Journal of Applied Statistics
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713428038

Competing risks, left truncation and late entry effect in A-bomb survivors cohort

J. Anzures-Cabrera[a]; J. L. Hutton[b]
[a] MRC Biostatistics Unit, Institute of Public Health, University Forvie Site, Cambridge, UK [b] Department of Statistics, University of Warwick, Coventry, UK

Online publication date: 15 April 2010

To cite this Article Anzures-Cabrera, J. and Hutton, J. L.(2010) 'Competing risks, left truncation and late entry effect in A-bomb survivors cohort', Journal of Applied Statistics, 37: 5, 821 — 831
To link to this Article: DOI: 10.1080/02664760902914417
URL: http://dx.doi.org/10.1080/02664760902914417

Journal of Applied Statistics
Vol. 37, No. 5, May 2010, 821–831

Taylor & Francis
Taylor & Francis Group

Competing risks, left truncation and late entry effect in A-bomb survivors cohort

J. Anzures-Cabrera[a] and J.L. Hutton[b]*

[a]*MRC Biostatistics Unit, Institute of Public Health, University Forvie Site, Robinson Way, Cambridge CB2 0SR, UK;* [b]*Department of Statistics, University of Warwick, Coventry CV4 7AL, UK*

(*Received 31 January 2008; final version received 19 March 2009*)

The cohort under study comprises A-bomb survivors residing in Hiroshima Prefecture since 1968. After this year, thousands of survivors were newly recognized every year. The aim of this study is to determine whether the survival experience of the late entrants to the cohort is significantly different from the registered population in 1968. Parametric models that account for left truncation and competing risks were developed by using sub-hazard functions. A Weibull distribution was used to determine the possible existence of a late entry effect in Hiroshima A-bomb survivors. The competing risks framework shows that there might be a late entry effect in the male and female groups. Our findings are congruent with previous studies analysing similar populations.

Keywords: competing risks; late entry effect; left truncation; sub-hazard function; Weibull distribution

1. Introduction

The end of the Second World War was marked by one of the most dreadful events in human history: the dropping of the atomic bomb on the cities of Hiroshima and Nagasaki in August of 1945. The Research Institute for Radiation Biology and Medicine (RIRBM) at Hiroshima University, Japan, records mortality for the entire population of atomic bomb (A-bomb) survivors residing in Hiroshima Prefecture since 1968.

An A-bomb survivor is administratively defined as an individual who has received the A-Bomb Survivor's Health Handbook (hereafter the Health Handbook) in Hiroshima Prefecture [9]. These Health Handbooks were issued to survivors after administrative verification of their personal declaration as A-bomb survivors. Health Handbook holders receive free medical care for certain designated diseases, allowances and welfare benefits.

The mortality observation at the RIRBM for these survivors began in 1968. The initial study subjects were survivors who received the Handbooks before 1 January 1968, referred to as the registered population (RP). After 1968, several thousand survivors were newly recognized every year, because of the availability of medical services and a change in the attitude of the survivors.

*Corresponding author. Email: j.l.hutton@warwick.ac.uk

ISSN 0266-4763 print/ISSN 1360-0532 online

DOI: 10.1080/02664760902914417
http://www.informaworld.com

These survivors are called left truncated observations in survival analysis. For this analysis we will call them "late entries" (LE).

It is believed that a stigma associated with being an A-bomb survivor might influence the timing of registration. Thus late entrants might contribute with a late entry effect to the survival rate if they register only when they need medical care. As the need for medical care might motivate registration, the type of illnesses might affect the timing of registration. For this reason, and as different types of disease have different associations with age, we use a competing risks analysis that considers two major disease groups, cardio-vascular and cancer, separately from other causes. Matsuura and Eguchi [8] proposed a semi-parametric model to analyse the entry effect in the Hiroshima data. As an alternative, we present a parametric model that explores late entry effect, and examines whether such effect might be attributable to particular causes of death. Thus, the aim of the article is to determine whether the survival experience of the late entrants is significantly different from the survival experience of the RP while considering a parametric competing risks framework.

2. Materials and methods

2.1 *Study populations*

The cohort comprises 49,765 A-bomb survivors registered in the RIRBM on or after 1 January 1968 and followed up for 29 years. There are 38,624 individuals registered on January 1968 and 11,141 late entrants. For each individual belonging to the RIRBM cohort the entry time, death or censoring time, censoring indicator, sex, age on 6 August 1945, estimated radiation dose and cause of death are recorded. As the beginning of the cohort is 23 years after the bomb, we are therefore looking at long-term survival rather than any short-term effects of bomb blast or radiation.

The estimated radiation dose is known as the shielded kerma dose. It depends on the location of a subject, shielded condition and age at time of the bombing. For further details see Hoshi *et al*. [5].

The cohort contains 27 different causes of death, which were reduced to three categories, cardio-vascular, cancer and other causes, in order to focus on the two major causes of death and to simplify the presentation. A quarter of the total deaths were from cancer, and a third were from heart disease. Figure 1 shows a diagram of the competing risks model for the A-bomb cohort.

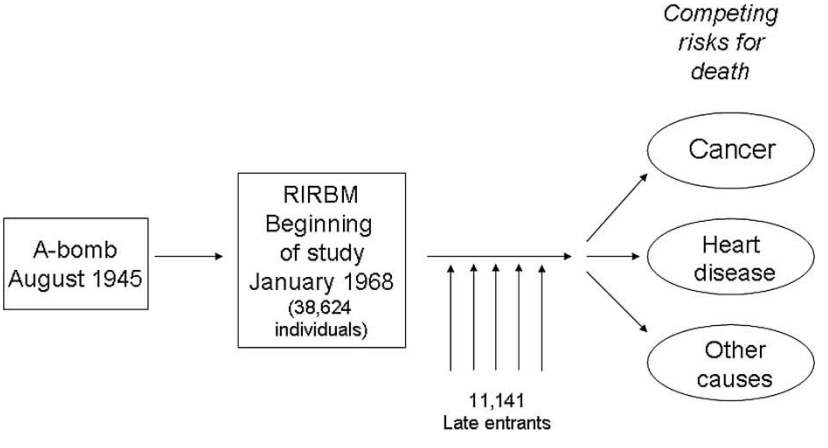

Figure 1. Time line and competing risks model for the A-bomb survivors cohort.

Table 1. Characteristics of individuals registered at the RIRBM cohort.

	Male		Female		
	RP	LE	RP	LE	Total
ATB in years					
Mean	45.5	36.2	47.8	39.9	45.0
SD	16.5	11.9	15.5	12.9	15.7
Radiation (cSv)					
Median	21.6	18.9	21.6	21.3	21.4
IQR	52.6	45.8	40.8	38.8	44.3
Distance < 3 km (%)†	66.0	11.9	67.1	10.2	54.4
Number of Deaths	7221	1133	8960	1257	18,571
% Deaths†	46.2	23.7	39.0	19.8	37.3
Total number of subjects	15,632	4790	22,992	6351	49,765

Notes: LE, late entries; RP, registered population; cSv, centi-Sievert; IQR, interquartile range; ATB-age at time of the bombing.
†As a proportion of the total in each group.
Distance < 3 km - percentage of individuals located at less than 3 km away from the hypocentre.

Women accounted for 59% of the registered cohort, and they are on average older than men (Table 1). Late entrants formed a quarter of the population. They were younger than the RP suggesting that the majority of late entrants would have been children in 1945, and therefore the decision of being registered before 1968 would be dependant on their parents. Late entrants have a higher percentage of censored observations (79%) than registered individuals. The rate of later entrants averaged 4% individuals per year (data not shown). The estimated radiation dose is positively skewed. The variation in radiation dose received is higher in men than in women. Cancer deaths were more exposed to radiation. The majority of male and female late entrants were located at a distance greater than 3 km from the hypocentre.

The following is a descriptive summary of the covariates included in the parametric models. The distance from the hypocentre is an indicator variable with the value of one for those individuals exposed at a ground distance of less than 3 km of the hypocentre, and zero for the remaining. Radiation is a positively skewed variable, we estimated its effect in the log-dose scale. Individuals located at distances greater than 3 km from the hypocentre were included in the database as having zero centi-Sieverts of radiation dose, [5]. Therefore, we incremented the radiation dose by one unit to avoid the logarithm of zero. To assess whether there is a late-entry effect, an indicator variable is included in the model. Its estimated coefficient shows the magnitude of deviation of any possible effect that late entrants can have in the mortality of the cohort.

3. Statistical methods

Assume that the survival experience of an individual is represented by (K, T, δ), where K and T are random variables representing the entry and failure times, respectively, and δ is a censoring indicator. Left truncation arises when there are individuals in the study for whom the beginning of the observation period does not coincide with the entry time origin of study [1]. Left truncated survival models consider the time in which the existence of an individual is known to the researcher, i.e. the time in which the A-bomb survivors decided to enroll in the cohort. These type of models use a conditional approach because the contributions to the likelihood function are conditional on individuals being alive over the truncated time [3]. Therefore the contribution to the likelihood from a failure is $f(t)/S(k)$ and from a censored observation is $S(t)/S(k)$, where $f(\cdot)$ and $S(\cdot)$ are

the density and survival functions, respectively, and t and k are the death/censoring time and entry time.

First, we will define the typical left-truncated model, and then we will extend this model to the competing risks case. If we assume that the failure times, T, follow a Weibull distribution, the late entrants incorporate into the cohort independently from one another, and we take into account covariate information, $\underset{\sim}{x}$, then the truncated survival function is

$$S(t|k) = \exp[-\lambda(t^{\alpha} - k^{\alpha})e^{\beta' \underset{\sim}{x}}], \tag{1}$$

where $\lambda > 0$, $\alpha > 0$, $k \geq 0$ and β is a vector of coefficients. We can think of this approach as a left-truncated proportional hazards model that includes the effect of covariates. The Weibull distribution performs a key role when modelling survival data. Depending on the value of the shape parameter, the hazard rate can be monotone decreasing ($\alpha < 1$) or increasing ($\alpha > 1$) [6]. The conditional hazard function for the Weibull model is given by the following equation:

$$h(t|k) = \lambda e^{\beta' \underset{\sim}{x}} \alpha t^{\alpha - 1}, \quad t > k. \tag{2}$$

To assess the effect of late entry in the hazard estimation, we will include in the vector of covariates an indicator variable that represents those individuals whose entrance to the study was after 1 January 1968. The effect of this dummy variable will also be assessed by considering interactions with other variables.

A competing risks situation arises when the individuals under study can experience more than one type of event [11]. Competing risks models address the differences in probability of death from different causes at different times. For example, the annual risk of death from sudden infant death syndrome, or motor vehicle accidents, does not follow the same pattern as risk of death from respiratory failure. For the competing risks models, we will follow the notation proposed by Crowder [4]. Assume that there are $j = 1, \ldots, m$ different causes of failure and the late entrants are independent from one another. The sub-hazard function $h(j, t|k)$, also known as the "cause-specific hazard function" or the "crude hazard function", represents the instantaneous failure rate from cause j at time t in the presence of the other types of failures, given that the entry time to the study was k. It is defined by the following expression:

$$h(j, t|k) = \lim_{\varepsilon \to 0} \frac{P(C = j, T \leq t + \varepsilon | T > k, T > t)}{\varepsilon}, \quad t > k.$$

The overall hazard rate is defined by the sum of each one of the sub-hazard functions for all the causes of failure:

$$h(t|k) = \sum_{j=1}^{m} h(j, t|k).$$

Expressing the overall hazard rate in terms of sub-hazard functions allows us to define a model that accounts for all the competing causes of death.

Suppose that the values $x_1, x_2, \ldots, x_p$ of p explanatory variables $X_1, X_2, \ldots, X_p$ are recorded for each of n individuals in the sample. Assume that each one of the sub-hazard functions is distributed as a truncated Weibull with cause-specific scale parameter λ_j and shape parameter α_j. Then the sub-hazard function for the cause of failure j can be written as

$$h(j, t|k) = \lambda_j e^{\beta_j' \underset{\sim}{x}} \alpha_j t^{\alpha_j - 1}, \quad t > k, \tag{3}$$

where $e^{\beta_j' \underset{\sim}{x}} = e^{\beta_{j1}x_1 + \beta_{j2}x_2 + \cdots + \beta_{jp}x_p}$ is the linear component of the model. Thus the model is specified such that the scale and shape parameters, as well as the effects of covariates, may differ with

each cause of death. Consequently, the overall hazard function is

$$h(t|k) = \sum_{j=1}^{m} \lambda_j \alpha_j t^{\alpha_j - 1} \mathrm{e}^{\beta_j' \underset{\sim}{x}}, \quad t > k. \tag{4}$$

As for the sub-hazard functions, the truncated time only affects the time in which the marginal hazard is observed. The truncated-marginal survivor function can be written in terms of the sub-hazards as

$$S(t|k) = \exp\left(- \sum_{j=1}^{m} \int_{k}^{t} h(j, u|k)\, \mathrm{d}u \right), \tag{5}$$

note that the truncated time only affects the limits of the integral.

Assume that for each of the n individuals involved in the A-bomb cohort, we can observe the data $(t_i, k_i, j_i, \underset{\sim}{x}_i)$ where t_i is the failure or censored time, k_i the entry time to the study, j_i the cause of death and $\underset{\sim}{x}_i$ a vector of covariates. Then, the expression for the likelihood in terms of sub-hazard functions for model given by Equation (3) is

$$L_n = \prod_{i=1}^{n} \left\{ \exp\left(- \sum_{j=1}^{m} \lambda_j \mathrm{e}^{\beta_j' \underset{\sim}{x}_i} (t_i^{\alpha_j} - k_i^{\alpha_j}) \right) \prod_{j=1}^{m} (\lambda_j \mathrm{e}^{\beta_j' \underset{\sim}{x}_i} \alpha_j t_i^{\alpha_j - 1})^{I(c_i = j)} \right\}, \tag{6}$$

where $I(\cdot)$ is an indicator function. The likelihood function can be maximized by using the function *optim* from the library `MASS` in `S-plus`. By using the gradient of the log-likelihood function within the *optim* function, one can obtain the standard errors of the estimators.

There are various advantages for using a sub-hazard approach when competing risks are present. As shown above, the likelihood function can be completely specified in terms of sub-hazard functions. Prentice *et al*. [12] point out that this property enables factorization of the likelihood function into separate components for each cause of death. The factorization together with standard survival techniques makes it clear that the sub-hazard functions are identifiable, i.e. they can be directly estimated from the data without introducing assumptions concerning the relationship between the different causes of death [12, pp. 544–546]. This is particularly useful in medical applications where it is questionable to assume that the different risks of death act independently.

4. Results

The use of time from entry to study as the time scale for modelling occurrence of events is reasonable in randomized controlled trials. Some authors have shown that using this time scale can bias estimates in epidemiological studies [7,15]. As actual age is the natural time scale when mortality is the focus, we model on this scale, and consider factors that modify natural aging. Delayed entry to the study is of interest. In the A-bomb survivors study, there is no length of exposure to a risk factor, so time-on-study does not have a link to duration or dose of exposure. All the individuals in the cohort are left truncated according to their age at the entry time to the study: the RP are left truncated at the age that they were in January 1968, whereas for the late entrants the truncation arises from their age at entry time.

If we validate the assumption that the distribution of failure times for the A-bomb survivors has a Weibull distribution, then our parameter estimates will be more precise than the ones given by semi-parametric models [2, p. 107]. A log-cumulative hazard plot of both the RP and the LE gave a fairly straight line, suggesting that the failure times for both groups have a Weibull distribution.

We split the data in male and female individuals, as there are analyses that demonstrate that there are differences in mortality between genders [13].

Table 2. Parameters estimate for truncated Weibull models for men and women, all causes of death.

Model	Description	$\lambda \times 10^{14}$	α	ln(rad + 1)	Dist. < 3 km	Late	Log-like
2.A Models for men							
1	RP	4.61	7.00	0.038	− 0.145	–	− 29,682
2	LE	24.60	6.61	*0.076*	*− 0.092*	–	− 5266
3	RP + LE	5.93	6.94	0.039	− 0.130	–	− 34,952
4	RP + LE	5.24	6.94	0.039	0.132	*− 0.006*	− 34,952
2.B Models for women							
1	RP	0.006	8.38	0.072	− 0.206	–	− 38,152
2	LE	0.042	7.91	*0.141*	*0.457*	–	− 5960
3	RP + LE	0.007	8.33	0.074	− 0.180	–	− 44,121
4	RP + LE	0.008	8.31	0.073	− 0.212	*− 0.113*	− 44,116

Notes: LE, late entries; RP, registered population; rad, radiation; Dist., distance; Log-like, loglikelihood; ln, natural logarithm.
Non-significant parameters are in *italics*.

For the purpose of understanding the late entry effect, we fitted the truncated Weibull proportional hazards model described by Equations (1) and (2) to the male and the female populations (Table 2). We used the deviance (minus twice the logarithm of the likelihood function) to compare differences between models. We modelled late entrants and RP in separate models (Models 1 and 2 for men and women), in a single model (Model 3) and in a single model differentiating late entrants by an indicator variable (Model 4). For each gender, we added up the log-likelihood values of Models 1 and 2 and compared it with the resultant log-likelihood of the full model (Model 3). As Model 3 is nested within Model 4, we were able to compare both models to assess the

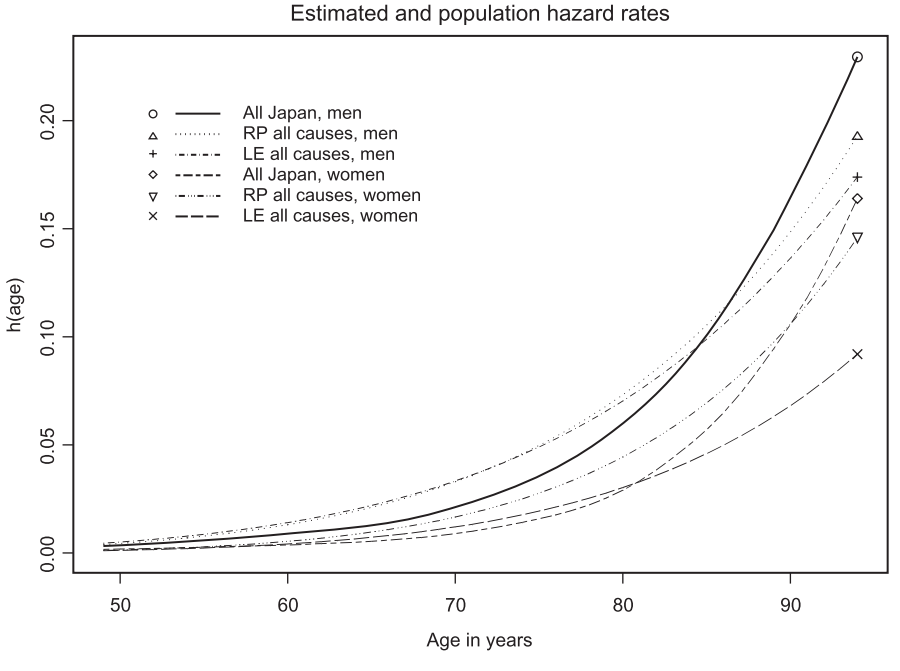

Figure 2. Hazard rates for Japanese population compared with estimated hazard rates from truncated Weibull model for men and women from the A-bomb Hiroshima survivors cohort. RP = Registered population. LE = Late entries.

effect of the late entry indicator. For men, there is no statistically significant difference between the combined Models 1 and 2 and Model 3 (d.f. = 4, p-value = 0.05). Also, there is no significant evidence to suggest that adding the late entry indicator to Model 3 improves the explanatory power of the model (d.f. = 1, p-value = 0.84). For women, the comparisons show that there is a difference between the two populations, as the combined Models 1 and 2 and Model 3 are statistically significantly different (d.f. = 4, p-value = 0.001). Also, there is a statistically significant difference between the nested Models 3 and 4 (d.f. = 4, p-value = 0.001). However, Model 4 is incapable of fully capturing the effect of late entrants as the indicator variable is not statistically significant. These results suggest the use of a competing risk framework, in men to rule out any possible late entry effect, whereas in women the competing risks approach will allow us to establish if differences in survival between female late entrants and RP are attributable to different causes of death. Figure 2 shows that the RP has higher hazard, or annual death rate, than the general female population of Japan until age 90. The late entry population's hazard is lower, and becomes less than the general population hazard about age 80. For men, there is less difference between RP and LE overall, and the population rate exceeds the A-bomb survivor rates after about 85 years.

Tables 3–5 present the results for the truncated competing risks Weibull models, fitted to men and women. The model is defined by Equations (3)–(5) and the likelihood function is given by Equation (6). For men, models with and without interaction terms are very similar. In the model without interactions, late entry and being within 3 km of the hypocentre are associated with a small increase in cancer mortality (Figure not shown). Figure 3 illustrates the difference in hazard

Table 3. Competing risks model for men – main terms only.

	Cancer β (SE)	Heart disease β (SE)	Other causes β (SE)
$\lambda \times 10^{14}$	665.260 (4.43E − 23)	0.001 (5.83E − 35)	4.560 (4.18E − 28)
α	5.565 (0.011)	8.604 (0.016)	6.782 (0.010)
Rad.	0.078 (0.019)	*0.021* (*0.019*)	*0.027* (*0.017*)
Dist.	− 0.147 (0.080)	− 0.173 (0.074)	*− 0.073* (*0.068*)
LE	*0.100* (*0.063*)	*− 0.102* (*0.064*)	*0.009* (*0.056*)
log likelihood		− 43,876.85	

Notes: LE, late entries; Rad., ln(radiation + 1); Dist., distance; SE, standard error.
Non-significant parameters are in *italics*.

Table 4. Competing risks model for men – including interaction terms.

	Cancer β (SE)	Heart disease β (SE)	Other causes β (SE)
$\lambda \times 10^{14}$	745.920 (5.56E − 23)	0.001 (5.43E − 35)	4.690 (4.42E − 28)
α	5.540 (0.011)	8.617 (0.016)	6.780 (0.010)
Rad.	0.073 (0.020)	*0.026* (*0.019*)	*0.029* (*0.017*)
Dist.	*− 0.134* (*0.083*)	− 0.217 (0.076)	*− 0.112* (*0.071*)
LE	*0.100* (*0.071*)	− 0.172 (0.071)	*− 0.053* (*0.063*)
Rad.*LE	*0.196* (*0.469*)	*0.021* (*0.106*)	*− 0.038* (*0.083*)
Dist.*LE	− 0.735 (0.385)	*0.141* (*0.369*)	*0.434* (*0.282*)
log likelihood		− 43,871.37	

Notes: LE, late entries; Rad., ln(radiation + 1); Dist., distance; SE, standard error.
Non-significant parameters are in *italics*.

Table 5. Competing risks model for women – main terms only.

	Cancer β (SE)	Heart disease β (SE)	Other causes β (SE)
$\lambda \times 10^{14}$	1919.560 (3.68E − 22)	0.000001 (4.43E − 41)	0.0002 (1.35E − 36)
α	5.143 (0.013)	10.092 (0.014)	8.893 (0.014)
Rad.	0.172 (0.022)	0.042 (0.017)	0.050 (0.018)
Dist.	− 0.347 (0.091)	− 0.175 (0.064)	− 0.173 (0.070)
L.E.	0.132 (0.066)	− 0.231 (0.053)	− 0.148 (0.056)
log likelihood		− 54,513.91	

Notes: LE, late entries; Rad., ln(radiation + 1); Dist., distance; SE, standard error.

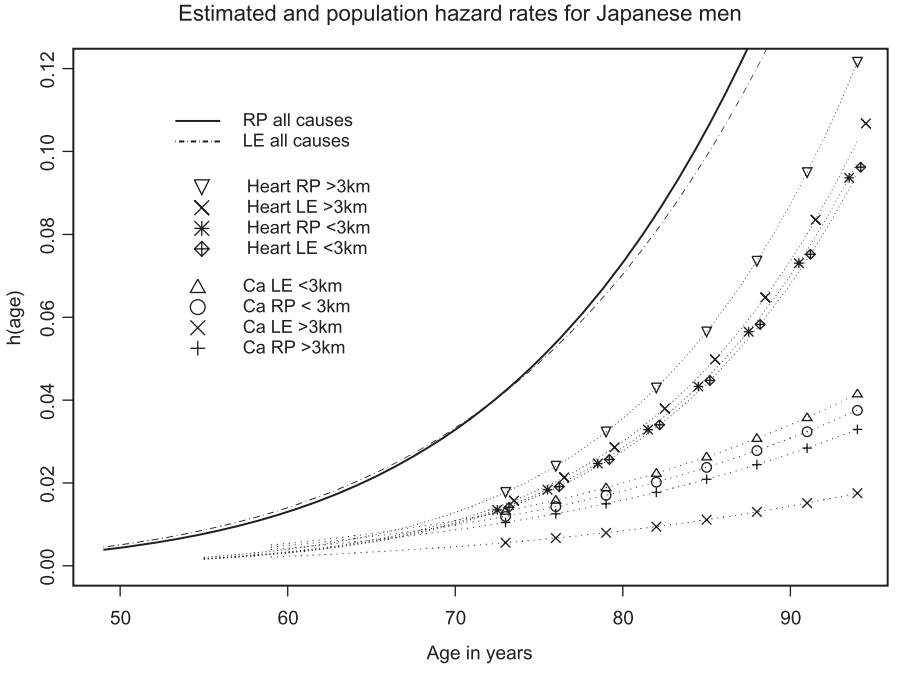

Figure 3. Hazard rates for Japanese men compared with estimated hazard rates from truncated Weibull competing risks model for men from the A-bomb Hiroshima survivors cohort. RP = Registered population. LE = Late entries; CA = Cancer.

rates for cancer and heart disease among men for the model with interactions. From about age 60, the risk of dying from heart disease increases more rapidly than cancer risk. Radiation exposure increases cancer mortality slightly, about the median radiation dose (20 cSv) the risk of cancer increases in 25%. These differences are reversed for heart disease: late entrants have slightly lower mortality, as do those within 3 km of the hypocentre. The interaction model estimates late entry, more distant men as having substantially lower-cancer mortality than the other three groups.

Late entry interaction terms are not significant for women, but the main effects are significant for cancer, heart and other causes (Table 5). For cancer, late entry is associated with increased risk ($e^{0.132} = 1.14$), and decreased risk for heart ($e^{-0.231} = 0.79$) and other causes ($e^{-0.148} = 0.86$). Radiation exposure is associated with increased mortality for all three causes, with the effect on cancer mortality three times greater than on heart and other causes and much higher than men.

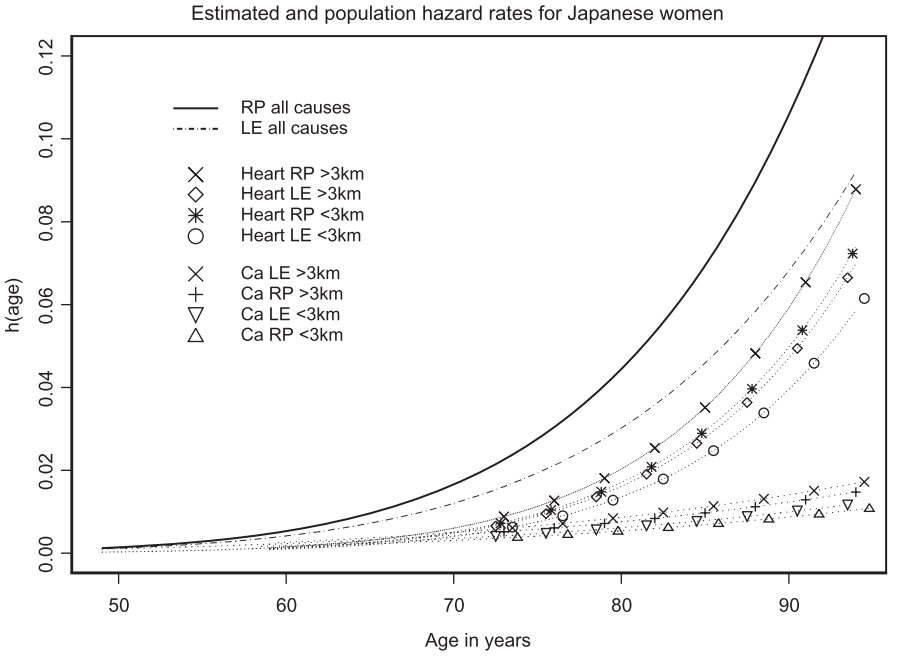

Figure 4. Hazard rates for Japanese women compared with estimated hazard rates from Weibull competing risks model for women from the A-bomb Hiroshima survivors cohort. RP = Registered population. LE = Late entries; CA = Cancer.

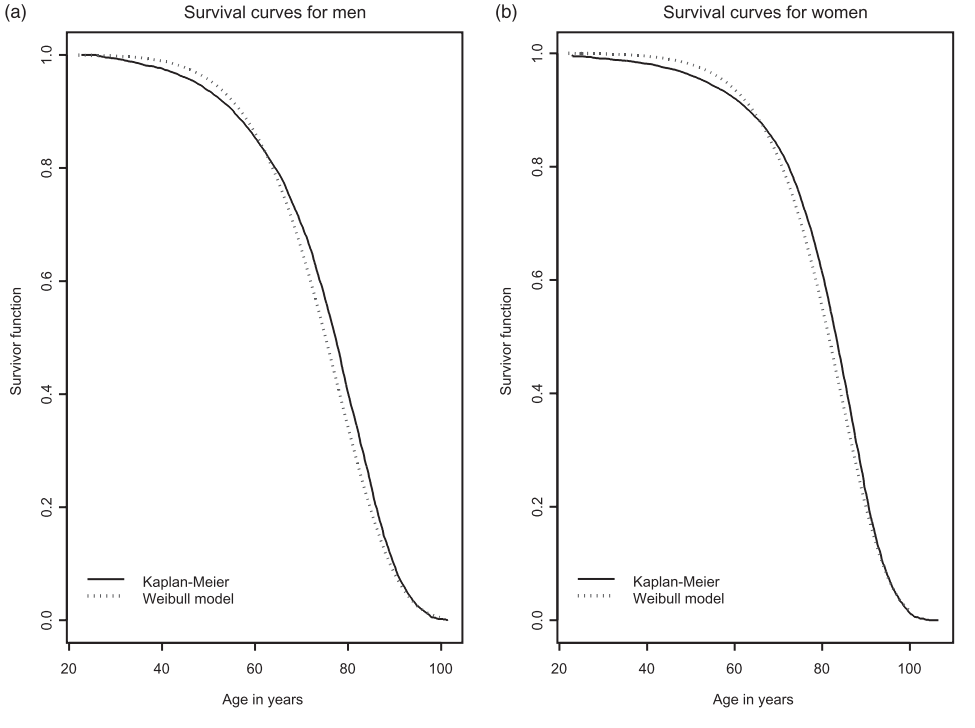

Figure 5. Kaplan–Meier curve and overall survivor function estimated by Weibull model (a) for men, and (b) for women. All the curves are for all causes of death.

Women within 3 km of the hypocentre have reduced mortality, by 29% for cancer and 16% for heart disease. Figure 4 shows that women have lower cancer and heart mortality rates than men. As for men, RP, distant women have the highest heart mortality rate, and late entrants within 3 km have the lowest rates. Cancer mortality increases very slowly with age.

The goodness of fit of the model can be assessed by visual inspection of Figure 5. Assuming that the Kaplan–Meier survivor curve adjusted for truncation estimated for men and women shows the real distribution of the survival times, then one can see that the Weibull model fits the data relatively well. In both cases, the model overestimates slightly the survival probabilities up to age 65, then there is a minimal underestimation.

5. Discussion

The aim of this study was to investigate if there was a difference in the survival experience of late entrants and RP in the Hiroshima cohort. We used a competing risks framework based on sub-hazard functions because it is not necessary to make assumptions with respect to the interrelation among the causes of failure, and direct estimation of the sub-hazard functions from the data allows the model to be identifiable.

A log-cumulative hazard plot validated a distributional assumption for the failure times: a straight line suggested that they follow a Weibull distribution. The analysis developed in this article was inspired by Matsuura and Eguchi [8]. Our aim was to propose a different approach to modelling the late entry effect in the Hiroshima data. Matsuura and Eguchi used a semi-parametric model to show that there is a late entry effect in the male population. In contrast to Matsuura and Eguchi, we selected a parametric approach rather than a semi-parametric proportional hazards model because it was possible to demonstrate that the survival times followed a particular distribution. Therefore, our estimates of hazard ratios and median survival times will have smaller standard errors than they would in the absence of a distributional assumption.

Matsuura and Eguchi only presented results for men considering separate models for all causes of death and cancer. We used a competing risks framework that accounts for the difference in probability of death from different causes at different times. Our models clearly show that the hazard functions differ between the three causes of death and how they differ, for example, the value of the shape parameter for heart disease demonstrates that the hazard increases faster for this condition than for the other two causes of death.

Matsuura and Eguchi found a late entry effect in the male group; result that is congruent with our competing risks analysis as our model suggests the presence of a late entry effect for the male cancer and heart disease groups. Our model also suggested the presence of a late entry effect in the male cancer group located at less than 3 km from the hypocentre, while Matsuura and Eguchi concluded that there was no entry effect in the cancer group. We were also able to explore the late entry effect in the female population. Late entry decreases the risk of heart disease and other causes while it has the opposite effect with cancer risk. The late entry effects are more statistically significant for women than for men. The effect of late entry for women is in the main effects and shows the increased hazard in cancer rates, at the expense of the two other causes of death. These effects are statistically significant with respect to all causes of death.

This is only a descriptive study of the Hiroshima population as the generalization of the results would be difficult to all the A-bomb survivors. This study only includes individuals that declared themselves as A-bomb survivors. All the individuals in this study were willing to ignore the Japanese stigma that surrounds the A-bomb survivors, and decided to make use of the health and welfare benefits that this status gave them. Thus the observational study ignores those individuals who were never registered as A-bomb survivors. We cannot assume that their survival experience is the same as those declaring themselves to be A-bomb survivors.

We have shown that a straight-forward approach to assess late entry effect can provide valuable insight into the magnitude of the effect in the survival probabilities. Furthermore, our competing risk models allowed us to evaluate more precisely the nature of this late entry effect. The models showed the possible existence of a late entry effect in the male and female groups. The significance of this variable may be explained by difference in population attitudes towards healthcare behaviour and the stigma of being an A-bomb survivor. The competing risks framework was also useful to understand the effect of the remaining covariates for each one of the causes of death.

Our study population is different from the one used by the Radiation Effects Research Foundation in the Life Span Study. The latter is widely known as an informative epidemiological study for establishing radiation risks [10,14]. The main differences between the two study populations are the start of follow-up, definition of the cohort and study area. Besides these differences, our results are similar to the ones presented by Shimizu *et al*. [14] in the sense that they also found a statistically significant association between radiation dose and non-cancer disease mortality. Shimizu *et al*. [14] provide a very clear analysis of the biases that might have led to this spurious association, concluding that even though it is difficult to explain the association, it might be wrong to conclude that such association is nonexistent.

Acknowledgements

The data for this research were obtained from the records of the RIRBM at Hiroshima University, Japan. Professor N. Hayakawa at this institute is the principal investigator of the database. We would like to thank Dr. Shinto Eguchi and Dr. Masaaki Matsuura for preparing for us the database and their helpful comments, and Prof. John B. Copas and Dr. Brian Tom for their constructive comments. We are also thankful to three anonymous referees for their helpful remarks. This research was supported partly by MRC Grant number U.1052.00.011 and the Mexican Council for Science and Technology (CONACYT), Grant #160987.

References

[1] K. Bull and D.J. Spiegelhalter, *Survival analysis in observational studies*, Stat. Med. 16 (1997), pp. 1041–1074.
[2] D. Collett, *Modelling Survival Data in Medical Research*, Chapman and Hall, London, 1994.
[3] D.R. Cox and D. Oakes, *Analysis of Survival Data*, Chapman and Hall, London, 1984.
[4] M. Crowder, *Classical Competing Risks*, Chapman and Hall/CRC, Boca Raton, 2001.
[5] M. Hoshi, M. Matsuura, N. Hayakawa, C. Ito, and N. Kamada, *Estimation of radiation doses for atomic bomb survivors in the Hiroshima university registrey*, Health Phys. 70 (1996), pp. 735–740.
[6] D. Kalbfleish and L. Prentice, *The Statistical Analysis of Failure Time Data*, John Wiley and Sons, New York, 1980.
[7] E.L. Korn, B.I. Graubard, and D. Midthurne, *Time-to-event analysis of longitudinal follow-up of a survey: Chaoice of the time-scale*, Amer. J. Epidemiol. 145 (1997), pp. 72–80.
[8] M. Matsuura and S. Eguchi, *Modeling late entry bias in survival analysis*, Biometrics 61 (2005), pp. 559–556.
[9] M. Matsuura, N. Hayakawa, and H. Shimokata, *Survival analyses of atomic bomb survivors in Hiroshima prefecture, Japan, 1968–1982. – Cancer mortality risk among early entrants*, Hiroshima J. Med. Sci. 44 (1995), pp. 29–38.
[10] D.A. Pierce, Y. Shimizu, D.L. Preston, M. Vaeth, and K. Mabuchi, *Studies of the mortality of A-bomb survivors. Report 12, Part I. Cancer: 1950–1990*, Radiat. Res. 146 (1996), pp. 1–27.
[11] M. Pintilie, *Competing Risks. A Practical Perspective*, John Wiley and Sons, Chichester, 2006.
[12] R.L. Prentice, J.D. Kalbfleisch, A.V. Peterson Jr., N. Flournoy, V.T. Farewell, and N.E. Breslow, *The analysis of failure times in the presence of competing risks*, Biometrics 34 (1978), pp. 541–554.
[13] G. Scambler, *Sociology as Applied to Medicine*, Saunders, Edinburgh, 2003.
[14] Y. Shimizu, D.A. Pierce, D.L. Preston, and K. Mabuchi, *Studies of the mortality of A-bomb survivors. Report 12, Part II. Noncancer mortality: 1950–1990*, Radiat. Res. 152 (1999), pp. 374–389.
[15] A.C.M. Thiébaut and J. Bénichou, *Choice of time-scale in Cox's model analysis of epidemiologic cohort datea: A simulation study*, Stat. Med. 23 (2004), pp. 3803–3820.