

# Journal of Dental Research

<http://jdr.sagepub.com>

---

## Use of the Weibull Hazard Model to Estimate Age-specific Probability of Permanent Tooth Loss

K. Otani, M. Ohtaki and M. Fujita

*J DENT RES* 1996; 75; 1458

DOI: 10.1177/00220345960750070501

The online version of this article can be found at:  
<http://jdr.sagepub.com/cgi/content/abstract/75/7/1458>

---

Published by:



<http://www.sagepublications.com>

On behalf of:

International and American Associations for Dental Research

**Additional services and information for *Journal of Dental Research* can be found at:**

**Email Alerts:** <http://jdr.sagepub.com/cgi/alerts>

**Subscriptions:** <http://jdr.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations** <http://jdr.sagepub.com/cgi/content/refs/75/7/1458>

# Use of the Weibull Hazard Model to Estimate Age-specific Probability of Permanent Tooth Loss

K. Otani<sup>1\*</sup>, M. Ohtaki<sup>2</sup>, and M. Fujita<sup>1</sup>

<sup>1</sup>Department of Oral and Maxillofacial Radiology, Hiroshima University School of Dentistry, 1-2-3, Kasumi, Minami-ku, Hiroshima 734, Japan; <sup>2</sup>Department of Environmetrics and Biometrics, Research Institute for Radiation Biology and Medicine, Hiroshima University, 1-2-3, Kasumi, Minami-ku, Hiroshima 734, Japan; \*to whom correspondence and reprint requests should be addressed

**Abstract.** In analysis of the probability of tooth loss with age, the exact time of tooth loss is often unknown, although it is clear whether a tooth remains or has been lost. That is, left censoring is inevitable during data sampling. This may provide a biased estimate if such data are dealt with by the product-limit method, which is a common method of survival analysis. To reduce such a bias in estimating the age-specific probability of tooth loss, we developed a survival analysis method taking left censoring into consideration. Four hundred and forty-six panoramic radiographs obtained in a daily clinical practice were used. The frequency of tooth loss with age was assumed to follow the Weibull hazard model, and a likelihood function taking left censoring into consideration was defined to estimate the probability of tooth loss. The estimate obtained from our method was compared with that from the product-limit method to examine whether the effect of left censoring was reduced. We found that the probability of tooth loss estimated by the product-limit method was biased by left-censored data, and that the bias was reduced when our method was used. A Monte Carlo simulation study, in which the true tooth loss time was given, also showed that our method provided an estimate closer to the true value. Our method is considered to be more accurate in estimating the probability of tooth loss, since it reduces the bias caused by left-censored data.

**Key words:** tooth loss, epidemiology, Weibull hazard model, survival analysis, left censoring.

## Introduction

The probability of tooth loss is one of the indicators used in estimating the oral health condition of a patient. Dental disease has been surveyed every six years since 1957 by the Ministry of Health and Welfare (1988) in our country. In the 1987 report, the life span of each permanent tooth was estimated by the life-table method (Cutler and Ederer, 1958), which is a non-parametric method of survival analysis. This method was developed to deal with a large sample size. However, its major shortcoming is that it is time-consuming and expensive for reliable samples to be collected. Another shortcoming is that it is difficult to investigate the relationships between tooth loss and factors relating to tooth loss. In survival data analysis, when a sample size is small, the product-limit method (Kaplan and Meier, 1958) is often used as an alternative non-parametric inference method.

In a survival study on the life span of teeth, the basic observation for each sample is the time between eruption of the tooth and tooth loss. This is called the waiting time (life span of the tooth). If tooth loss has not occurred before oral examination, the observation is called "right-censored". When tooth loss has occurred before oral examination, and the time of tooth loss is unknown, the observation is called "left-censored". Furthermore, an observation is called "failed" if the exact time of tooth loss is known (Armitage, 1971; Kalbfleisch and Prentice, 1980; Lee, 1980). In general, data sets used in survival analysis consist of right-censored cases and failed cases. In the study of the life span of teeth, it is difficult to determine the time of tooth loss accurately, and generally we can obtain information about only whether a tooth is present at the time of examination. Therefore, such data sets consist of left-censored as well as right-censored cases. Left censoring is inevitable in any study on the life span of teeth. If left-censored data are used in the product-limit method, the result might be biased. However, there have been no tooth survival analysis methods which take this into consideration.

Thus, we developed a tooth survival analysis method for a small sample size, which reduces the bias produced by left-censored data. We describe here our procedure and its

Received July 19, 1995; Accepted March 25, 1996

**Table 1.** Patients' age and sex distributions at the time of radiographic examination

Age (years)	No. of Patients (%)	
	Male	Female
15-19	33 (18.2)	31 (12.2)
20-24	11 (5.8)	26 (10.2)
25-29	11 (5.8)	19 (7.4)
30-34	18 (9.4)	17 (6.7)
35-39	12 (6.3)	23 (9.0)
40-44	23 (12.0)	29 (11.4)
45-49	16 (8.4)	26 (10.2)
50-54	25 (13.1)	21 (8.2)
55-59	13 (6.8)	31 (12.2)
60-64	11 (5.8)	9 (3.5)
65-69	8 (4.2)	14 (5.5)
70-74	8 (4.2)	5 (2.0)
75-79	2 (1.0)	4 (1.6)
Total	191 (100)	255 (100)

clinical application to estimating the probability of tooth loss with age. We examined the usefulness of our method by comparing it with the product-limit method.

**Materials and methods**

**Materials**

Between January, 1991, and June, 1994, 9112 patients received a panoramic radiographic examination for a consultation and/or treatment of their illness at our hospital. All radiographs are kept in our hospital, and 500 panoramic radiographs from among them were randomly selected for this study. However, 34 panoramic radiographs were excluded because those patients had deciduous and/or mixed dentition. Twenty panoramic radiographs were also excluded because they were not suitable for

observing teeth. The remaining 446 radiographs were used in this study. The patients' age and sex distributions are shown in Table 1. We examined a panoramic radiograph from each patient to determine whether a permanent tooth, except a third molar, was present. The format of tooth loss data is illustrated in Table 2. In this study, our interest was in the time of tooth loss, not the duration of existence of a tooth. The starting point of this study was the birth of the subject. The end point was the time of tooth loss. Tooth loss probability was examined for the left maxillary central incisor, canine, 1st molar, and 2nd molar, although tooth loss data were taken on 28 permanent teeth.

**Statistical analysis designed for estimating tooth loss rate**

In this study, panoramic radiographs were used to obtain data; thus, the exact time of tooth loss was difficult to determine. Therefore, we assumed that the frequency of tooth loss with age followed the Weibull hazard (Weibull, 1951; Aitken and Clayton, 1980). The hazard function,  $h(t)$ , of tooth loss at the time, or age,  $t$ , is given by:

$$h(t) = \frac{\beta}{Y^\beta} t^{\beta-1}$$

where  $\beta$  and  $Y$  are the unknown Weibull parameters to be estimated. The distribution function,  $F(t)$ , is then given by:

$$F(t | \beta, Y) = 1 - \exp\{-(t/Y)^\beta\}$$

Since every observation should be either left- or right-censored cases, the likelihood function,  $L(\beta, Y)$ , can be expressed as:

$$L(\beta, Y) = \prod_{i=1}^n F(t_i | \beta, Y)^{\delta_i} \{1 - F(t_i | \beta, Y)\}^{1-\delta_i} \\ = \prod_{i=1}^n [1 - \exp\{-(t_i/Y)^\beta\}]^{\delta_i} [\exp\{-(t_i/Y)^\beta\}]^{1-\delta_i}$$

where  $i$  is the indicator of  $i$ th observation, and  $\delta_i$  is the indicator such that

**Table 2.** Format of tooth loss data

Obs. $i^a$	Age	Sex	Left Maxillary						
			Central Incisor	Lateral Incisor	Canine	First Premolar	Second Premolar	First Molar	Second Molar
1	45	Male	1 <sup>c</sup>	0 <sup>b</sup>	0	1	0	0	1
2	37	Female	1	0	0	0	0	0	1
3	19	Female	0	0	0	0	0	0	0
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
446	28	Female	0	0	0	0	1	0	0

<sup>a</sup> $i$  = identification number (sequential).  
<sup>b</sup>0 = tooth present at radiographic examination.  
<sup>c</sup>1 = tooth not present at radiographic examination.

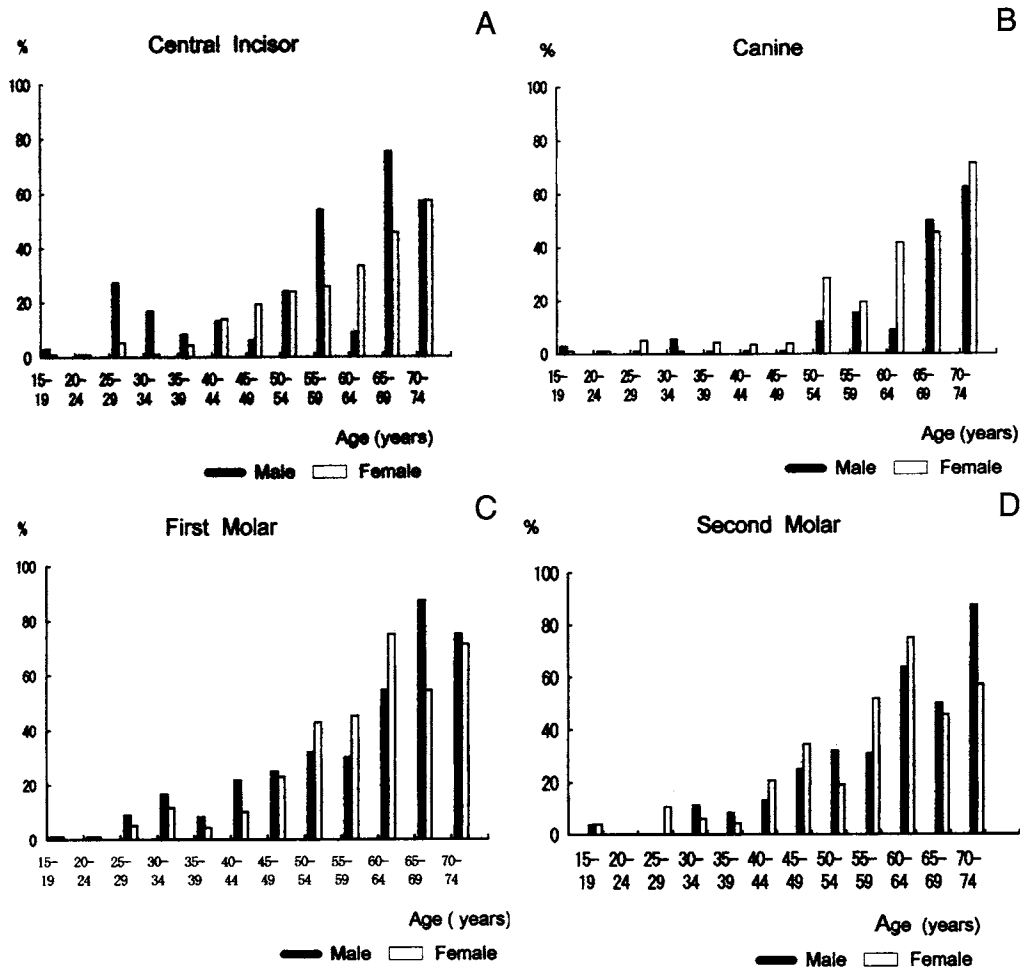


Figure 1. Histograms of sex- and age-specific tooth loss of the central incisor (a), canine (b), 1st molar (c), and second molar (d).

$$\delta_i = \begin{cases} 1 & \text{if the } i\text{th observation is the left-censored case,} \\ 0 & \text{if the } i\text{th observation is the right-censored case.} \end{cases}$$

Unknown Weibull parameters,  $\beta$  and  $\gamma$ , in the likelihood function were estimated by the maximum likelihood method (Cohen, 1965). This calculation was performed by means of a software program written in FORTRAN. We also used the product-limit method for analyzing our data on the assumption that the age at tooth loss was identical to the age at radiographic examination. Results obtained from these two methods were compared.

## Results

Histograms of the tooth-loss rate of the left maxillary central incisor, canine, 1st molar, and 2nd molar for men and women at different ages (Fig. 1) showed increased tooth-loss rates with age. Two curves of the probability of tooth loss for each tooth, one of which was obtained by our method

and the other by the product-limit method, respectively, are shown in Fig. 2. The estimated Weibull parameters ( $\beta$ ,  $\gamma$ ) of a maxillary central incisor, canine, 1st molar, and 2nd molar were (1.80, 101.7), (2.87, 105.8), (2.82, 70.5), and (2.84, 73.4), respectively, in males, and (2.85, 87.6), (1.30, 91.4), (2.80, 70.7), and (2.82, 70.3), respectively, in females. The curves produced by our method showed that the probabilities of having lost a left maxillary central incisor, canine, 1st molar, or 2nd molar by the age of 80 were 48%, 37%, 76%, and 72%, respectively, in males, and 55%, 49%, 75%, and 75%, respectively, in females. Molars had the highest probabilities of being lost at any age. The curve produced by our method was located under that produced by the product-limit method until the two crossed between the ages of 60 and 70.

## Monte Carlo simulation study

We applied the Monte Carlo simulation method (Morgan, 1984) to determine whether our method reduced the bias caused by left censoring. The Monte Carlo simulation data

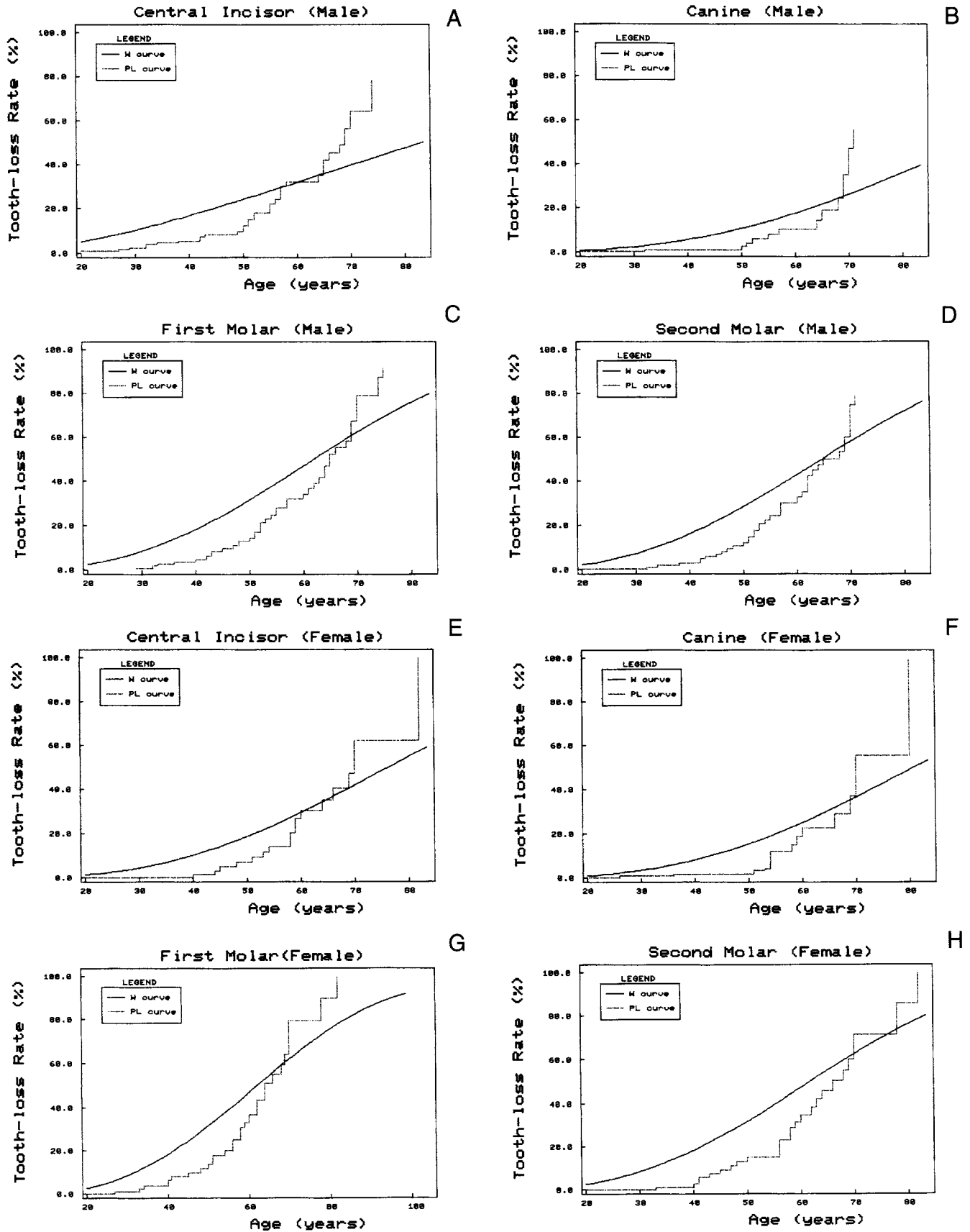
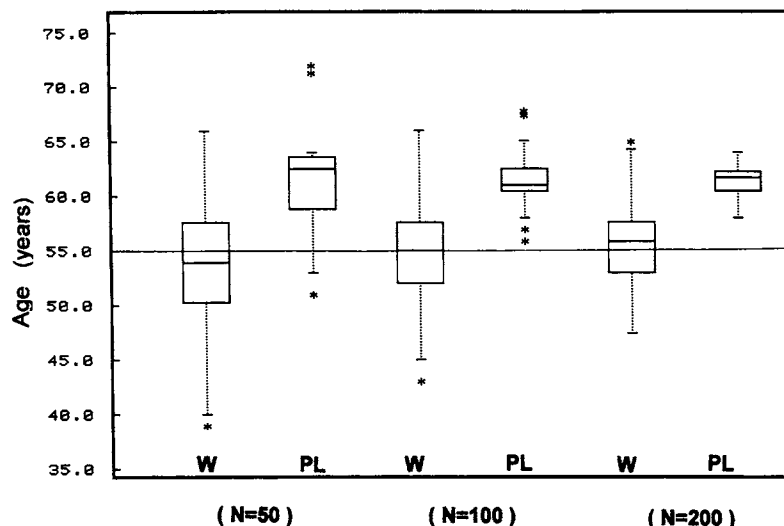


Figure 2. Tooth-loss rate curves for the central incisor, canine, 1st molar, and 2nd molar [males, (a) through (d); females, (e) through (h)] of the maxilla obtained by our method (—) and by the product-limit method (---).



**Figure 3.** Box plots of medians estimated by our method (W) and by the product-limit method (PL) for three series of 500 simulated data sets. The median for the true age distribution was 55 years. N denotes sample size.

were obtained as follows: The Weibull distribution with a median age of 55 years ( $\beta = 2.3$ ,  $\gamma = 92.5$ ) was assumed to show the true time of tooth loss. The age when the probability of tooth loss was 50% was defined as the median. The age of the patients at radiographic examination was assumed to be normally distributed, with a mean age of 50 years and a standard deviation of 20 years. These data were generated from the Weibull and normal random tables. For every individual, a tooth was judged to exist if the true age at tooth loss was higher than the patient's age at the time of radiographic examination, and a tooth was judged to be lost if the true age at tooth loss was lower than the patient's age at the time of radiographic examination. Three series of 500 data sets were made. Each data set consisted of 50, 100, or 200 simulated data. Illustrated in Fig. 3 are the distributions of the estimated medians obtained from our method and the product-limit method with three series of 500 simulated data sets in box plots (Tukey, 1977).

The medians obtained from our method were distributed around the median of the Weibull distribution, 55 years of age. The subjects were 54 years of age in a series of 500 data sets each consisting of 50 simulated data, 55 years of age in a series of data sets each consisting of 100 simulated data, and 56 years of age in a series of data sets each consisting of 200 simulated data. However, when the product-limit method was used, medians were distributed around 62 years of age. The subjects were 62 years of age in a series of 500 data sets each consisting of 50 simulated data, 61 years of age in a series of data sets each consisting of 100 simulated data, and 62 years of age in a series of data sets each consisting of 200 simulated data. Thus, bias was reduced when our method was used.

## Discussion

In tooth survival analysis, it is necessary to know the exact time of tooth loss for the tooth survival rate to be estimated

correctly. A data set for a survival analysis, however, is often made up of two types of censored data. One case is when an incident occurs before the time of examination, and the occurrence time is not available (left-censored data). The other is when an incident has not yet occurred at the time of the examination (right-censored data). A data set used to analyze the probability of tooth loss consists of both left-censored and right-censored data. Left-censored data, however, have not been taken into consideration in conventional survival analysis methods, and this may lead to biased estimates.

In this study, we developed a tooth survival analysis method for estimating the probability of tooth loss from a small data set containing left-censored data, and examined its clinical usefulness. We also used a Monte Carlo simulation study in comparison with the product-limit method.

The tooth loss probability curves for each tooth, obtained with our method and with the product-limit method, crossed between 60 and 70 years of age. The curve estimated by the product-limit method was located below the curve estimated by our method in a younger age group, and this was reversed in an older age group. Analysis by the product-limit method disregards the left censoring of data, and the age at radiographic examination when a tooth was found to be missing was regarded as the age at tooth loss. This means that the failure time assessment is too old. Therefore, the interval between failure times is assessed longer than the actual time interval in the younger age group, and the probability of tooth loss in the younger age group is then estimated to be lower than its actual value. In the older age group, the interval between failures is assessed for too short a time, and thus the probability of tooth loss was estimated to be higher. Such bias was reduced by our method.

A Monte Carlo simulation study showed that the product-limit method gave biased estimates, and that such a bias was reduced when our method was used. This indicates that our method provides more reliable estimates for tooth survival analysis of data sets including left-censored data.

The frequency of tooth loss with age was assumed to follow the Weibull hazard in this study. The hazard function,  $h(t)$ , at the time, or age,  $t$ , is given by:

$$h(t) = \frac{\beta}{\gamma^\beta} t^{\beta-1}$$

The parameter  $\beta$  is called the shape parameter of the Weibull distribution, because it determines the shape of the survival curve by specifying the change in the hazard as a function of time. Specifically, if  $\beta = 1.0$ , the hazard is constant; if  $\beta > 1.0$ , the hazard increases in time; and if  $\beta < 1.0$ , the hazard decreases with time.

The Weibull model is sufficiently flexible to provide a generally useful tool for the analysis of sets of waiting time

of an incident, life span of a machine, or some kinds of cancer data (Byar, 1982), and it should not be viewed as a specialized parametric model of limited usefulness. The hazard of tooth loss is considered to increase monotonically with age; therefore, it is not inappropriate to apply the Weibull model to tooth loss data. Our interest was thus focused on the degree of performance of the Weibull model. However, no direct method for comparing the actual and predicted curves is available, since observations contain left-censored data.

A semi-parametric model, rather than a full-parametric model, may be more suitable for the analysis of the frequency of tooth loss with age. As yet, no useful semi-parametric model is available for analyzing tooth survival data that include left-censored data. This type of survival analysis procedure is required.

In this study, the probability of tooth loss was assumed to correlate with age alone, and single-variate analysis for age at tooth loss was performed. However, the probability of tooth loss may also be affected by background factors, such as sex, and period. Thus, a multivariate analysis method that takes other factors into consideration should be designed.

### Acknowledgments

This study was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education (No. 06302015), Japan.

### References

- Aitken M, Clayton D (1980). The fitting of exponential, Weibull and extreme value distributions to complex censored survival data using GLIM. *Appl Statist* 29:156-163.
- Armitage P (1971). *Statistical methods in medical research*. New York: Wiley.
- Byar DP (1982). Analysis of survival data: Cox and Weibull models with covariates. In: *Statistics in medical research*. Valerie M, Stanley KE, editors. New York: John Wiley & Sons, pp. 365-401.
- Cohen AC (1965). Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples. *Technometrics* 7:579-588.
- Cutler SJ, Ederer F (1958). Maximum utilization of the life table method in analyzing survival. *J Chron Dis* 8:699-712.
- Kalbfleisch J, Prentice RL (1980). *The statistical analysis of failure time data*. New York: John Wiley & Sons.
- Kaplan EL, Meier P (1958). Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 53:457-481.
- Lee ET (1980). *Statistical methods for survival analysis*. Belmont, CA: Lifetime Learning Publications.
- Ministry of Health and Welfare of Japan (1988). Report of the survey of dental disease in 1987. Tokyo: Medical Affairs Bureau.
- Morgan BJT (1984). *Elements of simulation*. London: Chapman and Hall.
- Tukey JW (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.
- Weibull W (1951). A statistical distribution function of wide applicability. *J Appl Mechan* 18:293-297.