Analysis of Complex Survival Data: a tutorial using the Shiny MSM.app application

Gustavo Soutinho¹ and Luís Meira-Machado²

¹EPIUnit - University of Porto Rua das Taipas 135, 4050-600 Porto - Portugal, gdsoutinho@gmail.com ²Centre of Mathematics, University of Minho, lmachado@math.uminho.pt

Keywords: R language, Shiny package, Survival analysis, Multi-state models.

The development of applications for obtaining interpretable results in a simple and summarized manner in multi-state models is a research field with great potential, namely in terms of using open source tools that can be easily implemented in biomedical applications. In this tutorial, we introduce MSM.app, an interactive web application using the Shiny package for the R language. In the following sections, we present the main functionalities of the MSM.app and an explanation of the outputs obtained for better understanding, independent of the statistical knowledge of users.

1. What is the Shiny MSM.app?

The appearance of the shiny R package allowed to automatically share results obtained from the R language to be analyzed for users without any prior knowledge in terms of informatics via the internet [1, 2, 3].

In this context, the MSM.app web application arises with the goal of carrying out an analysis of multi-state survival data sets. Among its functionalities, we can highlight the possibility to conduct a traditional survival analysis regarding the following topics: (i) estimation of survival functions (using the classical Kaplan-Meier estimator); (ii) comparison of survival functions between groups; and (iii) use of semi-parametric and parametric regression models to study the relationship between explanatory variables and survival time. The MSM. app can also be used for the analysis of multi-state survival data that can be seen as a generalization of survival analysis in which survival is the ultimate outcome of interest but where information is available about intermediate events that individuals may experience during the study period [4, 5, 6].

To this statistical analysis, the Shiny MSM.app application combines shiny with some other packages such as survival [7], mstate [8] and survidm [9].

2. Features of the Shiny MSM.app

The MSM.app is a web application that was developed to perform dynamic analysis through a set of dynamic web forms, tables, and graphics.

It is built upon two components: the user-interface scripts for the layout of the application where the outputs are displayed (ui.R); and the other given by the server scripts with the instructions of the application (server.R) [10].

Among the main aspects of this Shiny web tool that improve the flexibility and productivity, we could highlight the easy rendering of the contents without multiple reloads, the feature to add computed (or processed) outputs from R scripts, or interactively add reports and visualizations.

The MSM.app also provides integration with other R packages, javascript libraries or CSS customization, being under the GPL-2 open source license. The communication between the client and server is done over the normal TCP connection [11].

3. How is the Shiny MSM.app different from other applications?

There exists a set of available web tools specifically aimed at carrying out some parts of multi-state analysis.

Among them, we can stand out the MSDshiny which provides a useful and streamlined way to plan and power clinical trials with multi-state outcomes such as a view of the multi-state structure, treatment effects, or the results of different types of simulations. Another example is MSM-shiny application. This tool uses a CSV file containing multi-state data and provides the modeling and comparison of transition hazard models and the prediction of occupation probabilities [12]. Recently, MSMplus provides a flexible visualization of the transition probabilities, transition intensities, or probability of visiting a particular state [13]. After analyzing existing web applications, we have concluded that their use by non-statisticians has been limited. A possible reason for this is the lack of friendly software that covers the main goals involving survival analysis and multistate models on the same platform.

The MSM.app allows users to explore various types of multi-state models and perform regression inference as well as obtain several predictive measures of interest, such as the occupation probabilities, the transition probabilities, and the cumulative incidence functions. Recent methods for checking the Markov assumption are also implemented.

4. How to get the Shiny MSM.app

The Shiny MSM.app is available for free open access at the Shiny Apps repository https://gsoutinho.shinyapps.io/appmsm/.

5. Structure of the Shiny MSM.app

The web application consists of three parts representing different aspects of the survival analysis and its extension to complex multi-state models.

The first one allows to perform the survival analysis from mainly of most common functions of the survival R packages.

The second enables one to obtain some of the main goals of a multi-state analysis, such as the inference of regression models and the estimation of transition probabilities, through the **survidm** and **mstate** R packages.

Finally, MSM.app also includes local and global statistical tests to check the Markov assumption for multi-state using the markovMSM package.

6. Type of input data files and requirements

The MSM.app only requires CSV files as input. By default, in terms of structure, the values of the data set should be separated by a comma. However, files that use a semicolon to separate the values are also accepted.

Three examples of data sets are available for consultation in the MSM.app from which it is possible to check the requirements that the files must have to pursue the data analysis. Each one corresponds to a specific type of structure of data: survival, illness-death model, or more general multi-state models.

The first data set corresponds to survival data in patients with acute myelogenous leukemia [14]. Figure 1 shows some registers and the three columns that the file must have. "Time1" and "status" correspond to the time to the event and the status of the censoring, respectively. "x" is the only covariate in this data that can be used for regression or obtaining the survival estimates for groups.

	A	В	С
1	time1	status1	x
2	9	1	Maintained
11	48	1	Maintained
12	161	0	Maintained
13	5	1	Nonmaintained
14	5	1	Nonmaintained
15	8	1	Nonmaintained
16	8	1	Nonmaintained

Figure 1: Structure of the CSV files for survival analysis (given by the *aml* data set).

The second corresponds to a data set from a clinical trial on colon cancer, which can be modeled using the progressive illness-death model [15]. Figure 2 shows the schematic diagram of transitions involved in the model. Among the variables of the *colonIDM* data set, the first four correspond to the times or the status indicator for the illness-death model. "time1" represents the sojourn time in the initial state and "Stime" and "event1" and "event" the corresponding censoring indicates. The other variables could be used for transition probabilities or to check the effect on the transition intensities (Figure 3).

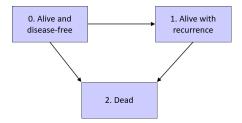


Figure 2: Illness-death model for the colon cancer study.

Finally, extensions to progressive processes beyond the three-state illnessdeath model are discussed using data from the European Group for Blood

1	A	В	С	D	E	F	G	Н	L	J	К	L
1	time1	event1	Stime	event	rx	sex	age	obstruct	perfor	adhere	nodes	differ
2	968	1	1521	1	Lev+5FU	1	<mark>43</mark>	0	0	0	5	2
3	3087	0	3087	0	Lev+5FU	1	63	0	0	0	1	2
4	542	1	963	1	Obs	0	71	0	0	1	7	2
5	245	1	293	1	Lev+5FU	0	66	1	0	0	6	2
6	523	1	659	1	Obs	1	69	0	0	0	22	2
7	904	1	1767	1	Lev+5FU	0	57	0	0	0	9	2
8	229	1	420	1	Lev	1	77	0	0	0	5	2

Figure 3: Structure of the CSV files for Illness-death models analysis (given by the *colonIDM* data set).

and Marrow Transplantation (EBMT) [4]. The movement of the patients among the six states can be modelled through the multi-state model with the following six states: 'Alive and in remission, no recovery or adverse event' (State 0); 'Alive in remission, recovered from the treatment' (state 1); 'Alive in remission, occurrence of the adverse event' (state 2); 'Alive, both recovered and adverse event' (state 3); 'Alive, in relapse' (treatment failure) (state 4) and 'Dead (treatment failure)' (state 5). In total there are 12 transitions, three intermediate events given by recovery (Rec), adverse event (AE) and a combination of the two (AE and Rec), and two absorbing states: Relapse and Death (Figure 4).

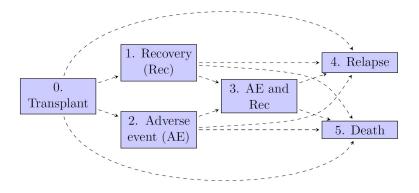


Figure 4: A six-states model for leukemia patients after bone marrow transplantation.

In terms of format, the CSV file should be in wide format. In the case of the *ebmt4* data (Figure 5), the transition times from the initial to the ultimate state ('srv') or to the intermediate states ('rec', 'ae', 'recae' and 'rel'), and the corresponding censoring variables are rec.s', 'ae.s', 'recae.s', 'rel.s' and 'srv.s'. Other variables correspond to the covariates for regression models.

	A	В	С	D	E	F	G	Н	1	J	Κ	L	M	N	0
1	id	rec	rec.s	ae	ae.s	recae	recae.s	rel	rel.s	srv	srv.s	year	agecl	proph	match
2	1	22	1	995	0	995	0	995	0	995	0	1995-1998	20-40	no	no gender mismatch
3	2	29	1	12	1	29	1	422	1	579	1	1995-1998	20-40	no	no gender mismatch
4	3	1264	0	27	1	1264	0	1264	0	1264	0	1995-1998	20-40	no	no gender mismatch
5	4	50	1	42	1	50	1	84	1	117	1	1995-1998	20-40	no	gender mismatch
6	5	22	1	1133	0	1133	0	114	1	1133	0	1995-1998	>40	no	gender mismatch
7	6	33	1	27	1	33	1	1427	0	1427	0	1995-1998	20-40	no	no gender mismatch
8	7	29	1	28,5	1	29	1	775	0	775	1	1995-1998	>40	no	no gender mismatch
9	8	31	1	1618	0	1618	0	1618	0	1618	0	1995-1998	20-40	no	no gender mismatch
10	9	87	1	29	1	87	1	1111	0	1111	0	1995-1998	20-40	no	gender mismatch

Figure 5: Structure of the CSV files for multi-state models analysis (given by the *ebmt4* data set).

For all these data sets, the names of the variables could be different. In these cases, the user must correctly indicate the corresponding name in the web forms of the application. The data files can be accessed at https://w3.math.uminho.pt/~lmachado/shiny/.

7. How to select the input files?

On the "input files" page we can find an interactive form. For each type of data set ("survival data", "illness-death model", or "multi-state model"), a new web form appears below the radio buttons, in which we indicate the times to the events or the status indicator (Figure 6, left hand side and center). In case of the more complex multi-state models, it is also necessary to indicate the transition schema as well as the number of states, for instance (Figure 6, right hand side).

After submitting a data set, a table appears on the right hand side of the page which can be dynamically. This table can be changed using filters or by searching for specific words in the table. Figure 7 shows a partial view the *veteran* data set that can be found in the **survival** package.

Input files	Input files	Input files
Input files	Browse colonIDM.csv	Browse ebmt4.csv
Browse veteran_comma.csv Uplaad complete	Type of data: © Survival data : # Illness-death model (IDM) © Other multi-state model (MSM) Names of the states (e.g. Healthy/Illness.Death)':	Uptood comprise Type of data: Survival data IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
Type of data:	Healthy,Illness,Death	Number of states:
\circledast Survival data \odot Illness-death model (IDM) \odot Other multi-state model (MSM)	Select variables for time and status: Time to the intermediate state:	6 Transitions schema (e.g. for ebmt4 data set: 2,3,5,6,4,5,6,5,6,5,6,0)*:
Select variables:	time1 •	2,3,5,6;4,5,6;4,5,6;5,6,0
Event time:	The status indicator of entering in the intermediate state (0-right censored, 1-event):	Name of states (e.g. for ebmt4 data set: Tx,Rec,AE,Rec+AE,Rel,Death)*:
time 🔹	Time to the ultimate state:	Tx.Rec.AE.Rec+AE.Rel,Death Variables for event times (e.g. for ebmt4 data set: rec.ae.recae.rel.srv)*:
Status (The status indicator, normally 0=right censored, 1=event):	Stime	rec,ae,recae,rel,srv
status	The status indicator of entering in the ultimate state (0-right censored, 1-event): event	Variables for events status (e.g. for ebmt4 data set; rec.s,ae.s,recae.s,rel.s,srv.s)*: rec.s,ae.s,recae.s,rel.s,srv.s

Figure 6: The event time and status variables for the survival analysis from the *veteran* data set (left); an indication of the two event times and their corresponding status for the illness-death model given by the *colonIDM* data (center); and a description of the MSM model through the number and the state names, the transition schema, and the event times and corresponding status for the *ebmt4* data (right).

MSM.app: a Web-Based Tool for the Analysis of Multi-state Survival Data

MSM.App Introduction Input files Survival Analysis • IDM - Analysis • MSM - A	nalysis 👻 Test	s for the Markov condition 👻								
	NULL									
Input files	Data set currently loaded:									
Browse veteran_comma.csv	Show 25	v entries								
Upload complete	trt	celltype	0 time	status	karno karno	diagtime				
Type of data:	1	squamous	72	1	60	7				
Survival data Illness-death model (IDM) Other multi-state model (MSM)	1	squamous	411	1	70	5				

Figure 7: The input file page with the data table with some results for the *veteran* data set and three radio buttons for each type of model.

8. How to carry out a Survival analysis?

From the "survival analysis" button, we can carry out the classical methods for survival analysis. In this section, we analyze the main aspects of the outputs displayed on the right sides of the pages: Kaplan-Meier estimator, Compare survival curves, Cox PH models and parametric models, by using the *veteran* data set.

Kaplan-Meier estimator

Non-parametric estimation of the survival function is traditionally performed using the Kaplan-Meier estimator, which can be desegregated for different groups of categorical variables. By default, a summary of the estimates is presented on the right side. Figure 8 shows, respectively, times with events, the number of individuals at risk at that time, and the survival estimates with corresponding standard errors and confidence intervals. It is also possible to display an interactive plot of the survivals with (or without) the confidence interval intervals (Figure 9).

ISM.App	Introduction	Input files	Survival Analysis 👻	IDM - An	alysis 👻	MSN	I - Analysi	s v T	ests for t	ne Markov cond	lition -
					Call:	survfit	(formula	= formu:	la.km, d	ata = db)	
Kaplan	-Meier es	timator					cellt	ype=aden			
					time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
Select varia	ble				3	27	1	0.9630	0.0363	0.89430	1.000
					7	26	1	0.9259	0.0504	0.83223	1.000
celltype				-	8	25	2	0.8519	0.0684	0.72786	0.997
					12	23	1	0.8148	0.0748	0.68071	0.975
					18	22	1	0.7778	0.0800	0.63576	0.952
View plot					19	21	1	0.7407	0.0843	0.59259	0.926
KM estin	nates				24	20	1	0.7037	0.0879	0.55093	0.899
					31	19	1	0.6667	0.0907	0.51059	0.870
+ Daumlaa	d n df				35	18	1	0.6296	0.0929	0.47146	0.841
L Download	a par				36	17	1	0.5926	0.0946	0.43344	0.810
					45	16	1	0.5556	0.0956	0.39647	0.778

Figure 8: Summary of the survival estimation for the categorical covariate "celltype" of the *veteran* data set using the Kaplan-Meier estimator.

Compare survival curves

Statistical tests can be used to compare survival rates between groups. The null hypothesis states that there is no difference in survival between groups. The log-rank test and the Gehan-Wilcoxon test are the most commonly used. Both tests are available in the MSM.App. The output of the tests comprises the chi-squared statistics, degrees of freedom, and the corresponding p-value. The number of individuals for each group and the number of observed and expected values are also presented. Results of Figure 10 show significant differences in the survival curves among groups given by the cell types (p < 0.0001).

Cox PH models

The semi-parametric Cox proportional hazards model [16] is usually used to evaluate the effects of several factors on survival. The output for the Cox model is updated as the user selects the variables to be included in the model in the web form. Figure 11 shows, respectively, the estimate of the coefficients

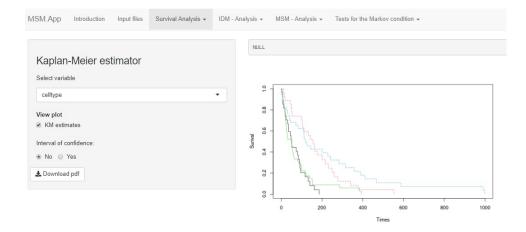


Figure 9: The output of the survival estimation for the categorical covariate "celltype" of the *veteran* data set using the Kaplan-Meier estimator. Survival curves for each group with confidence intervals could be shown at the bottom by choosing "Yes".

for the covariates "celltype", Karnofsky performance score ("karno") and "age"; the hazard ratio given by exp(coef) and the respective standard error; and the statistical significance of the model. Result of the global test for the proportional hazards assumption confirms that this requirement is no fulfilled and consequently a parametric accelerated failure time (AFT) model must be preferable to be fitted to this survival data.

Parametric AFT models

The idea behind the outputs for parametric survival models is quite similar to those provided by the Cox PH models. In this case, six possibilities of distributions are available to model the baseline hazard function of the models: exponential, weibull, gaussian, logistic, lognormal or loglogistic. Besides the summary of the models, the Akaike Information Criterion (AIC) values are also presented to make it easier to compare which model fits better to assess the effect of several risk factors on survival time. Figures 12, 13 and 14 show, respectively, the outputs of the AFT models given by the exponential, weibull and loglogistic distributions. Through the AUC values we can assume that loglogistic parametric model is preferable. From this model, we can observe that only "age" and "small cells" do not have a significant effect

SM.App Introduction Input files Survival Analysis - IDM	M - Analysis + MSM - Analysis + Tests for the Markov condition +
Compare survival curves	Call: survdiff(formula = formula.LR, data = db, rho = 0)
Compare curvia curves	N Observed Expected (0-E)^2/E (0-E)^2/V
Type of method:	celltype=adeno 27 26 15.7 6.77 8.19
ijpo or mourou.	celltype=large 27 26 34.5 2.12 3.02
Log-rank O Gehan-wilcoxon	celltype=smallcell 48 45 30.1 7.37 10.20
	celltype=squamous 35 31 47.7 5.82 10.53
Select variable	Chisq= 25.4 on 3 degrees of freedom, p= 1e-05
celltype -	

Figure 10: The output of the Log-rank test for the categorical covariate "cell-type" of the *veteran* data set for testing the null hypothesis of no difference in survival between the two groups.

on survival.

9. How to carry out a illness-death model Analysis?

From the "IDM-Analysis" button, we can carry out a data analysis of a progressive illness-death (IDM) model given by two events, three states, and three transitions.

Number of events

First, we are interested in having an idea of the movement of individuals among the three states. It is also possible to see the proportions of transitions (Figure 15).

Regression models

In the multi-state models, there are so many transition intensities as there are transitions. For each one, we can then check the effect of the individual characteristics of the individuals by fitting separate intensities using semiparametric Cox proportional hazard regression models [16]. In the inference of the regression models, we must take into account if the Markov condition, in which the past and future are independent given the present state, is verified. Regarding the dependence of the transition intensities and time, in the case of the failure of the markovianity, we can use a semi-Markov model in which the future of the process does not depend on the current time but rather on the duration of the current state. These two methods are available in MSM.app. In terms of interpretation, the outputs of the

MSM.App Introduction	Input files Survival Ana	IDM - Analysis • MSM - Analysis • Tests for the Markov condition •
Cox PH model Select variables trt Celltype Karno diagtime age prior		Call: coxph(formula = formula.cox, data = db) coef exp(coef) se(coef) z p celltypelarge -0.84993 0.427418 0.293910 -2.892 0.00383 celltypesmallcell -0.447778 0.639047 0.256061 -1.749 0.08034 celltypesquamous -1.171907 0.309776 0.293738 -3.990 6.62e-05 karno -0.032016 0.968492 0.00544 -5.924 3.14e-09 age -0.006034 0.993984 0.009054 -0.666 0.50512 Likelihood ratio test=59.81 on 5 df, p=1.331e-11 n= 137, number of events= 128 Test the Proportional Hazards Assumption of a Cox Regression: chisq df p celltype 14.15 3 0.00271 karno 14.77 1 0.00012 age 1.99 1 0.15823 GLOBAL 30.18 5 1.4e-05

Figure 11: Results of the Cox model for the variables "celltype", "karno" and "age" of the *veteran* data set. Result of the test for the proportional hazard assumption shows that a Cox PH model

regression, for each transition, are very similar to those presented for the survival analysis. Results of Figures 16 indicate that save the treatment "Lev(amisole)+5-FU" all the other covariates have no effect for recurrence transition and only age could be considered important on mortality without recurrence. At least, only the covariates "perfor", "sex" and the treatment "Obs" have no association for the transition from recurrence to mortality.

It is also possible to obtain the outputs of ANOVA tests and the p-values of the tests for nonlinearity. For both outputs, the summaries show the chisquared statistics and the p-values. For ANOVA tests, the log-likelihood values for each parameter are also presented (Figures 17 and 18).

Transition probabilities

The transition probabilities are quantities of particularly interest since they allow for long-term predictions of the multi-state process. The MSM.app allows estimating these quantities using the Aalen-Johansen estimator [17], the landmark methods (LM) proposed by [18], as well as its presmoothed version (PLM) proposed by [19], and the landmark Aalen-Johansen [20]. Related references can also be seen in the papers by [21] and [22]. Categorical covariates can be included using all four of these methods by splitting the sample for each level of the covariate and repeating the described procedures for

MSM.App Introduction Input files Survival Analysis - IDM - Analysis - N	ISM - Analysis - Tests for the Markov condition -
Parametric models Type of distributions: exponentia velobul gaussia of logistic of logionormal of logiogistic of logitype of karno of logitype of logitype of karno of logitype of karno of logitype of logitype of karno of logitype of karno of logitype of l	Call: survreg(formula = formula.parametric, data = db, dist = input\$type.radio.parametric) Value Std: Error z p (Intercept) 2.03146 0.6999 2.90 0.0037 celltypelarge 0.79483 0.28302 2.81 0.0050 celltypesquamous 1.10491 0.26330 4.10 4.1e-05 karno 0.0336 0.06050 6.00 1.9e-09 age 0.00431 0.00895 0.48 0.6300 Scale fixed at 1 Exponential distribution Loglik(model)= -716.9 Loglik(intercept only)= -751.2 Chisq= 68.73 0.5 degrees of freedom, p= 1.9e-13 Number of Newton-Raphson Iterations: 5 n= 137
	AIC value of the model: [1] 1445.714

Figure 12: The outputs of the parametric models page with the results of the fitted model for the variables "celltype", "karno" and "age" of the *veteran* data set using the exponential distribution.

each subsample. Through the IPCW estimator proposed by [23] is also possible to estimate transition probabilities conditional on one single continuous covariate. Finally, the MSM.app also provides the estimation of transition probabilities conditional on several covariates through Breslow's method for estimating the baseline hazard function of the Cox models fitted marginally to each transition. The outputs for all these methods are identical. As an example, Figure 19 shows the estimates of the transition probabilities, using the landmark approach, from the initial single time s = 365 days to the next four years (730, 1095, 1460, and 1825 days). Results are presented combining the values of the corresponding times and transitions which are labeled from "00" to "12". For instance, "01" corresponds to the transition from the initial state (State 0) to intermediate state (Recurrence, State 1) and, in similar way, "12" the transition from the intermediate state (Recurrence, State 1) to death (State 2). Plots for each five transition probabilities are shown in Figure 20. $\hat{p}_{00}(s = 365, t)$ corresponds to the probability of a individual to occupy the initial state at time t conditional to be in the same state at time s = 365. In a similar way, $\hat{p}_{11}(s,t)$ represents the conditional probability of those individuals observed in State 1 at time s = 365 to remain in the intermediate state at a later time t. Plots for these transition probabilities report, respectively, survival fractions along times among the individuals that belong to initial state and the intermediate state at time s = 365, being represented

MSM.App Introduction Input files Survival Analysis • IDM - Analysis • M	ASM - Analysis • Tests for the Markov condition •
Parametric models Type of distributions: exponential @ weibuil @ gaussian @ logistic @ lognormal @ logiogistic select variables If tr v cetitype kamo diagtime g age prior	Call: survreg(formula = formula.parametric, data = db, dist = input\$type.radio.parametric) Value 5td. Error z p (Intercept) 2.08165 0.66538 3.13 0.0018 callype.mallcell 0.40514 0.2328 1.74 0.0819 callypesmallcell 0.40514 0.2328 1.74 0.0819 callypesmallcell 0.40514 0.2328 1.74 0.0819 callypesmallcell 0.40524 0.42 9.7-066 karno 0.02204 0.424 0.7-06 karno 0.02468 0.00422 0.59 0.6199 Log(scale) -0.0466 0.06593 -0.98 0.3267 Scale= 0.937 Weibull distribution Loglik(model)716.4 Loglik(intercept only)= -748.1 Chiqa 63.4 on 5 degrees of freedom, p= 2.4e-12 Number of Neuton-Raphson Iterations: 5 n = 137
	AIC value of the model: [1] 1446.786

Figure 13: The outputs of the parametric models page with the results of the fitted model for the variables "celltype", "karno" and "age" of the *veteran* data set using the weibull distribution.

by monotone non-increasing functions. Plots for $\hat{p}_{02}(s,t)$, report one minus the survival fraction along time, among the individuals in the initial state at time s. In this case, the plot is given by a monotone non-decreasing function. Finally, plots for $\hat{p}_{01}(s,t)$ allows for an inspection along time of the probability of being in State 1 for the individuals who belong to State 0 at time s. As expected the confidence bands become wider with greater lags times t-s. Finally, Figure 21 depict the estimates of the transition probabilities for individuals with covariate values ("rx", "sex", and "age") as Obs,1,48, respectively.

Cumulative Incidence Function (CIF)

The cumulative incidence of the illness (intermediate state) is another quantity of interest in IDM models [24]. This quantity denotes the probability of the individual or item being or having been in the intermediate 'diseased' state at some particular time t. It can be estimated conditional on a covariate, continuous or categorical. Figure 22 shows the estimates, and respective bounds of confidence, of the CIF conditional to "age" at 50 years for three specific times.

10. How to carry out a multi-state analysis for other models?

By clicking on the "MSM-analysis", we can extend some of the methods

SM.App Introduction Input files Survival Analysis - IDM - Analysis - N	ASM - Analysis + Tests for the Markov condition +
Parametric models Type of distributions: exponential webuil gaussian logistic lognormal logiogistic select variables trt celitype carmo diagtime cugage prior	Call: survreg(formula = formula.parametric, data = db, dist = input\$type.radio.parametric Value Std. Error z p (Intercept) 1.22436 0.60959 2.01 0.0447 celltypeIarge 0.75516 0.26034 2.90 0.0037 celltypesquanous 0.74433 0.26742 2.78 0.0634 karno 0.06353 0.06440 8.26 (2=16 age 0.06013 0.06078 0.93 0.3547 Log(scale) -0.54535 0.07414 -7.36 1.9=13 Scale= 0.50 Log logistic distribution Logik((model)= -712.2 Logik((intercept only)= -750.3 Chisg 76.2 on 5 degrees of freedom, p= 5.2e-15 Number of Interon.Teaptons: 4

Figure 14: The outputs of the parametric models page with the results of the fitted model for the variables "celltype", "karno" and "age" of the *veteran* data set using the loglogistic distribution.

MSM.App	Introduction	Input files	Survival Analysis +	IDM - Analysis -	MSP	M - Analysis 👻	Tests for the Markov co	ndition +			
	f event: Proportion					Illness	(state 0): e state (state 1): ate (state 2):		Healthy Illness Death	Illness Death 0.504 0.041 0.058 0.446 0.000 0.487	
						Death					

Figure 15: The number of transitions among the three states of the *colonIDM* data set.

addressed for the IDM models to more complex multi-state models (MSM) with more than three states and possible reversible transitions.

Number of events

Figure 23 shows the movement of the individuals among the six states of the multi-state model represented by the data set of the European Group for Blood and Marrow Transplantation (*ebmt4* data set).

Regression models

Figure 24 shows the results of the Cox regression model for the transition $1 \rightarrow 2$ which includes the covariates "year", "age" and "proph". As we can see, in terms of interpretation, the output is quite similar to those of IDM models but also shows the global significance of the model through different

tegression models ype of model* Markovian © semi-Markovian ee related menu in Introduction tab	Initial state (state 0): Healthy Intermediate state (state 1): Illness Utilimate state (state 2): Death	Cox Harkov Hodeli transition 0 > 1 cof exp(cof) how 0.5% upper 0.5% Pr()[2]) rx(ev454 -0.49944020 0.606376 0.4809022 0.7670442 2.922274-05 rx0bs 0.00972208 1.0005101 0.8185304 1.2457029 9.274051-01 tex -0.10358020 0.015400 0.781051 1.015102 0.44580-01 age -0.01557055 0.914476 0.0480811 1.0101406 1.575420-01 obtrvt-0.108157077 1.707220 0.0450811 1.511051 1.01554-01 perfor 0.25990015 1.540855 0.245533 2.245420 2.300714-01
ype of Regression		Cox Markov Model: transition 0 -> 2 coef $exp(coef)$ lower 0.95 upper 0.95 $Pr(> z)$
Cox models O ANOVA O PH		rxLev+5FU 0.19928575 1.220531e+00 0.5462840 2.726961 6.270584e-0 rxObs 0.31675801 1.372670e+00 0.6015221 3.132427 4.517620e-0
rx		sex 0.49427936 1.639316e+00 0.8466932 3.173946 1.425704e-0 age 0.08625773 1.090087e+00 1.0493666 1.132388 8.966451e-0
sex		obstruct 0.72410896 2.062892e+00 0.9656328 4.406979 6.153014e-0
		perfor -16.10675048 1.011410e-07 0.0000000 Inf 9.959598e-0
age obstruct perfor		Cox Markov Model: transition 1 -> 2
adhere		coef exp(coef) lower 0.95 upper 0.95 Pr(> z)
nodes		rxLev+5FU 0.282533676 1.3264864 1.0311061 1.706484 0.02792756
differ		rxObs -0.060556027 0.9412410 0.7516577 1.178641 0.59771279 sex 0.160110302 1.1736403 0.9645093 1.428117 0.10980608
extent		sex 0.100110302 1.1736403 0.9645093 1.428117 0.10900008 age 0.009482975 1.0095281 1.0016386 1.017480 0.01783758
surg		obstruct 0.217866731 1.2434213 0.9737161 1.587831 0.08073275

Figure 16: Results of the application of the Cox PH model with the following covariates: "rx", "sex", "age", "obstruct" and "perfor". Results for each of the three transition intensities of the *colonIDM*. A Markovian process is assumed.

tests.

Transition probabilities

The steps for obtaining the transition probabilities are identical to those used for the illness-death model. In Figure 25, the output shows all the estimates from the indicated start and last states of the transition probabilities as well as the corresponding confidence intervals. A plot with the transition probability can also be presented for each transition.

11. How to check the Markov condition?

Traditionally the Markov assumption is checked by including covariates depending on the history through a proportional hazards model. Since the landmark methods of the transition probabilities are free of the Markov assumption, they can also be used to introduce such tests by measuring their discrepancy to Markovian (AJ) estimators.

The MSM. app web application offers two types of tests for checking this assumption using recent literature methods: (i) local tests, which are obtained by fixing a specific time value, s and are particularly useful for estimating transition probabilities; and (ii) global tests, which may be preferable for

	Initial state (state 0):	Cox Markov Model: transition 0 -> 1
Regression models	Healthy Intermediate state (state 1): Illness	Analysis of Deviance Table Cox model: response is Surv(mydata[, 1], s01) Terms added sequentially (first to last)
Type of model*	Ultimate state (state 2):	loglik Chisq Df Pr(> Chi) NULL -3040.3
Markovian semi-Markovian See related menu in Introduction tab	Death	NULL -3040.3 rx -3028.1 24.3435 2 5.175e-06 *** sex -3027.4 1.3212 1 0.25038
Type of Regression		age -3026.2 2.5766 1 0.10845 obstruct -3024.6 3.1202 1 0.07733 . perfor -3023.9 1.3170 1 0.25112
Cox models ANOVA PH		Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '
/ariables to include:		Cox Markov Model: transition 0 -> 2
i sex 1 age		Analysis of Deviance Table Cox model: response is Surv(mydata[, 1], s02)
e obstruct		Terms added sequentially (first to last)
e perfor		loglik Chisq Df Pr(> Chi) NULL -232.42
nodes		<pre>rx -232.18 0.4783 2 0.7873 sex -231.62 1.1270 1 0.2884 ae -219.81 23.6075 1 1.181e-06 ***</pre>
extent		obstruct -218.52 2.5803 1 0.1082 perfor -217.58 1.8865 1 0.1696
node4		Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '

Figure 17: ANOVA results for transitions $0 \rightarrow 1$ and $0 \rightarrow 2$ with the following covariates: "rx", "sex", "age", "obstruct" and "perfor" of the *colonIDM*. A Markovian process is assumed.

regression purposes.

Local tests

Two types of methods are available for checking the local tests of the Markov condition: (i) the AUC method, which is based on measuring the discrepancy between the AJ estimator of the transition probabilities (which provides consistent estimates when the process is Markovian) and the landmark estimators (which are free of the Markov condition). In this case, the web tool uses the LM estimator for the progressive illness-death models and LMAJ in the case of more complex MSM models; (ii) the Log-rank method, which considers summaries from families of log-rank statistics where patients are grouped by the state occupied at different times. For both types of models (IDM and MSM), the output using the AUC is the same. As an example, in Figure 26, we obtained the p-values for the local tests based on the s times 365, 730, 1095, 1460, and 1825 for the transition from state 2 to state 3. Results were obtained for an IDM model based on the colon cancer data using 100 replicas. Even though the web form is quite similar to the AUC local test, the output of the log-rank test only provides the results for a specific transition and the times chosen (Figure 27).

MSM.App Introduction Input files Survival Analysis - IDM - Analysis - I	MSM - Analysis - Tests for the Markov condition -
Regression models Type of model* Markovian @ semi-Markovian "see related menu in Introduction tab Type of Regression Cox models @ ANOVA @ PH Variables to include: If nc If inc	$ \begin{array}{llllllllllllllllllllllllllllllllllll$

Figure 18: The proportional hazards assumption was tested for the transitions $0 \rightarrow 1$ and $0 \rightarrow 2$ with the following covariates: "rx", "sex", "age", "obstruct" and "perfor" of the *colonIDM*. A Markovian process is assumed.

Global tests

In the MSM. app application, three global tests are available for both IMD and more complex MSM models: (i) the first one is based on Cox models, from which it is possible to evaluate the effect of history on the process. In this case, this can be done by checking the significance of the covariate time until entering the first state of a particular transition. As illustrated in Figure 28, we can conclude that there is no effect of the time spent in the initial state on the transition $2 \rightarrow 3$ (p-value = 0.1543195), which does not induce the failure of the Markovianity. (ii) the recent proposed global test propose [25], based on the area under the curves (AUC), can be used. This test is based on the (AUC) local test results for specific percentiles. The outputs on the right hand of Figure 29 shows the proportion of rejections of the test for all possible transitions between state 1 and state 5. (iii) it is also possible to use the global test based on the log-rank statistics [26] throughout the similar steps of the previous methods, after selecting Log-rank in the radio button HTML element. The outputs provide only the results of the tests for each transition (Figure 30).

Transition probabilities Pij(s,t) s (Single time value) 365 Times (e.g. 400,600): 730,1095,1460,1825 Type of methods: © Nonsparametric © One single categonical covariate © One single continuous covariate © More than one covariate More than one covariate Morparametric:	00: transition 0-50 01: Estimation of pij(s=365,t) 01: to 00: 00: 00: 00: 00: 00: 00: 00: 00: 00
Interval of confidence: No • Yes View plots & Download par	

Figure 19: Results of the estimates of the transition probabilities and the corresponding confidence intervals for s = 365 and times 730, 1090, 1460, and 1825 days for the *colonIDM* data set using the landmark estimators.

Acknowledgements

This research was financed within the research grants PTDC/MAT-STA /28248/2017 and PD/BD/142887/2018.

References

- Wojciechowski, J., Hopkins, A. and Upton, R. Interactive pharmacometric applications using R and the Shiny package. CPT: pharmacometrics & systems pharmacology, 4(3), 146–159; 2015
- [2] Chang, W. Web Application Framework for R, CRAN; 2017
- [3] Kaushik, S. Creating Interactive data visualization using Shiny App in R (with examples), Analytics Vidhya; 2016
- [4] Putter, H., Fiocco, M. and Geskus, R.B. Tutorial in biostatistics: Competing risks and multi-state models, Statistics in Medicine, 26(11), 2389– 2430.

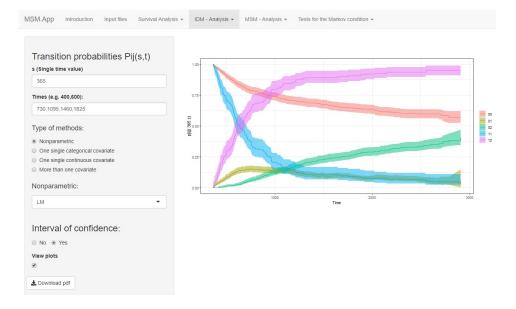


Figure 20: Transition probability estimates with confidence intervals for each transition for s = 365 using the landmark estimators.

- [5] Meira-Machado, L., de Uña-Álvarez, J. and Cadarso-Suárez, C. and Andersen, P.K. Multi-state models for the analysis of time to event data. Statistical Methods in Medical Research, 18, 195–222; 2009.
- [6] Meira-Machado, L. and Sestelo, M. Estimation in the progressive illnessdeath model: A nonexhaustive review. Biometrical Journal, 61 (2), 245– 263; 2019.
- [7] Therneau, T.M. A Package for Survival Analysis in R. https://CRAN.Rproject.org/ package=survival; 2021.
- [8] Putter, H., de Wreede, L.C., Fiocco, M., Geskus, R.B., Bonneville, E.F., Manevski, D. mstate: Data Preparation, Estimation and Prediction in Multi-State Models. The R Journal; 2020.
- [9] Soutinho, G., Sestelo, M. and Meira-Machado, L. survidm: An R package for Inference and Prediction in an Illness-Death Model. The R Journal, *Accept for publication*; 2021.
- [10] Govan, P. eAnalytics: Dynamic Web-based Analytics for the Energy Industry. Journal of Open Research Software, 4: e45; 2016. DOI: http://dx.doi.org/10.5334/jors.144

MSM.App Introduction Input files Survival Analysis - IDM - Analysis -	MSM - Analysis - Tests for the Ma	arkov condition 👻
Transition probabilities Pij(s,t) s (Single time value) 365 Times (e.g. 400,600): 730,1095,1460,1825 Type of methods: Nonparametric One single continuous covariate More than one covariate More than one covariate Breslow method: Variables to include in the model: Variables to include in the model:	00: transition 0->1 02: transition 0->1 02: transition 0->2 11: transition 1->1 12: transition 1->2	Estimation of pij(s-365,t) t 00 1 02 11 12 700 0.70775 0.176605 0.6025325 0.4355623 0.650438 1055 0.6723394 0.1743610 0.1532951 0.1267067 0.6732950 1060 0.6302575 0.120470 0.47964637 0.0693725 0.050470 1025 0.5882736 0.1133271 0.29859931 0.06930107 0.9306990 2.5% t 00 0 1 02 11 12 730 0.7057259 0.1377642 0.0325186 0.24357470 0.5346002 1055 0.6094602 0.1371243 0.1082248 0.6259318 0.747705 1060 0.500528 0.032518 0.2458747 0.5346002 1055 0.6094602 0.1371243 0.1082248 0.6259318 0.747705 1060 0.500530 0.280525 0.082512 0.353518 0.625650 105 0.025666 0.2364646 0.65681224 0.551237 0.855900 105 0.2266648 0.324646 0.6658124 0.5512360 0.776274 170 0.0206648 0.324646 0.6686124 0.551240 0.750274 170 0.0206648 0.324646 0.6685124 0.5513830 0.762274 170 0.0206648 0.324646 0.6585124 0.551380 0.762274 170 0.0206648 0.324646 0.6585124 0.551340 0.750274 170 0.0206648 0.324646 0.6585124 0.551380 0.750274 170 0.0206648 0.324646 0.6585124 0.551380 0.750274 170 0.0206648 0.324646 0.6585124 0.551340 0.750274 170 0.0206648 0.324646 0.6585124 0.551340 0.750274 170 0.0206648 0.324646 0.551340 0.750274 170 0.0206648 0.324646 0.6585124 0.551340 0.750274 170 0.0206648 0.324626 0.555554 0.55554 0.55554 0.750765 0.750000 0.55555 160 0.710552 0.100219 0.113676 0.710900 0.55555 160 0.710552 0.100219 0.113676 0.710900 0.55555 160 0.710552 0.100219 0.113676 0.710900 0.55555 160 0.710552 0.100219 0.555554 0.55555
extent surg node4 Value(s) Obs.1.48 Interval of confidence: No ® Yes View plots ■ Δownload pdf		

Figure 21: Results of the estimates of the transition probabilities and the corresponding confidence intervals for s = 365 and times 730, 1090, 1460, and 1825 days for the *colonIDM* data set using the Breslow estimator.

- [11] Seal, A., Wild, D.J. Netpredictor: R and Shiny package to perform Drug-Target Bipartite network analysis and prediction of missing links. Cold Spring Harbor Laboratory; 2016. doi: https://doi.org/10.1101/080036
- [12] Lacy, S. msm-shiny, https://stulacy.shinyapps.io/msm-shiny/; 2021
- [13] Skourlis, N., Crowther, M.J., Andersson, T., Lambert, P.C. MSMplus, https://nskbiostatistics.shinyapps.io/MSMplus/; 2021
- [14] Miller, R.G. Survival Analysis. John Wiley & Sons; 1997
- [15] Moertel, G., Fleming, T.R., Macdonald, J.S., Haller, D.G., Laurie, J.A., Goodman, P.J., Ungerleider, J.S., Emerson, W.A., Tormey, D.C., Glick, J.H., Veeder, M.H. and Mailliard, J.A., Levamisole and fluorouracil for adjuvant therapy of resected colon carcinoma. New England Journal of Medicine, 322(6):352–358; 1990
- [16] Cox, D.R. Regression models and life tables. Journal of the Royal Statistical Society Series B, 34, 187–220; 1972.

Cumulative Incidence Function	Estimation of CIF(t) t CIF
	365 0.2213297
Times	730 0.3970167
365, 730, 1090	1090 0.4278694
	2.5%
nterval of confidence:	t CIF
	365 0.1539464
No 💿 Yes	730 0.3223258
	1090 0.3509827
Select variable	97.5%
age	t CIF
	365 0.2857126
/alue	730 0.4694265
	1090 0.4979088
50	
/iew plots	

Figure 22: Cumulative recurrence incidence with 95% bootstrap confidence intervals. Data from a colon cancer study.

- [17] Aalen, O., Johansen, S. An empirical transition matrix for non homogeneous Markov and chains based on censored observations. Scandinavian Journal of Statistics, 5: 141–150; 1978.
- [18] de Uña-Álvarez, J. and Meira-Machado, L. Nonparametric estimation of transition probabilities in the non-markov illness-death model: A comparative study. Biometrics, 71(2), 364–375; 2015.
- [19] Meira-Machado, L. Smoothed landmark estimators of the transition probabilities. SORT-Statistics and Operations Research Transactions,

Type of event: Tx Rec AE Rec+AE Rel Death no event total entering Tx 0 755 907 0 95 160 332 2279 ® Count Proportion Rec 0 0 227 112 39 407 785 Rec 0 0 0 212 112 39 407 785 Rec 0 0 0 123 56 112 397 416 660 Rec+AE 0 0 0 137 416 660 8e1 0 0 0 137 56 570	MSM.App	Introduction	Input files	Survival Analysis +	IDM - Analysis 👻	MSM -	Analysis 🔻	Te	ests fo	r the I	vlarkov o	condi	ion 👻			
							from Tx Rec AE	Tx 0 0 0	785 0 0 0	907 0 0 0	0 227 433 0	95 112 56 107	160 39 197	332 407 221	2279 785 907	

Figure 23: The number of transitions among the states of the ebmt4 data set.

M.App Introduction Input files Survival Analysis - IDM - Analysis	MSM - Analysis
Regression models - MSM /ariables to include: 8 year 8 agecl 9 proph 9 match 1 match	Call: coxph(formula = formula, data = msebmt) n= 15512, number of events= 3255 coef exp(coef) se(coef) z Pr(> z) year1.1 0.40377 1.49745 0.10015 4.032 5.54e-05 *** year2.1 0.52710 1.69401 0.10300 5.114 3.16e-07 *** agecl1.1 0.03651 1.23320 0.10342 2.047 0.0407 * agecl1.2 0.0565 1.05590 0.08857 0.063 0.5662 proph.1 -0.36967 0.69996 0.09292 -3.578 6.94e-05 ***
	Signif. codes: 0 (**** 0.001 (*** 0.01 (** 0.05 (.) 0.1 ())
1->2	<pre>exp(coef) exp(-coef) lower .95 year1.1 1.497 0.6678 1.2306 1.822</pre>
	year2.1 1.694 0.5903 1.3841 2.073
	agecl1.1 1.233 0.8109 1.0089 1.507
	agecl2.1 1.055 0.9478 0.8864 1.256
	proph.1 0.691 1.4473 0.5759 0.829
	Concordance= 0.525 (se = 0.003)
	Likelihood ratio test= 72.39 on 5 df, p=3e-14
	Wald test = 66.1 on 5 df, p=7e-13 Score (logrank) test = 67.93 on 5 df, p=3e-13

Figure 24: The output of the Cox regression model for the transition $1 \rightarrow 2$ that include the 'year", "age" and "proph" covariates. *ebmt4* data set.

ISM.App Introduction Input files Survival Analysis - IDM - Analysis	s + MSM - Analysis + Tests for the Markov condition +
Transition probabilities - MSM s (single time value)	1 1.01 2.02 2.04 2.05 267 70 80.06027 20.06027 20.06027 111 1005 8.00035715 0 0.0506237 123 1406 8.00035715 0 0.05062459 123 1406 8.00035715 0 0.0506459 142 1825 8.00035770 0 0.0506459
365	2.5%
Times (e.g. 400,600):	$\begin{array}{cccccccccccccccccccccccccccccccccccc$
730,1095,1460,1825	3 1460 0.8164799 0 0 0 0.000846 4 1825 0.806689 0 0 0.0049983
Type of methods:	97.5%
AJ	t 1->1 1->2 1->3 1->4 1->5 1 730 1 0.1090932 0.007560 0 0.0054029
Interval of confidence:	2 1095 0.9755765 0.039842 0.0885558 0 0.0534534 3 1440 0.9560579 0.0412021 0.086505 0 0.0468646 4 1225 0.9564469 0.0424291 0.086519 0 0.0468983
⊙ No ⊛ Yes	
From (first state).	1>2
	•
	0.09
To (last state):	
5	•
View plots	4
8	0.00-
Transitions for plots:	
1->2	•
	Ativar o Win
Lownload pdf	Times Aceda a Denniço

Figure 25: Estimates of all possible transition probabilities from the state 1 to 5, for s = 365 and times equal to 730, 1095, 1460, and 1825 using the AJ estimators. *ebmt4* data set.

VISM.App Introduction Input files Survival Analysis - IDM - Analysis -	MSM - Analysis + Tests for the Markov condition +
Local tests Type of data: Iliness-death model (IDM) Other multi-state model (MSM)	5 2->2 2->3 1 365 0.003706224 0.0049639 2 730 0.123566948 0.1505162 3 1095 0.171679220 0.1736436 4 1460 0.5933134995 0.6029005 5 1825 0.245509888 0.4472789
Type of method: • Area Under the two curves (AUC)* Log-rank	
*Computationally demanding. Times**	
365,730,1095,1460,1825 **At least one time is necessary.	
From:	
2 •	
3	
***Provide results of the local tests for all transitions and times. Replicas	
100	
View	

Figure 26: Results of the local test for the illness-death model using the colon cancer data set, for s = 365, 730, 1095, 1460, and 1825 days, using the AUC test, from state 1 to state 3.

40: 375–398; 2016.

- [20] Putter, H. and Spitoni, C. Non-parametric estimation of transition probabilities in non-markov multistate models: The landmark aalenjohansen estimator. Statistical Methods in Medical Research, 27, 2081– 2092; 2018.
- [21] Moreira, A., de Uña-Álvarez, J. and Meira-Machado, L. Presmoothing the aalen-johansen estimator in the illness-death model. Electronical Journal of Statistics; 2013, 7:1491–1516. DOI: https://doi.org/10.1214/13-EJS816.
- [22] Araújo A, Meira-Machado L, Roca-Pardiñas J (2014). "TPmsm: Estimation of the Transition Probabilities in 3-State Models." Journal of Statistical Software, 62(4), 1–29. doi: 10.18637/jss.v062.i04.

	[1]	0.036 0.685 0.509
Local tests		
Type of data:		
Illness-death model (IDM) O Other multi-state model (MSM)		
Type of method:		
 Area Under the two curves (AUC)* 		
ø Log-rank		
*Computationally demanding.		
Times**		
365, 730, 1460		
**At least one time is necessary.		
From:		
2	•	
To***:		
3	•	
***Provide results of the local test for the specific transition and times.		
Replicas		
1000		

Figure 27: Results of the local test for the illness-death model using the colon cancer data set, for s = 365, 730 and 1460 days, using the Log-rank test, for state 2 to state 3.

- [23] Meira-Machado, L., de Uña-Álvarez, J. and Datta, S. Nonparametric estimation of conditional transition probabilities in a non-Markov illness-death model. Computational Statistics, 2015, 30(2), 377–397. doi: https://doi.org/10.1007/s00180-014-0538-6
- [24] Kalbfleisch, J. D. and Prentice R. L. The statistical analysis of failure time data. John Wiley & Sons; 1980.
- [25] Soutinho, G. and Meira-Machado, L. Methods for checking the Markov condition in multi-state survival data. Comput Stat; 2021. https://doi.org/10.1007/s00180-021-01139-7
- [26] Titman, A. and Putter, H. General tests of the Markov property in multi-state models. Biostatistics; 2020. doi:10.1093/biostatistics/kxaa030.

MSM.App	Introduction	Input files	Survival Analysis •	IDM - Analysis 👻	MSM - An	nalysis 👻	Tests for the Markov condition -
Type of n Dependi Area Uni Log-rank 	ata: eath model (IDM) nethod: ng on History (thr der the two curves	ough CPHM) s (AUC)*	ulti-state model (MSM)		ľ	1] 0.1543:	195

Figure 28: Results for this global test given by the Cox PH model to our data indicated that the effect of the time spent in State 1 is not significant (p-value of 0.154), revealing no evidence against the Markov model for the colon data set.

Global tests	1->1 1->2 1->3 1->4 1-> 1 0.08650676 0.01105257 0.02312809 0.152981 0.000802568
Type of data:	
Illness-death model (IDM) Other multi-state model (MSM) 	
Type of method:	
Depending on History (through CPHM)	
 Area Under the two curves (AUC)* Log-rank 	
Computationally demanding.	
From:	
1	•
To***:	
5	•
***Provide results of the global tests for all transitions.	
Replicas	
100	

Figure 29: Outputs of the global test for the illness-death model based on the ebmt data set using the AUC test, from state 1 to state 5. Results for the AUC local test are also shown.

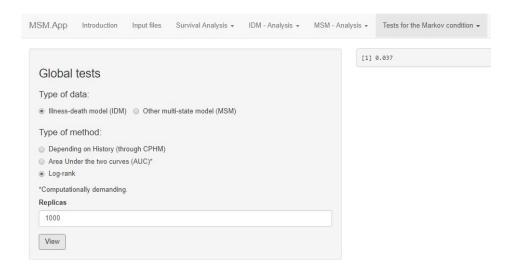


Figure 30: Results of the global test for the illness-death model using the colon cancer data set using the Log-rank test.