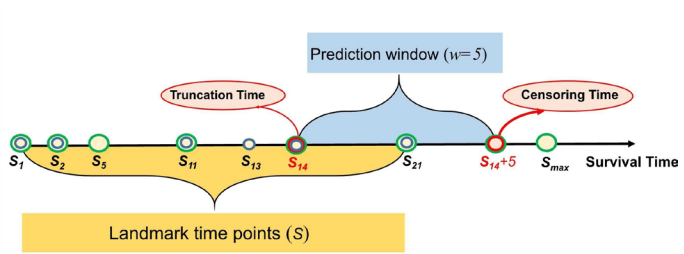
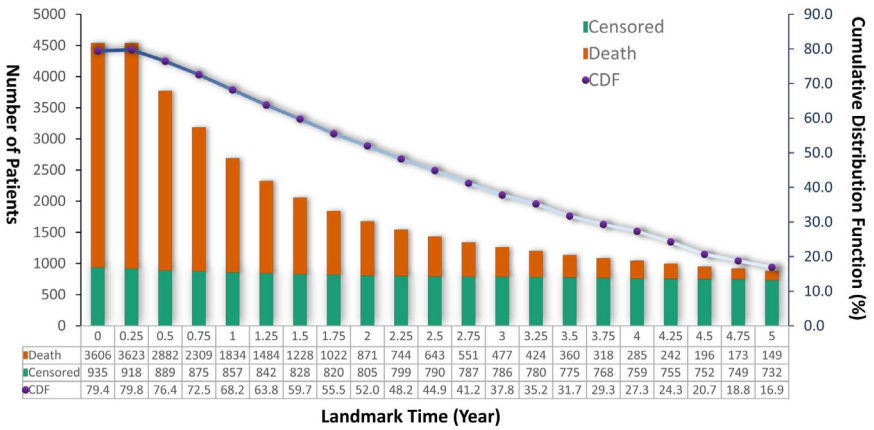


Figure S1 Structure of dynamic prediction using proportional baselines landmark supermodel. The green circle represents the patients' survival time. The blue circle represents the landmark time points. The red circle represents the two endpoints (Truncation time and Censoring time) of the prediction window. Different types of circles overlapping at the same time indicate that the time points occur at the same time.



	0	0.25	0.5	0.75	1	1.25	1.5	1.75	2	2.25	2.5	2.75	3	3.25	3.5	3.75	4	4.25	4.5	4.75	5
Death	3606	3623	2882	2309	1834	1484	1228	1022	871	744	643	551	477	424	360	318	285	242	196	173	149
Censored	935	918	889	875	857	842	828	820	805	799	790	787	786	780	775	768	759	755	752	749	732
CDF	79.4	79.8	76.4	72.5	68.2	63.8	59.7	55.5	52.0	48.2	44.9	41.2	37.8	35.2	31.7	29.3	27.3	24.3	20.7	18.8	16.9

Figure S2 Number of patients in the landmark datasets since the diagnosis of EC in relation to Death Censored status. The green bar shows the censored patients in each landmark time point, the red bar shows the deaths in each landmark time point, and the purple dot shows the cumulative distribution function (CDF) in each landmark time point.

https://dx.doi.org/10.21037/atm-21-4964

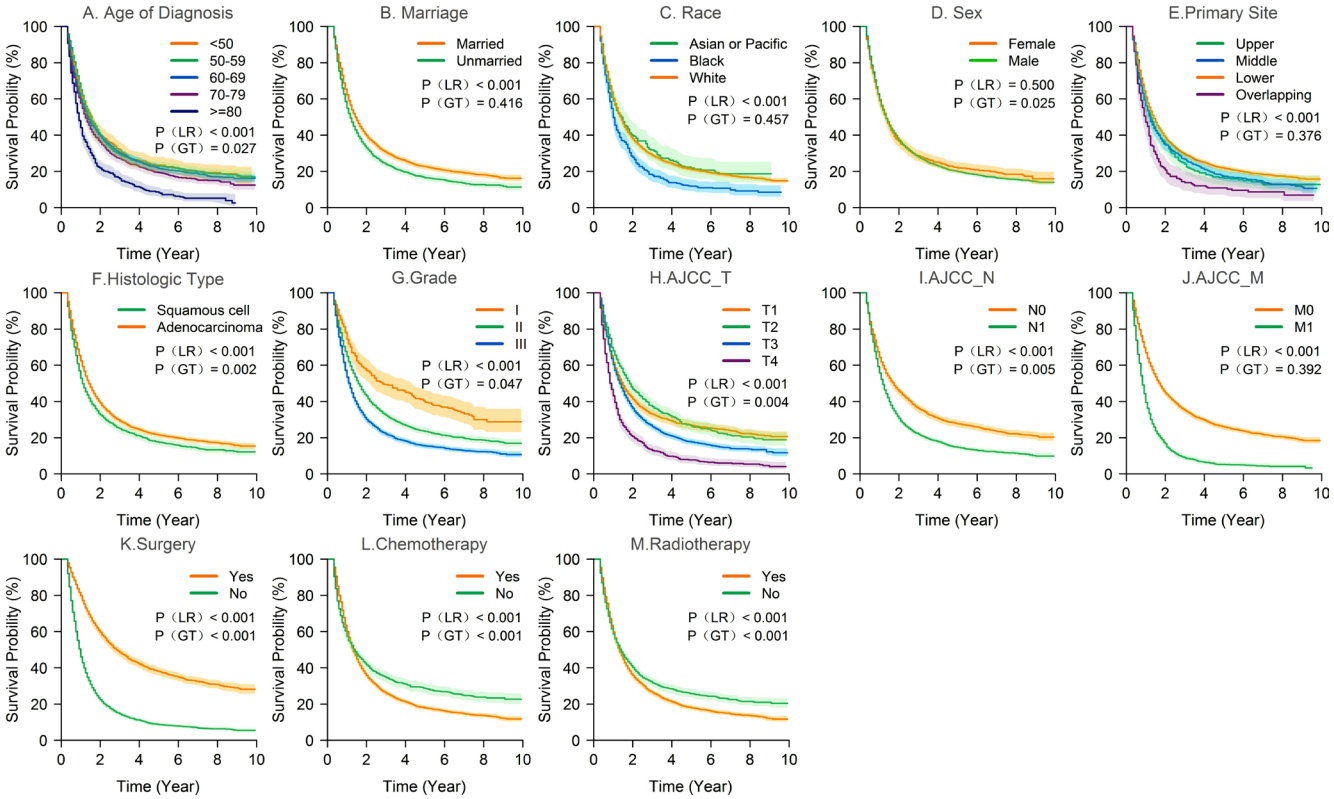


Figure S3 The overall survival curves of different factors. The Kaplan-Meier curves of OS were first compared with LR, of which a value <0.05 suggested that there were significant differences in survival rates among the different groups. The colored line exhibits the overall survival of a group of patients, and the corresponding shaded area is the confidence interval. The proportional hazards assumption was also checked by GT, of which a value <0.05 implied that the proportional hazard assumption was not satisfied. The Kaplan-Meier survival curves for age at diagnosis, sex, T stage, chemotherapy, and radiotherapy had intersecting evidence, which also implied that the proportional hazard assumption was not satisfied in this case. AJCC, American Joint Committee on Cancer; LR, Log-rank test; GT, Grambsch-Therneau proportional hazards test.

https://dx.doi.org/10.21037/atm-21-4964

Supplemental method I The construction of proportional baselines landmark supermodel

The proportional baselines landmark supermodel (PBLS model) was originated from the Landmark method (landmark analysis), a debate on the effect of response to chemotherapy on survival 33 years ago in the first volume, the *Journal of Clinical Oncology* (28). If we count the patients with guarantee time bias (immortal time bias) in the response group, then we will underestimate the death rate for responders, overestimate the death rate in the early month for nonresponders (29). So Anderson (28) provided an alternative method to consider the guarantee time bias in the early months by fixing time after the initiation of therapy as a landmark. People whose survival time less than the landmark time will be excluded from the analysis. This approach effectively removes the bias present in the early time. In 2007, Houwelingen (13) designed the PBLS model, one way of dynamic prediction, combining Cox proportional hazards model and landmark method, having advantages to solve the covariates with time-varying effect, simply operating and easily understanding.

Suppose that a data based on a sample of size n, consists of the triple (T_i, δ_i, Z_i), where $T_i\,(i=1,2,...,n)$ is the time to event for the i-th patient. δ_i is the event indicator (δ_i=1 if the event has occurred, δ_i=0 if the time is censored). Z_{ij} is the i-th patient's j-th covariate. Then the Cox proportional hazards (PH) model can be constructed as follows:

$$h(t \mid \boldsymbol{Z})=h_0(t)\exp(\boldsymbol{\beta}'\boldsymbol{Z})=h_0(t)\exp(\sum \beta_j Z_j) \quad [1]$$

where j is the coefficient for the j-th covariate (Z_j) and the $h_0(t)$ is the baseline hazard function

However, the Cox PH model was found evidence of the non-proportionality of hazards for many reasons. The time-varying effect model is denoted as

$$h(t \mid \boldsymbol{Z})=h_0(t)\exp(\boldsymbol{\beta}(t)'\boldsymbol{Z}) \quad [2]$$

Exploring the prognosis factors is important for clinical research, but a good prognostic prediction model might be more important to patients. Patients after diagnosis of cancer may pay more attention to the probability of "*w*" years' survival or mortality. Moreover, patients maybe not only pay attention at the start of the treatment but also any time of the follow-up visits. This emphasizes the dynamic use of time in the prediction model. In van Houwelingen [2007], landmarking is introduced as a tool to obtain predictions from s up to $t_{hor}=s+w$. And then use the Cox PH model in that interval as follows:

$$h(t \mid \boldsymbol{Z}, s_l, w)=h_0\left(t \mid s_l, w\right)\exp(\boldsymbol{Z}'\boldsymbol{\beta}_{LM}),\ s_l \le t \le s_l + w \quad [3]$$

We can use this simple Cox PH model at any landmark time point $s \in [s_1, s_L]$. Maybe the PH assumption is violated in some intervals. But it is a very convenient and useful way to obtain a dynamic prediction without having to fit a model with complicated time-varying effects.

The analysis is based on a "super prediction dataset" (more information see the supplemental method 2 "*The construction of a super prediction data set*"). After obtaining the data set, we can establish the proportional baselines landmark supermodel. The first step is to let the regression coefficients β_{LM} depend on $\left\{ s_l = \frac{s}{s_L - s_1}, l=1,2,3,...,L \right\}$ in a smooth way and to model that in a linear way. The form can be

$$h(t \mid \boldsymbol{Z}, s, w)=h_0\left(t \mid s, w\right)\exp(\boldsymbol{Z}'\boldsymbol{\beta}_{LM}(s)),\ s \le t \le s + w \quad [4]$$

where $\boldsymbol{\beta}_{LM}(s_l)=\sum_{j=1}^{m_h}\gamma_j f_j(s_l)$ and m_h is the sum of the terms of the coefficients. An easy way is to use the spline or a parametric model like: $\beta_{LM}(s_l)=\gamma_0+\gamma_1 s_l+\gamma_2 s_l^2$. Every dataset can be viewed as a "strata." There are some patients repeated in each strata in the super dataset. So the baseline hazard function depends on s_l via the smooth functions modeled directly by letting

$$h_0(t \mid s, w)=h_0(t)\exp(\theta(s))=h_0(t)\exp(\sum_{j=1}^{m_h}\eta_j g_j(s)) \quad [5]$$

Where $g_j(s_l)=f_{j+1}(s_l)$ and m_h is the sum of the terms of the baselines. Then the proportional baselines landmark supermodel (PBLS) can be

$$h(t \mid Z, s, w)=h_0(t)exp(Z'\beta_{LM}(s)+\theta(s)), s \le t \le s+w \quad [6].$$

https://dx.doi.org/10.21037/atm-21-4964

And the integrated partial log-likelihood (*ipl**) is

$$ipl^*(\gamma,\eta)=\sum_{i=1}^{n} d_i \ln\left(\frac{\sum_{\{s|s\le t_i\le s+w\}}\exp(Z_j'\beta_{LM}(s\mid\gamma)+\theta(s\mid\eta))}{\sum_{\{s|s\le t_i\le s+w\}}\sum_{j\in R(t_j)}\exp(Z_j'\beta_{LM}(s\mid\gamma)+\theta(s\mid\eta))}\right), s_1\le t_i\le s_L+w \quad [7]$$

The baseline hazards are:

$$\hat{h}_0^*(t_i)=\frac{\int_0^{t_i}\theta(s)ds}{\sum_{\{s|s\le t_i\le s+w\}}\sum_{j\in R(t_j)}\exp(Z_j'\beta_{LM}(s\mid\gamma)+\theta(s\mid\eta))}, s_1\le t_i\le s_L+w \quad [8]$$

Let $\hat{H}_0^*(t_i)=\sum_{t_i\le t}\hat{h}_0^*(t_i)$ be the corresponding cumulative hazard, then the simple predictive landmark model is given by

$$\hat{H}(t\mid Z,t_{LM}=s)=\exp(Z'\hat{\beta}_{LM}(s)+\hat{\theta}(s))(H_0^*(t)-H_0^*(s^-)) \quad [9]$$

The predicted w-year dynamic survival rate at any prediction time was obtained by

$$S(s+w\mid s,Z)=P(T>s+w\mid T>s,Z)=\exp\left(-\int_s^{s+w}h(t|\mathrm{Z,s})dt\right) \quad [10]$$

Then the w-year dynamic HR at different time point can be calculated as following equations:

$$HR^w(s_l)=\exp(\gamma_0+\gamma_1 s_l+\gamma_2 s_l^2), \left\{s_l=\frac{s}{s_L-s_1}, l=1,2,3,...,L\right\} \quad [11]$$

References

28. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response. J Clin Oncol 1983;1:710-9.
29. Giobbie-Hurder A, Gelber RD, Regan MM. Challenges of guarantee-time bias. J Clin Oncol 2013;31:2963-9.

https://dx.doi.org/10.21037/atm-21-4964

Supplemental method II The construction of a super prediction data set

To better understand the construction of a super prediction data set, we fabricate a toy data set containing two variables for six single-record-per-patient subjects. The data is shown in Table S1.

Then we construct the super prediction data set for six subjects as the following steps:

Fix the prediction window w (see the blue part in Figure S1). The selection of w relies on the severity of the cancer, w=5 or w=10 for the milder cancers, w=1 or w=2 for the severe cancers. Suppose the six patients want to know the probability of mortality of 5 years at any prediction time point (s), then the prediction window was fixed at w=5.

Select a set of prediction landmark time points ($s \in [s_1, s_L]$). (see the blue circles and the yellow parts in Figure S1). The value of L defines a weighting of the prediction time points in the model to be developed. The simplest approach is taking an equidistant grid of points on an interval $\{s_1, s_L\}$. For six subjects, a set of prediction landmark time points $\{s_1, s_2, s_3, s_4\} = \{0,1,2,3\}$ was selected at every 1 year between 0 and 3 years. In reality, the value of L between 20 and 100 will be sufficient and should not depend on the actual event times.

Create a prediction data subset for each landmark timepoint (s), who was still alive at s and administrative censored at $s+w$ (see the red circles and the blue parts in Figure S1). In the first landmark time point ($s_1 = 0$) from Table S2, all the subjects were included and administrative censored at $s_1 + w = 0 + 5$, so **Time** for the subject (ID=5) and the subject (ID=6) became 6 to 5 and 8 to 5 and **Status** for the subject (ID=6) became 1 to 0 because subject 6 still alive after the diagnosis at the beginning. In the second landmark time point ($s_2 = 1$), the subject (ID=1) was excluded from the interval $[1,6]$ due to the event time ($0.9 < 1$). **Time** for the subject (ID=5) and the subject (ID=6) became 6 to 6 and 8 to 6 and **Status** for the subject (ID=6) became 1 to 0. In the third landmark time point ($s_3 = 2$), the subject (ID=1) and the subject (ID=2) were excluded from the interval $[2,7]$ due to the event time ($0.9 < 2$ and $1.8 < 2$). **Time** for the subject (ID=6) became 8 to 7 and **Status** became 1 to 0. In the last landmark time point ($s_4 = 3$), only three subjects (4,5,6), who were still alive at 3 years after following-up, were included in the interval $[3,8]$.

Stack all-created small subsets into a super prediction data set. In this large data set, the subsets corresponding to a given prediction time(s) are labeled as "strata". But it couldn't be the stratified proportional hazards model. Because there are some patients repeated in each strata in the super dataset.

Table S1 Six single-record-per-patient subjects

ID	Time	Status	Age at diagnosis	Gender
1	0.9	1	50	0
2	1.8	0	42	1
3	2.9	1	30	0
4	4.2	1	36	1
5	6	0	20	0
6	8	1	45	1

The unique subject identifier is ID. The variable Status takes on a value of 1 if the subject dies and 0 if the subject is censored. The time of death or censoring is captured by Time. The predictors of interest are Age at diagnosis and Gender.

https://dx.doi.org/10.21037/atm-21-4964

Table S2 A super prediction data set for six subjects

Landmark ID	ID	Time	Status	Age at diagnosis	Gender	Landmark time
1	1	0.9	1	50	0	0
2	2	1.8	0	42	1	0
3	3	2.9	1	30	0	0
4	4	4.2	1	36	1	0
5	5	5	0	20	0	0
6	6	5	0	45	1	0
21	2	1.8	0	42	1	1
31	3	2.9	1	30	0	1
41	4	4.2	1	36	1	1
51	5	6	0	20	0	1
61	6	6	0	45	1	1
32	3	2.9	1	30	0	2
42	4	4.2	1	36	1	2
52	5	6	0	20	0	2
62	6	7	0	45	1	2
43	4	4.2	1	36	1	3
53	5	6	0	20	0	3
63	6	8	1	45	1	3

https://dx.doi.org/10.21037/atm-21-4964