

Package: hvtiRlifetables (via r-universe)

October 20, 2026

Type Package

Title Age, Sex and Race Matched US Reference Survival

Version 0.1.5

Date 2026-10-12

Maintainer John Ehrlinger <john.ehrlinger@gmail.com>

License GPL-3

Encoding UTF-8

Language en-US

URL <https://ehrlinger.github.io/hvtiRlifetables/>,
<https://github.com/ehrlinger/hvtiRlifetables>

BugReports <https://github.com/ehrlinger/hvtiRlifetables/issues>

Description Age, sex and race matched United States reference survival, as consumed by clinical outcomes reports within The Heart and Vascular Institute at the Cleveland Clinic. Reproduces the 'SAS' macro 'usmatchd', which evaluates a stored three-phase parametric hazard fit of Blackstone, Naftel and Turner (1986) <doi:10.1080/01621459.1986.10478314> on the age axis rather than interpolating a life table. Ships the fitted parameter blocks, one set per vintage per stratum, together with a thin evaluator over the 'TemporalHazard' package. The reference vintage is never defaulted, so an analysis must state which reference population it used.

Depends R (>= 4.4.0)

Imports TemporalHazard (>= 1.2.0), utils

Suggests haven, knitr, quarto, rmarkdown, testthat (>= 3.0.0), withr

Remotes TemporalHazard=ehrlinger/TemporalHazard

VignetteBuilder quarto

SystemRequirements Quarto command line tool
(<<https://github.com/quarto-dev/quarto-cli>>)

Config/testthat/edition 3

```
Roxygen list(markdown = TRUE)
Config/roxygen2/version 8.1.0
LazyData true
Repository https://ehrlinger.r-universe.dev
Date/Publication 2026-10-14 22:56:27 UTC
RemoteUrl https://github.com/ehrlinger/hvtiRlifetables
RemoteRef HEAD
RemoteSha 93117c87d5b19d3ea59b3647d9f06913704c3886
```

Contents

us_cohort_curve	2
us_lifetable_model	3
us_lifetable_models	4
us_lifetable_vintages	5
us_matched	6
Index	9

us_cohort_curve	<i>Cohort expected-survival curve, optionally within groups</i>
-----------------	---

Description

Reduces the per-patient output of [us_matched](#) to the curve a figure actually carries: the cohort’s expected survival at each time, and its hazard. With by, one such curve per group.

Usage

```
us_cohort_curve(x, by = NULL)
```

Arguments

- | | |
|----|--|
| x | A data frame as returned by us_matched with individual = TRUE – columns id, time, smatched and hmatched. An already-aggregated frame is rejected rather than averaged twice. |
| by | How to group the result. One of: <ul style="list-style-type: none">• NULL (default), for a single cohort-wide curve;• a vector or factor with one value per <i>row</i> of x, reported in a column called group;• a named list or data frame of such vectors, for a crossed grouping, reported under its own names. It carries values, not column names: group by a column of x with by = x\$arm. Because x holds one row per patient per time, a patient-level grouping must be expanded to it – see the examples. |

Value

A data frame with columns `time`, `smatched` and `hmatched`, one row per time – or, with `by`, the grouping columns first and one row per group per time. Groups and times appear in order of first appearance in `x`, so a factor grouping does not reorder to its levels.

The averaging, and why it is here

The reduction follows the SAS macro %usmatchd (lines 338-350): the hazard is converted to a density (`hmatched * smatched`), the **unweighted** cohort means of survival and of the density are taken at each time, and the mean hazard is recovered by division. It is deliberately **not** the mean of the individual hazards; in a heterogeneous cohort the two differ, and the difference is not cosmetic.

The mean is over **survival**, not over cumulative hazard, and no patient is weighted by follow-up. Missing values are not dropped: every value `us_matched` returns is finite by construction, so an NA arriving here is a defect rather than a datum, and averaging around it would hide it.

Grouping is not stratification

`us_matched`'s `table` argument chooses which strata the *life table* is built from – "`sexrace`", "`race`", "`sex`", "`overall`". `by` chooses how the resulting curves are grouped for *reporting* – age band, treatment arm, era. The two axes are independent: a report broken down by age band is normally still matched on "`sexrace`".

See Also

`us_matched`, whose `individual = FALSE` is the ungrouped case of this function.

Examples

```
x <- us_matched(age = c(55, 60, 78, 82), male = c(1, 0, 1, 0),
               other = c(0, 0, 1, 1), times = c(0, 5, 10),
               vintage = "table84")

# The cohort curve.
us_cohort_curve(x)

# The same, broken down by an age band of the study's own choosing. The
# grouping is per patient, so it is expanded to x's one-row-per-patient-
# per-time shape by id.
band <- c("55-79", "55-79", "55-79", "80+")
us_cohort_curve(x, by = factor(band[x$id], levels = c("55-79", "80+")))
```

us_lifetable_model	<i>One fitted parameter set</i>
--------------------	---------------------------------

Description

Returns the raw fitted model for a single vintage and stratum, for inspection. `us_matched()` is the function most callers want.

Usage

```
us_lifetable_model(vintage, stratum)
```

Arguments

vintage	Character scalar. One of the identifiers returned by <code>us_lifetable_vintages()</code> . There is no default ; omitting it is an error.
stratum	Character scalar. One of "all", "f", "m", "w", "wf", "wm", or the vintage's non-white codes — "o", "of", "om" for table84, "b", "bf", "bm" otherwise.

Value

A list with elements vintage (character), stratum (character), params (named numeric of length 11), status (named integer of length 11, the `_STATUS_` gate), flags (named numeric of length 6, metadata only), and vcov (11 by 11 numeric matrix, unused in this version).

Examples

```
m <- us_lifetable_model("table84", "wm")
m$params[["THALF"]]

# The _STATUS_ gate: table84's non-white male stratum has no constant
# phase. Its C0 is 0 *and absent*, not 0 *and meaning exp(0) = 1*.
om <- us_lifetable_model("table84", "om")
om$params[["C0"]]
om$status[["C0"]]
```

us_lifetable_models	<i>Fitted US life-table hazard model parameters</i>
---------------------	---

Description

The fitted three-phase parametric hazard models that the Cleveland Clinic SAS macro `%usmatchd` evaluates, one record per vintage per stratum. These are fitted models, not life tables: `%usmatchd` runs PROC HAZPRED against a stored fit on the **age** axis, with time origin at birth.

Usage

```
us_lifetable_models
```

Format

A data frame with 27 rows and 6 columns:

vintage character. One of "table84", "table2008", "table2023". Vintages are not interchangeable: model *structure* differs, not only fitted values.

stratum character. One of "all", "f", "m", "w", "wf", "wm", plus the vintage's non-white codes – "o", "of", "om" for table84; "b", "bf", "bm" for table2008 and table2023.

params list column of named numeric vectors, length 11: DELTA, THALF, NU, M, TAU, GAMMA, ALPHA, ETA, E0, C0, L0.

status list column of named integer vectors, length 11. The `_STATUS_` gate. $\mu = \exp(E0 \mid C0 \mid L0)$ **only** where the corresponding status is 1; status 0 means the phase is absent from the model, not that $\log(\mu)$ is zero.

flags list column of named numeric vectors, length 6: G1FLAG, FIXDEL0, FIXMNU1, G3FLAG, FIXGE2, FIXGAE2. These record which sub-form was fitted. They are metadata only and are never read during evaluation.

vcov list column of 11 by 11 numeric matrices, the parameter covariance blocks. Nothing in this version reads them; they ship because they are the only route to a confidence band later.

Race semantics

table84 names its non-white strata o, honestly "other". table2008 and table2023 name the same strata b, and the macro's own comment asserts these are Black-race estimates. For 2023 that assertion is **false**: the category is a risk-weighted average of Black, Asian, American Indian and Hispanic death rates, weighted by number at risk, stored under B only to keep the macro's naming consistent. See [us_lifetable_vintages\(\)](#), which reports this per vintage.

Source

Fitted by The Cleveland Clinic Foundation. Built into this package by `data-raw/build-models.R`, which is not run at install time. That script's header names the share location to restore the inputs from — deliberately kept out of R/, which must carry no literal share path.

`us_lifetable_vintages` *Available US life-table vintages*

Description

Reports the fitted-model vintages this package ships, with the provenance and race semantics needed to choose one. Vintages are **not** interchangeable — model structure differs, not only fitted values.

Usage

```
us_lifetable_vintages()
```

Value

A data frame with one row per usable vintage and columns: `vintage` (character identifier), `n_strata` (integer, always 9), `nonwhite_code` (character, the stratum code this vintage uses for its non-white category), `nonwhite_meaning` (character, what that category actually contains), `added` (character date the fits were added, where known, otherwise NA).

There is deliberately no source column: R/ carries no literal share path. `data-raw/build-models.R` names the location the fits are restored from.

Read nonwhite_meaning before citing a stratum

The table2023 non-white category is stored under code b but is a risk-weighted average of Black, Asian, American Indian and Hispanic death rates. It is not Black. The macro’s own comment says otherwise and is wrong.

Examples

```
us_lifetable_vintages()
```

us_matched	<i>Age, sex and race matched US reference survival</i>
------------	--

Description

Reproduces the Cleveland Clinic SAS macro %usmatchd: for each patient, a US reference survival curve matched on age, sex and race, together with its hazard. This is the dashed comparison line on a clinical survival figure.

Usage

```
us_matched(  
  age,  
  male,  
  other,  
  times,  
  id = seq_along(age),  
  vintage,  
  table = c("sexrace", "race", "sex", "overall"),  
  scale = c("years", "months", "days"),  
  individual = TRUE  
)
```

Arguments

age	Numeric vector of ages in years , always in years regardless of scale.
male	Numeric vector, 1 male, 0 female. The macro’s coding, unchanged.
other	Numeric vector, 1 non-white, 0 white. The macro’s coding, unchanged. What "non-white" contains differs by vintage – see us_lifetable_vintages() .
times	Numeric vector of follow-up times in the units of scale, non-negative.
id	Optional vector of patient identifiers, recycled into the output. Defaults to seq_along(age).
vintage	Character scalar naming the fitted-model vintage. There is no default. Omitting it is an error. See Details.

table	How finely to stratify patients, mirroring the macro's TABLE= modes. "sexrace" uses all four crossings, "race" uses white against the vintage's non-white category, "sex" uses male and female, "overall" sends every patient to the combined stratum and ignores male and other.
scale	Units of times. Applies the macro's SCALEF of 1, 1/12 and 1/365.2425 respectively.
individual	If TRUE (default), one row per patient per time. If FALSE, the cohort mean curve, one row per time – the ungrouped case of us_cohort_curve , which is what to use for a curve within groups of the caller's own choosing.

Details

%usmatchd is not a life-table lookup. It evaluates a stored three-phase parametric hazard fit on the **age** axis, time origin birth, and reads conditional survival off that one smooth curve twice. That is why the resulting hazard is smooth *within* a one-year age bin where a life table would be flat.

hmatched is **per year regardless of scale**. This matches the macro's source, which assigns _HAZARD without applying SCALEF, notwithstanding the macro's own header comment to the contrary.

When individual = FALSE, the cohort mean follows the macro's arithmetic: the hazard is converted to a density (hmatched * smatched), the cohort means of survival and of the density are taken at each time, and the mean hazard is recovered by division. It is deliberately **not** the mean of the individual hazards.

Value

A data frame. When individual = TRUE, columns id, time, agesurv (survival from birth to the patient's current age), smatched (reference survival over times, conditional on having reached age) and hmatched (the reference hazard), with one row per patient per time. When individual = FALSE, columns time, smatched and hmatched only, with one row per time.

Why vintage has no default

The macro's default silently moved twice: from table84 to table2008 between 2010 and 2024, then to table2023 on 2025-12-23. Every job re-run across either move got different numbers with no signal. An analysis that does not state its reference vintage is not reproducible, so this package refuses to guess. State it literally in analysis code.

See Also

[us_cohort_curve](#) to reduce the individual output to a cohort curve, or to one curve per group; [us_lifetable_vintages\(\)](#) for the available vintages and what their non-white stratum actually contains; [us_lifetable_model\(\)](#) for the raw fitted parameters.

Examples

```
# A 70-year-old white male, ten years of follow-up, 1984 fits.
r <- us_matched(age = 70, male = 1, other = 0,
               times = seq(0, 10, by = 2), vintage = "table84")
r[, c("time", "smatched", "hmatched")]
```


Index

* **datasets**

- `us_lifetable_models`, [4](#)
- `us_cohort_curve`, [2](#), [7](#)
- `us_lifetable_model`, [3](#)
- `us_lifetable_model()`, [7](#)
- `us_lifetable_models`, [4](#)
- `us_lifetable_vintages`, [5](#)
- `us_lifetable_vintages()`, [4–7](#)
- `us_matched`, [2](#), [3](#), [6](#)
- `us_matched()`, [3](#)