

EDITORIAL



A SAS macro for estimating direct adjusted survival functions for time-to-event data with or without left truncation

© The Author(s), under exclusive licence to Springer Nature Limited 2021, corrected publication 2022

SERIES EDITORS INTRODUCTION

Since we published our last article on adjusted survival curves many readers wrote asking how to calculate these for left-truncated, right-censored data and what confidence intervals and p-value should be included when comparing adjusted curves. Responding, we developed a SAS macro to address these questions. The macro, `%adjSurvlt()` is available online, reduces computing time and is easy to use. (If you are having difficulty ask any 5 year old with an iPad.) Some terms we use such as do-loops, Monte Carlo simulations and confidence bands rather than intervals may be foreign to many readers but do not be put off. The message should be clear and the macro useful. Many readers will want to share this article with their bio-statistician. We realize `%adjSurvlt()` is not a catchy name for our macro and are hosting a competition for suggestions. The winner will receive a Keuffel and Esser slide rule. We are standing by to announce the lucky reader.

Robert Peter Gale MD, PhD & Mei-Jie Zhang PhD.

ABSTRACT

There are several statistical programmes to compute direct adjusted survival estimates from results of the Cox proportional hazards model. However, when used to analyze observational databases with large sample sizes or highly stratified treatment groups such as in registry-related datasets, these programmes are inefficient or unable to generate confidence bands and simultaneous p values. Also, these programmes do not consider potential left-truncation in retrospectively collected data. To address these deficiencies we developed a new SAS macro `%adjSurvlt()` able to produce direct adjusted survival estimates based on a stratified Cox model. The macro has improved computational performance and is able to handle left-truncated and right-censored *time-to-event* data. Several mechanisms were implemented to improve computational efficiency including choosing matrix operations over do-loops and reducing dimensions of co-variate matrices. Compared to the latest SAS macro, `%adjSurvlt()` used $< 0.1\%$ computational time to process a dataset with 100 treatment cohorts and a sample size of 20,000 and showed similar computational efficiency when analyzing left-truncated and right-censored data. We illustrate use of `%adjSurvlt()` to compare retrospectively collected survival data of 2 transplant cohorts.

Bone Marrow Transplantation (2022) 57:6–10; <https://doi.org/10.1038/s41409-021-01435-2>

INTRODUCTION

Clinical trialists often use the Cox proportional hazards model to estimate hazard ratios between different treatment cohorts whilst controlling potential confounding by other co-variables from retrospective observational datasets [1]. Several approaches produce estimates of adjusted survival functions which consistently represent results of the Cox model [2–4]. For example, the *direct adjusted survival function method* relies on directly averaging predicted survival probabilities of subjects in pooled samples [2, 3, 5]. This approach overcomes limitations of the alternative *average covariate method* and better represents the underlying populations [4]. Estimates of direct adjusted survival functions can be interpreted as the survival probabilities of populations with similar prognostic co-variables [3]. Using a stratified Cox model also allows time-varying effects between treatment groups [6]. Confidence bands for the survival estimates can be constructed using a Monte Carlo method for pair-wise comparisons over given time periods [7, 8]. Similarly, p values can be derived for simultaneous hypothesis testing [8].

Lee et al. [9] developed several statistical programmes to compute direct adjusted survival estimates and Ghali et al. [10].

Zhang et al. [6] developed a SAS macro `%adjSurv()` able to produce direct adjusted survival estimates and estimates of standard errors and point-wise confidence limits. This macro was incorporated by the SAS Institute Inc. into SAS/STAT 12.1. More recently, Wang and Zhang improved the Zhang/SAS macro by adding the ability to produce confidence bands and simultaneous p values [8].

Despite these improvements there are limitations to the latest macro. First, it is inefficient for datasets with large sample sizes or highly stratified treatment groups, features common in multi-center observational studies like those of the *Centre for International Blood and Marrow Transplant Research* (CIBMTR) and *European Bone Marrow Transplant Group* (EBMT). Second, none of the programmes we discuss can handle left-truncated retrospective data. (Left truncation occurs when certain subjects from the underlying populations are unknown to the observers when their event-time fails to surpass certain time threshold such as subjects relapsing before a transplant can be done.) Choosing a starting time before the enrollment time in retrospective observational datasets introduces left-truncation. It also occurs when age is used as the time-scale instead of *time-on-study* [11, 12].

Received: 18 July 2021 Revised: 26 July 2021 Accepted: 3 August 2021
Published online: 19 August 2021

Table 1. Results of simulation study of CPU time between %adjsurvlt() macro and Wang and Zhang (W&Z) macro.

Subjects	N strata	Co-variables	Time (h:m:s)		% of Time
			New	W&Z	New/W&Z
1000	10	10	0:00:01	0:02:57	0.36
5000	10	10	0:00:07	0:21:10	0.52
	25	10	0:00:08	1:55:05	0.12
	50	10	0:00:14	6:40:35	0.06
10,000	10	10	0:00:22	0:56:00	0.65
	25	10	0:00:26	4:28:14	0.16
	50	10	0:00:28	15:13:30	0.07
20,000	100	10	0:04:06	120:06:19	0.06

Table 2. Simulation study of CPU time of %adjsurvlt() to analyze data with or without left-truncation.

Subjects	N strata	N co-variables	Time (h:m:s)	
			Left-truncated	Not left-truncated
1000	10	10	0:00:01	0:00:01
5000	10	10	0:00:09	0:00:07
10,000	10	10	0:00:31	0:00:22

To address these issues we developed a new SAS macro, %adjsurvlt() which is more computationally efficient and which handles left-truncated and right-censored data. Additional features include capabilities to produce confidence bands for survival differences and directly output survival curve plots in user-defined formats. An R package using the same algorithm is currently being developed and will be uploaded to the *Comprehensive R Archive Network* (CRAN).

COMPUTATIONAL METHODS

Detailed formulae to estimate direct adjusted survival functions and their variances are published [6]. Constructing confidence bands and computing simultaneous p values is described by Wang and Zhang [8]. For left-truncated data, the risk-set needs to be appropriately adjusted at estimating time individually.

Improvements in computational performance

%adjsurvlt() has several computational enhancements. First, instead of repeating large blocks of codes within do-loops it mostly relies on matrix operations which greatly reduces computing time. Second, the co-variate matrix is constructed with only the distinct co-variate vectors of the pooled sample. Doing so reduces the dimensions of the co-variate matrix because numbers of distinct co-variate vectors is often less than numbers of subjects.

Comparison of computational performance

Using %adjsurvlt() we repeated Wang and Zhang's simulation study [8]. Data sets with a sample size of 100 and two therapy cohorts were created in the original study using constant baseline hazard functions of three different parameter settings. Confidence bands for the differences between the two adjusted survival functions were constructed based on 1000 simulation processes with a 5% significance level. A thousand replications were performed for each setting. The rejection rate for each setting was calculated from the percentage of replications where the confidence bands failed to fully cover the zero reference line. Rejection rates achieved by %adjsurvlt() were consistent with those achieved using the Wang and Zhang macro.

Next, we compared performance of %adjsurvlt() and the Wang and Zhang macro. The original simulation study used a sample size of only 100 with two co-variables and a therapy-assignment variable with two strata. To better illustrate performance differences we increased sample sizes to 1000, 5000, 10,000 and 20,000 with 10 co-variables and 10, 25, 50 and 100 therapy strata.

Treatment assignment was based on a discrete uniform distribution of K values where K was the total number of strata. We constructed 10 co-variables. One-half were continuous variables based on standard normal distributions independent of the therapy-assignment and the other half dichotomous categorical co-variables dependent on therapy assignment. Baseline hazard functions were constant with a hazard rate of 0.1. Right-censoring times were generated from exponential distributions with varying hazard rates to keep the censoring rate close to 30%. 95% confidence bands were constructed based on 1,000 simulations. Performance tests were conducted on a server equipped with two Intel Xeon E5-2670 v3 CPUs at 2.30 GHz with 24 cores, 48 logic processors and 64GB of memory. SAS 9.4 (TS1M4) with SAS/IML 14.2 was used for the simulation.

Table 1 shows improved performance of %adjsurvlt() over the Wang and Zhang macro. As the sample sizes and the number of comparison strata increased, the percentage of CPU time saved by %adjsurvlt() increased. To complete the analysis of a simulated data set with a sample size of 10,000 and 50 strata, %adjsurvlt() used only 0.07% of the CPU time required by Wang and Zhang macro, 15 h less. With a sample size of 20,000 and 100 strata %adjsurvlt() required only 0.06% of the CPU time or 120 h less.

We also evaluated the performance of %adjsurvlt() for analyzing data with and without left-truncation. To simulate left-truncated data initial sample sizes were increased by 50%. Left-truncation times were generated from exponential distributions with varying hazard rates to keep the truncation rate near 33.3%. Left-truncated observations were dropped from the final samples. As displayed in Table 2 performance of %adjsurvlt() was similar for analyzing data sets with the same sample sizes regardless of left-truncation. It took 22 s to process a data set of 10,000 subjects which was not left-truncated versus 31 s for a left-truncated dataset. The source code of %adjsurvlt() and the specifications of the macro are listed in the supplemental part.

An example

A CIBMTR analysis compared outcomes of allotransplants in persons with myeloproliferative neoplasms (MPN) in blast phase compared with persons with *de novo* acute myeloid leukemia (AML) or with myelodysplastic syndromes (MDS) transformed to AML [13]. A significant interaction was found for survival between the diseases and disease state pretransplant.

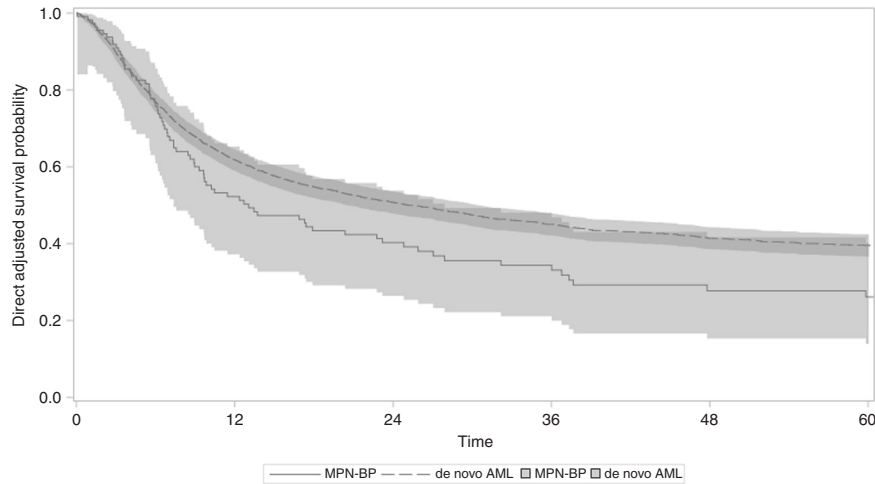


Fig. 1 Direct adjusted survival curves for subjects with blast phase MPN (blue solid line with 95% confidence band) and those with *de novo* AML in remission pretransplant (red solid line with 95% confidence band).

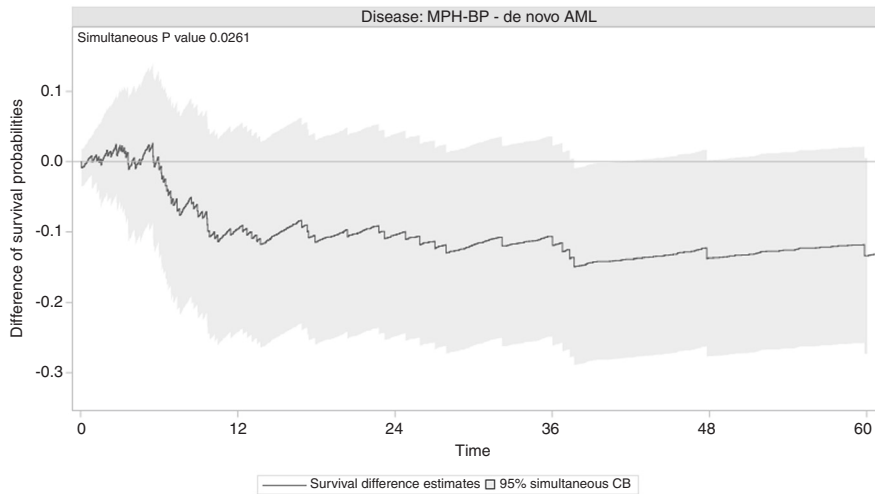


Fig. 2 Differences in direct adjusted survival curves for between subjects with blast phase MPN and those with *de novo* AML in remission pretransplant (solid line with 95% confidence band).

We show use of `%adjsurvlt()` on a subset of 96 subjects with blast phase MPN and 2825 subjects with *de novo* AML in remission pretransplant with death the outcome of interest and with end of follow-up considered right-censoring. Subject- and disease-related co-variables associated with survival were age at transplant (40–49 years, 50–59 years, 60–69 years or ≥ 70 years), recipient sex, Karnofsky performance score (90–100, < 90 , or missing), cytogenetics (favorable/intermediate, poor, or missing), donor-type (HLA-matched sibling, other HLA-matched relative, HLA-well-matched unrelated or HLA-partially/mismatched unrelated donor), conditioning regimen (total body irradiation-based myeloablative, chemotherapy-based myeloablative or reduced-intensity/non-myeloablative) and year of transplant (2001–2005, 2006–2010 or 2011–2015). Compared with subjects with *de novo* AML, subjects with blast phase MPN were older (chi-square $p < 0.01$), more likely male ($p = 0.04$), more likely to have a low Karnofsky performance score ($p = 0.02$) and more likely to have been transplanted recently ($p = 0.04$). Using a Cox regression model to adjust these prognostic co-variables showed a hazard ratio of 1.40 (95% Confidence Interval [CI], 1.11, 1.76; $p < 0.01$) for death in subjects with blast phase MPN compared with subjects with *de novo* AML.

A SAS data set `inlib.final` was prepared for the example. The data set contains these categorical variables: `disgp` for

disease indication, `agegp` for age at transplant, `sex` for sex, `karnofcat` for Karnofsky score at transplant, `cytoab` for cytogenetics, `donorgp` for donor type, `condint` for conditioning regimen and `yeartxgp` for year of transplant. It also includes death as the event indicator and `intxsurv onths`.

We used this statement to evoke the macro:

```
%adjsurvlt(indata=inlib.final, event=dead, time=
intxsurv, strata=disgp, covlst=agegp sex kpsgp cytoab
donorgp condint yeartxgp, seed=86311, nsim=10000, time-
list=12 36 60, maxtime=60, outsurvplot=2, outdiffplot=
1, showc=0, tickvalues=0 12 24 36 48 60, width=800px,
height=450px, imagefmt=pdf);
```

By specifying `strata=disgp` and `covlst=agegp sex kpsgp cytoab donorgp condint yeartxgp`, we estimated survival differences between diseases whilst adjusting for prognostic co-variables. By setting `timelist=12 36 60` and `maxtime=60`, we shortened the output results to 12, 36, and 60 months and set the upper boundary of confidence bands to 60 months. Without specifying a value for `mintime`, we let the macro take its default value of the minimum observed event time as the lower boundary of confidence bands. We also specified `outsurvplot=2` and `outdiffplot=1` to generate the adjusted survival plot and the survival difference plot. Using

Table 3. Output results of %adjsurvlt().

1. Summary of event and censoring								
					Band boundaries			
Disease	Total	Event	Censor	Lower	Upper			
MPN-BP	96	77	19	0.0658	60			
de novo AML	2825	1781	1044	0.0329	60			
2. Direct adjusted survival estimates								
				Conf limits		Conf bands		
Time	Disease	Surv	Std Err	Lower	Upper	Lower	Upper	N at risk
12	MPN-BP	0.5223	0.0480	0.4244	0.6112	0.3722	0.6525	46
36	MPN-BP	0.3437	0.0466	0.2542	0.4349	0.2107	0.4808	25
60	MPN-BP	0.2614	0.0455	0.1773	0.3533	0.1393	0.4013	16
12	de novo AML	0.6181	0.0090	0.6002	0.6354	0.5896	0.6452	1732
36	de novo AML	0.4503	0.0092	0.4321	0.4683	0.4215	0.4786	1133
60	de novo AML	0.3954	0.0093	0.3772	0.4136	0.3666	0.4241	797
3. Direct adjusted survival difference estimates								
				Conf limits		Conf bands		
Time	Disease 1	Disease 2	Surv Diff	Std Err	Lower	Upper	Lower	Upper
12	MPN-BP	de novo AML	-0.0958	0.0488	-0.1914	-0.0003	-0.2422	0.0505
36	MPN-BP	de novo AML	-0.1066	0.0474	-0.1995	-0.0136	-0.2490	0.0359
60	MPN-BP	de novo AML	-0.1340	0.0463	-0.2247	-0.0432	-0.2730	0.0051
4. Hypothesis tests								
Disease 1		Disease 2	N of Sim	Crit value		P value		
MPN-BP		de novo AML	10000	3.0018		0.0261		

showci = 0 we suppressed the display of confidence limits in the plots. We manually set the values of the tick marks on the x axis using tickvalues = 0 12 24 36 48 60. Lastly, we used width = 800px, height = 450px and imagefmt = pdf to set the resolution and file format of the plots.

Table displays output results of %adjsurvlt(). Adjusted survival estimates for the myeloproliferative neoplasms in blast phase cohort at 1, 3 and 5 years were 52% (42, 61%), 34% (25, 44%) and 26% (18, 35%). Similar survival estimates for the *de novo* AML cohort are 62% (60, 64%), 45% (43, 47%) and 40% (38, 41%). Adjusted survival difference estimates were 10% (0, 19%), 11% (1.4, 20%) and 13% (4, 23%) at 1, 3 and 5 years favoring the *de novo* AML cohort. Adjusted survival plots and the difference plot were saved as PDF files (Figs. 1 and 2). There was a time-varying effect between the two cohorts before and after the first 6 months (proportionality test $p = 0.10$). Consequently, a simultaneous p value derived from the direct adjusted survival estimates would be more appropriate than a p value of a Cox regression. As indicated in Table 3 the difference in adjusted survival between the cohorts was significant with a simultaneous p value of 0.03.

DISCUSSION

We show %adjsurvlt() is computationally efficient for estimating direct adjusted survival functions for datasets with large sample sizes and highly stratified therapy cohorts and can handle left-truncation in right-censored data. %adjsurvlt() provides estimates of standard errors, confidence limits and confidence bands for adjusted survival functions, for pair-wise differences and p values for simultaneous hypothesis testing of survival differences over specified time intervals.

%adjsurvlt() has some limitations. First, although it can conduct pair-wise comparisons between therapy cohorts, further development is needed to support comparisons across three or more therapy cohorts. Second, the underlying Cox model only applies when no competing risk precludes the outcome(s) of

interest. When there are competing risks we need to use alternative multiple regression models such as the Fine-Gray model and its corresponding direct adjusted cumulative incidence [14, 15]. However, compared with other statistical programmes %adjsurvlt() provides improved computational performance and more flexibility in handling right-censored *time-to-event* data with or without left-truncation. We hope readers will find it useful.

Zhen-Huan Hu¹✉, Hai-Lin Wang², Robert Peter Gale³ and Mei-Jie Zhang¹

¹Medical College of Wisconsin, Milwaukee, WI, USA. ²Everest Clinical Research Services, Inc., Little Falls, NJ, USA. ³Imperial College London, London, UK. An updated version of the SAS macro is available at:

<https://github.com/zhen-huan-hu/adjsurvlt>
✉email: novelkeny@gmail.com

REFERENCES

- Cox DR. Regression models and life-tables. *J R Stat Soc Ser B Stat Methodol.* 1972;34:187–220.
- Chang IM, Gelman R, Pagano M. Corrected group prognostic curves and summary statistics. *J Chronic Dis.* 1982;35:669–74.
- Makuch RW. Adjusted survival curve estimation using covariates. *J Chronic Dis.* 1982;35:437–43.
- Nieto FJ, Coresh J. Adjusting survival curves for confounders: a review and a new method. *Am J Epidemiol.* 1996;143:1059–68.
- Gail MH, Byar DP. Variance calculations for direct adjusted survival curves, with applications to testing for no treatment effect. *Biom J.* 1986;28:587–99. <https://doi.org/10.1002/bimj.4710280508>.
- Zhang X, Loberiza FR, Klein JP, Zhang MJ. A SAS macro for estimation of direct adjusted survival curves based on a stratified Cox regression model. *Comput Methods Prog Biomed.* 2007;88:95–101. <https://doi.org/10.1016/j.cmpb.2007.07.010>.
- Zhang MJ, Klein JP. Confidence bands for the difference of two survival curves under proportional hazards model. *Lifetime Data Anal.* 2001;7:243–54.
- Wang H, Zhang X. Confidence band for the differences between two direct adjusted survival curves. *J Stat Softw.* 2016;70:17. <https://doi.org/10.18637/jss.v070.c02>.

9. Lee J, Yoshizawa C, Wilkens L, Lee HP. Covariance adjustment of survival curves based on Cox's proportional hazards regression model. *Comput Appl Biosci*. 1992;8:23–27. <https://doi.org/10.1093/bioinformatics/8.1.23>.
10. Ghali WA, Quan H, Brant R, van Melle G, Norris CM, Faris PD, et al. Comparison of 2 methods for calculating adjusted survival curves from proportional hazards models. *JAMA*. 2001;286:1494–7. <https://doi.org/10.1001/jama.286.12.1494>.
11. Lamarca R, Alonso J, Gomez G, Munoz A. Left-truncated data with age as time scale: an alternative for survival analysis in the elderly population. *J Gerontol A Biol Sci Med Sci*. 1998;53:M337–343. <https://doi.org/10.1093/gerona/53a.5.m337>.
12. Pencina MJ, Larson MG, D'Agostino RB. Choice of time scale and its effect on significance of predictors in longitudinal studies. *Stat Med*. 2007;26:1343–59. <https://doi.org/10.1002/sim.2699>.
13. Gupta V, Kim S, Hu ZH, Liu Y, Aljurf M, Bacher U, et al. Comparison of outcomes of HCT in blast phase of BCR-ABL1- MPN with de novo AML and with AML following MDS. *Blood Adv*. 2020;4:4748–57. <https://doi.org/10.1182/bloodadvances.2020002621>.
14. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *J Am Stat Assoc*. 1999;94:496–509. <https://doi.org/10.2307/2670170>.
15. Zhang X, Zhang MJ. SAS macros for estimation of direct adjusted cumulative incidence curves under proportional subdistribution hazards models. *Comput Methods Programs Biomed*. 2011;101:87–93. <https://doi.org/10.1016/j.cmpb.2010.07.005>.

ACKNOWLEDGEMENTS

The data analyzed were supplied by the *Centre for International Blood and Marrow Transplant Research* (CIBMTR) which was supported in part by *National Institutes of Health* [grant number U24-CA076518-20]. The findings presented here are the responsibility of the authors and are not the opinion of the CIBMTR. The authors thank Jennifer Motl, of Milwaukee, Wisconsin, for editorial support, funded by the

Medical College of Wisconsin, Milwaukee, Wisconsin, in accordance with *Good Publication Practice* (GPP3) guidelines (<http://www.ismpp.org/gpp3>). RPG acknowledges support from the National Institute of Health Research (NIHR) Biomedical Research Centre funding scheme.

AUTHOR CONTRIBUTIONS

ZH, HW, and MZ conceived and designed the study. ZH and HW developed the SAS and R programmes and collected and assembled the testing data. ZH, HW, and MZ participated in data analyses and interpretations. All authors wrote the typescript and agreed to submit for publication.

COMPETING INTERESTS

RPG is a consultant to BeiGene Ltd., Fusion Pharma LLC, LaJolla NanoMedical Inc., Mingsight Pharmaceuticals Inc. and CStone Pharmaceuticals; advisor to Antegene Biotech LLC, Medical Director, FFF Enterprises Inc.; partner, AZAC Inc.; Board of Directors, Russian Foundation for Cancer Research Support; and Scientific Advisory Board: StemRad Ltd.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41409-021-01435-2>.

Correspondence and requests for materials should be addressed to Z.-H.H.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.