Development of a dynamic interactive web tool to enhance understanding of multi-state model analyses: MSMplus

Nikolaos Skourlis (✉ nikolaos.skourlis@ki.se)

Karolinska Institutet <https://orcid.org/0000-0002-5003-9611>

Michael J. Crowther

Karolinska Institutet

Therese M-L. Andersson

Karolinska Institutet

Paul C. Lambert

University of Leicester Department of Health Sciences

Software

Keywords: Web tool, Multi-state models, Interactive, Rshiny

Posted Date: April 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-382863/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at BMC Medical Research Methodology on November 27th, 2021. See the published version at <https://doi.org/10.1186/s12874-021-01420-9>.

RESEARCH

Development of a dynamic interactive web tool to enhance understanding of multi-state model analyses: MSMplus

Nikolaos Skourlis^{1*}, Michael J. Crowther¹, Therese M-L Andersson¹ and Paul C Lambert^{1,2}

*Correspondence:

nikolaos.skourlis@ki.se

¹ Department of Medical

Epidemiology and Biostatistics,
Karolinska Institutet, Nobels Väg
12A, Stockholm, Sweden

Full list of author information is
available at the end of the article

Abstract

Background: Multi-state models are used in complex disease pathways to describe a process where an individual moves from one state to the next, taking into account competing states during each transition. In a multi-state setting, there are various measures to be estimated that are of great epidemiological importance. However, increased complexity of the multi-state setting and predictions over time for individuals with different covariate patterns may lead to increased difficulty in communicating the estimated measures. The need for easy and meaningful communication of the analysis results motivated the development of a web tool to address these issues.

Results: MSMplus is a publicly available web tool, developed in RShiny, with the aim of enhancing the understanding of multi-state model analyses results. The results from any multi-state model analysis are uploaded to the application in a pre-specified format. Through a variety of user-tailored interactive graphs, the application contributes to an improvement in communication, reporting and interpretation of multi-state analysis results as well as comparison between different approaches. The predicted measures that can be supported by MSMplus include, among others, the transition probabilities, the transition intensity rates, the length of stay in each state, the probability of ever visiting a state and user defined measures. Representation of differences, ratios and confidence intervals of the aforementioned measures are also supported. MSMplus is a useful tool that enhances communication and understanding of multi-state model analyses results.

Conclusions: Further use and development of web tools should be encouraged in the future as a means to communicate scientific research.

Keywords: Web tool; Multi-state models; Interactive; Rshiny

¹Background

²Multi-state models are used in a variety of epidemiological settings, enabling the²
³study of individuals transitioning through different disease states. Studying acute [1]³
⁴or chronic disease progression [2], recurrent events such as repeated hospitalizations⁴
⁵[3] typical examples of multi-state models use. When studying such processes, multi-⁵
⁶state models are used both in the epidemiological area [2, 3] as well as in health⁶
⁷economics [4], in order to portray accurately, with sufficient complexity the real-⁷
⁸world issue under study and provide useful and meaningful predictions.⁸

⁹ Various measures are of interest in a multi-state process with transition prob-⁹
¹⁰abilities, transition specific covariate effects and transition intensities being very¹⁰
¹¹important in understanding the process under study. Other measures facilitate a¹¹
¹²more in-depth understanding of the multi-state process (Table 1), depending on¹²
¹³the models used and their assumptions. For more information on the modelling¹³
¹⁴approaches and assumptions in the multistate models framework, see Putter et al¹⁴
¹⁵[5], Cook [6], Andersen [7] and Hougaard [8].¹⁵

¹⁶ In contrast with standard survival analyses settings where a single estimate may¹⁶
¹⁷suffice (e.g a constant hazard ratio), in a multi-state process most of the measures¹⁷
¹⁸of interest are functions of time. Therefore, graphical displays of these changes are¹⁸
¹⁹more suitable to display the time-dependent processes. For a certain transition, the¹⁹
²⁰effects of a covariate of interest on the hazard rate/transition intensity rate can be²⁰
²¹well defined and estimated. However, as the covariate of interest may have different²¹
²²effects on different transitions and due to accounting for competing states while the²²
²³process moves from one state to the next, the overall effect of the covariate cannot²³
²⁴be well defined. In this case, the graphical display of the estimated measures (e.g²⁴
²⁵probability of transitioning to a specific state) over the levels of the covariate of²⁵
²⁶interest across time can provide a visual assessment of the overall effects.²⁶

²⁷ The structural complexity of multi-state settings, in combination with the numer-²⁷
²⁸ous time-varying measures estimated during the multi-state analysis, and the need²⁸
²⁹for comparison of predictions for different individuals can make the communication²⁹
³⁰and interpretation of the results quite challenging. This challenge motivated the³⁰
³¹development of a tool, MSMplus that can present the results in an easy, compre-³¹
³²hensible and meaningful fashion, aiding the user to present the research findings to³²
³³the targeted audience.³³

¹ A systematic review from Trevena et al [9] on communicating evidence with pa-
²tients, supports communication tools that, among other traits, are interactive, using
³illustrations and graphs to convey the evidence, and are user-tailored. MSMplus has
⁴these traits, making it easy to use by researchers and health care professionals and
⁵powerful in communicating multi-state analysis results to the targeted audiences.
⁶In addition, the produced plots and graphs can be stored and used for potential
⁷publications. Examples of other successful web tools are InterPreT [10], INTEREST
⁸[11] and MetaDTA [12].

⁹ In the remainder of this paper, section 2 describes the multi-state measures that
¹⁰are supported by MSMplus, information about the development of the tool and how
¹¹multi-state analysis results should be uploaded to the app. In Section 3, we use a
¹²European Blood and Marrow Transplant registry (EBMT) dataset to illustrate the
¹³use of MSMplus via the presentation and interpretation of graphs created automat-
¹⁴ically by the app. In sections 4 and 5, we discuss the overall usefulness of MSMplus,
¹⁵reflect on its strengths and limitations and discuss potential future extensions.

¹⁹Implementation ¹⁹

²⁰ Multi-state model analysis can be performed via various statistical software. A re-
²¹view of software for multi-state models was published in 2014 [13], but new packages
²²have been developed since then. The MSMplus tool can use the results from any
²³multi-state analysis as long as the output is stored in a specific format and struc-
²⁴ture. It reads summary information and statistical predictions from a multistate
²⁵analysis without the need of uploading research data to the application. [MSMplus](https://nskbstatistics.shinyapps.io/MSMplus)
²⁶(<https://nskbstatistics.shinyapps.io/MSMplus>) was developed in Rshiny [14], a
²⁷branch of R specifically dedicated for building interactive applications. As a post-
²⁸analysis tool, MSMplus visualizes the results of the multi-state model analysis via
²⁹interactive graphs of predictions for different covariate patterns of interest. To il-
³⁰lustrate the MSMplus application we will use data from the European Blood and
³¹Marrow Transplant registry (EBMT), which are commonly used as an example
³²multi-state model dataset [5, 15] consisting of 2.204 patients who received bone
³³marrow transplantation and are available from the mstate package in R.

Table 1: Measures estimated from multi-state models that are supported for visualization from the application tool MSMplus.

Measure	Description	Tab inside the app
Transition probability	The probability of occupying state i at time t if in state j at times s	Probabilities
Transition intensity rate	The hazard rates of transitioning from state j to state i	Hazard
Hazard Ratios	Hazard ratios between two covariate patterns of experiencing a transition	Hazard
Transition intensity rate ratios	Ratio between the different transition intensity rates for a certain covariate pattern. The aim is to compare the rates between the different transitions.	Hazard
Restricted expected length of stay	Total length of stay of an individual in each state over time.	Length of state
Probability of ever visiting a state	Probability of ever visiting state i by time t	Visit
Expected number of visits	Expected number of visits in state i	Extra
Expected first passage times	Expected time for which the process will enter state i for the first time.	Extra
Mean sojourn times	Mean sojourn time in each transient state i for each covariate pattern of interest.	Extra
Probability that each state is next	Probability that each state j is the state that follows present state i	Extra

Multi-state model measures

There are various measures that can be estimated from a multi-state model analyses. The most common are the transition probabilities and the covariate effects in the form of hazard ratios. Many more measures can be estimated depending on the models and assumptions used. Table 1 gives a short description of the measures supported for visualization by the MSMplus tool. Most of the measures mentioned in the table are estimated as a function of time for the covariate patterns of interest. Detailed definitions and interpretation examples of the measures estimated in multi-state models can be found at the "Interpretation" tab of the application.

Tool development

Rshiny

RShiny [14] is a branch of R used for creating interactive applications either locally (user's server) or applications hosted on web platforms (shinyapps.io). At the core of each Rshiny application lie two parts: a user interface part and a server function part. The code is comprised of "reactive" expressions that react to the input of the user (e.g reactive values, datasets, and plots). Cascading style sheets (CSS) are also used supplementary in Rshiny, by calling a specific CSS theme from Bootswatch.com

¹but specific app design features can be furtherly adjusted within the Rshiny code¹

²with the use of HTML `<style>` Tag element. ²

³ ³

⁴*Interactive figures and graphs* ⁴

⁵*Summary information:* The user has the freedom to fully adjust the multi-state⁵
⁶model diagram attributes such as the coordinates, heights and widths of the MSM⁶
⁷box-states, colors and text size. The number of individuals being in each state⁷
⁸and the number of those that have undergone each transition up to a certain time⁸
⁹point since the beginning of the process are dynamically depicted via a slider that⁹
¹⁰continuously updates the frequencies on the graph. ¹⁰

¹¹*Estimated measures-Analysis results:* MSMplus reads and visualizes the estimated¹¹
¹²measures across the time points of prediction for the different covariate patterns of¹²
¹³interest. The plots and graphs are dynamic, derived using the plotly R function,¹³
¹⁴with options for smooth changes between different covariate patterns on the same¹⁴
¹⁵graph or a grid plot of the different covariate patterns. ¹⁵

¹⁶*User-tailored plot attributes:* The ranges of the plot axes can be manually adjusted¹⁶
¹⁷via the options provided or by just clicking and dragging the axes. The number of¹⁷
¹⁸ticks on each axis, the label and legend sizes can also be specified by the user via¹⁸
¹⁹provided options. In addition, due to the plotly attributes, features can easily be¹⁹
²⁰selected or deselected when specific comparisons are of interested, by simply clicking²⁰
²¹on them and all the plots can be downloaded as png files. There are further options²¹
²²for depiction of confidence intervals, provided that they are supplied by the user. A²²
²³log scale option is also available when studying the transition intensities. ²³

²⁴*Deriving the input files* ²⁴

²⁵MSMplus requires two json or csv files as input. The first file contains structural²⁵
²⁶and descriptive information of the multi-state model diagram while the second file²⁶
²⁷contains the various predictions from the multistate model. The json files are auto-²⁷
²⁸atically derived through functions developed in Stata and R. The csv files have²⁸
²⁹to be prepared by the user but allow for the use of multi-state analysis results²⁹
³⁰originating from any statistical software and package. ³⁰

³¹*Summary information file:* If the json input options has been selected, the first³¹
³²json file to be uploaded should include states and transitions names, initial coordi-³²
³³nates, widths and heights for the multi-state model diagram boxes and additional³³

information for plotting the structure of the multi state model. It can also include¹
frequency of people in each state across time and frequency of people that have²
made each transition across time. This file can easily be produced in Stata via run-³
ning *msboxes* command by specifying options interactive and jsonpath (Additional⁴
file 1) or in R via function *msboxesR* (provided in Additional file 3). If the csv input⁵
options has been selected, the user has to specify the number of states and the tran-⁶
sition matrix on the app platform. The multi-state model box traits (coordinates,⁷
width, and height) take an initial default value that can be adjusted by the user⁸
in a subsequent tab. There is no need to upload a csv summary information file in⁹
this case. Optionally, a csv file with information about the frequency of individuals¹⁰
in each state and individuals that have made each transition across time can be¹¹
uploaded.¹²

13

13

14

14

¹⁵ *Analysis results file*: The analysis results file provides predictions from the multi-¹⁵
state model analysis to the application. It consists of some mandatory elements¹⁶
while the rest are optional, and depend on the analysis. The time points and a¹⁷
list of the covariate patterns for which the predictions were made is mandatory.¹⁸
The rest of the estimates are optional, meaning that the application will be able¹⁹
to run, visualizing only the information that it is being given. The user can supply²⁰
all the estimated measures from Table 1 over the time points of prediction and²¹
the different covariate patterns. Differences and ratios of the estimated measures²²
as well as confidence intervals are also supported by MSMplus. If the json input²³
options has been selected, the json file to be uploaded can be derived via running²⁴
predictms command in Stata by specifying options interactive and jsonpath, or in²⁵
R via manually developed functions compatible with packages *msm*, *flexsurv* and²⁶
mstate (Additional file 4). If the csv input option has been selected, the csv file to²⁷
be uploaded has to be made manually by the user. The csv file should have a dataset²⁸
structure and the variable names should follow specific naming rules (Additional²⁹
file 5). If the aim is to compare two different multistate modeling approaches then³⁰
the user will be asked to provide two files with prediction results (csv or json).³¹
Instructions and examples on how to build the csv results file is given in Additional³²
file 5 as well as in a [tutorial](#) within MSMplus.³³

Results

The European Blood and Marrow Transplant registry (EBMT) example consists of 2204 patients who received bone marrow transplantation. The three states a patient can be in is 1) Post-transplant, 2) Platelet recovery 3) Relapse or Death as seen in Figure 1. A post-transplant patient can experience either platelet recovery (transition 1) or Relapse/Death (transition 2) while those with platelet recovery can experience Relapse/Death (transition 3). The covariate patterns used in this example are the 3 age categories of the patients, namely < 20 years old, 20-40 years old or > 40 years old at the time of the transplantation. We fit a separate flexible parametric survival model on the log hazard scale with 4 degrees of freedom for the baseline hazard (< 20 years old) for each transition. Then, the estimated parameters from the transition-specific models are fed to command *predictms* in Stata in order to perform the multi-state model analysis with Markov assumptions. We derive estimations for the measures of interest for a period of up to 10 years after the transplantation for time intervals of 0.1 year. Equivalent models can be fitted in R with the use of package *flexsurv* (See code in Additional file 2).

The EBMT example is immediately available in MSMplus by clicking "Yes" on the example button in the "Load" tab. By clicking "Yes", the two input json files derived from the multi-state model analysis using the Stata code in Additional file 1 are then internally uploaded to the app which is now ready to be traversed. We recommend readers run the example themselves whilst reading this manuscript.

Illustrations, figures, screenshots

Multi-state model diagram

By uploading the summary information json file, MSMplus is able to construct the multi-state graph at the "Model Structure" tab. Figure 1 illustrates the multi-state graphs produced based on the EBMT example, showing the number of patients in each state and number of patients having experienced each transition by the first half year after the transplantation ($t=5$). The boxes-states style and positions are adjusted either via sliders specifying the coordinates or fully interactively by clicking and dragging each box using R package *visNetwork*. Characteristics of the multi-state model diagram such as size, color and shapes of the boxes-states can be adjusted by the user.

Figure 1: Multi-state structure. a) static, b) interactive. The frequency of individuals that are in each state and transition can be presented over time with the use of a slider. A download button enables saving the multi-state graphs generated.

¹*Transition Probabilities*

²
³A variety of interactive plots and graphs are available for displaying predictions by
⁴covariate patterns for a variety of measures, including the transition probabilities
⁵("Transition probabilities" tab). Figure 2a is a bar plot that depicts the state tran-
⁶sition probabilities with different colors for the different covariate patterns, across
⁷the different states for time $t=0.5$ (it can easily be updated for any time point via
⁸the use of a slider). Figure 2b is a stacked bar plot, giving further insight concerning
⁹the state occupation probabilities among states and covariate patterns over time.
¹⁰Figure 2c is a stacked line plot portraying and comparing the transition probability
¹¹for each state for the different age categories as different frames over time with the
¹²present frame focusing on patients <20 years old. From Figure 2a and Figure 2b we
¹³can observe that 0.5 years after transplantation, the younger patients are slightly
¹⁴more likely to not have experienced neither platelet recovery nor relapse/death.
¹⁵Patients of the middle age category are more likely than the other age categories
¹⁶to be in the platelet recovery state, while patients >40 years old are more likely to
¹⁷have experienced relapse or death compared to the other age categories.

²⁰*Transition intensities*

²¹MSMplus provides graphs depicting the transition intensities by transition for dif-
²²ferent covariate patterns via the "Hazards" tab. It also supports graphs of hazard
²³ratios and differences between the covariate patterns for each transition, and tran-
²⁴sition intensity ratios among different transitions for the same covariate pattern.
²⁵Figure 3a is a line plot depicting the estimated transition intensity rates with a
²⁶different frame for each transition and different color for the different covariate pat-
²⁷terns, in this case the transition from post-transplant directly to Relapse/Death
²⁸(transition 2). The transition rate is greater for patients over 40 years of age at
²⁹transplantation followed by the younger patients (<20 years old) while the patients
³⁰of the middle age group appear to have the lower hazard for the specific transition.
³¹Figure 3b calculates the differences between the transition intensity rates of the
³²three age groups seen in Figure 3a together with 95% confidence intervals, using
³³the younger patients (< 20 years old) as the reference covariate pattern.

Figure 2: Graphs depicting the estimated transition probabilities for each state among the different age categories. The estimations can focus on a specific time point or over time. (a) Bar graphs changing over time, (b) Stacked bar plot depict the transition probabilities at 0.5 years after transplantation. (c) Stacked line plot of transition probabilities across all the predicted time points patients <20 years old.

Stacked probability of states across covariate patterns

Stacked probabilities of states among covariate patterns

Figure 3: Graphs of transition intensity rates. a) depicts the estimated transition intensity rates over time for transition 2 (Post transplant to Relapse/death) via frames. b) depicts transition intensity rates differences between the covariate patterns (different age categories) for transition 2 (Post transplant to Relapse/death).

29 Length of stay and probability of visiting a state

Two additional clinically useful measures are the length of stay in a state and the probability of ever visiting a state. Figure 4a depicts the length of stay in each state for the different age groups as a grid of line plots. Summing the estimated length of stay across states give the time the length of stay was evaluated at. The middle

age group appears to perform well by having a high length of stay in a platelet¹
 recovery state as opposed to the other age groups. Figure 4b shows the differences²
 of length of stay between the reference covariate pattern (< 20 years old) and the³
 rest of the covariate patterns of interest again as a grid. This plot shows that the⁴
 middle age group tends to stay more in the platelet recovery state and less in⁵
 the relapse/death state compared to the young age group. On the other hand the⁶
 old age group appears to stay less in the platelet recovery state and more in the⁷
 relapse/death state compared to the young age group. Regarding the probability of⁸
 ever visiting each state, Figure 4c depicts these probabilities for each state across⁹
 the covariate patterns as frames, in this case for the young patients group (< 20 ¹⁰
 years old). The probability of this age group of patients ever reaching the platelet¹¹
 recovery state flattens out at 0.5 in less than half a year after transplantation. The¹²
 aforementioned figures are visualized in tabs "Length of stay" and "Visit".¹³

User defined function

MSMplus also supports the presentation of a user-defined estimated measure across¹⁶
 the time points of prediction over the different covariate patterns ("User" tab).¹⁷
 For example, when using *predictms* in Stata, the user can calculate a function¹⁸
 consisted of transition probabilities and/or length of stays for each covariate pattern¹⁹
 of interest. In our example, the user function is defined as the sum of length of stays²⁰
 in state 1 (Post transplant) and state 2 (Platelet recovery). The sum of time that²¹
 an individual spends in these two states can be thought of as the total relapse²²
 free time. The user can include any function estimated from the multi-state model²³
 analysis, as long as it is predicted across the time points over the covariate patterns²⁴
 specified. Figure 5 depicts a user defined function (relapse free time) for two of the²⁵
 three age categories (younger than 20 years old and older than 40 years old). The²⁶
 user can select/deselect a subset of the covariate patterns for which the estimated²⁷
 measures will appear either at the "Settings" tab or interactively by clicking on the²⁸
 plot legends. The younger age group has greater relapse free time than the old age²⁹
 group over time.³⁰

Predictions comparison between two different approaches

MSMplus gives the user the option of comparing two different multi-state mod-³²
 elling approaches by supplying two separate prediction files. The interactive graphs³³

Figure 4: Length of stay and probability of ever visiting a state. Graphs depicting a) the estimated length of stay in each state for the different age categories in grid mode, b) the difference in length of stays between the age categories in grid mode and c) the estimated probability of ever visiting each state for individuals <20 years old.

available are then being created both for the first and the second approach and the user can optically compare the two approaches. Figure 6 shows the estimated

length of stay in each state and covariate pattern for two different multi-state approaches. The first approach has a Markov assumption (clock forward approach) while the second approach has a semi-Markov assumption (clock reset approach). The estimated length of stay appears quite robust in this example.

Notes

We have given an overview of some of the features of MSMplus. For space limitations we have not shown additional predictions. All the potential predictions can be seen in Table 1. The multi-state model analysis for this example was conducted in Stata via *msboxes* and *predictms*, using flexible parametric hazard models (Additional file 1). The corresponding analysis in R can be done via package *flexsurv* (Additional file 2). The source code of MSMplus is in Additional file 6 and github (See Availability of data and materials).

Figure 6: Comparison of two different modelling approaches. On the left are the estimated length of stay in each state for each covariate pattern assuming a Markov process (clock forward), while on the right are the same estimates assuming a semi-Markov process (clock reset approach).

Discussion

MSMplus is a web tool designed for communicating multi-state analysis results in a flexible and meaningful way. It encompasses graphs and plots that change over the different covariate patterns and across time, providing easy comparisons of the measures between individuals with different traits to describe how the disease process changes over time. MSMplus is primarily aimed for researchers who want to communicate the findings of their research with scientific or general audiences. It can also be used for generating plots and graphs when writing academic articles that use single event survival analysis, competing risks and multi-state models as these can be thought as special cases of multi-state analysis.

To the best of our knowledge, there is one other web-tool for multi-state models called MSM-Shiny [16]. MSM-Shiny requires the user to upload real data online. The analysis is then conducted within the app via the *mstate* or *flexsurv* package, and the estimated measures presently include, a multi-state model diagram, automatically derived by the app, a plot for the probability of being in each state,

¹and a survival plot. MSMplus does not require the raw research data to be up-¹
²loaded online, only the model predictions, as there may be data governance issues²
³in uploading research data to a website. MSMplus, is compatible with R (through³
⁴packages *msm*, *mstate* and *flexsurv*) and Stata via the command *predictms*. How-⁴
⁵ever, any statistical software can be used for the statistical analysis and the user⁵
⁶can then upload the results on MSMplus as csv files of an appropriate format. In⁶
⁷addition, MSMplus offers a variety of graphs and plots over various estimated mea-⁷
⁸sures, over different covariate patterns and across time as well as giving the user⁸
⁹increased control and flexibility in designing the multi-state model diagram and the⁹
¹⁰illustrations of the predicted measures. A potential small drawback is that MSM-¹⁰
¹¹plus requires the initial step of creating the input json/csv files. However, as it can¹¹
¹²be seen from the code in the additional files, that procedure is made easy for the¹²
¹³json files through the specific options (Stata) and functions (R), that we developed¹³
¹⁴for this purpose. 14

15

15

¹⁶**Conclusions** 16

¹⁷Thoughts about future extensions of MSMplus could be to increase the number of¹⁷
¹⁸measures that the app supports as well as adding the option of running the multi-¹⁸
¹⁹state analysis on MSMplus, using packages *msm*, *mstate* or *flexsurv*. In conclusion,¹⁹
²⁰we argue in favour of the further use and development of web tools in the future as²⁰
²¹a means to communicate scientific research. 21

22

22

²³**Availability and requirements** 23

24

24

Project name: MSMplus

25

25

Project home page:OS address: <https://nskbiostatistics.shinyapps.io/MSMplus/>,

26

26

Github:nskourlis/MSMplus

27

27

Operating system(s): shinyapps.io

28

28

Programming language: Rshiny

29

29

Other requirements: Compatible with all web browsers

30

30

License: Free commercial licence of Shinyapp.io

31

31

Any restrictions to use by non-academics: No restrictions

32

32

33

33

¹ Abbreviations	1
² Declarations	2
Ethics approval and consent to participate	
³ Not applicable. This article describes a web tool developed for presenting multi-state model results. The EBMT	3
⁴ data used for the illustration example are available for public use.	4
⁵ Consent for publication	5
Not applicable. This article describes a web tool developed for presenting multi-state model results. There are no	
⁶ participants.	6
⁷ Availability of data and materials	7
⁸ The example dataset EBMT supporting the conclusions of this article is available via <i>mstate</i> package in R. The	8
examples and R functions developed for the derivation of the json input files are available in the supplementary	
⁹ information files and the nsourlis/MSMplus github repository.	9
¹⁰ Competing interests	10
¹¹ The authors have no conflicts of interest to declare that are relevant to the content of this article.	11
¹² Funding	12
This work was supported by the Swedish Cancer Society (Cancerfonden) (Grant number 2018/744) and the Swedish	
¹³ Research Council (Vetenskapsrådet) (Grant number 2017-01591).	13
¹⁴ Authors' contributions	14
¹⁵ The contributions of each author to the work is as follows. NS, MJC, TMLA and PCL, equally contributed towards	15
idea generation. NS was responsible for the development, coding and dissemination of MSMplus. PCL, MJC and	
¹⁶ TMLA supervised the project and contributed comments and suggestions throughout the entire development stage	16
¹⁷ of the web-tool. All authors were involved in the drafting and submission of the final manuscript.	17
¹⁸ Acknowledgements	18
This work was supported by the Swedish Cancer Society (Cancerfonden) (Grant number 2018/744) and the Swedish	
¹⁹ Research Council (Vetenskapsrådet) (Grant number 2017-01591).	19
²⁰ Author details	20
²¹ ¹ Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Nobels Väg 12A, Stockholm,	21
Sweden. ² Biostatistics Research Group, Department of Health Sciences, University of Leicester, University Road,	
²² Leicester, UK.	22
²³ References	23
²⁴ 1. Van Houwelingen HC, Putter H. Dynamic predicting by landmarking as an alternative for multi-state modeling:	24
an application to acute lymphoid leukemia data. <i>Lifetime data analysis</i> . 2008;14(4):447.	
²⁵ 2. Conlon A, Taylor J, Sargent DJ. Multi-state models for colon cancer recurrence and death with a cured	25
²⁶ fraction. <i>Statistics in medicine</i> . 2014;33(10):1750–1766.	26
²⁷ 3. Ieva F, Jackson CH, Sharples LD. Multi-State modelling of repeated hospitalisation and death in patients with	27
Heart Failure: the use of large administrative databases in clinical epidemiology. <i>Statistical methods in medical</i>	
²⁸ research. 2017;26(3):1350–1372.	28
²⁹ 4. Williams C, Lewsey JD, Briggs AH, Mackay DF. Cost-effectiveness analysis in R using a multi-state modeling	29
survival analysis framework: a tutorial. <i>Medical Decision Making</i> . 2017;37(4):340–352.	
³⁰ 5. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. <i>Statistics in</i>	30
medicine. 2007;26(11):2389–2430.	
³¹ 6. Cook RJ, Lawless JF. Statistical issues in modeling chronic disease in cohort studies. <i>Statistics in Biosciences</i> .	31
2014;6(1):127–161.	
³² 7. Andersen PK, Perme MP. Inference for outcome probabilities in multi-state models. <i>Lifetime data analysis</i> .	32
2008;14(4):405.	
³³ 8. Hougaard P. Multi-state models: a review. <i>Lifetime data analysis</i> . 1999;5(3):239–264.	33

9. Trevena LJ, BPSych HMD, Barratt A, Butow P, Caldwell P. A systematic review on communicating with patients about evidence. *Journal of evaluation in clinical practice*. 2006;12(1):13–23.
10. Mozumder SI, Dickman PW, Rutherford MJ, Lambert PC. InterPreT cancer survival: A dynamic web interactive prediction cancer survival tool for health-care professionals and cancer epidemiologists. *Cancer epidemiology*. 2018;56:46–52.
11. Gasparini A, Morris TP, Crowther MJ. INTEREST: Interactive Tool for Exploring REsults from Simulation sTudies. *arXiv preprint arXiv:190903813*. 2019;.
12. Freeman SC, Kerby CR, Patel A, Cooper NJ, Quinn T, Sutton AJ. Development of an interactive web-based tool to conduct and interrogate meta-analysis of diagnostic test accuracy studies: MetaDTA. *BMC medical research methodology*. 2019;19(1):1–11.
13. Willekens F. Statistical Packages for Multistate Life History Analysis. In: *Multistate Analysis of Life Histories* with R. Springer; 2014. p. 135–203.
14. Shiny;. Accessed: 2021-02-01. <https://shiny.rstudio.com/>.
15. de Wreede LC, Fiocco M, Putter H, et al. mstate: an R package for the analysis of competing risks and multi-state models. *Journal of statistical software*. 2011;38(7):1–30.
16. Lacy S. Exploring Multi-State Models using Interactive Shiny Web Applications. *Survival Analysis for Junior Researchers*, Leicester, U.K; 2017. .

Figures

Figure 1—Multi-state structure

a) static, b) interactive. The frequency of individuals that are in each state and transition can be presented over time with the use of a slider. A download button enables saving the multi-state graphs generated.

Figure 2—Graphs depicting the estimated transition probabilities for each state among the different age categories

The estimations can focus on a specific time point or over time. (a) Bar graphs changing over time, (b) Stacked bar plot depict the transition probabilities at 0.5 years after transplantation. (c) Stacked line plot of transition probabilities across all the predicted time points

Figure 3—Graphs of transition intensity rates

Figure 3 a) depicts the estimated transition intensity rates over time for transition 2 (Post transplant – > Relapse/death) via frames. Figure 3 b) depicts transition intensity rates differences between the covariate patterns (different age categories) for transition 2 (Post transplant – > Relapse/death).

Figure 4—Length of stay and probability of ever visiting a state

Graphs depicting a) the estimated length of stay in each state for the different age categories in grid mode, b) the difference in length of stays between the age categories in grid mode and c) the estimated probability of ever visiting each state for individuals <20 years old.

Figure 5—Graphs depicting the estimated user function for each age category with confidence intervals.

User function in this example: Relapse free time

Figure 6—Comparison of two different modelling approaches.

On the left are the estimated length of stay in each stay for each covariate pattern assuming a Markov process (clock forward), while on the right are the same estimates assuming a semi-Markov process (clock reset approach).

Tables

Table 1 — Measures estimated from multi-state models that are supported for visualization from the application tool MSMplus.

Additional Files

Additional file 1 - Stata code for deriving the Summary information and the Analysis results input files for MSMplus
Stata code for fitting the multistate models on the EBMT dataset and for deriving the Summary information and the analysis results input files for the web tool MSMplus using the Stata commands msboxes and predictms. (PDF file 404KB)

¹ Additional file 2 — R code for deriving the Summary information and the Analysis results input files for MSMplus	1
² Examples of creating json files in R as input for MSMplus via the manually derived functions msboxesR (Additional	2
file 3) and functions flexjson (Additional file 4, part a), mstatejson (Additional file 4, part b) and msmjson	
³ (Additional file 4, part c). (PDF file 536KB)	3
⁴	4
Additional file 3 — msboxesR function to be used for deriving the Summary information file for MSMplus	
⁵ Manually developed function msboxesR for deriving the descriptives input file for MSMplus in R. Examples of	5
⁶ applying this function are presented in Web Appendix 2. (TXT file 15KB)	6
⁷ Additional file 4 — R functions to be used for deriving the Analysis results file for MSMplus when using flexsurv,	7
mstate or msm R packages	
⁸ Developed R functions flexjson (part a), mstatejson (part b) and msmjson (part c) for deriving the analysis json	8
⁹ input files for MSMplus after using flexsurv, mstate and msm packages for the statistical analysis respectively.	9
Examples of applying these functions are presented in Additional file 2. (TXT file 106KB)	
¹⁰	10
¹¹ Additional file 5 — Creating a csv Analysis results file to provide to MSMplus	11
Guidance in creating an Analysis results csv file to provide as input to MSMplus. (PDF file 522KB).	
¹²	12
¹³ Additional file 6 — Example JSON files and MSMplus source code	13
¹⁴ The summary information and analysis results json files based on the EBMT example dataset. The source code of	14
¹⁵ MSMplus application. (TXT file 1893KB).	15
¹⁶	16
¹⁷	17
¹⁸	18
¹⁹	19
²⁰	20
²¹	21
²²	22
²³	23
²⁴	24
²⁵	25
²⁶	26
²⁷	27
²⁸	28
²⁹	29
³⁰	30
³¹	31
³²	32
³³	33

Figures

Figure 1

Multi-state structure. a) static, b) interactive. The frequency of individuals that are in each state and transition can be presented over time with the use of a slider. A download button enables saving the multi-state graphs generated.

Stacked probability of states across covariate patterns

Figure 2

Graphs depicting the estimated transition probabilities for each state among the different age categories. The estimations can focus on a specific time point or over time. (a) Bar graphs changing over time, (b) Stacked bar plot depict the transition probabilities at 0.5 years after transplantation. (c) Stacked line plot of transition probabilities across all the predicted time points patients <20 years old.

Figure 3

Graphs of transition intensity rates. a) depicts the estimated transition intensity rates over time for transition 2 (Post transplant to Relapse/death) via frames. b) depicts transition intensity rates differences between the covariate patterns (different age categories) for transition 2 (Post transplant to Relapse/death).

Figure 4

Length of stay and probability of ever visiting a state. Graphs depicting a) the estimated length of stay in each state for the different age categories in grid mode, b) the difference in length of stays between the age categories in grid mode and c) the estimated probability of ever visiting each state for individuals <20 years old.

User specified measure- Relapse free time across covariate patterns

Figure 5

Graphs depicting the estimated user function for each age category with condence intervals. User function in this example: Relapse free time

Figure 6

Comparison of two different modelling approaches. On the left are the estimated length of stay in each state for each covariate pattern assuming a Markov process (clock forward), while on the right are the same estimates assuming a semi-Markov process (clock reset approach).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1.pdf](#)
- [Additionalfile2.pdf](#)
- [Additionalfile3.txt](#)
- [Additionalfile4.txt](#)
- [Additionalfile5.pdf](#)
- [Additionalfile6.txt](#)