eAppendix

Estimation of statistical cure

Flexible parametric cure models were used to model statistical cure by covariates.[1] For the melanoma and colon cancer cohorts we used 6 degrees of freedom (df) to model the cumulative baseline excess hazards where the internal knots placed at the 20^{th}, 40^{th}, 60^{th}, 80^{th} and 95^{th} centiles of the distribution of the uncensored log survival times. For the Acute Myeloid Leukemia (AML) cohort we included one additional internal knot at 9 years after diagnosis since the 95^{th} centile of the distribution of the uncensored log survival times for the AML cohort occurred already at 4.1 years after diagnosis. A complete summary of the location of all knots in the three cohorts is provided in eTable 1.

Estimation

The `stpm2` module in the Stata software (StataCorp. 2011. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP) was used to fit the flexible parametric cure model. We wrote a post-estimation command that was subsequently used to estimate equations (2)-(4). Numerical integration was used to evaluate the integrals. Each numerical approximation was done by dividing *t* using 1000 equidistant cut-points, and by calculating a weighted sum of the integrand in each interval. The 95% confidence intervals were constructed via the `predictnl` functionality in Stata. The `predictnl` command calculates non-linear predictions and uses the delta method to approximate the variances, standard errors and pointwise confidence intervals for the predictions (StataCorp. 2011. Stata: Release 12. Statistical Software. College Station, TX: StataCorp LP).

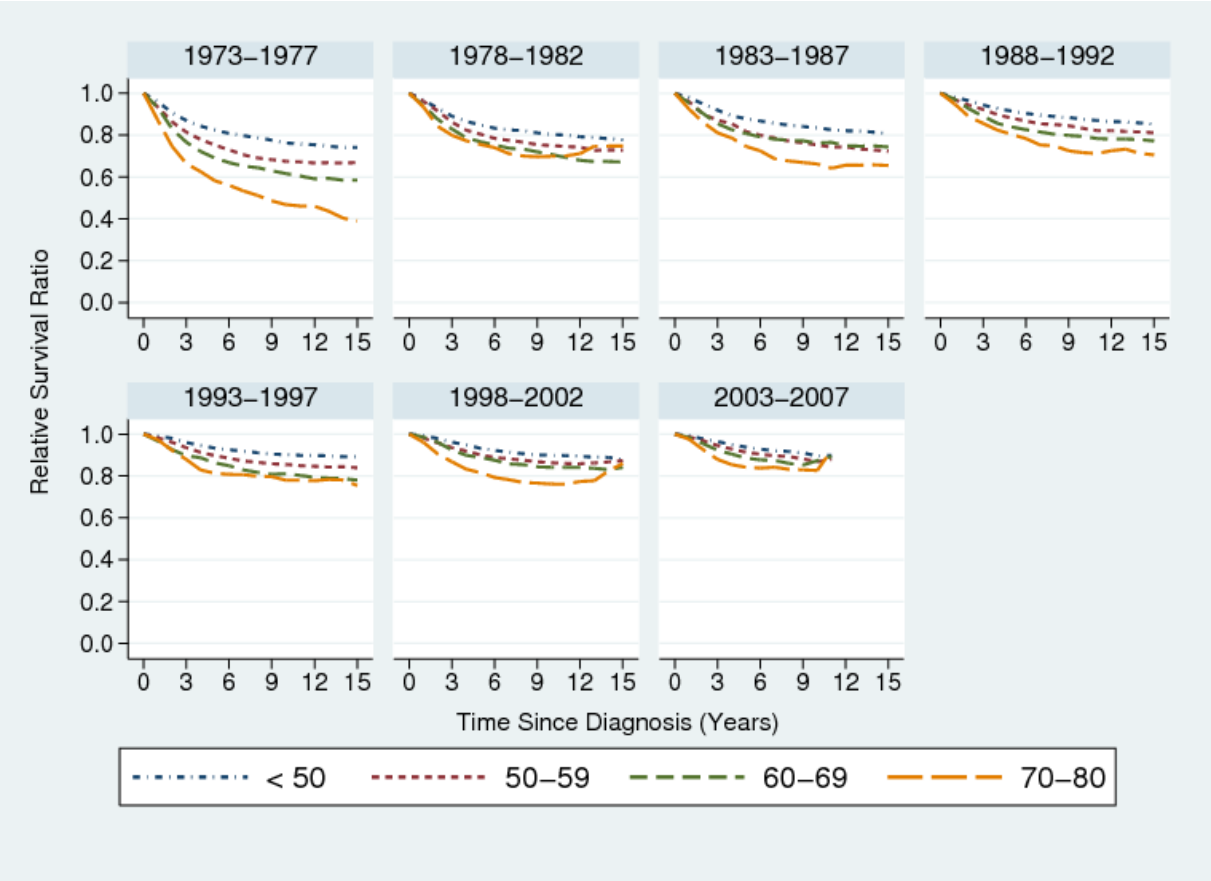
Assumptions and sensitivity analysis

In order to assess if statistical cure occurs in our data, we plotted the empirical estimates of relative survival, using the Ederer 2 method to estimate expected survival, using the same categorization for age at diagnosis and year of diagnosis as in Table 1 (eFigures 1-3).[2] There is currently no formal test available to determine if statistical cure is feasible, but plots of the cumulative relative survival curves can be used as a guideline when assessing the appropriateness of cure models for population-based data.[1] eFigures 1-3 show that the cumulative relative survival curves appear to flatten out for nearly all covariate patterns, indicating no, or little excess mortality after ten years of patient follow-up. The exceptions seem to occur among patients diagnosed with melanoma before the mid 1980´s and patients diagnosed with AML at ages 50-69 years in the most recent calendar period. In these groups model-based sensitivity analyses, similar to those shown in Figure 4 in the manuscript are warranted.

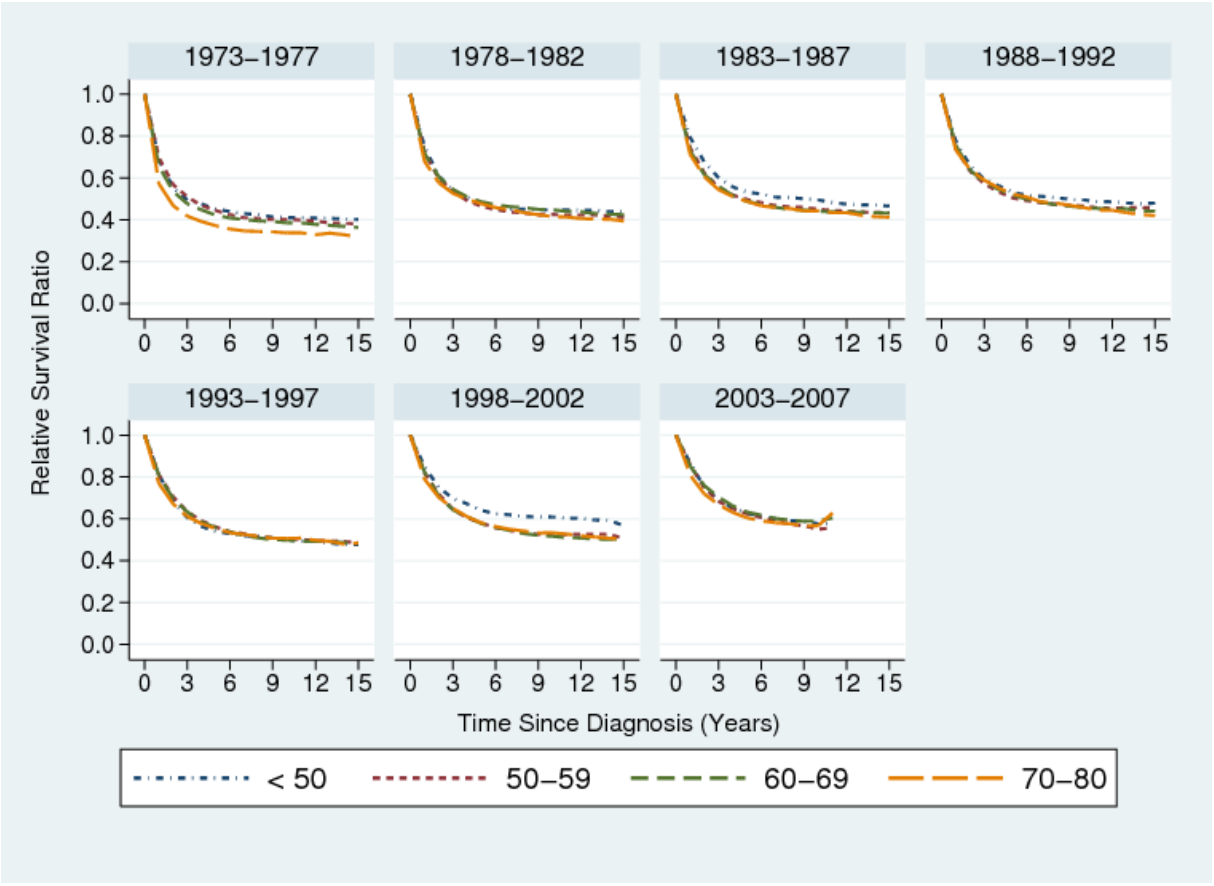
eTable 1: Placement of knots used to model the baseline excess hazard function in the flexible parametric cure models.

Cancer site	Degrees of freedom	Location of boundary knots (years since diagnosis)	Location of internal knots (years since diagnosis)
Melanoma	6	0 and 10	1.3, 2.6, 4.3, 6.6 and 9.1
Colon Cancer	6	0 and 10	0.3, 0.9, 2.1, 4.4 and 8.2
AML	7	0 and 10	0.1, 0.3, 0.7, 1.5, 4.1 and 9.0

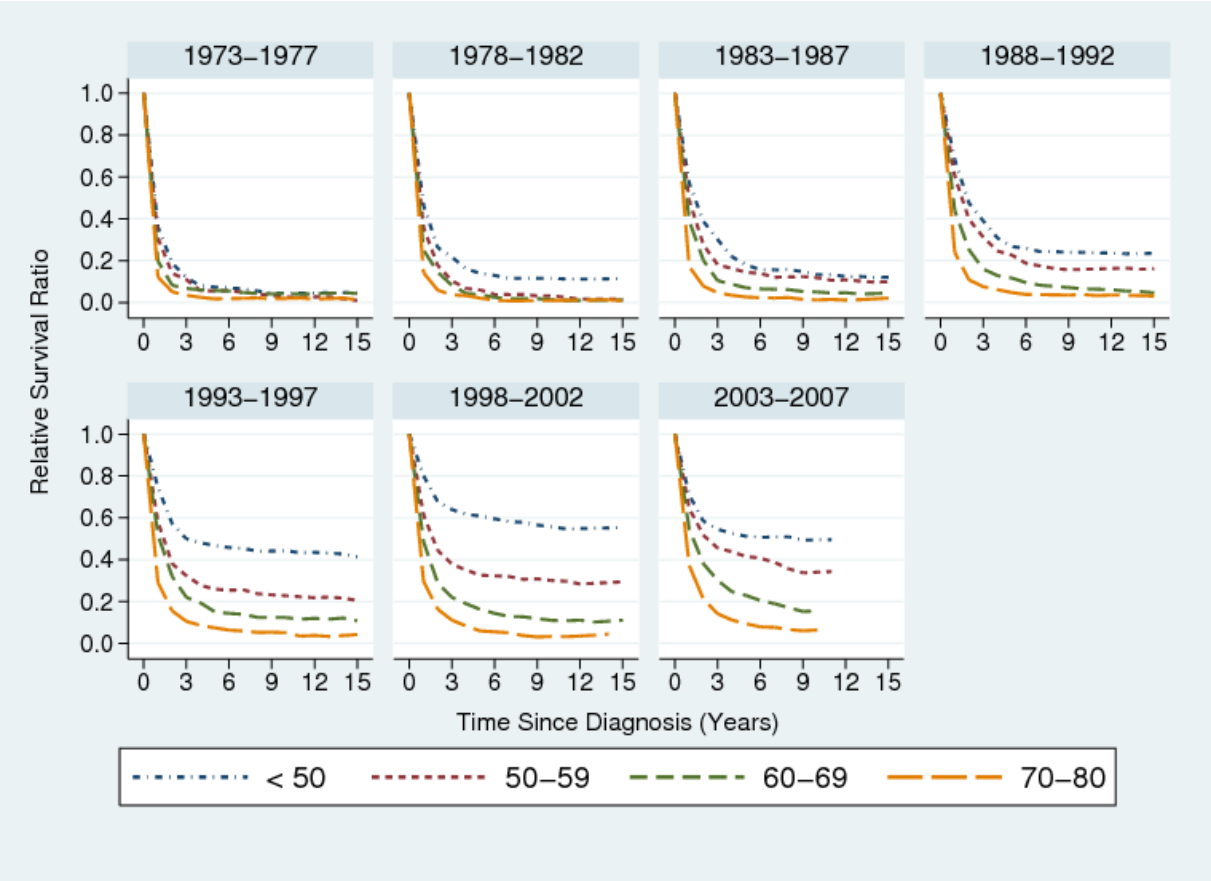
eFigure : Ederer II life table estimates of relative survival by age at diagnosis for patients diagnosed with melanoma in Sweden between 1973 and 2007.



eFigure 2: Ederer II life table estimates of relative survival by age at diagnosis for patients diagnosed with colon cancer in Sweden between 1973 and 2007.



eFigure 3: Ederer II life table estimates of relative survival by age at diagnosis for patients diagnosed with Acute Myeloid Leukemia (AML) in Sweden between 1973 and 2007.



References

1. Andersson TM, Dickman PW, Eloranta S, et al. Estimating and modelling cure in population-based cancer studies within the framework of flexible parametric survival models. BMC Med Res Methodol 2011;11:96. doi: 10.1186/1471-2288-11-96.

2. Ederer T, Heise H. Instructions to Ibm 650 Programmers in Processing Survival Computations. Methodological note no 10 End Results Evaluation Section, National Cancer Institute, Bethesda MD. 1959.