gbg.kaplan.meier.plot.v1

The function creates a Kaplan-Meier figure according to the working instructions “AA\_Statistic\_Kaplan-Meier curves and Forest-plots\_GBG-standard.pdf” (19.5.2016, rev. 1). The figure may contain one or more Kaplan-Meier curves. Main optional features: patients-at-risk table including a legend on the left, a text block containing additional information, lines indicating the median survival, a figure title.

The function returns a graphical object (grid::grob) containing the Kaplan-Meier figure, it does not perform the plotting.

Libraries: survival, grid

Dependencies: gbg.format.p.v1

Owner: Karsten Weber

Parameters:

| **Name** | **Values** | **Default** | **Description** |
| --- | --- | --- | --- |
| endpoint.time  endpoint.event | see description | <must be specified> | Sample data for the Kaplan-Meier curves. The data is always interpreted as a linear vector with one sample per index; if samples belong to different subgroups this is specified by parameter group. Several data formats are accepted:  (1) endpoint.time is a survival::Surv object of type=”right” and endpoint.event is not specified  (2) endpoint.time is numeric and endpoint.event is logical  (3) endpoint.time is numeric and endpoint.event is numeric and contains only values 0, 1, and NA  (4) endpoint.time is numeric and endpoint.event is character and contains only values “no”, “yes”, and NA  (5) endpoint.time is numeric and endpoint.event is a factor with levels “no” and “yes”  With the exception of format (1) both variables must have the same lengths and endpoint.time must not be negative. Missing values should be coded by NA values; samples with a NA value in at least one variable will be omitted (it will also not be counted in the patients-at-risk table). |
| group | see description |  | Denotes the subgroup of the samples assigning the samples to different Kaplan-Meier curves.  If group is not specified only one Kaplan-Meier curve is generated containing all samples. If group is specified it is interpreted as a linear vector and must have the same length as parameters endpoint.time and endpoint.event. Parameter group may contain NAs: These samples are omitted. Three data formats are accepted:  (1) group is a factor. In this case the groups displayed in the legend are displayed in the same order as the defined levels.  (2) group is of class character. In this case the order of groups displayed in the legend is defined by the lexicographic order of group names.  (3) group is numeric. The order of groups displayed in the legend is the numerical order; group names are generated by function as.character.  In all formats, subgroups without valid samples (e.g. because of NAs in endpoint.time) are omitted. |
| bw | logical scalar | F | If true a black-and-white figure is generated, if false a colored figure is generated. Parameter bw only changes the default values of color and line type parameter; it has no effect if colors and/or line types are explicitly specified. |
| groups.col | vector of color specifications (see R help, function par, section “Color Specification”) | colored figure:  c("blue", "red", "green4", "magenta", "cyan", " goldenrod4 ")  black-and-white: "black" | Defines the colors of the Kaplan-Meier curves. If not specified and a colored figure is requested the first colors from the default are used for the subgroups; if more than 6 subgroups are present an error is raised. If specified the number of colors in groups.col must either be one (same color for all subgroups) or match the number of subgroups; otherwise an error is raised. |
| groups.lty | numeric or character vector with:  1="solid" 2="dashed" 3="dotted" 4="dotdash" 5="longdash" 6="twodash" | colored figure:  1  black-and-white:  1:6 | Defines the line types of the Kaplan-Meier curves. If not specified and a black-and-white figure is requested the first line types from the default are used for the subgroups; if more than 6 subgroups are present an error is raised. If specified the number of line types in groups.lty must match the number of subgroups; otherwise an error is raised. |
| atrisk.shown | logical scalar | T | If true, the table containing the number of patients at risk including the legend on the left is shown. If false, it is not shown.  NOTE: If you have more than one group (Kaplan-Meier curves) and specify atrisk.shown=F then you should describe outside the figure which curve belongs to which group, because there will be no legend in the figure. |
| atrisk.colored | logical scalar | T | If true, the numbers of patients at risk are shown in the same color as the corresponding Kaplan-Meier curves. If false the numbers of patients at risk are shown in black color. |
| legend.colored | logical scalar | T | If true, the legend texts (i.e. group names; not the legend lines) are shown in the same color as the corresponding Kaplan-Meier curves. If false the legend texts are shown in black color. |
| xlab | character string | “time (months)” | X-axis label text. |
| xlab.in | logical scalar | T | If true, the x-axis label appears within the plot area above the x-axis. If false, the x-axis label appears below the x-axis and (if present) below the patients-at-risk table. |
| ylab | character string | “event free survival” | Y-axis label text. The text is rotated by 90° in the figure. |
| lab.cex | positive numeric scalar | 1.2 | Magnification factor of the font size for the x-axis and y-axis label texts. The absolute font size is lab.cex\*fontsize. |
| xlim | positive numeric scalar | 2% more than the last event or censored sample or xticks | The right limit of the x-axis (the left limit is always 0). |
| xticks | numeric vector with at least one element, non-negative | seq(0, xlim, ifelse(xlim < 12, 1, 12)) | The times on the x-axis where the time as a number is shown. These are also the times patients-at-risk are shown for. |
| yticks | numeric vector, values between 0 and 100 | seq(from = 0, to = 100, by = 20) | The percentages on the y-axis where the percentage as a number is shown. |
| title | character string |  | If specified the title is shown at the top of the figure. If not specified, no space is reserved at the figure top. |
| title.cex | positive numeric scalar | 1.2 | Magnification factor of the font size for the title. The absolute font size is title.cex\*fontsize. |
| fontsize | positive numeric scalar | 12 | Font size for the figure. The fontsize immediately determines the font size for the axes ticks, patients-at-risk and the legend. All other figure elements (axes labels, title, censored markers, block) have parameters for magnifying their font size. |
| fontfamily | Font family specification | NULL | The fontfamily determines the font family for all texts in the figure plot. Please note that the font family depends on the graphical device for output. For details see R help, function gpar, and function Hershey. |
| lineheight | positive numeric scalar | 1.2 | The height of a text line as a multiple of the size of text. The parameter impacts several distances between axes, axes tick labels, patients at risk text lines, and axis labels; this applies to the x-axis as well as to the y-axis.  The absolute text line heights are proportional to lineheight\*fontsize and also depend on lab.cex and title.cex.  For the block, see block.lineheight. |
| censor.pch | whole number or single character | 18 | Text symbol identifier for censored samples. For details about potential symbols see R help, function points, parameter pch.  Set censor.pch=NULL to suppress plotting of censored samples. |
| censor.cex | positive numeric scalar | 0.8 | Magnification factor for the font size for the censored samples symbols. The absolute font size is censor.cex\*fontsize. |
| curves.lwd | positive numeric scalar | 1 | Line width for Kaplan-Meier curves. |
| axes.lwd | positive numeric scalar | 1 | Line width for axes and plot area borders |
| median.shown | logical scalar | F | If true, lines describing the median survival are added. For each subgroup a horizontal line at y = 50% and a vertical line at the median survival are generated. |
| median.col | color specifications (see R help, function par, section “Color Specification”) | groups.col | Determines colors for median survival lines. If not specified, the same colors as for the Kaplan-Meier curves are used. If only one color is specified this color is used for all subgroups. If as many colors as subgroups are specified, theses colors are used respectively. Any other number of color specifications results in an error. |
| median.lty | line type specification, see parameter groups.lty | 6 (two-dashed) | Determines line types for median survival lines. If not specified two-dashed lines are plotted. If only one line type is specified this line type is used for all subgroups. If as many line types as subgroups are specified, theses line types are used respectively. Any other number of line type specifications results in an error. |
| median.lwd | positive numeric scalar | curves.lwd | Line width for median survival lines. |
| block.features | list, see below |  | If not specified or an empty list no block is shown. If specified each list element contains a block text line (either one or one per subgroup), see below. |
| block.cex | positive numeric scalar | 1.0 | Magnification factor of the font size for the block. The absolute font size is block.cex\*fontsize. The censor symbol is shown in the same size as in the Kaplan-Meier curves, thus it is not affected by block.cex. |
| block.lineheight | positive numeric scalar | 1.0 | The height of a block text line as a multiple of the size of text. The absolute line height is block.lineheight\*block.cex\*fontsize. |
| block.colored | logical scalar | T | If true, the text of block features depending on the subgroups are shown in subgroup colors (groups.col). If false, all block features are shown in black text color. |
| block.just | numeric vector, length 2, element range 0 to 1 | c(0, 0) | Justification of block as a whole. The first value specifies the horizontal justification (0=left, 1=right), the second value specifies the vertical justification (0=bottom, 1=top). |
| block.x | grid::unit scalar | unit(0.02, “npc”) | Horizontal position of the block with respect to the plot area / x-axis. |
| block.y | grid::unit scalar | unit(1.2 \* lab.cex, "lines") | Vertical position of the block with respect to the plot area / y-axis. |

**block.features:**

Parameter block.features is a list specifying the text lines of the block containing additional information. Using this type of specification very different blocks can be generated.

Each list element must either be a character string or a named list element; these character strings and names are equivalent and can be used interchangeably. Character strings denote simple block elements with no or default parameterization. Named list elements denote block elements with parameters which can either be atomic or lists; in the latter case the list recursively follows the same rules as the block.features list. Here is an example of a block.features specification demonstrating the different types of specification.

block.features = list(

"censored",

"median" = "years",

"HR" = list("CI", level=0.95, "Wald"))

The following table shows all available elements including their potential parameters:

| **Name** | **Value** | | **Default** | **Description** |
| --- | --- | --- | --- | --- |
| **Name** | **Default** |
| "censored" | <none> | | | Shows the symbol for censored samples and the word “censored”. This option is not allowed if censored samples are not shown (censor.pch=NULL). |
| "events" | <none> | | | Shows the number of events and samples for each subgroup. Each subgroup generates a separate line of text. |
| "HR" | list() with: | | list(  "CI") | Shows the hazard ratio between subgroups. Only valid for exactly two subgroups. |
| "ratio" | <none> | Shows the subgroup names compared to each other for the hazard ratio (second group vs first group): This notation and the HR value correspond to the output of ods table HazardRatios from proc phreg in SAS. |
| "CI" | <none> | Shows the confidence interval for the hazard ratio. |
| "level" | 0.95 | Probability level of the confidence interval; the confidence interval is always symmetric. |
| "Wald" | <none> | Shows the Wald p-value for the hazard ratio. |
| "LR" | <none> | Shows the likelihood ratio p-value for the hazard ratio. |
| "log-rank" | <none> | | | Shows a log-rank p-value. The log rank test can be performed for two as well as more subgroups. |
| "log-rank stratified" | list() with: | | <must be specified> | Shows a stratified log-rank p-value. The stratified log rank test can be performed for two as well as more subgroups. |
| "data" | <must be specified> | Data for stratification; must be a vector of the same length as endpoint.time. Supported data types are numeric, character and factor. NAs are not allowed and result in an error. |
| "name" | <must be specified> | Character string to be shown as the name of the stratification variable. |
| "median" | character string | | "months" | Shows the median survival times (see parameter median.shown) for each subgroup as text line. The value must be a character string which is printed after the calculated median survival time; it shall denote the physical unit of time. Please note that it is the responsibility of the user that the unit specified here corresponds to the unit shown in the x-axis label (see parameter xlab). |
| "user" | character string | | "" | Shows a line of text defined by the user. |

Example:

DFS <- c( 7, 11, 18, 22, 24, 31, 38, 42, 47)

eventDFS <- c( 1, 1, 0, 1, 0, 0, 1, 0, 0)

group <- c("a", "b", "b", "a", "b", "a", "b", "b", "a")

g <- gbg.kaplan.meier.plot.v1(DFS, eventDFS, group)

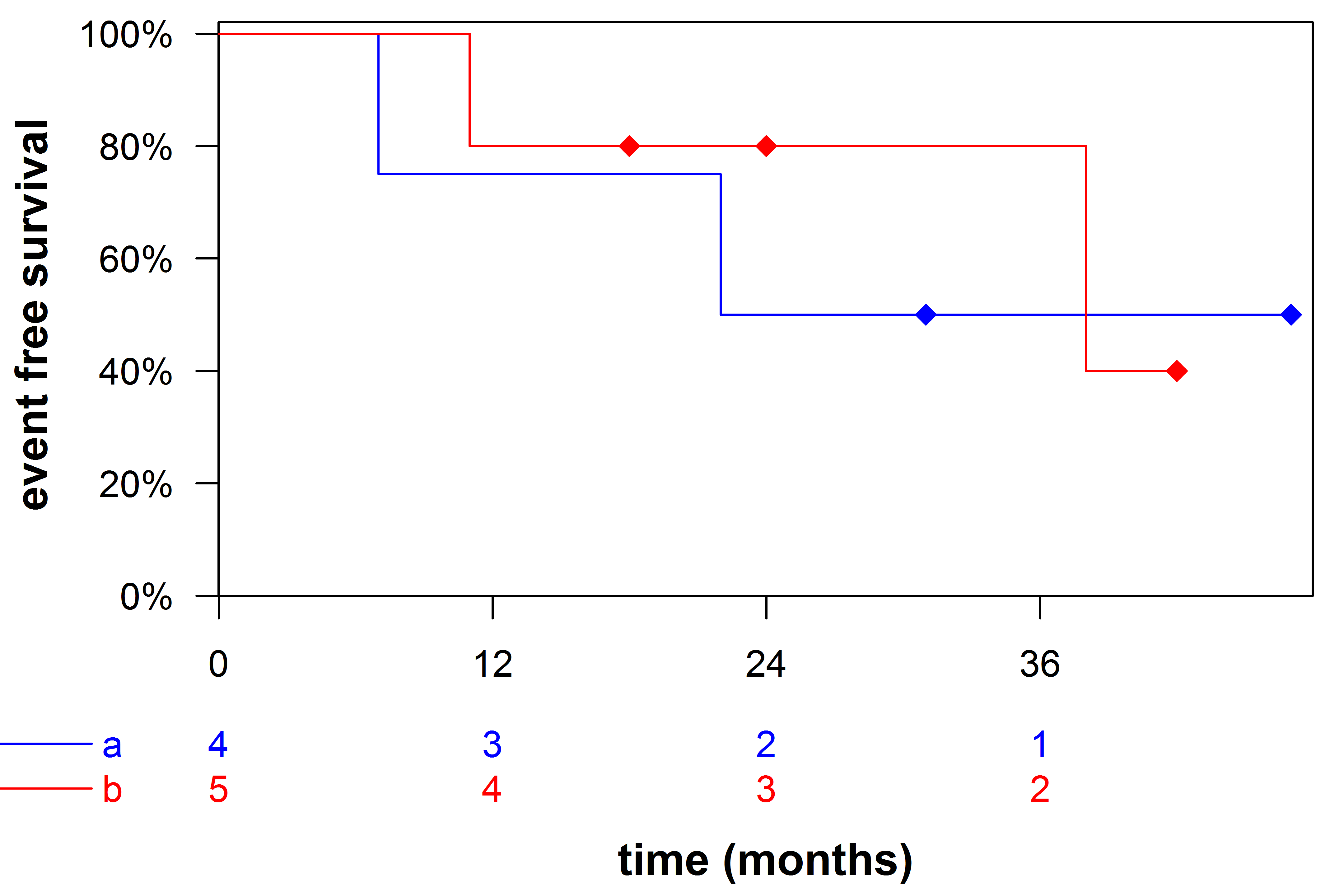
png(filename = "u:/temp/example.png", width = 6, height = 4,

unit = "in", res = 720)

grid.draw(g)

dev.off()

The following figure shows the result of the example above and also demonstrates the impact of some parameters on the resulting figure:



group

groups.col

groups.lty

curves.lwd

atrisk.shown

atrisk.colored

legend.colored

xlab

lab.cex

ylab

lab.cex

xlim

xticks

censor.pch

censor.cex

axes.lwd

yticks

title

title.cex

lineheight

lineheight

lineheight

lineheight