survival to the expected survival from a comparable group in the general population and provides a measure of the "excess mortality" experienced by patients diagnosed with the disease of interest, irrespective of whether the excess mortality is directly or indirectly attributable to the disease. The expected survival (or background mortality rate) can be obtained from national mortality statistics and is usually calculated after matching for age, sex, year of diagnosis, and possibly other covariates (Coleman *and others*, 1999a). Relative survival is often estimated using life tables with separate estimates for subgroups of interest. Recently, there has been growing interest in applying statistical models for relative survival. These models are on the hazard scale, which enables modeling of the "excess hazard (mortality) rate." Most models split the timescale in order to fit piecewise effects for the excess hazard (Dickman *and others*, 2004; Esteve *and others* 1990; Hakulinen and Tenkanen, 1987). Sasieni (1996) proposed a proportional excess hazards model with the baseline excess hazard estimated nonparametrically. Recently, there has been interest in modeling both the baseline excess hazard rate and time-dependent covariate effects continuously using splines (Bolard *and others*, 2002; Giorgi *and others*, 2003) or fractional polynomials (Lambert *and others*, 2005). However, none of the models described above assume that a proportion of the patients may be cured of disease.

For the majority of cancers, the relative survival curve often appears to plateau after a number of years. This plateau effect occurs when the mortality rate of the diseased individuals is the same as the mortality rate in the general population. This is essentially equivalent to the interval-specific relative survival being equal to one or the excess mortality rate being zero. At the point from which the diseased individuals no longer experience excess mortality, we refer to the group as being "cured" (or "statistically cured"). It is important to note that this definition of cure is from a population perspective and it does not provide information on individuals. An individual may be considered "medically cured" if he or she no longer displays symptoms of the disease. However, to be certain of medical cure is difficult and in population-based cancer studies such information is unlikely to be available and thus, in the models presented here, we are only interested in cure from a population perspective.

There have been 2 main types of standard cure models proposed. Most work has concentrated on the "mixture cure model" where it is assumed that a proportion π of patients are cured and are not at risk of experiencing the event of interest, with the remaining proportion 1-π being "uncured," and that these subjects will eventually experience the event of interest and thus the survival function will tend to zero for these subjects. The second type of cure fraction model is the "non-mixture" cure model, which defines an asymptote for the cumulative hazard and hence for the cure fraction. One of the advantages of the non-mixture cure fraction model is that it has a proportional hazards model as a special case.

Earlier work for the mixture model that incorporated background mortality rates was performed by Berkson and Gage (1952), who assumed a constant excess mortality rate for the uncured group, and Cutler and Axtell (1963), who used a piecewise estimate for the uncured group estimated at yearly intervals.

More recently, De Angelis *and others* (1999) developed mixture cure models incorporating background mortality using both exponential and Weibull survival distributions for the uncured group. Their approach allowed simultaneous modeling of the cure fraction and the mortality (hazard) rate in the uncured group. However, the shape parameter in the Weibull distribution was held constant and did not vary by covariates. These models have been extended to allow the background mortality rate to be higher than that of the general population (Phillips *and others*, 2002). The mixture cure fraction models of De Angelis *and others* have been incorporated into models that combine cancer incidence and survival to obtain estimates of prevalence (or cancer burden) (Heinavaara and Hakulinen, 2005; Verdecchia *and others*, 2002). Other recent work has used the mixture model for grouped data comparing a range of distributions for the uncured group including lognormal, log-logistic, Weibull, and Gompertz distributions (Gamel *and others*, 2000; Yu *and others*, 2004) and implemented in the CanSurv Software (Yu *and others*, 2005).

The non-mixture cure fraction model has not been extended to incorporate background mortality rates. However, it was noted by Tsodikov *and others* (2003) that the modified Gompertz model proposed by Haybittle (1965) could be rewritten as a non-mixture model incorporating background mortality rates and an exponential distribution.

This paper extends the non-mixture cure fraction model to incorporate background mortality rates so that the cure fraction can be modeled in population-based cancer studies. The approach is compared to the mixture cure model proposed by De Angelis *and others* (1999). The paper is laid out as follows. Section 2 describes the motivating data sets. Section 3 describes the mixture cure fraction model and describes how the non-mixture model can be extended to incorporate background mortality rates with Section 4 illustrating the methods. Section 5 discusses the models proposed.

2. Motivating data

Data were obtained from the public-use data set of all England and Wales cancer registrations between 1 January 1971 and 31 December 1990 with follow-up to 31 December 1995 (Coleman *and others*, 1999b). We present results for 33 874 females with cancer of the ovary and 56 525 males with cancer of the colon aged 50 years or over and diagnosed between 1 January 1981 and 31 December 1990 with follow-up to 31 December 1995. Maximum follow-up was restricted to 10 years as one would expect to observe the cure fraction within this timescale for many types of cancer. Interest lies in the effect of deprivation, defined in terms of the area-based Carstairs score (Coleman *and others* 1999a), and age at diagnosis on the cure fraction. There are 5 deprivation categories ranging from the least-deprived (affluent) to the most-deprived quintile in the population. Age is split into 4 groups, 50–59, 60–69, 70–79, and 80+. Background mortality rates were obtained from England and Wales national mortality statistics by age, geographical region, period of diagnosis, and deprivation group (Coleman *and others*, 1999b).

3. Methods

3.1 *Mixture cure fraction models*

Standard mixture cure fraction models are of the form

$$S(t) = \pi + (1 - \pi)S_{\mathrm{u}}(t), \tag{3.1}$$

where π is the proportion cured and $S_{\mathrm{u}}(t)$ is the survival function for the uncured individuals. $S_{\mathrm{u}}(t)$ is usually a standard parametric survival curve function, for example Weibull or lognormal. It is also possible to fit nonparametric (usually piecewise) models (Sy and Taylor, 2000; Taylor 1995), but these are not considered here.

De Angelis *and others* (1999) extended the standard parametric cure fraction model to incorporate background mortality. The all-cause survival can be written as the product of the expected survival, $S^*(t)$, and the disease-related survival functions

$$S(t) = S^*(t)(\pi + (1-\pi)S_u(t)), \tag{3.2}$$

where π is the cure fraction and $S_u(t)$ is the survival function for the uncured subjects. Similarly, the overall hazard rate is the sum of the background mortality rate and the excess mortality rate associated with the disease of interest

$$h(t) = h^*(t) + \frac{(1-\pi)\, f_u(t)}{\pi + (1-\pi)S_u(t)}, \tag{3.3}$$

where $h^*(t)$ is the expected mortality (hazard) rate and $f_u(t)$ is the probability density function associated with $S_u(t)$. Both $h^*(t)$ and $S^*(t)$ are assumed to be known, which is a reasonable assumption as these are based on the whole population of England and Wales. A further standard assumption of relative survival models is that the expected and excess mortality rates are independent. With the exception of smoking-based cancers, this is usually a reasonable assumption in population-based cancer studies (Ederer *and others*, 1961).

In population-based cancer studies, interest usually lies in modeling from the time of diagnosis. Application of the mixture model therefore assumes that at diagnosis there are a group of subjects who do not experience any excess mortality compared to the general population. In a discussion of cure models in childhood cancer clinical trials, Sposto (2002) argues that this interpretation is not applicable in an era when treatment can last many years and that cure may occur at any time during the treatment period. When modeling from the time of diagnosis, it seems even more unlikely that there is a group of individuals cured before any treatment has even been administered. This does not invalidate the use of this model as it may fit the data well and the model can be thought of as a useful mathematical function with an asymptote that can be used to estimate the proportion of patients cured of disease.

For survival models, the log-likelihood contribution for the ith subject with survival/censoring time t_i and censoring indicator d_i can be defined as

$$\ln L_i = d_i \ln(h(t_i)) + \ln(S(t_i)). \tag{3.4}$$

So the log-likelihood for the mixture model incorporating background mortality is

$$\ln L_i = d_i \ln\left(h^*(t_i) + \frac{(1-\pi)\, f_u(t_i)}{\pi + (1-\pi)S_u(t_i)}\right) + \ln(S^*(t_i)) + \ln(\pi + (1-\pi)S_u(t_i)). \tag{3.5}$$

This is equivalent to the log-likelihood given by De Angelis *and others* (1999). As noted by De Angelis *and others*, $S^*(t)$ is independent from the model parameters and so can be removed from the likelihood. Thus, the likelihood can be simply defined for any standard parametric distribution given the probability density function $f_u(t)$ and survival function $S_u(t)$ of the uncured group.

3.2 *Non-mixture cure fraction models*

The motivation behind non-mixture cure fraction models is that "after treatment," it is assumed that an individual is left with N_i "metastatic-competent" cancer cells, i.e. a tumor cell that has the potential of metastasizing. N_i is assumed to have a Poisson distribution with mean θ. The cure fraction is thus $P(\theta = 0)$. When θ is not equal to 0, let Z_j denote the time of the jth metastatic-competent cell to produce a detectable metastatic tumor with distribution function $F_Z(t) = 1 - S_Z(t)$. The change of subscript is

as each subject may have more than one metastatic-competent cell, but these are assumed independent within subject. The overall survival function can be written as

$$S(t) = \pi^{F_Z(t)}. \tag{3.6}$$

Alternative names for the non-mixture model are the "promotion time cure model" and the "bounded cumulative hazard model." For more details of the biological justification of the non-mixture model, see Tsodikov *and others* (2003).

In population-based cancer studies, the biological definition of the non-mixture cure fraction models is generally not appropriate. However, as for the mixture cure model, the non-mixture model can be considered as a useful mathematical function with an asymptote that can be applied to estimate the cure fraction (Sposto, 2002) and is useful for models that do not "fit" the biological definition as long as it is reasonable to assume cure (Ibrahim *and others*, 2001).

The survival function, $S(t)$, for the non-mixture model can also be expressed as

$$S(t) = \exp(\ln(\pi)F_Z(t)), \tag{3.7}$$

where π is the proportion cured and $F_Z(t)$ is a cumulative distribution function generally chosen to be $1 - S_Z(t)$, where $S_Z(t)$ is a standard parametric survival function, for example Weibull or lognormal distribution functions. Thus, the survival function has an asymptote at the cure fraction π, and the cumulative hazard has an asymptote at $-\ln(\pi)$. $F_Z(t)$ can also be nonparametric (Tsodikov, 2002; Tsodikov *and others*, 2003), but here we just concentrate on the application of parametric distribution functions.

The hazard function, $h(t)$, for a non-mixture model is

$$h(t) = -\ln(\pi)\, f_Z(t), \tag{3.8}$$

where $f_Z(t)$ is the probability density function for $F_Z(t)$. Note that if the parameters in $f_Z(t)$ do not vary by covariates, then the above is a proportional hazards model. This is a particular advantage of the non-mixture cure model over the mixture model, as the latter does not have a proportional hazards model for the whole group as a special case.

To incorporate background mortality rates to enable relative survival cure models to be fitted, let $S^*(t)$ and $h^*(t)$ be the expected survival and hazard functions, respectively. Overall survival can be expressed as the product of the expected survival and disease-related (relative) survival. Thus,

$$S(t) = S^*(t)\pi^{F_Z(t)} \tag{3.9}$$

or equivalently

$$S(t) = S^*(t)\exp(\ln(\pi) - \ln(\pi)S_Z(t)). \tag{3.10}$$

Similarly, the overall hazard rate can be expressed as

$$h(t) = h^*(t) - \ln(\pi)\, f_Z(t). \tag{3.11}$$

Thus, the overall mortality rate is made up of 2 components, the background mortality rate and the excess mortality rate associated with the disease of interest. The excess mortality rate has an asymptote at zero. If the parameters in $f_Z(t)$ do not vary by covariates, then the above is a proportional excess hazards model.

The non-mixture cure fraction model can be rewritten as

$$S(t) = S^*(t)\left(\pi + (1-\pi)\left(\frac{\pi^{F_Z(t)} - \pi}{1-\pi}\right)\right) \tag{3.12}$$

which is a mixture model and thus the survival distribution of the uncured patients can also be obtained from a non-mixture model by a simple transformation of the model parameters.

For non-mixture cure fraction models incorporating background mortality rates, the log-likelihood contribution for the ith subject with survival/censoring time t_i and censoring indicator d_i can be written as

$$\ln L_i = d_i \ln(h^*(t_i) - \ln(\pi) f_Z(t_i)) + \ln(S^*(t_i)) + (\ln(\pi) - \ln(\pi) S_Z(t_i)). \tag{3.13}$$

Thus, the likelihood can be simply defined for any standard parametric distribution given the probability density function $f(t)$ and survival function $S(t)$. As for the mixture model, $S^*(t_i)$ does not depend on any of the parameters and thus is not required in the likelihood for maximization and just the background mortality rate at death is required.

In addition, we do not have to split the data for analysis and use the binomial or Poisson approximations (Dickman *and others*, 2004) as we only need the expected hazard at death or censoring. An important advantage is that it is computationally much quicker than splitting the data.

3.3 *Parametric distribution*

There are a wide range of distribution functions to choose from, and here we consider the Weibull distribution for both the mixture and non-mixture cure fraction models. The Weibull distribution is flexible in that it will enable a monotonic increasing or decreasing mortality rate for the uncured group in the mixture model and either a monotonic decreasing or a positively skewed excess mortality rate in the non-mixture model (an increasing excess hazard mortality rate is not biologically plausible). The probability density function for the Weibull distribution is

$$f(t) = \gamma \lambda t^{\gamma - 1} \exp(-\lambda t^{\gamma}) \tag{3.14}$$

with survival function

$$S(t) = \exp(-\lambda t^{\gamma}). \tag{3.15}$$

Sposto (2002) discusses how the scale and shape parameters in the Weibull distribution for both the mixture and non-mixture models can be modeled as a function of covariates but states that "In my experience, it is rarely if ever necessary to include modelling of the shape parameter." De Angelis *and others* (1999) only consider the modeling of the scale parameter with only a constant estimated for the shape parameter. Here, we will investigate the effect of modeling the shape parameter as well as the scale parameter for both the mixture and non-mixture cure models. In Section 5, we discuss alternative choices for the parametric distribution.

3.4 *Link functions*

Both the cure fraction π and the parameters in the Weibull distribution may vary by covariates. Sposto (2002) described 3 link functions for the cure fraction when modeling covariates X each with advantages in certain situations. These were:

1) The identity link $\pi_i = \beta' X$. This has the advantage that covariate effects are relatively easy to interpret being in units of the cure fraction. However, there may be boundary problems for low or high cure fractions.
2) The logistic link $\log(\pi_i/(1-\pi_i)) = \beta' X$. Covariate effects are expressed as (log) odds ratios, and thus covariate effects have a similar interpretation to those in logistic regression.
3) The log(-log) link $\log(-\log(\pi_i)) = \beta' X$. This link function is particularly useful for the non-mixture model as covariate effects are expressed as (log) excess hazard ratios if the parameters

within the distribution function do not vary by covariates, that is, if proportional excess hazards can be assumed.

For both the mixture and non-mixture cure fraction models, it may be (and usually is) necessary to model the parameters in the parametric distribution function, that is, the scale (λ) and shape (γ) parameters in the case of the Weibull distribution. In this paper, both parameters are modeled using a log link function.

3.5 *Estimation*

The likelihood functions for both the mixture and non-mixture cure fraction models were maximized using the Newton–Raphson technique with the first and second derivatives estimated numerically as implemented using the "ml" command in Stata (Gould *and others*, 2003). Variance estimates were obtained using the inverse of the negative Hessian matrix evaluated at the maximum likelihood estimates. All models presented in this paper converged in 7 or fewer iterations. The programs for fitting these models are available from the authors on request.

3.6 *Comparisons*

We compare the estimates derived from the mixture and non-mixture cure fraction models with 2 commonly used methods in relative survival. First, we compare the estimates of relative survival from the cure fraction models applied to single groups with the empirical life table estimates derived using Hakulinen's (1982) method. Second, we use the piecewise proportional excess hazards model described by Dickman *and others* (2004) for comparison with the proportional excess hazards non-mixture cure fraction model, that is, when the parameters in $f_z(t)$ do not vary by covariates. In the piecewise model, let K denote the number of time intervals, d_{ik} a censoring indicator, and y_{ik} the time at risk for the ith subject in the kth interval. The model uses a Poisson likelihood and Dickman *and others* (2004) show that the model can be estimated as a generalized linear model with outcome d_{ik}, link function $\ln(\mu_{ik} - d^*_{ik})$, and offset $\ln(y_{ik})$. For the piecewise model, we defined 5 equally spaced intervals within the first year of follow-up, two 6-month intervals in the second year, followed by yearly intervals until the maximum follow-up of 10 years.

4. Results

Figures 1 and 2 show the estimated relative survival curves for both the mixture and non-mixture cure fraction models fitted to the 4 age groups separately for cancer of the ovary and cancer of the colon (males), respectively. Also shown on the graphs are the estimated cure fractions from both models and the life table estimates of relative survival, with 95% confidence intervals, obtained using Hakulinen's (1982) method. Although, the Hakulinen estimates should not be regarded as the true estimates of relative survival, one would expect the estimated cure fraction to be similar to where these estimates appear to reach a plateau. Generally, both models appear to give reasonable estimates for most age groups for cancer of the colon and ovary with the estimated cure fractions being close to where the Hakulinen estimates of relative survival appear to reach a plateau. The main exceptions are for the oldest group for both types of cancer and to some extent the middle 2 age groups for cancer of the colon. For the oldest age group, there appears to be an overestimate of the cure fraction for both the mixture and non-mixture models for both types of cancer. These results are similar to those of De Angelis *and others* (1999) who found an overestimation of the cure fraction in mixture models in the oldest age groups. For cancer of the colon age groups 60–69 and 70–79, there appears to be some underestimation of the cure fraction, though one

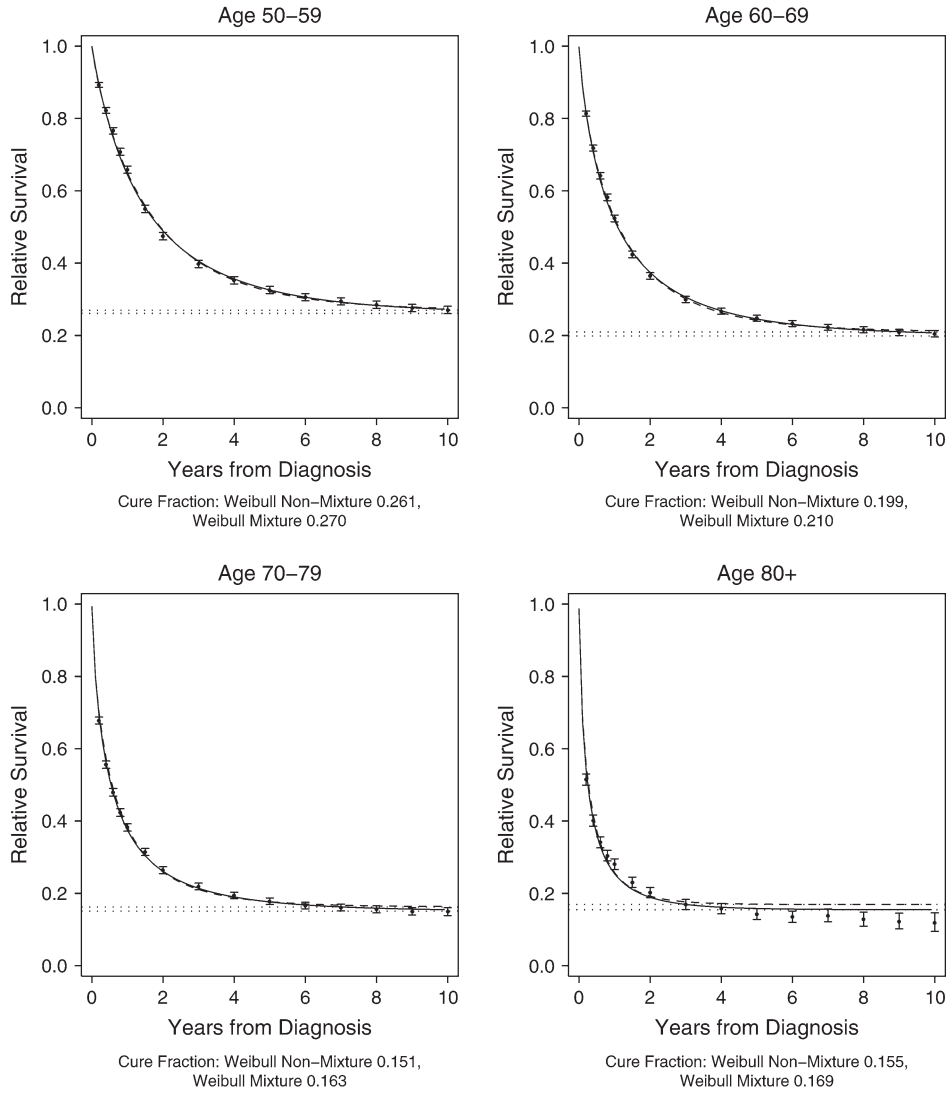

Fig. 1. Estimated relative survival curves for the non-mixture (solid line) and mixture (dashed line) cure fraction models applied to each age group for women with cancer of the ovary diagnosed between 1981 and 1990 with follow-up to 1995. Also shown are the Hakulinen estimates of relative survival with 95% confidence intervals. Note it is difficult to distinguish between the 2 estimated relative survival curves as they are in close agreement.

could argue whether the Hakulinen estimates of relative survival had actually reached a plateau within the range of the data. When the estimated cure fraction is noticeably lower than that expected, then for both the mixture and non-mixture cure models this indicates that the estimated excess hazard function has a long "tail" and does not approach the asymptote for the cure fraction until past the follow-up period. In these situations, the cure fraction is based on extrapolation of the parametric survival function and there needs to be a degree of caution in interpretation.

Table 1 shows the estimated excess hazard ratios for the non-mixture cure fraction model, with a log(-log) link function for the cure fraction, applied to both the cancer of the ovary and colon examples.

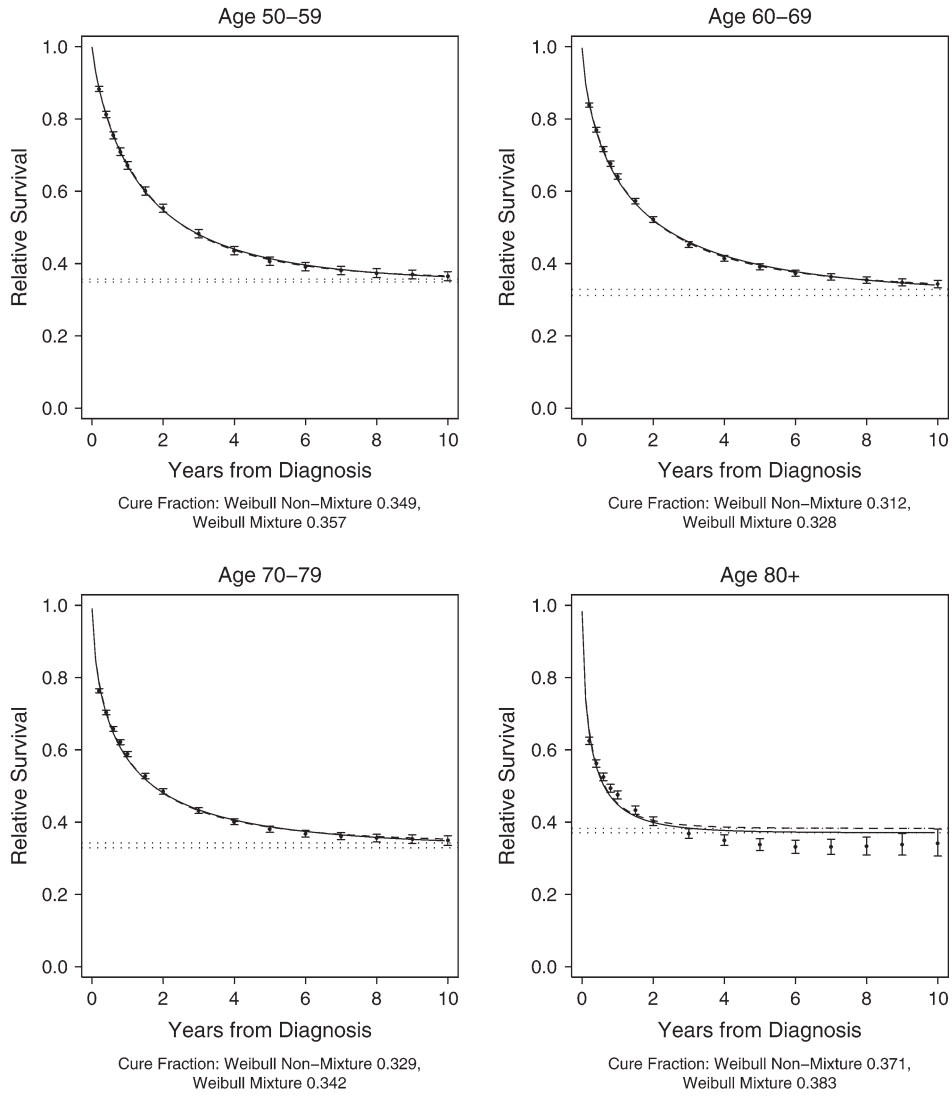

Fig. 2. Estimated relative survival curves for the non-mixture (solid line) and mixture (dashed line) cure fraction models applied to each age group for men with cancer of the colon diagnosed between 1981 and 1990 with follow-up to 1995. Also shown are the Hakulinen estimates of relative survival with 95% confidence intervals. Note it is difficult to distinguish between the 2 estimated relative survival curves as they are in close agreement.

The models include the effects of age group and deprivation group for the cure fraction. The scale parameter λ and the shape parameter γ of the Weibull distribution are not modeled and so these are proportional excess hazards models. Also shown are the excess hazard ratios obtained from fitting the piecewise model for the baseline excess hazard rate described in Section 3.6. The estimated excess hazard ratios, and their 95% confidence intervals, are extremely similar for both the cancer of the ovary and colon examples.

Table 2 shows the parameter estimates for 3 non-mixture models applied to the cancer of the ovary data. All models use an identity link for the cure fraction and include the effects of age group and deprivation group for the cure fraction but vary in how they model the scale parameter λ and the shape parameter

Table 1. *Estimated excess hazard ratios (with 95% confidence intervals) for a non-mixture cure fraction model with log(-log) link function and a piecewise model: (a) cancer of the ovary, (b) cancer of the colon*

Parameter	Non-mixture cure fraction model with proportional excess hazards	Piecewise model with proportional excess hazards
(a) Cure fraction		
Age group		
50–59	—	—
60–69	1.30 (1.26 to 1.35)	1.31 (1.27 to 1.35)
70–79	1.81 (1.75 to 1.87)	1.82 (1.76 to 1.89)
80+	2.55 (2.45 to 2.67)	2.60 (2.48 to 2.71)
Deprivation group		
1 (affluent)	—	—
2	1.04 (1.00 to 1.07)	1.03 (0.99 to 1.07)
3	1.07 (1.03 to 1.12)	1.08 (1.03 to 1.12)
4	1.09 (1.05 to 1.13)	1.09 (1.05 to 1.14)
5 (deprived)	1.14 (1.09 to 1.19)	1.14 (1.09 to 1.19)
(b) Cure fraction		
Age group		
50–59	—	—
60–69	1.07 (1.03 to 1.11)	1.08 (1.04 to 1.12)
70–79	1.20 (1.16 to 1.24)	1.22 (1.18 to 1.26)
80+	1.72 (1.65 to 1.79)	1.75 (1.68 to 1.82)
Deprivation group		
1 (affluent)	—	—
2	1.02 (0.98 to 1.05)	1.02 (0.98 to 1.06)
3	1.09 (1.05 to 1.12)	1.09 (1.50 to 1.13)
4	1.09 (1.05 to 1.13)	1.09 (1.05 to 1.13)
5 (deprived)	1.15 (1.11 to 1.20)	1.16 (1.11 to 1.20)

γ of the Weibull distribution. The models are as follows:

Model NMIX1: No modeling of λ or γ, that is, they do not vary by covariates. In this model, proportional excess hazards are assumed.

Model NMIX2: Modeling of the effects of age and deprivation group on the scale parameter λ but not on the shape parameter γ.

Model NMIX3: Modeling of the effects of age and deprivation group on the scale parameter λ and shape parameter γ.

Table 3 shows similar results for the mixture cure model, with models MIX1, MIX2, and MIX3 having the same definitions in terms of modeling the Weibull parameters; however, these parameters now apply to the survival function of the uncured group. Model MIX2 is equivalent to model 2 of De Angelis *and others* (1999), while model MIX3 allows more flexibility in the effect of covariates on the shape of the excess hazard/relative survival functions. Tables 4 and 5 show the corresponding results for cancer of the colon.

For both cancer of the ovary and cancer of the colon, there is a substantial difference in the cure fraction parameter estimates between the 3 non-mixture models. For example, the estimated cure fraction (and standard error) for the reference group (age group 50–59, affluent group) for cancer of the ovary is

Table 2. *Parameter estimates (standard errors) for non-mixture cure models for cancer of the ovary*

Parameter	Non-Mixture model		
	Model NMIX1 Scale (λ) constant Shape (γ) constant	Model NMIX2 Scale (λ) modeled Shape (γ) constant	Model NMIX3 Scale (λ) modeled Shape (γ) modeled
Cure fraction			
Intercept	0.303 (0.006)	0.243 (0.008)	0.268 (0.008)
Age group			
50–59	—	—	—
60–69	−0.091 (0.006)	−0.045 (0.008)	−0.061 (0.007)
70–79	−0.183 (0.005)	−0.078 (0.008)	−0.107 (0.008)
80+	−0.246 (0.005)	−0.072 (0.010)	−0.106 (0.010)
LRT	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$
Deprivation group			
1 (affluent)	—	—	—
2	−0.009 (0.005)	−0.008 (0.009)	−0.011 (0.008)
3	−0.019 (0.005)	−0.007 (0.009)	−0.011 (0.008)
4	−0.019 (0.005)	0.002 (0.009)	−0.005 (0.009)
5 (deprived)	−0.031 (0.005)	−0.007 (0.009)	−0.018 (0.009)
LRT	$P < 0.0001$	$P = 0.68$	$P = 0.34$
Scale parameter (λ)			
LRT (age group)	—	$P < 0.0001$	$P < 0.0001$
LRT (deprivation group)	—	$P < 0.0001$	$P < 0.0001$
Shape parameter (γ)			
LRT (age group)	—	—	$P < 0.0001$
LRT (deprivation group)	—	—	$P < 0.0001$
Log likelihood	−43 632.30	−43 105.18	−42 994.88

LRT, likelihood ratio test.

0.303 (0.006) for model NMIX1, 0.243 (0.008) for model NMIX2, and 0.268 (0.008) for model NMIX3. Similarly, the estimated difference in the cure fraction between the youngest and oldest age groups for cancer of the ovary varies considerably between the 3 models at −0.246 (0.005) in model NMIX1, −0.072 (0.010) in model NMIX2, and −0.106 in model NMIX3. Similar differences in the age effects can be seen for cancer of the colon. The effect of deprivation also changes with a highly significant effect of deprivation in model NMIX1 and a nonsignificant effect in models NMIX2 and NMIX3 for both cancer of the colon and cancer of the ovary. Thus, the different models can lead to different conclusions. Importantly for both cancer of the ovary and cancer of the colon, the effects of age and deprivation are highly significant ($P < 0.0001$) for both the scale and shape parameters as assessed using the likelihood ratio test.

The mixture models for both cancer sites show a very similar pattern to the 3 non-mixture models with strong evidence that both the scale and shape parameters of the Weibull distribution vary by both deprivation and age group. There is fairly a good agreement for the estimates of the cure fraction between the non-mixture cure models and the corresponding mixture cure models.

The different parameter estimates for the various models will lead to different estimates of the cure fraction in each subgroup. Figure 3 shows the estimated cure fractions for the non-mixture model for each of the age and deprivation subgroups estimated from a separate model for each subgroup, and for models NMIX1, NMIX2, and NMIX3. Figure 4 shows similar results for the mixture cure model. These

Table 3. *Parameter estimates (standard errors) for mixture cure models for cancer of the ovary*

Parameter	Mixture model		
	Model MIX1 Scale (λ) constant Shape (γ) constant	Model MIX2 Scale (λ) modeled Shape (γ) constant	Model MIX3 Scale (λ) modeled Shape (γ) modeled
Cure fraction			
Intercept	0.292 (0.007)	0.255 (0.007)	0.277 (0.007)
Age group			
50–59	—	—	—
60–69	−0.074 (0.007)	−0.045 (0.007)	−0.058 (0.007)
70–79	−0.148 (0.007)	−0.077 (0.007)	−0.106 (0.007)
80+	−0.206 (0.008)	−0.064 (0.010)	−0.101 (0.010)
LRT	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$
Deprivation group			
1 (affluent)	—	—	—
2	−0.009 (0.007)	−0.008 (0.008)	−0.010 (0.008)
3	−0.015 (0.007)	−0.007 (0.008)	−0.011 (0.008)
4	−0.012 (0.007)	0.001 (0.008)	−0.006 (0.008)
5 (deprived)	−0.025 (0.008)	−0.009 (0.009)	−0.021 (0.009)
LRT	$P = 0.030$	$P = 0.69$	$P = 0.19$
Scale parameter (λ)			
LRT (age group)	—	$P < 0.0001$	$P < 0.0001$
LRT (deprivation group)	—	$P < 0.0001$	$P < 0.0001$
Shape parameter (γ)			
LRT (age group)	—	—	$P < 0.0001$
LRT (deprivation group)	—	—	$P < 0.0001$
Log likelihood	−44 386.46	−43 244.63	−43 107.79

LRT, likelihood ratio test.

plots clearly demonstrate the variation in estimates of the cure fraction from the various models. For the majority of subgroups, there is generally fairly a good agreement between the cure fraction estimate from the separate model and the NMIX3 and MIX3 models, which are the models where both parameters in the Weibull distribution are modeled. The single-group models have a wider 95% confidence interval as there is no "borrowing of strength" across subgroups.

5. DISCUSSION

Estimation of the cure fraction in population-based cancer studies is important in providing information to patients and monitoring trends in cancer survival over time. We have extended the standard non-mixture cure fraction model to incorporate background mortality for use with population-based cancer data and compared these to the mixture models proposed by De Angelis *and others* (1999).

Although both types of models have a "biological" definition for why the asymptote may exist, it is unlikely that the biological definition for either model can be justified in population-based cancer studies. This does not invalidate their use as both the mixture and non-mixture cure fraction models can be considered useful mathematical functions for modeling relative survival when it is believed that a cure fraction exists. Even when it is not reasonable to assume a cure fraction, for example in breast cancer, the models

Table 4. *Parameter estimates (standard errors) for non-mixture cure models for cancer of the colon (males)*

Parameter	Non-mixture model		
	Model NMIX1 Scale (λ) constant Shape (γ) constant	Model NMIX2 Scale (λ) modeled Shape (γ) constant	Model NMIX3 Scale (λ) modeled Shape (γ) modeled
Cure fraction			
Intercept	0.383 (0.008)	0.309 (0.011)	0.359 (0.010)
Age group			
50–59	—	—	—
60–69	−0.026 (0.007)	−0.006 (0.010)	−0.033 (0.010)
70–79	−0.067 (0.007)	0.042 (0.010)	−0.018 (0.010)
80+	−0.189 (0.007)	0.071 (0.011)	0.016 (0.011)
LRT	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$
Deprivation group			
1 (affluent)	—	—	—
2	−0.007 (0.006)	−0.006 (0.011)	−0.006 (0.011)
3	−0.029 (0.006)	−0.012 (0.010)	−0.020 (0.011)
4	−0.031 (0.006)	0.016 (0.010)	−0.000 (0.011)
5 (deprived)	−0.050 (0.007)	0.003 (0.011)	−0.019 (0.012)
LRT	$P < 0.0001$	$P = 0.090$	$P = 0.18$
Scale parameter (λ)			
LRT (age group)	—	$P < 0.0001$	$P < 0.0001$
LRT (deprivation group)	—	$P < 0.0001$	$P < 0.0001$
Shape parameter (γ)			
LRT (age group)	—	—	$P < 0.0001$
LRT (deprivation group)	—	—	$P < 0.0001$
Log likelihood	−76 270.90	−75 642.17	−75 506.18

LRT, likelihood ratio test.

may still fit the data well for the observed follow-up period. However, the cure fraction estimate in these situations may be inappropriate as it will be based on extrapolation of the parametric distribution beyond the range of the data.

For both the non-mixture and mixture models, we have used the Weibull distribution. The Weibull distribution can provide a fairly flexible relative survival or excess hazard functions. However, it may be desirable to fit more general distributions, for when the Weibull distribution is not flexible enough, for example for both cancer of the ovary and cancer of the colon, it appeared that the Weibull distribution overestimated the cure fraction for the eldest age group. One potential solution is the extended generalized gamma distribution (Prentice, 1974). The advantage of this distribution is its flexibility in that it has the standard gamma, lognormal, Weibull, and exponential distributions as special cases. However, initial investigation of this model revealed that in some situations, these models failed to converge with parameters "wandering to infinity." This was a particular problem when there was a very high excess hazard rate in the first few weeks, which is common in population-based cancer data, particularly in older age groups. The lognormal distribution has been used for both the non-mixture and mixture cure models (Koti, 2001; Sposto 2002). However, we found that the very high excess hazard rate typically seen in the first few weeks after diagnosis in population-based cancer studies forced the lognormal distribution to have a

Table 5. *Parameter estimates (standard errors) for mixture cure models for cancer of the colon (males)*

Parameter	Mixture model		
	Model MIX1 Scale (λ) constant Shape (γ) constant	Model MIX2 Scale (λ) modeled Shape (γ) constant	Model MIX3 Scale (λ) modeled Shape (γ) modeled
Cure fraction			
Intercept	0.375 (0.008)	0.325 (0.010)	0.366 (0.009)
Age group			
50–59	—	—	—
60–69	−0.016 (0.008)	−0.004 (0.011)	−0.028 (0.009)
70–79	−0.038 (0.008)	0.048 (0.011)	−0.014 (0.009)
80+	−0.151 (0.009)	0.078 (0.012)	0.021 (0.011)
LRT	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$
Deprivation group			
1 (affluent)	—	—	—
2	−0.006 (0.008)	−0.006 (0.009)	−0.006 (0.010)
3	−0.028 (0.008)	0.036 (0.009)	−0.018 (0.010)
4	−0.022 (0.008)	0.073 (0.010)	−0.003 (0.010)
5 (deprived)	−0.041 (0.009)	0.001 (0.010)	−0.019 (0.011)
LRT	$P < 0.0001$	$P = 0.16$	$P = 0.18$
Scale parameter (λ)			
LRT (age group)	—	$P < 0.0001$	$P < 0.0001$
LRT (deprivation group)	—	$P < 0.0001$	$P < 0.0001$
Shape parameter (γ)			
LRT (age group)	—	—	$P < 0.0001$
LRT (deprivation group)	—	—	$P < 0.0001$
Log likelihood	−76 597.82	−75 735.75	−75 592.92

LRT, likelihood ratio test.

long tail and thus underestimated the cure fraction. This is illustrated in Figure 5 which shows the estimated relative survival curves for the Weibull and lognormal non-mixture models for women aged 70–79 years with cancer of the ovary. The estimated cure fraction is 0.059 for the lognormal non-mixture model and 0.151 for the Weibull non-mixture model. The life table relative survival estimates using Hakulinen's method are also shown. Visually, the Weibull distribution appears to give a better estimate of the cure fraction. However, this example illustrates a very important aspect of model comparison as the lognormal model has a much lower akaike information criterion (AIC). This is because during the follow-up period of the study, the lognormal model provides a better fit to the observed data. However, the cure fraction is based on estimation of an asymptote and because the lognormal distribution has a long tail, the relative survival curve will not approach the asymptote until long after the follow-up period. The relative survival curves have been extrapolated to 15 years and it can be seen that the lognormal model predicts an excess mortality notably in excess of zero even at this time point. This example demonstrates that using measures such as the AIC for model selection can be dangerous if interest lies in estimation of the cure fraction. A similar difference between the Weibull and lognormal mixture cure fraction models was observed.

Some of these problems occur as the cure fraction is an asymptote and the time of cure does not actually occur until time infinity, which can lead to the estimate of the cure fraction being sensitive to the tail of the parametric distribution. If excess mortality rate is still noticeably greater than zero, then

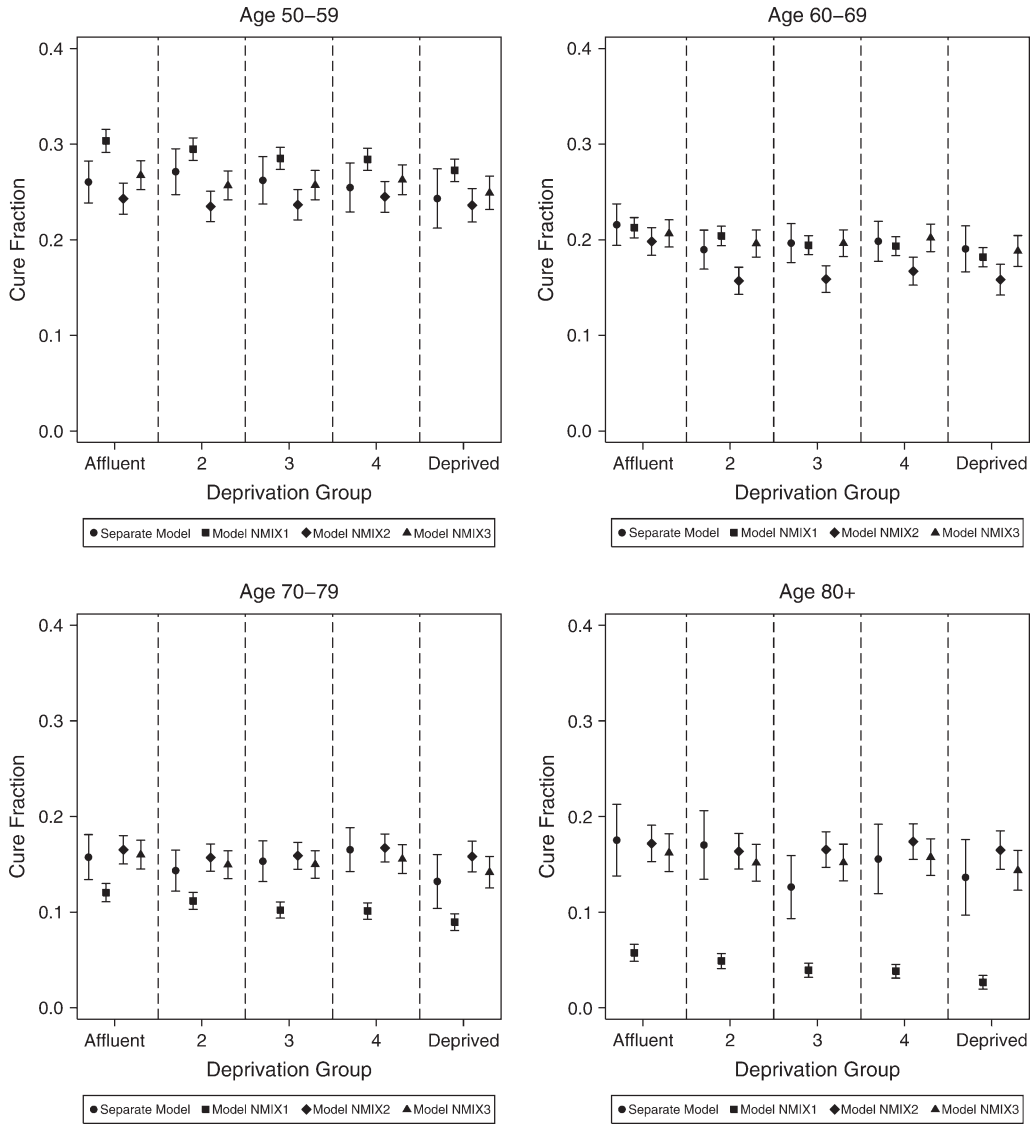

Fig. 3. Estimated cure fraction with 95% confidence intervals by deprivation and age group for women with cancer of the ovary obtained with a separate non-mixture cure fraction model for each subgroup and from the NMIX1, NMIX2, and NMIX3 models.

the estimate of the cure fraction will be based on extrapolation beyond the range of observed follow-up. The estimate of the cure fraction can be unstable when there are a small number at risk toward the end of follow-up, but in population-based cancer studies, the sample sizes are usually sufficiently large for this not to be a problem. Semi-parametric models have been developed for models not incorporating background mortality, but it should be realized that the issues regarding the tail of the distribution do not disappear for these models as they require an artificial constraint such as assuming that cure occurs at the last event time. In addition, proportional excess hazards is often not appropriate in population-based cancer studies and it may be more appropriate to model the nonproportionality parametrically, as

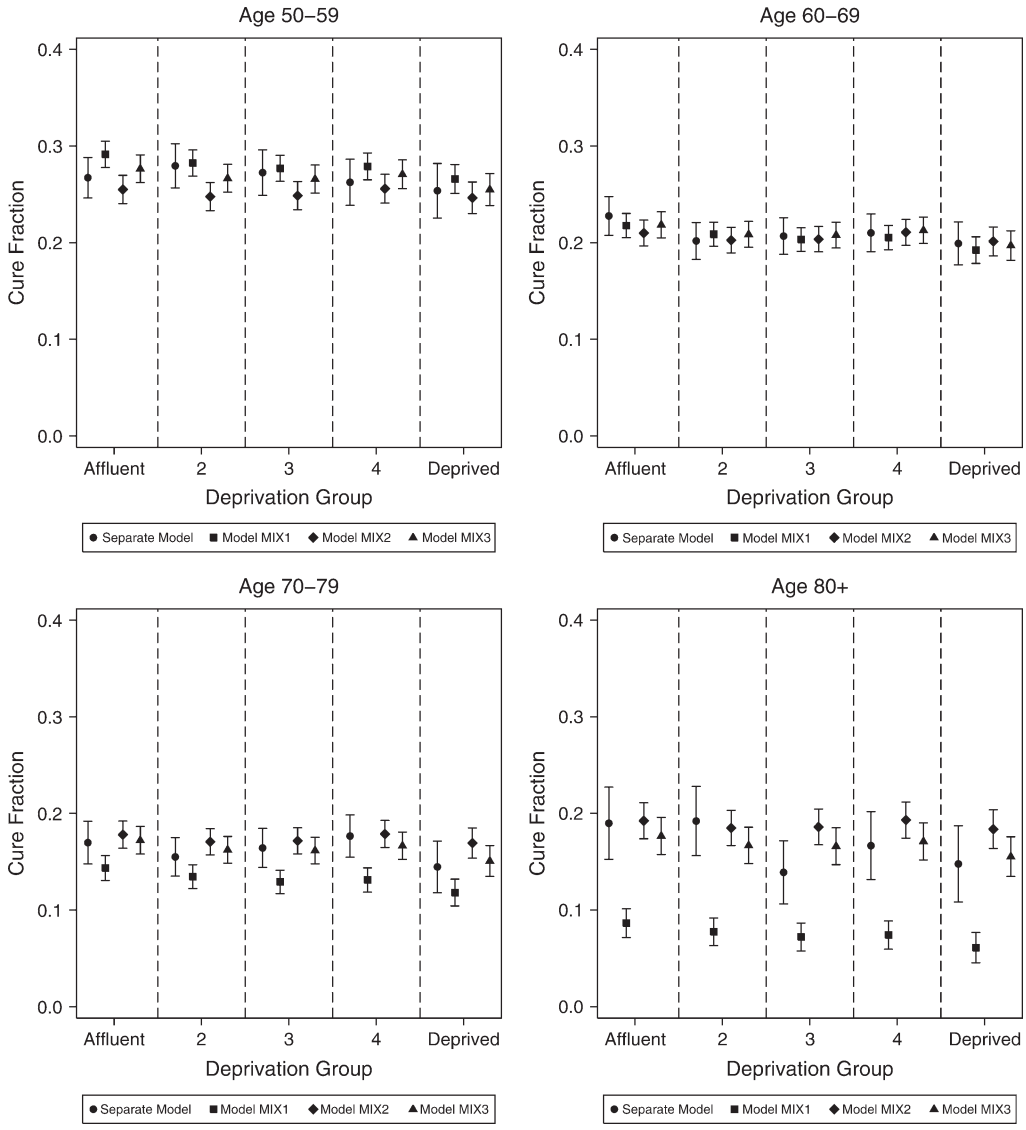

Fig. 4. Estimated cure fraction by deprivation and age group for women with cancer of the ovary obtained with a separate mixture cure fraction model for each subgroup and from the MIX1, MIX2, and MIX3 models.

done in the models described here. Estimation of the time of cure is not possible in these models as the asymptote occurs at time infinity. However, similar clinically useful measures are available such as the time at which the survival distribution for the uncured reaches a certain value, for example 90% or 95%. A nonparametric method to estimate the lower bound of the time of cure has been proposed by Rabinowitz and Ryan (1993) but does not provide an estimate of the cure fraction and does not appear suitable in a modeling framework as is applied to single groups.

When modeling the cure fraction incorporating covariates for the effect of age and deprivation group, allowing the scale and shape parameters of the Weibull distribution to vary was crucial and not including

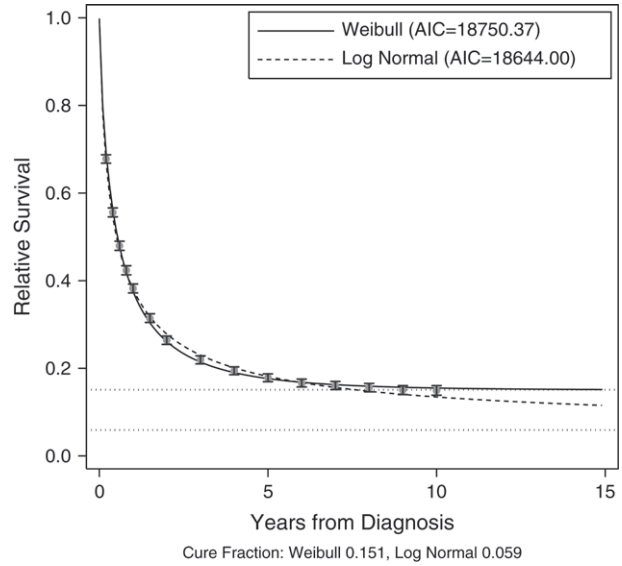

Fig. 5. Estimated relative survival and cure fractions for Weibull and lognormal non-mixture models for women aged 70–79 years with cancer of the ovary diagnosed between 1981 and 1990. Also shown are the Hakulinen estimates of relative survival with 95% confidence intervals.

covariate effects for both could lead to different model conclusions. De Angelis *and others* (1999) did not consider modeling of the shape parameter in mixture cure models, but in the models shown here, for both the mixture and non-mixture cure models, this can lead to incorrect model conclusions. For example for cancer of the ovary, the covariate effects for age group are appreciably different between MIX2 and MIX3 and also between NMIX2 and NMIX3.

The choice of link function is important in that it leads to different assumptions regarding the joint effect of covariates. When there are no interactions, then use of the identity link makes the assumption of additive covariate effects, while use of the logistic link makes the assumption of multiplicative covariate effects. We generally prefer the use of the identity link if the assumptions are plausible as covariate effects are on the cure fraction scale and easier to interpret. However, there could be boundary problems with low (or high) cure fractions.

The models presented here could be extended in a number of ways. First, as discussed above, alternative parametric distributions could be considered that provide more flexibility in the shape of the excess mortality/relative survival functions, while still giving reliable estimates of the cure fraction. Second, some measure of model fit when the cure fraction is the main outcome of interest would be beneficial. Third, the use of "period analysis" to obtain up-to-date estimates of relative survival has grown popular in the last few years (Brenner and Gefeller, 1997). Period analysis can be incorporated by artificially left truncating the survival times at a set date and fitting delayed entry models (Smith *and others*, 2004).

Acknowledgments

Part of this work was carried out while on study leave from the University of Leicester visiting the Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden, a visit partially funded by the Swedish Cancer Society (Cancerfonden). *Conflict of Interest:* None declared.

REFERENCES

BEGG, C. B. AND SCHRAG, D. (2002). Attribution of deaths following cancer treatment. *Journal of the National Cancer Institute* **94**, 1044–1045.

BERKSON, J. AND GAGE, R. P. (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association* **47**, 501–515.

BOLARD, P., QUANTIN, C., ABRAHAMOWICZ, M., ESTEVE, J., GIORGI, R., CHADHA-BOREHAM, H., BINQUET, C. AND FAIVRE, J. (2002). Assessing time-by-covariate interactions in relative survival models using restrictive cubic spline functions. *Journal of Cancer Epidemiology and Prevention* **7**, 113–122.

BRENNER, H. AND GEFELLER, O. (1997). Deriving more up-to-date estimates of long-term patient survival. *Journal of Clinical Epidemiology* **50**, 211–216.

COLEMAN, M. P., BABB, P., DAMIECKI, P., GROSCLAUDE, P., HONJO, S., JONES, J., KNERER, G., PITARD, A., QUINN, M. J., SLOGGETT, A. *and others* (1999a). *Cancer Survival Trends in England and Wales, 1971-1995: Deprivation and NHS Region*. London: Office for National Statistics.

COLEMAN, M. P., BABB, P., MAYER, D., QUINN, M. J. AND SLOGGETT, A. (1999b). *Cancer Survival Trends in England and Wales, 1971-1995: Deprivation and NHS Region (CDROM)*. London: Office for National Statistics.

CUTLER, S. J. AND AXTELL, L. M. (1963). Partitioning of a patient population with respect to different mortality patterns. *Journal of the American Statistical Association* **58**, 701–712.

CUTLER, S. J. AND AZTELL, L. M. (1969). Adjustments of long-term survival rates due to intercurrent disease. *Journal of Chronic Diseases* **22**, 485–495.

DE ANGELIS, R., CAPOCACCIA, R., HAKULINEN, T., SODERMAN, B. AND VERDECCHIA, A. (1999). Mixture models for cancer survival analysis: application to population-based data with covariates. *Statistics in Medicine* **18**, 441–454.

DICKMAN, P. W., SLOGGETT, A., HILLS, M. AND HAKULINEN, T. (2004). Regression models for relative survival. *Statistics in Medicine* **23**, 41–64.

EDERER, F., AXTELL, L. M. AND CUTLER, S. J. (1961). The relative survival rate: a statistical methodology. *National Cancer Institute Monograph* **6**, 101–121.

ESTEVE, J., BENHAMOU, E., CROASDALE, M. AND RAYMOND, L. (1990). Relative survival and the estimation of net survival: elements for further discussion. *Statistics in Medicine* **9**, 529–538.

GAMEL, J. W., WELLER, E. A., WESLEY, M. N. AND FEUER, E. J. (2000). Parametric cure models of relative and cause-specific survival for grouped survival times. *Computer Methods and Programs in Biomedicine* **61**, 99–110.

GIORGI, R., ABRAHAMOWICZ, M., QUANTIN, C., BOLARD, P., ESTEVE, J., GOUVERNET, J. AND FAIVRE, J. (2003). A relative survival regression model using B-spline functions to model non-proportional hazards. *Statistics in Medicine* **22**, 2767–2784.

GOULD, W., PITBLADO, J. AND SRIBNEY, W. (2003). *Maximum Likelihood Estimation with Stata*. Texas, TX: Stata Press.

HAKULINEN, T. (1982). Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics* **38**, 933–942.

HAKULINEN, T. AND TENKANEN, L. (1987). Regression analysis of relative survival rates. *Applied Statistics* **36**, 309–317.

HAYBITTLE, J. L. (1965). A two-parameter model for the survival curve of treated cancer patients. *Journal of the American Statistical Association* **60**, 16–26.

HEINAVAARA, S. AND HAKULINEN, T. (2006). Predicting the lung cancer burden: accounting for selection of the patients with respect to general population mortality. *Statistics in Medicine* **25**, 2967–2980.

IBRAHIM, J. G., CHEN, M. H. AND SINHA, D. (2001). *Bayesian Survival Analysis*. New York: Springer.

KOTI, K. M. (2001). Failure-time mixture models: yet another way to establish efficacy. *Drug Information Journal* **35**, 1253–1260.

LAMBERT, P. C., SMITH, L. K., JONES, D. R. AND BOTHA, J. L. (2005). Additive and multiplicative covariate regression models for relative survival incorporating fractional polynomials for time-dependent effects. *Statistics in Medicine* **24**, 3871–3885.

PHILLIPS, N., COLDMAN, A. AND MCBRIDE, M. L. (2002). Estimating cancer prevalence using mixture models for cancer survival. *Statistics in Medicine* **21**, 1257–1270.

PRENTICE, R. L. (1974). A log gamma model and its maximum likelihood estimation. *Biometrika* **61**, 539–544.

RABINOWITZ, R. AND RYAN, L. (1993). Lower confidence bounds for time to cure. *Biometrika* **80**, 681–687.

SASIENI, P. D. (1996). Proportional excess hazards. *Biometrika* **83**, 127–141.

SMITH, L. K., LAMBERT, P. C., BOTHA, J. L. AND JONES, D. R. (2004). Providing more up-to-date estimates of patient survival: a comparison of standard survival analysis with period analysis using life-table methods and proportional hazards models. *Journal of Clinical Epidemiology* **57**, 14–20.

SPOSTO, R. (2002). Cure model analysis in cancer: an application to data from the Children's Cancer Group. *Statistics in Medicine* **21**, 293–312.

SY, J. P. AND TAYLOR, J. M. (2000). Estimation in a Cox proportional hazards cure model. *Biometrics* **56**, 227–236.

TAYLOR, J. M. (1995). Semi-parametric estimation in failure time mixture models. *Biometrics* **51**, 899–907.

TSODIKOV, A. (2002). Semi-parametric models of long- and short-term survival: an application to the analysis of breast cancer survival in Utah by age and stage. *Statistics in Medicine* **21**, 895–920.

TSODIKOV, A. D., IBRAHIM, J. G. AND YAKOVLEV, A. Y. (2003). Estimating cure rates from survival data: an alternative to two-component mixture models. *Journal of the American Statistical Association* **98**, 1063–1078.

VERDECCHIA, A., DE ANGELIS, G. AND CAPOCACCIA, R. (2002). Estimation and projections of cancer prevalence from cancer registry data. *Statistics in Medicine* **21**, 3511–3526.

YU, B., TIWARI, R. C., CRONIN, K. A. AND FEUER, E. J. (2004). Cure fraction estimation from the mixture cure models for grouped survival data. *Statistics in Medicine* **23**, 1733–1747.

YU, B., TIWARI, R. C., CRONIN, K. A., MCDONALD, C. AND FEUER, E. J. (2005). CANSURV: a Windows program for population-based cancer survival analysis. *Computer Methods and Programs in Biomedicine* **80**, 195–203.

[*Received March 9, 2006; revised August 11, 2006; accepted for publication September 29, 2006*]