Department of Global Public Health
Master Programme in Public Health Sciences
Public Health Epidemiology
Degree Project, 30 credits
Spring term 2020

Extrapolating cancer patient survival: a comparison of the flexible parametric model and the rolling-over algorithm

Date: 04 June 2020

Master thesis for Degree of Master of Medical Science (120c) with a Major in Public Health Sciences

Author: Enoch Yi-Tung Chen

Supervisors: Paul Dickman, Department of Medical Epidemiology and Biostatistics, Karolinska Institutet
Jung-Der Wang, Department of Public Health, College of Medicine, National Cheng Kung University

The Master program in Public Health Sciences at KI is carried out in collaboration between mainly three departments: The Department of Global Public Health, the Institute of Environmental Medicine and the Department of Medicine, Solna.

Jette Möller
Program Director

Department of Global Public Health
Master Programme in Public Health Sciences
Public Health Epidemiology
Degree Project, 30 credits
Spring term 2020

Extrapolating cancer patient survival: a comparison of the flexible parametric model and the rolling-over algorithm

Abstract

Background: With short-term survival data, extrapolation is often required to estimate survival over follow-up time. Extrapolating all-cause survival may not always give good estimate. The flexible parametric model and the rolling-over algorithm method have both been applied to extrapolate survival within relative survival framework, but no study has been done yet to compare these two approaches by using the same dataset.
Aim: To increase understanding of difference and sensitivity of extrapolating cancer patient survival between the flexible parametric model and the rolling-over algorithm.
Setting: Nationwide registry-based cancer cohort study, Sweden.
Methods: Four cohorts of cancer patients (colon cancer and breast cancer, diagnosis during 1981-1990 and 2001-2010) were identified from the Swedish Cancer Register. The follow-up period was limited and applied with the flexible parametric model and the rolling-over algorithm to extrapolate survival. The extrapolation results were compared to the empirical Kaplan-Meier survival functions.
Results: With longer follow-up periods, these two methods gave similar and precise extrapolation on survival functions. In contrast, with shorter follow-up data, the extrapolation curves potentially deviated more from the empirical Kaplan-Meier's curves; however, if approaching constant excess hazard within the available follow-up period, extrapolation by either method would potentially have good performance.
Conclusion: With sufficiently long follow-up data, extrapolating cancer patient survival carried out by either method can obtain reasonable accuracy. Uncertainty in extrapolation should be taken into consideration.

Keywords: extrapolation, breast cancer, colon cancer, survival analysis, flexible parametric model, rolling-over algorithm, relative survival, population-based

Contents

List of abbreviations

AIC	Akaike information criterion
BIC	Bayesian information criterion
AUC	Area under the curve
CI	Confidence interval
d.f.	Degrees of freedom
FPM	Flexible parametric model
K-M	Kaplan-Meier
LE	Life expectancy
LLE	Loss of life expectancy
ROA	Rolling-over algorithm

1 Background

1.1 Estimating loss of life expectancy

Life expectancy (LE) provides a summary of the survival experience of a cohort and is calculated as the integral of the survivor function (area under the survival curve) [1, 2]. For cancer patients, the loss of life expectancy (LLE) due to cancer could be interpreted as the difference between LE of cancer patients and cancer-free individuals [2, 3], illustrated in Figure 1. These quantities of LE are of interest to patients, clinicians and also offer resource-planning evidence for cancer prevention to policymakers [4, 5]. For instance, to estimate the societal impact of the burden of cancer, the lifetime survival function can be multiplied with a consequence function in order to quantify lifetime duration of quality of life, employment, disability, and healthcare expenditure [6, 7]. Furthermore, in health technology assessment studies, the amount of LLE saved is quantified as successful prevention of disease [7]. Estimating LE necessitates the survival function to reach zero probability, which typically requires a long follow-up to capture all the relevant future health outcomes [7, 8]. Short-term survival data will then require extrapolation[1] to estimate the lifetime survival [5, 8], especially when survival data is heavily censored [5]. Censoring happens when a patient have not happened an event (e.g., still alive, lost to follow-up, or exit from the study) at the end of follow-up [10].

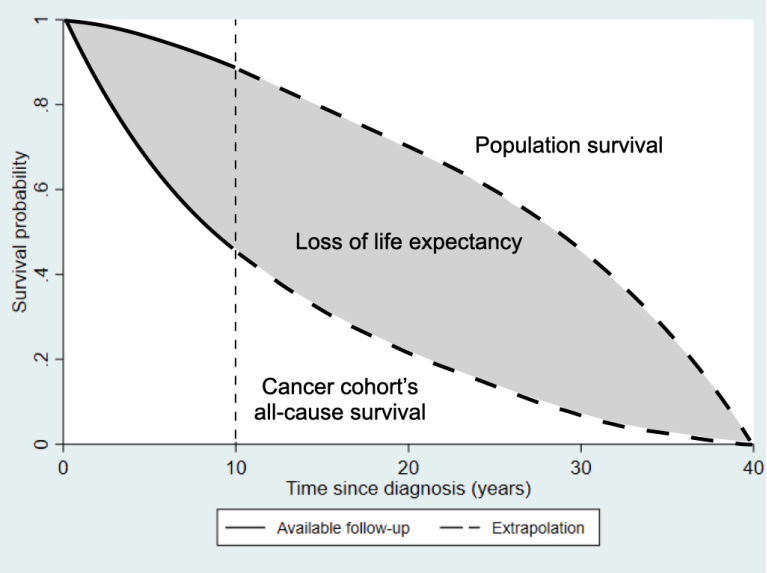

Figure 1: Graphical illustration of loss of life expectancy. The upper curve stands for population survival whereas the lower curve is cancer cohort's all-cause survival. The shaded area between two curves is interpreted as loss of life expectancy due to cancer. The available follow-up here is 10 years.

Not all models perform well with regard to extrapolating survival beyond available follow-up time [11], even though a variety of distributions can be used to fit the the data to the end of follow-up, such as exponential distribution [12], the Weibull distribution [13], the Gompertz

[1] Extrapolation means prediction outside the range of observed values [9].

distribution [14], the log-normal distribution [15], and the loglogistic distribution [16]. Given that the maximum follow-up of survival data is limited, external information is often required to aid extrapolation [8]. Relative survival is a typical framework used in integrating external information into population-based studies, such as national cancer registry studies or other patient registry studies [17, 18]. Among different extrapolation methods, the flexible parametric model (FPM) and the rolling-over algorithm (ROA) adopted relative survival framework in their respective extrapolation approaches.

1.2 Relative survival

The FPM and the ROA fit the relative survival function on other scales, the log cumulative excess hazard scale and the logit relative survival scale respectively [1, 4]. Before introducing the two extrapolation approaches, we would like first to introduce relative survival framework. Due to limited follow-up, extrapolation of both the expected survival (of the general population) and the all-cause survival (of the cancer patient cohort) is generally the approach to estimate LLE [1, 7]. Extrapolating all-cause survival is straightforward, but it generally requires assumptions on the survival function over the follow-up period, and it does not always give good prediction [1]. A more plausible alternative approach is extrapolating the relative survival, and transform it back to the all-cause survival by multiplying the expected survival [19], since the transformation of the relative survival function often approaches a straight line after a certain number of years since diagnosis suitable [1, 4], which enables predicting survival more precisely based on a linear trend [20].

Relative survival, $R(t)$, is defined as the ratio of the all-cause survival among a cohort of patients, $S(t)$, and the expected survival of a comparable population, $S^*(t)$ [18], written as

$$R(t) = \frac{S(t)}{S^*(t)} \tag{1}$$

where t is time since diagnosis. The analogue of Equation (1) on a hazard scale is

$$h(t) = h^*(t) + \lambda(t), \tag{2}$$

where $h(t)$ is the all-cause hazard, $h^*(t)$ the expected hazard, and $\lambda(t)$ the excess hazard.

By integration, the result is

$$H(t) = H^*(t) + \Lambda(t) \tag{3}$$

showing $H(t)$ is the cumulative all-cause hazard, $H^*(t)$ the cumulative expected hazard, and $\Lambda(t)$ the cumulative excess hazard. Individual survival data (e.g., from a patient registry) can provide data on $S(t)$, $h(t)$, and $H(t)$, whereas population mortality rates (e.g., from nationwide life tables) can offer information on $S^*(t)$, $h^*(t)$, and $H^*(t)$, by matching individuals to

corresponding age, sex, and calendar year [2, 21, 22].

1.3 Flexible parametric models

Flexible parametric models were developed by Royston and Parmar [16, 23] and extended to relative survival between the index cohort and the reference population [19, 24]. The FPM for relative survival is first fitted on the log cumulative hazard scale to estimate baseline hazard functions [1, 19]. In order to capture the shape of the hazard function and impose continuity restrictions on it, restricted cubic splines[2] were applied into the models [1, 17]. Generally, the log cumulative hazard is modeled in a full likelihood approach as a function of time t, which indicates time since diagnosis, along with covariates, z. It then gives a proportional hazards model as shown below:

$$\ln(H(t;z)) = \ln(-\ln S(t;z)) = s\left(x;\gamma_0\right) + z\beta \tag{4}$$

where $x = \ln(t)$, and $s\left(x;\gamma_0\right)$ is defined as a restricted cubic spline function[2].

To allow non-proportional hazards (time-dependent covariate effects), we include interactions between covariates and restricted cubic splines for log time into the model:

$$\ln(H(t;z)) = s\left(x;\gamma_0\right) + z\beta + \sum_{i=1}^{D} s\left(x;\gamma_i\right) z_i, \tag{5}$$

where D is the number of time-dependent covariate effects, and $s\left(x;\gamma_i\right)$ is the restricted cubic spline function for the i th time-dependent effect [1].

Within the follow-up period, the FPM for all-cause survival was fitted well. Nevertheless, the transformation of all-cause survival on the log cumulative all-cause hazard scale does not approach a linear line, so extrapolation from the function would result in higher inaccuracy [1]. Rather than extrapolating all-cause survival, Andersson et al. proposed extrapolating survival within relative survival framework using the FPM [1], by breaking the cumulative all-cause hazard into two components: the cumulative expected hazard, and the cumulative excess hazard, same as Equation (3).

Andersson et al. then extended Equation (3) to include covariates [1], shown as

$$H(t;z) = H^*\left(t;z'\right) + \Lambda(t;z), \tag{6}$$

where z includes patient characteristics, such as age, sex, year of diagnosis, or other collected information, and z' is a subset of z. By matching the mortality rates of the patient cohort to which of the general population, by age, sex, and calendar year, the cumulative expected hazard can be obtained [25]. On the other hand, the cumulative excess hazard, $\Lambda(t;z)$, can be modeled

[2]For illustration of restricted cubic splines, please refer to Appendix A.

in the FPM on log scale, shown as

$$\ln\left(\Lambda\left(t;z\right)\right)=s\left(x;\gamma_0\right)+z\beta+\sum_{i=1}^{D}s\left(x;\gamma_i\right)z_i \tag{7}$$

The analogue of cumulative excess hazard is relative survival can be shown as

$$R\left(t;z\right)=\exp\left(-\exp\left(\ln\left(\Lambda\left(t;z\right)\right)\right)\right)=\exp\left(-\exp\left(s\left(x;\gamma_0\right)+z\beta+\sum_{i=1}^{D}s\left(x;\gamma_i\right)z_i\right)\right) \tag{8}$$

where x, $s\left(x;\gamma_0\right)$, and $s\left(x;\gamma_i\right)$ have the same definitions in Equations (4)-(5). The all-cause survival function can then be calculated by multiplying the relative survival function with the expected survival function estimated by the Ederer I method [26], basically the same equation as Equation (1), and here shown as

$$S(t)=R(t)S^{*}(t) \tag{9}$$

1.4 Rolling-over algorithm

Hwang and Wang proposed fitting a simple linear regression to logit $R(t)$, the logit transformation of relative survival, during the available follow-up using partial likelihood [5, 22], written as

$$\text{logit}\,R(t)=\alpha+\beta t+\varepsilon_t, \tag{10}$$

where Monte Carlo methods were used to gain external information from life tables to match age, sex, and calendar year with an index patient cohort to obtain the relative survival [22].

In order to fit the flattened curve of logit $R(t)$ within the follow-up period, Hwang and Wang adopted the restricted cubic spline function from Andersson et al. to impose linearity on the estimated function after the boundary knot [1, 4, 27]. That is,

$$\text{logit}\,R(t)=s\left(t;\gamma_0\right), \tag{11}$$

where $s\left(t;\gamma_0\right)$ is a restricted cubic spline function. However, if logit $R(t)$ has not yet approached a straight line at the end of follow-up, the extrapolated values of logit $R(t)$ may only be close to the true survival curve in the first few points but deviate from it in the longer prediction period. Therefore, to solve this problem of uncertainty in extrapolation, Hwang et al. took advantage of using restricted cubic spline functions in short-term extrapolation to predict the survival step by step, aided by model-updated data [4]. This so-called rolling-over algorithm method uses a moving time-window of a fixed size (H points) to forecast the survival through the data points, illustrated in the following steps, and aided by Figure 2:

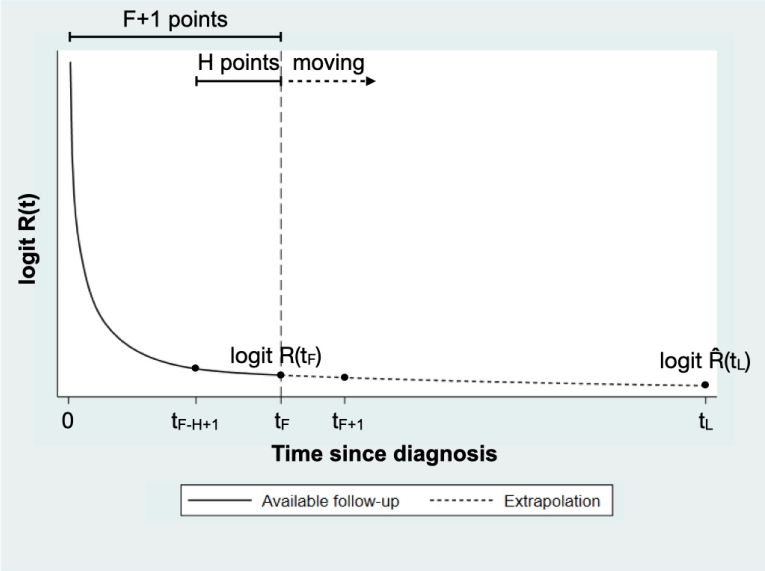

Figure 2: Logit transformation curve of relative survival used in the rolling-over algorithm method for extrapolating survival. A moving time-window of H points aids the logit $R(t)$ model to fit over and over until logit $\widehat{R}(t_L)$ is obtained.

1. Assume originally the available follow-up period has $F + 1$ time points of t_j, where $j = 0, 1, 2, ..., F$, corresponding to logit $R(t_j)$.

2. For the first extrapolation, use a time-window of H points of the data, logit $R(t_{F-H+1})$,..., logit $R(t_F)$, where $H < F$, to fit the model logit $R(t_k)$, where $k = F - H + 1, ..., F$.

3. Based on the fitted model logit $R(t_k)$, estimate only one predicted value logit $\widehat{R}(t_{F+1})$.

4. For the i th extrapolation, where $i > 1$, similar to the concept of Step 2 and 3 above, but drop logit $R(t_{F-H+i-1})$ and include logit $\widehat{R}(t_{F+i-1})$ as an observed value to refit the model, logit $R(t_k)$, where $k = F - H + i, ..., F + i - 1$, with still H points of the moving time-window and obtain logit $\widehat{R}(t_{F+i})$.

5. As the time-window moves toward t_L, we repeat Step 4 until the $(L - F)$ th extrapolation and get logit $\widehat{R}(t_L)$, where t_L is the last time point for extrapolation.

In summary, the extrapolation is aided by a moving time-window of H points to step-by-step predict the logit $\widehat{R}(t)$ to the final time point. The model for the first extrapolation in Step 2 only uses the observed values to predict the next survival point. The second to the H th times of extrapolation use both observed values within the follow-up period and model-updated values predicted by the extrapolation to estimate the next predicted values. For the $(H+1)$ th to $(L-F)$ th times of extrapolation, the model only uses predicted values to conduct extrapolation. In addition, bootstrap methods were used to tackle statistical uncertainty in estimation [4]. After completing the extrapolation of the relative survival function in logit transformation, we first transform it back to the relative survival function $R(t)$, and then use the same approach shown in Equation (9) to obtain the all-cause survival function of the patient cohort.

1.5 Knowledge gap

The FPM and the ROA have both been used to extrapolate patients' survival over the observed time [1, 4, 19], and they were also further applied to estimate LLE due to cancer [3, 6, 28–33]. Previous literatures have shown how sensitive the extrapolated survival is to tunable parameters, such as minimum time used for extrapolation [1, 33], the number of knots in the model [1], assumption on population mortality projection [29], and adjusting expected mortality rates by different matching approach [5, 34].

However, it remains unclear how sensitive the extrapolated survival is to limited follow-up data from patients of different cancer sites and diagnosed within different diagnosis periods. Additionally, previous studies which applied the ROA did not investigate how influential the size of the moving time-window would be in extrapolation results [4, 6, 30–32]. Furthermore, no study has been done yet to apply these two methods to extrapolate survival and evaluate their performance by using the same dataset.

2 Aim and Research questions

2.1 Aim

To increase understanding of the difference and sensitivity of extrapolating cancer patient survival between the flexible parametric model and the rolling-over algorithm by comparing with the empirical Kaplan-Meier's estimate.

2.2 Research questions

1. What is the all-cause survival, estimated using the Kaplan-Meier's method, for patients diagnosed with colon cancer and breast cancer in Sweden during 1981-1990 and 2001-2010 with follow-up until the end of 2017?

2. What is the difference compared with the Kaplan-Meier's estimate of applying the flexible parametric model and the rolling-over algorithm method to extrapolate the survival with limited follow-up data?

3. What is the difference compared with the Kaplan-Meier's estimate of using different years of follow-up data to extrapolate the survival by using the flexible parametric model and the rolling-over algorithm method?

4. How sensitive are the flexible parametric model and the rolling-over algorithm method to the tunable parameters?

3 Methods

3.1 Study design

This study is a nationwide registry-based cancer cohort study. Cohort study design was chosen to evaluate the performance of the FPM and the ROA in extrapolating survival. A cohort means a group of people who share the same condition at baseline [35], here characterized by the same diagnosis in cancer site and within a certain period of time. Routinely-collected registry data enabled us to follow up the patients and capture their survival information until the event (death) happened, or they became censored.

3.2 Data sources and population

Totally four different cohorts of colon cancer patients (n = 24 683, diagnosis in 1981-1990; n = 34 350, 2001-2010) and breast cancer patients (n = 43 490, 1981-1990; n = 63 209, 2001-2010) were identified from the Swedish Cancer Register, by using International Classification of Diseases Version 7 (code 153.x for colon cancer; code 170.x for breast cancer). The sample size calculation charts can be found in Appendix B. The follow-up information of all the patient cohorts was available until 31 Dec, 2017. If no event, namely no death or lost to follow up, happened to the patients by the end of follow-up, they were classified as censored. Only microscopically-verified adenocarcinoma patients were included. Cases diagnosed at autopsy were excluded. Individuals with multiple records of adenocarcinoma were only included with their first recorded diagnosis. For colon cancer patients, we included both sex in the analysis, while for breast cancer patients, males (<1%) were excluded from the analysis.

The underlying reason for choosing different cancer sites and diagnosis periods is to evaluate the extrapolation performance of the FPM and the ROA in different settings, where two cancer sites were included in this study: colon cancer and breast cancer. Colon cancer is a classic example of cancer survival with statistical cure point, defined as when cancer patients die at the same mortality as the reference population [36]. Therefore, the log cumulative baseline-excess-hazard function will have zero slope at cure point [37]. On the other hand, breast cancer is the most frequent female cancer worldwide [38]. For females, the 5-year relative survival of breast cancer is one of the highest among all cancer sites [39, 40], which leads to overall a better survival for recently diagnosed patients, and the survival data would be highly-censored.

Life tables of Sweden were downloaded from the Human Mortality Database [41]. Life table projections on death rate were downloaded from Statistics Sweden [42]. The Swedish Cancer Register was established in 1958, and by law, clinicians and pathologists are required to register all incident cancer cases [43]. As a result, the completeness and reliability of the cancer registry data in Sweden are high [43].

3.3 Variables

In the models, the exposure variable is age groups, which were obtained from categorizing patients' age at diagnosis into 5 groups (< 50, 50-59, 60-69, 70-79, and ≥ 80). The outcome variable is vital status at exit (death or censoring). Survival time of individuals was calculated from the following variables: date of diagnosis, date of death, or date of exit/censoring. In addition to age and survival time, sex and calendar year were also required in the analysis when matching the patients' data to the population mortality data (life tables) to accquire the expected survival function.

3.4 Statistical analyses

We first abstracted 37-year follow-up data (1981-2017) of colon cancer and breast cancer patients by diagnosis periods (1981-1990 and 2001-2010). The maximum and minimum follow-up for the patients diagnosed during 1981-1990 and 2001-2010 were 27 to 37 years and 7 to 17 years respectively, shown in Figure 3. In this study, the Kaplan-Meier (K-M) method, which is a non-parametric statistic to estimate the empirical survival function of a cohort of people [44], was applied to estimate the mean observed survival of patients within the follow-up time.

To evaluate the extrapolation results, we restricted the follow-up period up to 10 years for the patients diagnosed during 1981-1990, and up to 7 years for those diagnosed during 2001-2010. We then compared the predicted survival to the empirical K-M estimate along with 95 % confidence intervals (CI) with respective 37 years or 17 years of follow-up data. In addition to restricting the follow-up time to 7 years or 10 years, shorter follow-up time (3 and 5 years) or longer follow-up time (20 years) were also applied to extrapolate from the end of follow-up to evaluate the difference of extrapolation.

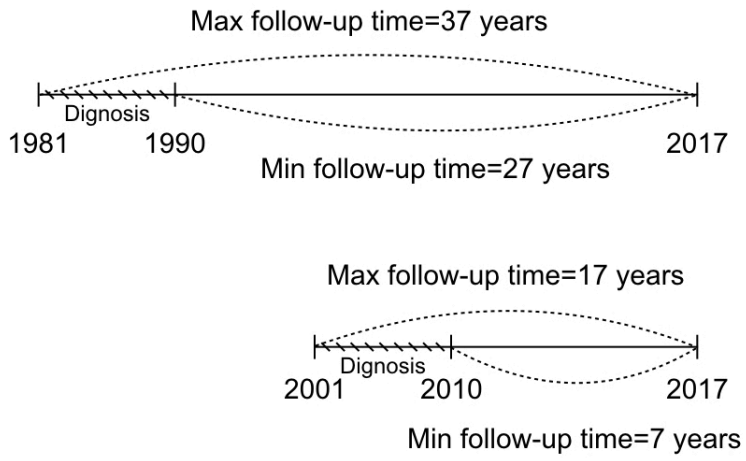

Figure 3: Maximum and minimum follow-up time for both colon cancer and breast cancer patients diagnosed during 1981-1990 and 2001-2010 in Sweden, with available follow-up until the end of 2017.

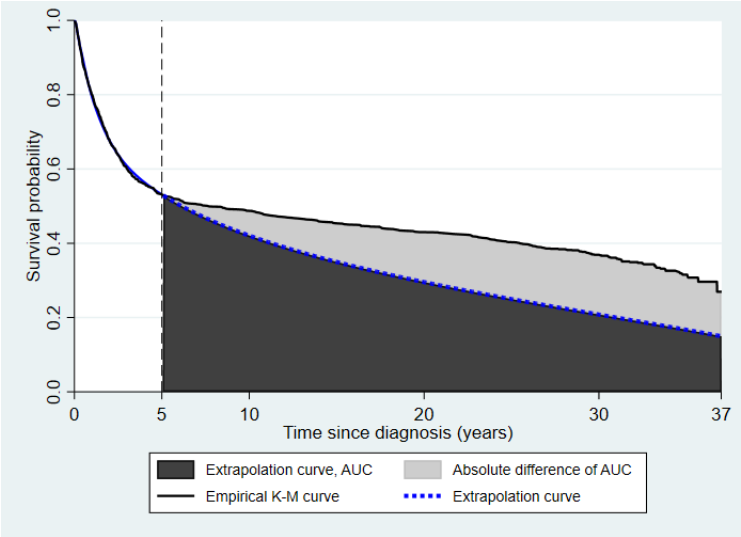

Figure 4: Graphical illustration of area under the curve (AUC) calculated from the extrapolation year. An example of survival extrapolated from 5 years, colon cancer patients aged < 50, diagnosis during 1981-1990, Sweden (n=1 148).

Absolute difference and relative difference in this study were used as a proxy for evaluating how approximate the extrapolation curve is to the empirical K-M curve. Absolute difference (the light grey area in Figure 4) was defined as the difference of area under the curve (AUC) between the K-M survival function and the extrapolated survival curves, both of which were calculated from the extrapolation year. For example, patients diagnosed during 1981-1990, the absolute difference of extrapolation from i years is calculated as

$$\left|\text{sum of AUC of extrapolation curve } i \text{ to 37 years} - \text{sum of AUC of K-M curve } i \text{ to 37 years}\right| \quad (12)$$

Relative difference was defined as absolute difference (the light grey area) divided by AUC of the K-M survival function (sum of the light grey area and the dark grey area), also illustrated by Figure 4. For instance, patients diagnosed during 1981-1990, the relative difference of extrapolation from i years is estimated as

$$100\% \times \left|\frac{\text{sum of AUC of extrapolation curve } i \text{ to 37 years} - \text{sum of AUC of K-M curve } i \text{ to 37 years}}{\text{sum of AUC of K-M curve } i \text{ to 37 years}}\right| \quad (13)$$

As a result of Akaike information criterion (AIC) and Bayesian information criterion (BIC) estimator, the FPMs were set with 7 degrees of freedom (d.f.) (8 knots) for baseline, and 2 d.f. (3 knots) for age group with time-dependent effects, where the knots are placed in centile positions of the survival times, as default in the Stata command `stpm2` [19]. In the FPMs, the predicted values beyond the last boundary knot were set to follow the linear trend, as where restricted cubic splines force the data in the tail of the distribution [17], without making other assumptions, such as statistical cure or constant excess hazard. For the ROA, the time window size, H, was set as 6 months for extrapolation from 3 or 5 years, 12 months for 7 years, and 18 months for 10 or 20 years. The number of knots of restricted cubic splines used in the ROA

was set as 5, as the default setting in the R package `iSQoL2` [4]. To obtain better estimation, one hundred bootstrap samples of the same sample size were calculated from the original survival data. The extrapolated survival value was then calculated as the mean of the estimated value and 100 bootstrapped values.

In sensitivity analysis, in addition to the original FPMs using 8 knots (7 d.f.) for baseline hazard function, we refitted models of the FPM with 7 to 9 knots (6 to 8 d.f.). For the ROA method, other than the main analysis, we tested $H = 6$, 9, and 12 for extrapolation from 3 or 5 years; $H = 6$, 12, and 18 for extrapolation from 7 years; $H = 12$, 18, and 24 for extrapolation from 10 or 20 years.

All the main computations were done with `stpm2` and its sub-packages [19] in Stata (Statacorp, College Station, TX, USA) and `iSQoL2` [4] and its sub-packages in R (R Foundation for Statistical Computing, Vienna, Austria).

4 Ethical Considerations

Individuals were not identified in the data used in this study. Informed consent was waived in this study, since the study used routinely-collected data from the Swedish Cancer Register. All data analyses and processing were done without personal numbers, and all presentations were carried out in aggregated form. There is no direct benefit for the subjects in the study population. However, our results may benefit future cancer patients through improving the healthcare system. In a registry-based study, people may experience integrity intrusion due to their data used in research. Nevertheless, we believe that the scientific benefit and potential benefit for the general population overweigh the risk. The study has been approved by the ethical review board at Karolinska Institutet (PI: Paul Dickman, Approval number: 2017/641-31/1).

5 Results

5.1 Descriptive statistics

Table 1 shows descriptive statistics for microscopically verified colon and breast adenocarcinoma patients diagnosed during 1981-1990 or 2001-2010 in Sweden, with follow-up (observed survival) until 2017. Patients who were diagnosed during 2001-2010 had higher proportions of censoring due to the follow-up time up to 2017. Furthermore, young breast cancer patients overall have good prognosis, resulting in high proportions of censoring among recently diagnosed patients. The largest patient cohort in this study was the breast cancer patients diagnosed during 2001-2010 (n = 63 209), whereas the smallest was the colon cancer patients diagnosed between 1981-1990 (n = 24 683). Among the colon cancer patients, the most common age group at diagnosis was 70-79 years old, and in each age group of the diagnosis period 1981-1990, females were slightly more than males; however, in the diagnosis period 2001-2010, only age group ≥ 80

years had more females than males. For the breast cancer patients, people aged 60-69 years old were the most common age group at diagnosis. Males (< 1%) were excluded from the analysis.

Table 1: Baseline characteristics of colon cancer patients and breast cancer patients diagnosed during 1981-1990 and 2001-2010, Sweden, with follow-up until 2017.

		Age group					
		<50	50-59	60-69	70-79	≥80	All
Colon cancer							
1981-1990	Patient size (n)	1 148	2 485	6 227	9 381	5 442	24 683
	Female (%)	51.74	52.43	50.09	51.61	59.28	53.01
	Proportion of censoring[a] (%)	34.41	15.69	3.12	0.22	0.09	4.07
	Mean diagnosis age (years)	42.11	55.38	65.08	74.56	83.81	70.77
	Median observed survival (years)	7.69	4.64	4.10	3.12	1.90	3.06
2001-2010	Patient size (n)	1 442	3 277	7 757	11 900	9 974	34 350
	Female (%)	48.47	49.74	46.80	49.68	56.61	51.00
	Proportion of censoring[a] (%)	56.45	53.49	47.43	29.91	10.32	31.54
	Mean diagnosis age (years)	41.75	55.41	64.93	74.71	84.37	72.08
	Median observed survival (years)	8.35	8.19	7.69	5.59	2.91	5.18
Breast cancer							
1981-1990	Patient size (n)	8 292	7 637	10 929	10 578	6 054	43 490
	Female (%)	100	100	100	100	100	100
	Proportion of censoring[a] (%)	44.03	26.63	6.41	0.28	0.12	14.77
	Mean diagnosis age (years)	42.49	54.76	64.65	74.28	84.36	63.77
	Median observed survival (years)	26.81	20.27	13.76	7.63	3.66	10.36
2001-2010	Patient size (n)	11 183	14 883	17 035	10 786	9 322	63 209
	Female (%)	100	100	100	100	100	100
	Proportion of censoring[a] (%)	81.28	80.45	73.37	46.59	11.85	62.79
	Mean diagnosis age (years)	43.09	54.81	64.25	74.20	85.24	63.08
	Median observed survival (years)	10.77	11.24	10.11	8.52	4.07	9.56

[a] Patients were classified as censored at the end of follow-up, 2017, if no event (death) happened to them.

5.2 Evaluating extrapolation

Figure 5 and Figure 9-11 in Appendix C show the extrapolated survival functions of the FPM and the ROA, along with the K-M estimate of empirical all-cause survival function with 95% CI, for patients of each cancer cite, diagnosis period and age group. Table 2 aids to show the absolute difference of AUC between the K-M survival curves and the extrapolated survival curves. Overall, the FPM and the ROA have good performance on extrapolating survival of age group ≥ 60, whereas for younger age groups overestimation and underestimation happen by using either method.

Figure 5 shows results of colon cancer patients diagnosed during 1981-1990, where 10 years of follow-up data were used for extrapolated survival functions. The ROA slightly overesti-

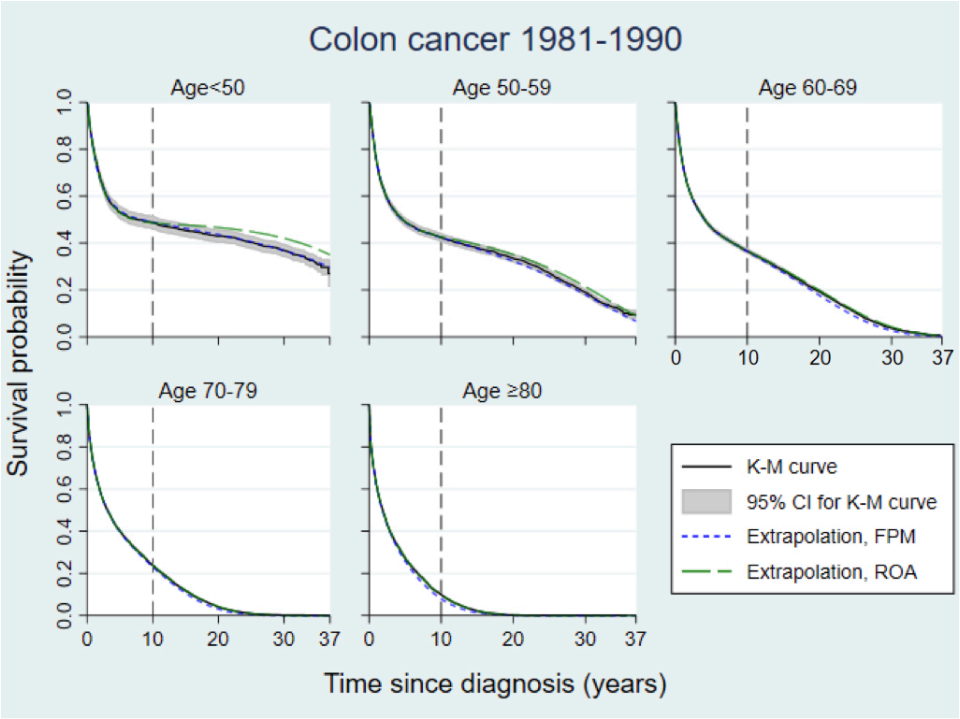

Figure 5: Kaplan-Meier (K-M) survival function with 95% confidence interval (CI) using 37 years of follow-up, in comparison with extrapolated survival functions using 10 years of follow-up, by the flexible parametric model (FPM) and the rolling-over algorithm (ROA), for colon cancer patients diagnosed during 1981-1990, Sweden (n=24 683).

mates the survival of age group < 50 with 0.97 years of AUC difference. As for the other age groups, the two approaches give sufficiently precise prediction on the survival functions with AUC difference less than 0.5 years. Figure 9 (Appendix C) shows results of colon cancer patients diagnosed during 2001-2010. With 7 years of follow-up data, though the FPM inconsiderably underestimates the survival of each age group with the largest difference, 0.35 years, in age group < 50, the extrapolated survival functions are both close to the empirical K-M curve. Figure 10 (Appendix C) shows results of breast cancer patients diagnosed during 1981-1990. The two approaches give a good fit on the predicted survival in spite of underestimating patients aged < 50, where the ROA underestimates the survival with 0.87 years and the ROA with 0.47 years. Figure 11 (Appendix C) shows results of breast cancer patients diagnosed during 2001-2010. Notwithstanding the high proportion of censoring (more than 80 % for age group > 50, 50-59, and 60-69), the survival functions extrapolated with only 7 years of follow-up data are considerably close to the empirical K-M curve with 17 years of follow-up time with difference less than 0.30 years in each age group.

Table 2: The absolute difference of area under the curve (AUC) between the Kaplan-Meier (K-M) curves and the extrapolated survival curves of the flexible parametric model (FPM) and the rolling-over algorithm (ROA), by cancer site and age group, using 7 or 10 years of follow-up time for extrapolation, diagnosis in 1981-1990 or 2001-2010, Sweden. (Colon 1981-1990, n=24 683; colon 2001-2010, n=34350; breast 1981-1990, n=43 490; breast 2001-2010, n=63 209.)

	Age group				
	<50	50-59	60-69	70-79	≥80
Colon cancer					
1981-1990, extrapolation from 10 years					
K-M curve, AUC[a] (years)	10.98	7.41	4.01	1.41	0.34
Proportion of censoring[b] (%)	48.78	42.33	36.60	23.89	9.81
FPM, difference (years)	0.04	0.25	0.30	0.18	0.10
ROA, difference (years)	0.97	0.46	0.11	0.00	0.01
2001-2010, extrapolation from 7 years					
K-M curve, AUC[a] (years)	5.73	5.36	4.68	2.88	0.96
Proportion of censoring[b] (%)	59.43	58.53	55.69	44.73	25.39
FPM, difference (years)	0.35	0.24	0.17	0.08	0.13
ROA, difference (years)	0.01	0.15	0.09	0.04	0.10
Breast cancer					
1981-1990, extrapolation from 10 years					
K-M curve, AUC[a] (years)	14.33	11.48	6.77	2.39	0.38
Proportion of censoring[b] (%)	68.37	66.64	60.55	38.90	11.27
FPM, difference (years)	0.87	0.22	0.33	0.20	0.12
ROA, difference (years)	0.47	0.13	0.10	0.13	0.04
2001-2010, extrapolation from 7 years					
K-M curve, AUC[a] (years)	8.14	8.11	7.28	4.57	1.15
Proportion of censoring[b] (%)	87.14	87.71	84.09	67.13	28.96
FPM, difference (years)	0.25	0.09	0.04	0.11	0.22
ROA, difference (years)	0.14	0.09	0.04	0.01	0.07

[a] AUC was calculated from the extrapolation year to the end of follow-up.
[b] Patients were made as censored at the end of limited follow-up, 7 or 10 years, if no event (death) happened to them.

Figure 6 demonstrates relative difference using limited follow-up time for extrapolation, for the patients of different cancer cites and diagnosis periods, by age group. For patients diagnosed during 1981-1990, 3, 5, 10, and 20 years of follow-up data were applied for extrapolation, whereas 3, 5, 7, and 10 years of follow-up data were used for extrapolation for patients diagnosed between 2001-2010. Overall for younger age groups, the relative difference is larger when using a shorter follow-up period of data for extrapolation. For younger colon cancer patients (< 60 years), the ROA performs better extrapolation with < 20 % relative difference given only 3 or 5 years of data. On the contrary, for younger breast cancer patients (< 60 years), the relative difference of using the FPM is lower than the ROA if 3 or 5 years of survival data was applied for extrapolation. The relative difference among older age groups is higher, especially when

using longer follow-up periods for extrapolation.

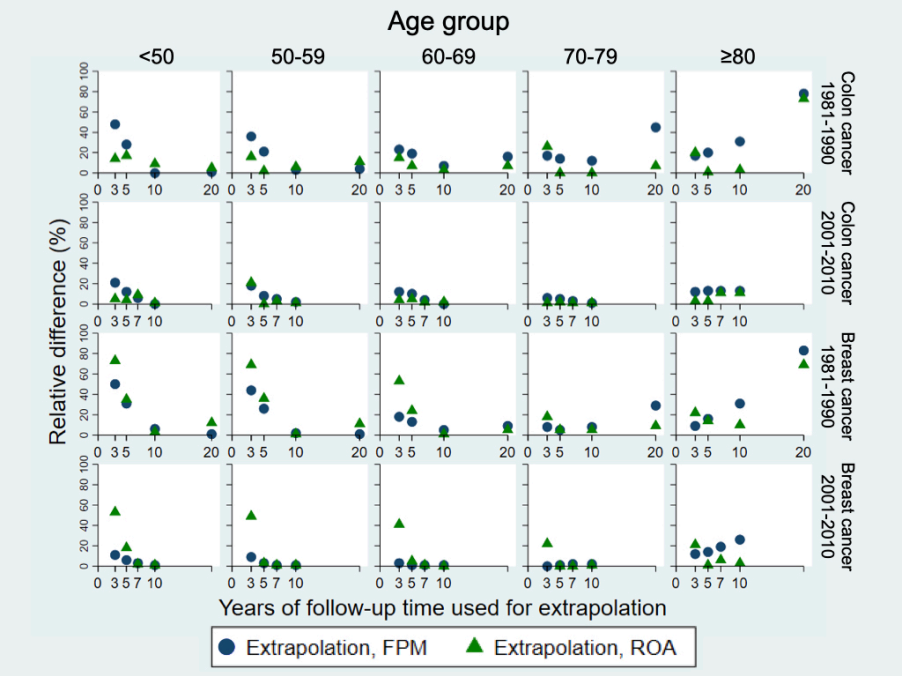

Figure 6: Relative difference of area under the curve (AUC) between the Kaplan-Meier (K-M) curves and the extrapolated survival curves of the flexible parametric model (FPM) and the rolling-over algorithm (ROA), using different limited follow-up time for extrapolation, by cancer site and age group, diagnosis in 1981-1990 or 2001-2010, Sweden. (Colon 1981-1990, n=24 683; colon 2001-2010, n=34 350; breast 1981-1990, n=43 490; breast 2001-2010, n=63 209.)

5.3 Sensitivity analysis

Table 3 along with Table 4-6 (Appendix D) display how sensitive the FPM is to the d.f. (the number of knots for restricted cubic splines), and the ROA is to the size of the moving time-window (H-month time points), by comparing the results of absolute difference of AUC between the K-M curves and the extrapolated survival curves of the FPM and the ROA.

For the FPM, the largest difference of AUC absolute difference is 0.23 years, which is observed between 6 and 8 d.f., using 10 years for extrapolating survival of age group < 50 of breast cancer patients diagnosed during 1981-1990 (Table 5). Contrarily, the smallest difference is 0 years.

For the ROA, the AUC difference varies more by using different sizes of H (months), especially extrapolating 3 or 5 years of follow-up time among younger age groups. For example, in age group < 50 of colon cancer patients diagnosed during 1981-1990 (Table 3), the difference of using 3 years for extrapolation is 2.10 years for $H = 6$ versus 11.68 years for $H = 12$, whereas the difference of using 20 years is 0.45 years for $H = 12$ versus 0.29 for $H = 24$. In contrast, the size of the time-window has less impact on using more than 5 years for extrapolating survival of older age groups 60-69, 70-79, and ≥ 80. For instance, age group 60-69 of breast cancer patients diagnosed during 2001-2010 (Table 6), using 5 years has difference of 0.49 years for

$H = 6$ versus 1.12 years for $H = 12$, and using 10 years has the AUC difference less than 0.05 years regardless of $H = 12$, 18, or 24.

Table 3: The absolute difference of area under the curve (AUC) between Kaplan-Meier (K-M) curve and the flexible parametric model (FPM) using different degrees of freedom in the models and the rolling-over algorithm (ROA) using different sizes (H months) of time-window , by age group, for colon cancer patients diagnosed during 1981-1990, Sweden (n=24 683).

			Age group				
	df	H	<50	50-59	60-69	70-79	≥80
Colon cancer 1981-1990							
Extrapolation from 3 years							
K-M curve, AUC[a] (years)			14.55	10.73	7.08	3.88	1.87
Proportion of censoring[b] (%)			60.10	56.06	55.52	50.71	39.21
FPM, difference (years)	6		7.04	3.92	1.68	0.68	0.33
	7		6.96	3.87	1.65	0.67	0.33
	8		6.83	3.77	1.59	0.64	0.32
ROA, difference (years)		6	2.10	1.70	1.08	1.03	0.38
		9	6.38	6.42	3.48	1.44	0.49
		12	11.68	7.64	3.91	1.62	0.51
Extrapolation from 5 years							
K-M curve, AUC[a] (years)			13.42	9.67	6.07	2.98	1.21
Proportion of censoring [b] (%)			53.14	48.93	46.62	40.20	27.12
FPM, difference (years)	6		3.71	1.98	1.10	0.41	0.24
	7		3.79	2.04	1.12	0.42	0.25
	8		3.83	2.06	1.14	0.43	0.25
ROA, difference (years)		6	2.24	0.22	0.44	0.00	0.01
		9	4.34	1.98	1.33	0.29	0.05
		12	5.25	3.16	1.89	0.44	0.06
Extrapolation from 10 years							
K-M curve, AUC[a] (years)			10.98	7.41	4.01	1.41	0.34
Proportion of censoring[b] (%)			48.78	42.33	36.60	23.89	9.81
FPM, difference (years)	6		0.02	0.26	0.30	0.18	0.10
	7		0.04	0.25	0.30	0.18	0.10
	8		0.07	0.23	0.29	0.17	0.10
ROA, difference (years)		12	1.26	1.00	0.34	0.03	0.00
		18	0.97	0.46	0.11	0.00	0.01
		24	0.76	0.11	0.00	0.02	0.01
Extrapolation from 20 years							
K-M curve, AUC[a] (years)			6.48	3.62	1.20	0.15	0.02
Proportion of censoring[b] (%)			43.03	33.40	19.40	4.06	0.28
FPM, difference (years)	6		0.04	0.11	0.19	0.07	0.01
	7		0.08	0.13	0.20	0.07	0.01
	8		0.06	0.12	0.19	0.07	0.01
ROA, difference (years)		12	0.45	0.45	0.11	0.01	0.01
		18	0.33	0.41	0.09	0.01	0.01
		24	0.29	0.39	0.07	0.01	0.01

[a] AUC of K-M curve was calculated from the extrapolation year to the end of follow-up.

[b] Patients were made censored at the end of limited follow-up, 3, 5, 10, or 20 years, if no event (death) happened to them.

6 Discussion

6.1 Summary of results

This study has demonstrated the difference between extrapolating cancer patient survival with the FPM and the ROA. Compared to the empirical K-M's estimate, we have displayed difference of AUC changes over time as extrapolation was carried out by using separate cohorts of cancer patient data. With a longer follow-up period of time, the FPM's and the ROA's extrapolation curves are both closer to the empirical K-M curve. In contrast, given shorter follow-up data, the absolute and relative difference of AUC between the empirical K-M curve and the extrapolation curves of the FPM and the ROA are still potentially higher, even though the ROA takes the advantage of using an adjustable moving time-window for extrapolation, and the FPM is not sensitive to the number of knots applied.

6.2 Interpretation of results

The study results in Figure 6 showed that when using longer follow-up periods for extrapolation, the relative difference among patients aged ≥ 70 was high. Nevertheless, in fact, both the FPM and the ROA predicted precise survival for older age groups given adequate follow-up time. The explanation is that the value of AUC of K-M curves (the denominator of relative difference) is already small for older age groups, so it becomes much smaller when applying longer follow-up data for extrapolation, resulting a high percentage in relative difference.

Uncertainty in extrapolation lies in both the selected model and the tunable parameter [11]. The sensitivity analysis showed that the extrapolation was not sensitive to the number of knots used in the FPM, which corresponds to the findings by Andersson et al. and Syriopoulou et al. [1, 45]. Another advantage of using restricted cubic splines in the FPM is that making assumption of the model is available. For instance, we can assume statistical cure by placing the knot to force the cumulative excess hazard to have zero slope after a certain time point [46], or assume constant excess hazard beyond the last boundary knot [1, 37]. On the other hand, the extrapolation results by the ROA were sensitive to the time window size, H (months). Given a bigger size of time-window, e.g., $H = 9$ or 12, for extrapolation from 3 or 5 years, the absolute difference would be considerably larger, whereas for extrapolation from 10 or 20 years, using different H did not have much impact on the absolute difference of AUC.

Conventionally, if the observed period captures sufficiently long follow-up survival data, suitable models can be used to extrapolate survival to estimate long-term survival [8, 33, 47]. Figure 12 (Appendix E) takes colon cancer patients aged 50-59 diagnosed during 1981-1990 as an example, showing that on the log cumulative excess hazard scale, $\ln \Lambda(t)$, how much the predicted curves extrapolated by the FPM deviates from the empirical curve by using short or long follow-up data. A drastic change over the first five years was observed from the empirical survival. Thus, given only 3 or 5 years of follow-up data used for modeling and extrapolation,

the predicted curve of log cumulative excess hazard apparently deviates from the true K-M curve, leading to a larger relative difference shown in Figure 6. Similarly, Figure 13 (Appendix E) uses the same group of people as an example, displaying that the logit transformation of relative survival, logit $\widehat{R}(t)$, extrapolated by the ROA, is closer to the empirical logit $R(t)$ if longer follow-up data was used for extrapolation. Both Figure 12 and Figure 13 show that with sufficiently long follow-up data, the curve of log cumulative excess hazard or logit relative survival would gain constant excess hazard and be approximate to a straight line. Under this condition, extrapolating survival would obtain reasonable accuracy.

6.3 Comparison with previous studies

The results of our study show at least 7 or 10 years of follow-up data depending on the selected patient cohort is needed to acquire constant excess hazard to perform robust extrapolation. Chu et al. validated using 7 years of follow-up data of 17 major cancers from Taiwan's National Cancer Registry, diagnosis between 1990 and 2001, to extrapolate survival and estimate LLE [5]. Andersson et al. recommended using at least 10 years of follow-up for extrapolating survival in population-based cohort studies; estimates were made for patients of colon cancer, breast cancer, melanoma, and bladder cancer diagnosed during 1961-70 with follow-up until 2010, in Sweden [1]. Another work by Andersson and her colleagues on LLE estimation for colon cancer patients also restricted the follow-up time up to 10 years to accurately extrapolate survival [3]. Also at least 10 years of follow-up were applied for estimating chronic myeloid leukemia patients' LE in Sweden and estimating LE of major cancer patients in Australia [48, 49]. Wright and Moorin identified 7 priority cancers from the Western Australia Cancer Registry and concluded that minimum time needed for estimating LLE differs between cancer sites and from population to population [33]. Fang et al. extrapolated the survival a cohort of patients with human immunodeficiency virus infection and found that if the excess hazard remains constant over time, the curve of the logit of relative survival will be approximate to a straight line, which is more plausible for prediction [20].

6.4 Strengths and limitations

Strengths of this study include showing how the choice of parameters within models influences the precision of survival extrapolation, and using the population-based registry data from the Swedish Cancer Register. Even though previous studies have already used the same nationwide cancer registry to estimate LLE due to cancer [1, 3, 28, 29, 50], by using the same FPM within relative survival framework, we increased the knowledge by extrapolating the survival of the patients identified from different diagnosis periods, even including a more recent period of time (diagnosis during 2001-2010). Our study aided the ROA a more holistic viewpoint of its sensitivity to the time-window selection and the follow-up time used, which was not emphasized by previous studies using the ROA method to estimate LLE and integrate into cost-effectiveness

analysis [4, 6, 30–32]. Additionally, to our understanding, this is the very first time that the ROA method was applied to extrapolate survival by using data from the Swedish Cancer Register.

This study has some limitations. Information on migration was not available in the data, so there might be some patients who were diagnosed with cancer but had already emigrated from Sweden, resulting in unknown lost to follow-up. For example, $\geq$ 80 year-old patients were diagnosed during 1981-1990 if they were censored at the end of the follow-up, 2017, and they were already at least 110 years old, which may be unreasonable. However, we assumed this would not have much impact on the results, since the Swedish Cancer Register has a high completeness of 96% [43], which overall captures all the other information needed in this study. Another limitation is that absolute difference and relative difference were interpreted as how close the extrapolation curve is to the empirical K-M curve. However, even though the absolute difference or relative difference is small, it is not absolutely equivalent to accurate extrapolation. If the extrapolation curve and the K-M curve cross over, the difference between their AUC could be canceled out, of which however only very few scenarios had this issue in our study.

6.5 Generalizability

For generalizability of this study, other factors, such as prognosis of the disease, clinical knowledge, advancement of treatment, and characteristics of the population should be taken into consideration when evaluating the minimum time of follow-up data needed for robustly extrapolating survival [11, 33]. In addition to cancer, the FPM or the ROA also enable us to predict survival of patients with other diseases, or from other smaller datasets, such as clinical trials. However, one potential exception in population-based studies is patients of a certain disease whose excess hazard is negative; namely, their survival is better than the general population. In this case, constant excess hazard is not attained, and the transformation of the relative survival may not converge to a straight line, which may acquire extra assumption to decrease inaccuracy in extrapolation.

6.6 Implication for public health and future research

Further research into different cancer patient cohorts would augment how sensitive the extrapolated survival of cancer patients is to the selection of model and parameter. Other than using the difference of AUC as a proxy of evaluating extrapolation, we recommend to calculate the actual difference between the extrapolation curve and the K-M curve in avoidance of cancelling out of AUC. Not only can we use patients of different age groups, cancer sites, diagnosis periods, but it is also available to extrapolate survival by stages of cancer, which the Swedish Cancer Register has information on from 2003, or by other covariates, such as sex [48], socio-economic status [29], education level [51], and so on. In addition, even though in this study we only investigated cancer patient survival extrapolation, both these two methods are not limited to be only applied in cancer survival but potentially also other diseases [17, 20, 52, 53].

For future research, it could be of value to examine the impact of sample size on survival extrapolation, for example, randomly selecting samples from the original dataset or extrapolating smaller clinical trial data. Potentially, extrapolation can be sensitive to a subset of the data due to a reduced sample size [54]. Another topic of interest is to what extent the survival rate of cancer would influence the model performance on extrapolation. If the survival of a patient cohort can be well captured, i.e. extremely good or poor survival rate (most patients survive or die after diagnosis), it is more likely to predict the survival precisely despite limited follow-up data. However, the survival rates of several cancer sites lie in-between, which indicates the importance of investigation into excess hazard to understand whether or not the survival attains statistical cure or constant excess hazard [20, 36, 55].

The study results have implications of predicting cancer patient survival over time and estimating LLE after cancer diagnosis to understand the burden of cancer between sub-groups, such as the general population and the cancer patient cohort, or varying different cancer types. The findings of this study can also be applied in cost-effectiveness studies on health technology assessment. In order to quantify the burden of cancer from both healthcare sector perspective and societal perspective [56], health outcomes (e.g., survival) and consequences of health (e.g., medical cost, quality of life, and productivity loss) will require prediction beyond the available time for justifying patients' lifetime healthcare resource use, where extrapolation is often carried out [7]. Accurate estimates of predicted survival could lead to reasonable quantification of expected life time survival, LLE, population mean medical costs, and other health benefits [4, 11].

7 Conclusion

This population-based study has demonstrated how sensitive extrapolated survival is to the tunable parameters by using the FPM and the ROA, in comparison with empirical K-M survival functions. Our study results suggest with sufficiently long survival data (e.g., 10 years), applying either the FPM or the ROA method to extrapolate survival is preferable; considering shorter follow-up data, if constant excess hazard has been approximated or reached within the available follow-up period, extrapolation carried out by the FPM or the ROA with an appropriate time-window size would have good performance. Further application of extrapolating survival could also be carried out in estimating LLE or health economics analysis. However, uncertainty in extrapolating survival should be clearly explained under different scenarios and taken into consideration in further studies.

8 Acknowledgments

First of all, I would like to thank my two supervisors. My main supervisor, **Paul Dickman**, thank you for all the kindness, generosity, and guidance you offered to me. Every time I see the light in your office, I feel assured because I know there is and will be support from you. As you said, ”Life is short.” And I did have fun with you! My co-supervisor, **Jung-Der Wang**, thank you for taking me as a summer intern student, which this extrapolation journey started from. Karolinska Institutet gave me proper training in Epidemiology and taught me to do things right, while you fully convinced me what other perspectives should be integrated into Epidemiology. You taught me doing the right things. Besides profession, I learned more from your belief and commitment to Jesus.

Thank you **Jing-Shiang Hwang** and **Therese Andersson** for joining the discussion with us. Without your shoulders, it would definitely be more challenging to climb this mountain. Thanks to **Chia-Ni Lin** for teaching me iSQoL2 and offering one-to-one lectures during my internship at NCKU. Thanks **Tsuey-Hwa Hu** for providing technical support and insight on the model settings on iSQoL2. I also want to thank **Mark Clements** and **Paul Lambert** for dropping in the discussion of restricted cubic splines and flexible parametric models. Sincerely, I hope that I can learn more from you in the future. Thanks to **Marie Jansson**, I appreciate the administration work you have done for MEB Biostat group.

To **Joshua Entrop**, the best-ever flexible parametric modeling buddy, without you, my journey in biostatistics would definitely be like life without muesli and guinea pigs. Thank you **Marlene Stratmann**, **Esa-Pekka Nykänen**, **Chiara Colizzi**, along with **Joshua** for being my best partners over the past two years. The beer we drank is as much as our friendship.

Thank you **Shan-Shan Cheng**, **Erh-Kang Fu**, **Chia-Wei Chang**, **Steve Pei**, and **Michael Jones** for being stars in the dark winter sky of Stockholm. The phone calls and prayers with you always refreshed my thesis life.

Last, I would like to thank my family and my dog for supporting me to study in Sweden. During the thesis period, it was the time of Covid-19 pandemic. Most of the flights between Sweden and Taiwan were suspended. It was the very first time when I realized my home is really far away from me. I hope I can see you all soon.

The aim of this thesis is to understand extrapolating survival, the next step for which is to estimate life expectancy. My deepest prayer is, ”**Teach us to number our days aright, that we may gain a heart of wisdom. (Psalms 90:12)**”

9 References

1. Andersson TML, Dickman PW, Eloranta S, Lambe M, Lambert PC. Estimating the loss in expectation of life due to cancer using flexible parametric survival models. Statistics in Medicine 2013;32:5286–5300.

2. Hakama M, Hakulinen T. Estimating the expectation of life in cancer survival studies with incomplete follow-up information. Journal of Chronic Diseases 1977;30:585–597.

3. Andersson TML, Dickman PW, Eloranta S, Sjövall A, Lambe M, Lambert PC. The loss in expectation of life after colon cancer: a population-based study. BMC Cancer 2015;15:412.

4. Hwang JS, Hu TH, Lee LJH, Wang JD. Estimating lifetime medical costs from censored claims data. Health economics 2017;26:e332–e344.

5. Chu PC, Wang JD, Hwang JS, Chang YY. Estimation of life expectancy and the expected years of life lost in patients with major cancers: extrapolation of survival curves under high-censored rates. Value Health 2008;11:1102–1109.

6. Wu TY, Chung CH, Lin CN, Hwang JS, Wang JD. Lifetime risks, loss of life expectancy, and health care expenditures for 19 types of cancer in Taiwan. Clinical Epidemiology 2018; 10:581–591.

7. Neumann PJ, Sanders GD, Russell LB, Siegal JE, Ganiats TG, et al. Cost-Effectiveness in Health and Medicine. Oxford University Press, 2016.

8. Jackson C, Stevens J, Ren S, Latimer N, Bojke L, Manca A, et al. Extrapolating Survival from Randomized Trials Using External Data: A Review of Methods. Medical Decision Making 2017;37:377–390.

9. National Institute for Health and Care Excellence. Developing NICE Guidelines: The Manual, 2015.

10. Kleinbaum DG. Survival Analysis: A Self-Learning Text. Springer, 1996.

11. Kearns B, Stevens J, Ren S, Brennan A. How Uncertain is the Survival Extrapolation? A Study of the Impact of Different Parametric Survival Models on Extrapolated Uncertainty About Hazard Functions, Lifetime Mean Survival and Cost Effectiveness. PharmacoEconomics 2020;38:193–204.

12. Beck JR, Kassirer JP, Pauker SG. A convenient approximation of life expectancy (the ”DEALE”). I. Validation of the method. American Jorunal of Medicine 1982;73:883–888.

13. Gaitatzis A, Johnson AL, Chadwick DW, Shorvon SD, Sander JW. Life expectancy in people with newly diagnosed epilepsy. Brain a Journal of Neurology 2004;127:2427–2432.

14. Cox C. The generalized F distribution: an umbrella for parametric survival analysis. Statistics in Medicine 2008;27:4301–4312.

15. Haybittle JL. Life expectancy as a measurement of the benefit shown by clinical trials of treatment for early breast cancer. Clinical oncology Royal College of Radiologists Great Britain 1998;10:92–94.

16. Royston P, Parmar MKB. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Statistics in Medicine 2002;21:2175–2197.

17. Royston P, Lambert PC. Flexible parametric survival analysis in Stata: Beyond the Cox model. Stata Press, 2011.

18. Dickman PW, Adami HO. Interpreting trends in cancer patient survival. Journal of Internal Medicine 2006;260:103–117.

19. Lambert PC, Royston P. Further development of flexible parametric models for survival analysis. The Stata Journal 2009;9:265–290.

20. Fang C, Chang Y, Hsu H, Twu S, Chen K, Lin C, et al. Life expectancy of patients with newly-diagnosed HIV infection in the era of highly active antiretroviral therapy. QJM 2006; 100:97–105.

21. Messori A, Trippoli S. A new method for expressing survival and life expectancy in lifetime cost-effectiveness studies that evaluate cancer patients (review). Oncology Reports 1999; 6:1135–1141.

22. Hwang JS, Wang JD. Monte Carlo estimation of extrapolation of quality-adjusted survival for follow-up studies. Statistics in Medicine 1999;18:1627–1640.

23. Royston P. Flexible parametric alternatives to the Cox model, and more. The Stata Journal 2001;1:1–28.

24. Nelson CP, Lambert PC, Squire IB, Jones DR. Flexible parametric models for relative survival, with application in coronary heart disease. Statistics in Medicine 2007;26:5486–5498.

25. Dickman PW, Coviello E. Estimating and modelling relative survival. The Stata Journal 2015;15:186–215.

26. Ederer F, Axtell L, Cutler S. The relative survival rate: A statistical methodology. National Cancer Institute Monograph 1961;6:101–121.

27. Durrleman S, Simon R. Flexible regression models with cubic splines. Statistics in Medicine 1989;8:551–561.

28. Bower H, Andersson TML, Bjorkholm M, Dickman PW, Lambert PC, Derolf AR. Continued improvement in survival of acute myeloid leukemia patients: an application of the loss in expectation of life. Blood Cancer Journal 2016;6:e390.

29. Andersson TML, Rutherford MJ, Lambert PC. Illustration of different modelling assumptions for estimation of loss in expectation of life due to cancer. BMC Medical Research Methodology 2019;19.

30. Lai WW, Lin CN, Chang CC, Wang JD. Lifetime risks, expected years of life lost, and cost-per-life year of esophageal cancer in Taiwan. Scientific reports 2020;10:3722.

31. Huang CC, Lin CN, Chung CH, Hwang JS, Tsai ST, Wang JD. Cost-effectiveness analysis of the oral cancer screening program in Taiwan. Oral Oncology 2019;89:59–65.

32. Hung MC, Ekwueme DU, White A, Rim SH, King JB, Wang JD, et al. Estimating health benefits and cost-savings for achieving the Healthy People 2020 objective of reducing invasive colorectal cancer. Preventive Medicine 2018;106:38–44.

33. Wright CM, Moorin RE. Does minimum follow-up time post-diagnosis matter? An assessment of changing loss of life expectancy for people with cancer in Western Australia from 1982 to 2016. Cancer epidemiology 2020;66:101705.

34. Bower H, Andersson TML, Crowther MJ, Dickman PW, Lambe M, Lambert PC. Adjusting Expected Mortality Rates Using Information From a Control Population: An Example Using Socioeconomic Status. American Journal of Epidemiology 2018;187:828–836.

35. Rothman K. Modern epidemiology. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2008.

36. Eloranta S. Development and Application of Statistical Methods for Population-Based Cancer Patient Survival. Phd thesis, Karolinska Institutet, 2013.

37. Eloranta S, Lambert PC, Andersson TML, Björkholm M, Dickman PW. The application of cure models in the presence of competing risks: a tool for improved risk communication in population-based cancer patient survival. Epidemiology 2014;25:742–748.

38. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA A Cancer Journal for Clinicians 2018;68:394–424.

39. Adami HO, Hunter DJ, Trichopoulos D. Textbook of cancer epidemiology. Oxford University Press, USA, 2008.

40. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Niksic M, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37,513,025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet 2018;.

41. University of California Berkeley and Max Planck Institute for Demographic Research. The Human Mortality Database. http://www.mortality.org/.

42. Statistics Sweden. The future population of Sweden 2019-2120, 2018.

43. Barlow L, Westergren K, Holmberg L, Talbäck M. The completeness of the Swedish Cancer Register: a sample survey for year 1998. Acta Oncologica 2009;48:27–33.

44. Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association 1958;53:457–481.

45. Syriopoulou E, Mozumder SI, Rutherford MJ, Lambert PC. Robustness of individual and marginal model-based estimates: A sensitivity analysis of flexible parametric models. Cancer Epidemiology 2018;58:17–24.

46. Andersson TML, Dickman PW, Eloranta S, Lambert PC. Estimating and modelling cure in population-based cancer studies within the framework of flexible parametric survival models. BMC Medical Research Methodology 2011;11:96.

47. Latimer NR. Survival analysis for economic evaluations alongside clinical trials–extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. Medical decision making an international journal of the Society for Medical Decision Making 2013;33:743–754.

48. Bower H, Björkholm M, Dickman PW, Höglund M, Lambert PC, Andersson TML. Life expectancy of chronic myeloid leukemia patients is approaching the life expectancy of the general population. Journal of Clinical Oncology 2016;34:2851–2857.

49. Baade PD, Youlden DR, Andersson TM, Youl PH, Walpole ET, Kimlin MG, et al. Temporal changes in loss of life expectancy due to cancer in Australia: a flexible parametric approach. Cancer causes control CCC 2016;27:955–964.

50. Ekberg S, Jerkeman M, Andersson PO, Enblad G, Wahlin BE, Hasselblom S, et al. Long-term survival and loss in expectancy of life in a population-based cohort of 7114 patients with diffuse large B-cell lymphoma. American Journal of Hematology 2018;.

51. Bower H, Andersson TML, Syriopoulou E, Rutherford MJ, Lambe M, Ahlgren J, et al. Potential gain in life years for Swedish women with breast cancer if stage and survival differences between education groups could be eliminated - three what-if scenarios. Breast 2019;45:75–81.

52. Lee J, Jeng J, Hwang J, Chang Y, Wang J. Life years lost and lifetime healthcare expenditures of stroke: A 14-year population-based cohort study in Taiwan. Journal of the Neurological Sciences 2017;381:880.

53. Yeh TS, Wang JD, Ku LJE. Estimating Life Expectancy and Lifetime Healthcare Costs for Alzheimer's Disease in Taiwan: Does the Age of Disease Onset Matter? Journal of Alzheimer's Disease 2020;73:307–315.

54. Kearns B, Stevenson MD, Triantafyllopoulos K, Manca A. Generalized Linear Models for Flexible Parametric Modeling of the Hazard Function. Medical Decision Making 2019; 39:867–878.

55. Andersson T. Quantifying cancer patient survival: extensions and applications of cure models and life expectancy estimation. Phd thesis, Karolinska Institutet, 2013.

56. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, et al. Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine. JAMA 2016;316:1093–1103.

Appendices

A Restricted cubic splines

Restricted cubic splines are an extension of cubic splines, so here we introduce cubic splines first. Splines enable us to impose continuity restrictions on functions, resulting in smooth fitted functions [17]. A spline function $s(x)$ without continuity restrictions is denoted as

$$s(x) = \sum_{j=0}^{n} \beta_{0j} x^j + \sum_{i=1}^{K} \sum_{j=0}^{N} \beta_{ij} \left(x - k_i\right)_+^j , \tag{14}$$

where the use of notion "+" in the following equations is defined as

$$u_+ = \begin{cases} u & \text{if } u > 0 \\ 0 & \text{if } u \leq 0 \end{cases} \tag{15}$$

and it is a function of order N, with covariate x, and K knots at $k_1 < \cdots < k_K$. The term $\beta_{ij} \left(x - k_i\right)_+^j$ permits a discontinuity at knot k_i for $s_{(j)}(x)$, which is defined as the j th derivative of $s(x)$. Contrarily, the absence of $\beta_{ij} \left(x - k_i\right)_+^j$ forces continuity of $s_{(j)}(x)$ at the knot k_i [17].

Cubic splines, where N=3, are the most common use of spline function, shown as

$$s(x) = \sum_{j=0}^{3} \beta_{0j} x^j + \sum_{i=1}^{K} \beta_{i3} \left(x - k_i\right)_+^3 \tag{16}$$

If a cubic spline function is forced to be linear before the first knot and after the last knot, then we obtain restricted cubic spline function, which is written as

$$s\left(x; \gamma_0\right) = \gamma_{00} + \gamma_{01} v_1(x) + \gamma_{02} v_2(x) + \ldots + \gamma_{0K-1} v_{K-1}(x), \tag{17}$$

where K is the number of knots and the j th basis function is

$$v_j(x) = \begin{cases} x, & \text{for } j = 1 \\ \left(x - k_j\right)_+^3 - \lambda_j \left(x - k_1\right)_+^3 - \left(1 - \lambda_j\right) \left(x - k_K\right)_+^3 , & \text{for } j = 2, \ldots, K-1 \end{cases} \tag{18}$$

and k_1 is defined as the position of the first knot, k_K the position of the last knot, k_j the position of the j th knot, and $\lambda_j = \frac{k_K - k_j}{k_K - k_1}$ [1, 17, 55].

B Sample size calculation

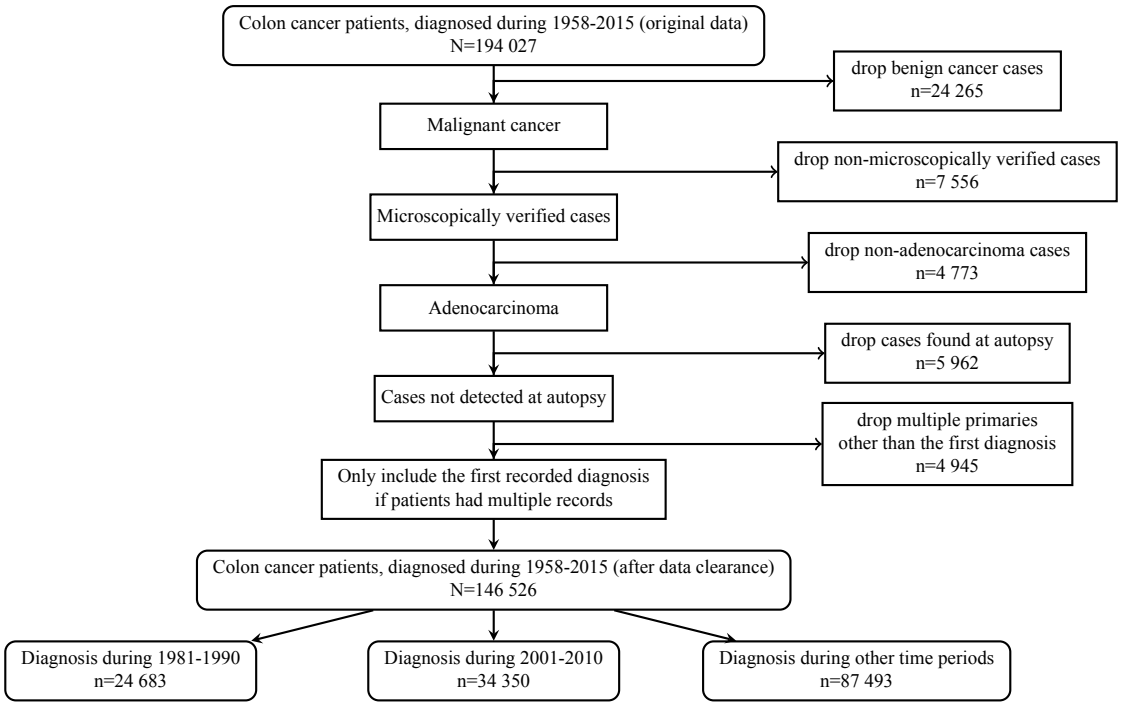

Figure 7: Sample size calculation of colon cancer patients, diagnosed during 1981-1990 (n=24 683) and 2001-2010 (n=34 350), Sweden.

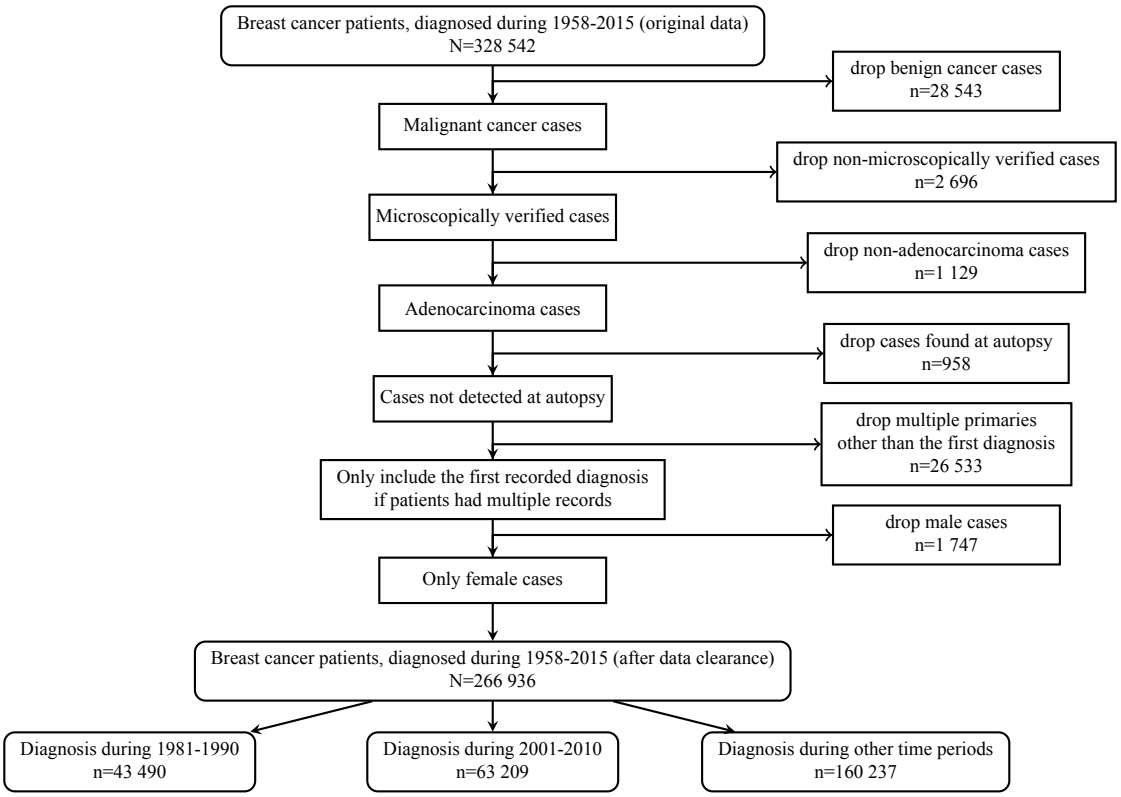

Figure 8: Sample size calculation of breast cancer patients, diagnosed during 1981-1990 (n=43 490) and 2001-2010 (n=63 209), Sweden.

C Extrapolation results

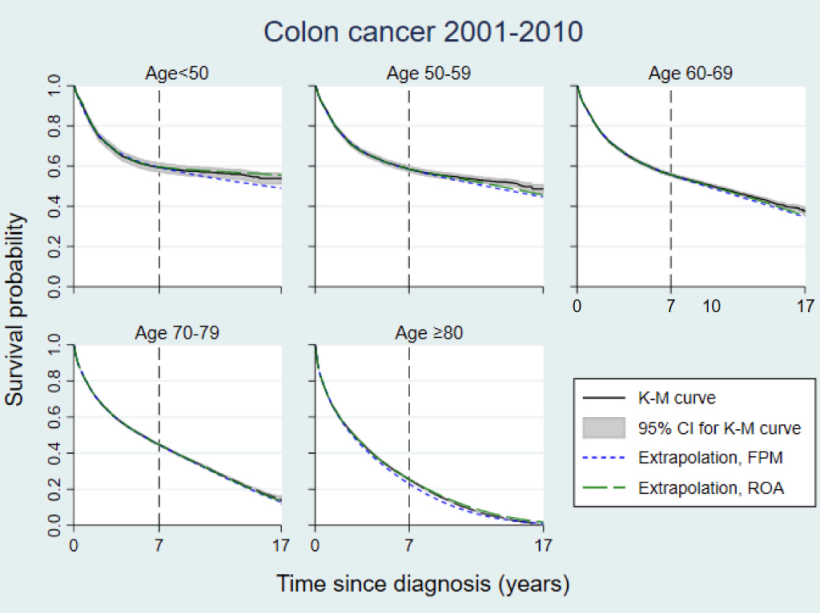

Figure 9: Kaplan-Meier (K-M) survival function with 95% confidence interval (CI) using 17 years of follow-up, in comparison with extrapolated survival functions using 7 years of follow-up, by the flexible parametric model (FPM) and the rolling-over algorithm (ROA), for colon cancer patients diagnosed during 2001-2010, Sweden (n=34 350).

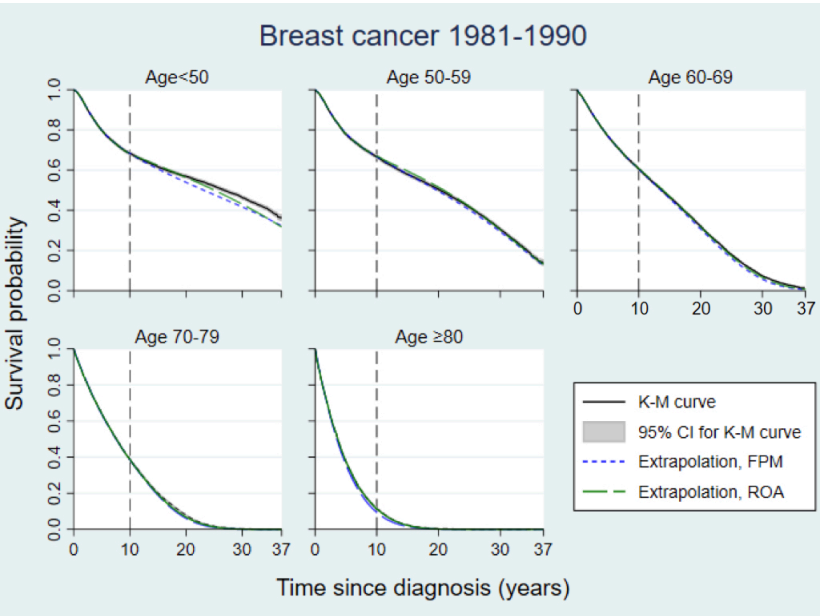

Figure 10: Kaplan-Meier (K-M) survival function with 95% confidence interval (CI) using 37 years of follow-up, in comparison with extrapolated survival functions using 10 years of follow-up, by the flexible parametric model (FPM) and the rolling-over algorithm (ROA), for breast cancer patients diagnosed during 1981-1990, Sweden (n=43 490).

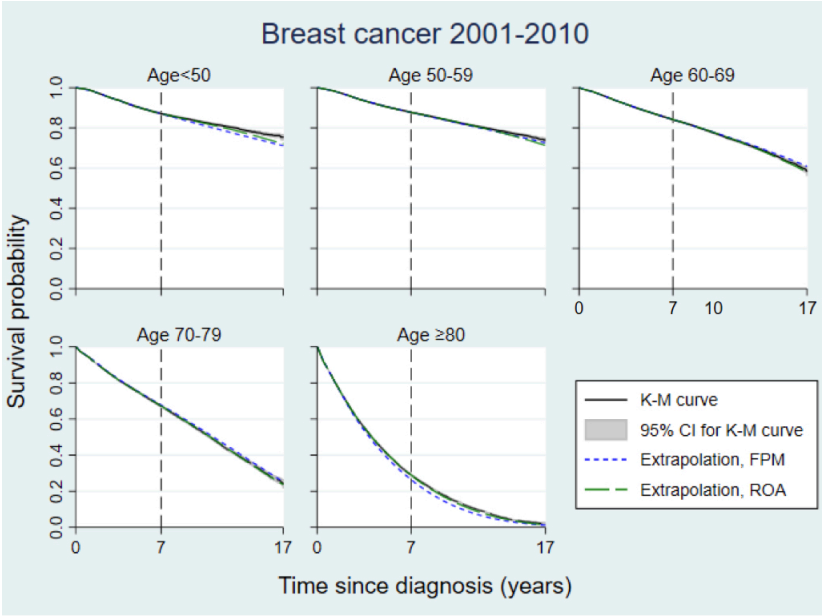

Figure 11: Kaplan-Meier (K-M) survival function with 95% confidence interval (CI) using 17 years of follow-up, in comparison with extrapolated survival functions using 7 years of follow-up, by the flexible parametric model (FPM) and the rolling-over algorithm (ROA), for breast cancer patients diagnosed during 2001-2010, Sweden (n=63 209).

D Sensitivity analysis

Table 4: The absolute difference of area under the curve (AUC) between Kaplan-Meier (K-M) curve and the flexible parametric model (FPM) using different degrees of freedom in the models and the rolling-over algorithm (ROA) using different sizes (*H* months) of time-window, by age group, for colon cancer patients diagnosed during 2001-2010, Sweden (n=34 350).

			Age group				
	df	H	<50	50-59	60-69	70-79	≥80
Colon cancer 2001-2010							
Extrapolation from 3 years							
K-M curve, AUC[a] (years)			8.18	7.89	7.14	4.99	2.40
Proportion of censoring[b] (%)			69.83	69.39	69.07	62.51	49.15
FPM, difference (years)	6		1.71	1.44	0.85	0.33	0.28
	7		1.69	1.43	0.83	0.32	0.28
	8		1.66	1.39	0.80	0.30	0.28
ROA, difference (years)		6	0.44	1.62	0.32	0.03	0.08
		9	2.67	3.68	2.57	1.37	0.22
		12	3.78	4.16	3.04	1.68	0.31
Extrapolation from 5 years							
K-M curve, AUC[a] (years)			6.87	6.57	3.85	3.85	1.57
Proportion of censoring[b] (%)			62.48	62.62	60.96	52.30	35.36
FPM, difference (years)	6		0.84	0.57	0.40	0.19	0.20
	7		0.82	0.55	0.39	0.18	0.20
	8		0.79	0.52	0.36	0.17	0.20
ROA, difference (years)		6	0.26	0.03	0.19	0.09	0.05
		9	0.47	0.44	0.62	0.16	0.01
		12	1.01	0.68	0.88	0.31	0.03
Extrapolation from 7 years							
K-M curve, AUC[a] (years)			5.73	5.36	4.68	2.88	0.96
Proportion of censoring[b] (%)			59.43	58.53	55.69	44.73	25.39
FPM, difference (years)	6		0.36	0.25	0.18	0.08	0.13
	7		0.35	0.24	0.17	0.08	0.13
	8		0.35	0.24	0.17	0.07	0.13
ROA, difference (years)		6	0.17	0.28	0.27	0.13	0.15
		12	0.51	0.15	0.09	0.04	0.10
		18	0.24	0.39	0.29	0.10	0.08
Extrapolation from 10 years							
K-M curve, AUC[a] (years)			3.91	3.66	3.09	1.69	0.39
Proportion of censoring[b] (%)			57.70	55.48	51.48	36.52	15.46
FPM, difference (years)	6		0.02	0.06	0.01	0.02	0.05
	7		0.02	0.06	0.01	0.02	0.05
	8		0.01	0.05	0.00	0.02	0.05
ROA, difference (years)		12	0.10	0.14	0.12	0.05	0.04
		18	0.05	0.04	0.06	0.02	0.04
		24	0.00	0.01	0.01	0.00	0.04

[a] AUC of K-M curve was calculated from the extrapolation year to the end of follow-up.

[b] Patients were made censored at the end of limited follow-up, 3, 5, 7, or 10 years, if no event (death) happened to them.

Table 5: The absolute difference of area under the curve (AUC) between Kaplan-Meier (K-M) curve and the flexible parametric model (FPM) using different degrees of freedom in the models and the rolling-over algorithm (ROA) using different sizes (*H* months) of time-window, by age group, for breast cancer patients diagnosed during 1981-1990, Sweden (n=43 490).

			Age group				
	df	H	<50	50-59	60-69	70-79	≥80
Breast cancer 1981-1990							
Extrapolation from 3 years							
K-M curve, AUC[a] (years)			19.67	16.70	11.81	6.34	2.44
Proportion of censoring[b] (%)			87.74	85.99	85.53	76.41	56.84
FPM, difference (years)	6		9.62	7.09	1.95	0.43	0.21
	7		9.85	7.30	2.07	0.49	0.23
	8		9.70	7.19	2.00	0.46	0.23
ROA, difference (years)		6	14.44	11.49	6.30	1.13	0.55
		9	15.60	12.28	7.20	2.24	0.70
		12	16.02	12.78	7.62	2.62	0.79
Extrapolation from 5 years							
K-M curve, AUC[a] (years)			17.99	15.07	10.19	4.94	1.50
Proportion of censoring[b] (%)			79.79	77.90	76.84	64.16	37.41
FPM, difference (years)	6		5.58	3.89	1.34	0.26	0.24
	7		5.54	3.86	1.33	0.26	0.24
	8		5.36	3.73	1.27	0.24	0.24
ROA, difference (years)		6	6.32	5.48	2.42	0.22	0.21
		9	9.42	7.39	3.38	0.64	0.27
		12	10.85	8.39	3.99	0.82	0.27
Extrapolation from 10 years							
K-M curve, AUC[a] (years)			14.33	11.48	6.77	2.39	0.38
Proportion of censoring[b] (%)			68.37	66.64	60.55	38.90	11.27
FPM, difference (years)	6		0.95	0.27	0.35	0.21	0.12
	7		0.87	0.22	0.33	0.20	0.12
	8		0.72	0.12	0.29	0.19	0.12
ROA, difference (years)		12	1.56	1.28	0.43	0.02	0.03
		18	0.47	0.13	0.10	0.13	0.04
		24	1.88	0.67	0.44	0.19	0.04
Extrapolation from 20 years							
K-M curve, AUC[a] (years)			8.12	5.67	2.13	0.26	0.02
Proportion of censoring[b] (%)			56.97	50.37	32.13	7.91	0.40
FPM, difference (years)	6		0.09	0.05	0.19	0.08	0.02
	7		0.11	0.06	0.18	0.08	0.02
	8		0.12	0.07	0.18	0.08	0.02
ROA, difference (years)		12	1.19	0.80	0.14	0.03	0.02
		18	0.95	0.64	0.10	0.02	0.02
		24	0.73	0.54	0.06	0.02	0.02

[a] AUC of K-M curve was calculated from the extrapolation year to the end of follow-up.

[b] Patients were made censored at the end of limited follow-up, 3, 5, 10, or 20 years, if no event (death) happened to them.

Table 6: The absolute difference of area under the curve (AUC) between Kaplan-Meier (K-M) curve and the flexible parametric model (FPM) using different degrees of freedom in the models and the rolling-over algorithm (ROA) using different sizes (*H* months) of time-window, by age group, for breast cancer patients diagnosed during 2001-2010, Sweden (n=63 209).

			Age group				
	df	H	<50	50-59	60-69	70-79	≥80
Breast cancer 2001-2010							
Extrapolation from 3 years							
K-M curve, AUC[a] (years)			11.78	11.75	10.82	7.59	2.87
Proportion of censoring[b] (%)			94.77	94.61	93.02	84.22	60.28
FPM, difference (years)	6		1.43	1.13	0.41	0.07	0.36
	7		1.34	1.04	0.35	0.02	0.34
	8		1.28	0.99	0.32	0.01	0.33
ROA, difference (years)		6	6.27	5.78	4.48	1.71	0.59
		9	6.66	6.50	4.98	2.39	0.77
		12	7.09	6.99	5.51	2.80	0.89
Extrapolation from 5 years							
K-M curve, AUC[a] (years)			9.92	9.90	9.01	6.00	1.86
Proportion of censoring[b] (%)			90.51	90.80	88.34	75.40	42.42
FPM, difference (years)	6		0.62	0.34	0.09	0.07	0.26
	7		0.61	0.33	0.09	0.07	0.26
	8		0.58	0.31	0.06	0.09	0.25
ROA, difference (years)		6	1.79	0.30	0.49	0.02	0.03
		9	2.38	0.77	0.86	0.18	0.12
		12	2.66	1.18	1.12	0.38	0.16
Extrapolation from 7 years							
K-M curve, AUC[a] (years)			8.14	8.11	7.28	4.57	1.15
Proportion of censoring[b] (%)			87.14	87.71	84.09	67.13	28.96
FPM, difference (years)	6		0.24	0.08	0.04	0.12	0.22
	7		0.25	0.09	0.04	0.11	0.22
	8		0.25	0.09	0.04	0.11	0.22
ROA, difference (years)		6	0.41	0.21	0.20	0.15	0.00
		12	0.14	0.09	0.04	0.01	0.07
		18	0.50	0.28	0.20	0.13	0.10
Extrapolation from 10 years							
K-M curve, AUC[a] (years)			5.58	5.54	4.85	2.75	0.49
Proportion of censoring[b] (%)			83.98	84.06	78.83	56.55	17.52
FPM, difference (years)	6		0.07	0.03	0.05	0.06	0.12
	7		0.07	0.04	0.05	0.05	0.13
	8		0.08	0.04	0.04	0.05	0.13
ROA, difference (years)		12	0.15	0.04	0.04	0.05	0.02
		18	0.07	0.03	0.00	0.02	0.02
		24	0.02	0.07	0.03	0.01	0.01

[a] AUC of K-M curve was calculated from the extrapolation year to the end of follow-up.

[b] Patients were made censored at the end of limited follow-up, 3, 5, 7, or 10 years, if no event (death) happened to them.

E Constant excess hazard

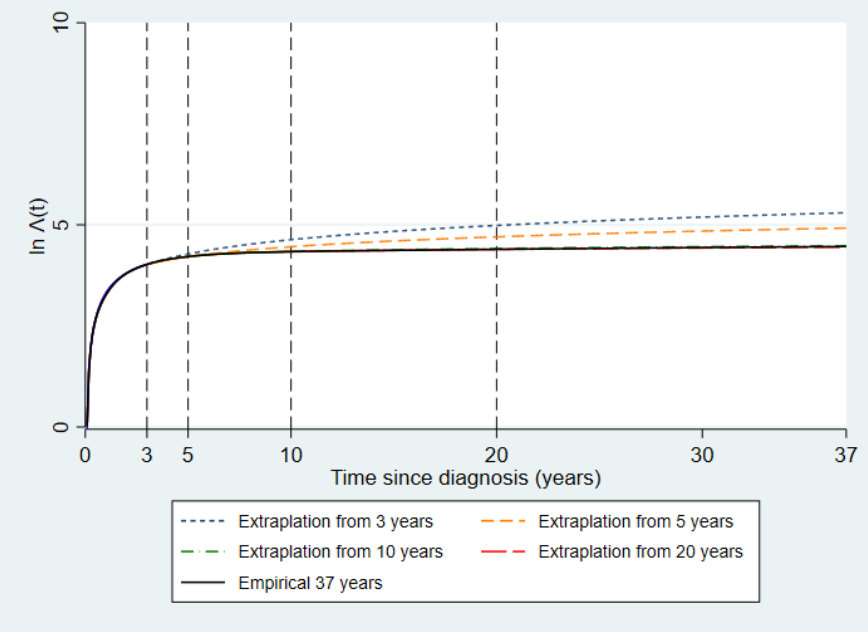

Figure 12: Extrapolating survival from 3, 5, 10, and 20 years of follow-up data on log cumulative excess hazard scale by the flexible parametric model, in comparison with the empirical 37 years of follow-up data, for colon cancer patients aged 50-59, diagnosis during 1981-1990, Sweden (n=2 485).

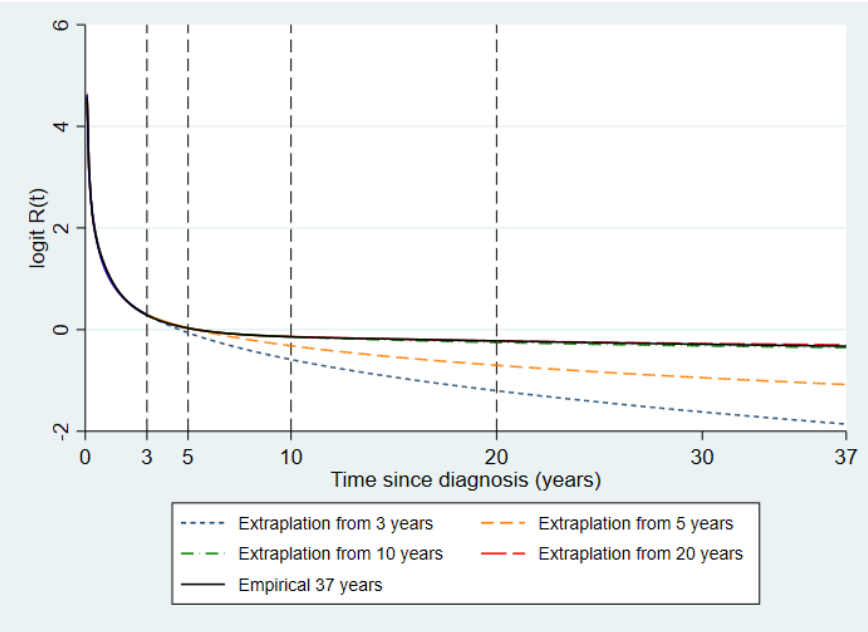

Figure 13: Extrapolating survival from 3, 5, 10, and 20 years of follow-up data on the logit transformation of relative survival, logit $\widehat{R}(t)$, by the rolling-over algorithm method, in comparison with the empirical 37 years of follow-up data, for colon cancer patients aged 50-59, diagnosis during 1981-1990, Sweden (n=2 485).