

# Package: **hdcuremodels** (via **r-universe**)

September 12, 2024

**Title** Penalized Mixture Cure Models for High-Dimensional Data

**Version** 0.0.1

**Date** 2024-06-11

**Description** Provides functions for fitting various penalized parametric and semi-parametric mixture cure models with different penalty functions, testing for a significant cure fraction, and testing for sufficient follow-up as described in Fu et al (2022)<[doi:10.1002/sim.9513](https://doi.org/10.1002/sim.9513)> and Archer et al (2024)<[doi:10.1186/s13045-024-01553-6](https://doi.org/10.1186/s13045-024-01553-6)>. False discovery rate controlled variable selection is provided using model-X knock-offs.

**License** MIT + file LICENSE

**Encoding** UTF-8

**Depends** R (>= 4.2.0)

**Imports** doParallel, flexsurv, flexsurvcure, foreach, ggplot2, ggpubr, glmnet, knockoff, mvnfast, parallel, plyr, methods, survival

**Roxygen** list(markdown = TRUE)

**RoxygenNote** 7.3.1

**Suggests** knitr, rmarkdown

**VignetteBuilder** knitr

**LazyData** true

**Repository** <https://kelliejarcher.r-universe.dev>

**RemoteUrl** <https://github.com/kelliejarcher/hdcuremodels>

**RemoteRef** HEAD

**RemoteSha** f15c26a85058bdfe0b0bc7f5d55e0546b7880348

## Contents

amltest . . . . .	2
amltrain . . . . .	11

AUC . . . . .	20
coef.mixturecure . . . . .	22
concordance_mcm . . . . .	23
cureem . . . . .	24
curegmifs . . . . .	27
cure_estimate . . . . .	29
cv_cureem . . . . .	30
cv_curegmifs . . . . .	35
generate_cure_data . . . . .	38
nonzerocure_test . . . . .	40
plot.mixturecure . . . . .	41
predict.mixturecure . . . . .	42
print.mixturecure . . . . .	43
sufficient_fu_test . . . . .	44
summary.mixturecure . . . . .	45

<b>Index</b>	<b>47</b>
--------------	-----------

---

amltest	<i>AML test data</i>
---------	----------------------

---

## Description

Duration of complete response for 40 cytogenetically normal AML patients and a subset of 320 transcript expression from RNA-sequencing.

## Usage

```
amltest
```

## Format

A data frame with 40 rows (subjects) and 322 columns:

**crry** duration of complete response in years

**relapse.death** censoring indicator: 1 = relapsed or died; 0 = alive at last follow-up

**ENSG0000001561** normalized expression for indicated transcript

**ENSG0000005249** normalized expression for indicated transcript

**ENSG0000006757** normalized expression for indicated transcript

**ENSG0000007062** normalized expression for indicated transcript

**ENSG0000007968** normalized expression for indicated transcript

**ENSG0000008283** normalized expression for indicated transcript

**ENSG0000008405** normalized expression for indicated transcript

**ENSG0000008441** normalized expression for indicated transcript

**ENSG0000010295** normalized expression for indicated transcript

**ENSG00000011028** normalized expression for indicated transcript  
**ENSG00000011198** normalized expression for indicated transcript  
**ENSG00000012779** normalized expression for indicated transcript  
**ENSG00000012817** normalized expression for indicated transcript  
**ENSG00000013306** normalized expression for indicated transcript  
**ENSG00000013725** normalized expression for indicated transcript  
**ENSG00000018189** normalized expression for indicated transcript  
**ENSG00000022267** normalized expression for indicated transcript  
**ENSG00000023171** normalized expression for indicated transcript  
**ENSG00000023909** normalized expression for indicated transcript  
**ENSG00000029639** normalized expression for indicated transcript  
**ENSG00000047634** normalized expression for indicated transcript  
**ENSG00000049192** normalized expression for indicated transcript  
**ENSG00000053524** normalized expression for indicated transcript  
**ENSG00000058056** normalized expression for indicated transcript  
**ENSG00000060138** normalized expression for indicated transcript  
**ENSG00000061918** normalized expression for indicated transcript  
**ENSG00000065809** normalized expression for indicated transcript  
**ENSG00000065923** normalized expression for indicated transcript  
**ENSG00000068489** normalized expression for indicated transcript  
**ENSG00000069020** normalized expression for indicated transcript  
**ENSG00000070404** normalized expression for indicated transcript  
**ENSG00000071894** normalized expression for indicated transcript  
**ENSG00000072422** normalized expression for indicated transcript  
**ENSG00000073605** normalized expression for indicated transcript  
**ENSG00000076555** normalized expression for indicated transcript  
**ENSG00000080823** normalized expression for indicated transcript  
**ENSG00000089723** normalized expression for indicated transcript  
**ENSG00000090382** normalized expression for indicated transcript  
**ENSG00000090975** normalized expression for indicated transcript  
**ENSG00000100068** normalized expression for indicated transcript  
**ENSG00000100077** normalized expression for indicated transcript  
**ENSG00000100299** normalized expression for indicated transcript  
**ENSG00000100376** normalized expression for indicated transcript  
**ENSG00000100418** normalized expression for indicated transcript  
**ENSG00000100448** normalized expression for indicated transcript  
**ENSG00000100596** normalized expression for indicated transcript

**ENSG00000100916** normalized expression for indicated transcript  
**ENSG00000102409** normalized expression for indicated transcript  
**ENSG00000102760** normalized expression for indicated transcript  
**ENSG00000104689** normalized expression for indicated transcript  
**ENSG00000104946** normalized expression for indicated transcript  
**ENSG00000105518** normalized expression for indicated transcript  
**ENSG00000105808** normalized expression for indicated transcript  
**ENSG00000106367** normalized expression for indicated transcript  
**ENSG00000106526** normalized expression for indicated transcript  
**ENSG00000106546** normalized expression for indicated transcript  
**ENSG00000106780** normalized expression for indicated transcript  
**ENSG00000107104** normalized expression for indicated transcript  
**ENSG00000107742** normalized expression for indicated transcript  
**ENSG00000107798** normalized expression for indicated transcript  
**ENSG00000107816** normalized expression for indicated transcript  
**ENSG00000107957** normalized expression for indicated transcript  
**ENSG00000109674** normalized expression for indicated transcript  
**ENSG00000110076** normalized expression for indicated transcript  
**ENSG00000110237** normalized expression for indicated transcript  
**ENSG00000110492** normalized expression for indicated transcript  
**ENSG00000110799** normalized expression for indicated transcript  
**ENSG00000111275** normalized expression for indicated transcript  
**ENSG00000112773** normalized expression for indicated transcript  
**ENSG00000113504** normalized expression for indicated transcript  
**ENSG00000114268** normalized expression for indicated transcript  
**ENSG00000114737** normalized expression for indicated transcript  
**ENSG00000115183** normalized expression for indicated transcript  
**ENSG00000115414** normalized expression for indicated transcript  
**ENSG00000115457** normalized expression for indicated transcript  
**ENSG00000115525** normalized expression for indicated transcript  
**ENSG00000116574** normalized expression for indicated transcript  
**ENSG00000117480** normalized expression for indicated transcript  
**ENSG00000119280** normalized expression for indicated transcript  
**ENSG00000120594** normalized expression for indicated transcript  
**ENSG00000120675** normalized expression for indicated transcript  
**ENSG00000120832** normalized expression for indicated transcript  
**ENSG00000120913** normalized expression for indicated transcript

**ENSG00000121005** normalized expression for indicated transcript  
**ENSG00000121039** normalized expression for indicated transcript  
**ENSG00000121274** normalized expression for indicated transcript  
**ENSG00000123080** normalized expression for indicated transcript  
**ENSG00000123836** normalized expression for indicated transcript  
**ENSG00000124019** normalized expression for indicated transcript  
**ENSG00000124882** normalized expression for indicated transcript  
**ENSG00000126822** normalized expression for indicated transcript  
**ENSG00000127152** normalized expression for indicated transcript  
**ENSG00000129824** normalized expression for indicated transcript  
**ENSG00000130702** normalized expression for indicated transcript  
**ENSG00000131188** normalized expression for indicated transcript  
**ENSG00000131370** normalized expression for indicated transcript  
**ENSG00000132122** normalized expression for indicated transcript  
**ENSG00000132530** normalized expression for indicated transcript  
**ENSG00000132819** normalized expression for indicated transcript  
**ENSG00000132849** normalized expression for indicated transcript  
**ENSG00000133401** normalized expression for indicated transcript  
**ENSG00000133619** normalized expression for indicated transcript  
**ENSG00000134531** normalized expression for indicated transcript  
**ENSG00000134897** normalized expression for indicated transcript  
**ENSG00000135074** normalized expression for indicated transcript  
**ENSG00000135245** normalized expression for indicated transcript  
**ENSG00000135272** normalized expression for indicated transcript  
**ENSG00000135362** normalized expression for indicated transcript  
**ENSG00000135363** normalized expression for indicated transcript  
**ENSG00000135916** normalized expression for indicated transcript  
**ENSG00000136026** normalized expression for indicated transcript  
**ENSG00000136193** normalized expression for indicated transcript  
**ENSG00000136231** normalized expression for indicated transcript  
**ENSG00000136997** normalized expression for indicated transcript  
**ENSG00000137193** normalized expression for indicated transcript  
**ENSG00000137198** normalized expression for indicated transcript  
**ENSG00000138722** normalized expression for indicated transcript  
**ENSG00000139318** normalized expression for indicated transcript  
**ENSG00000140287** normalized expression for indicated transcript  
**ENSG00000144036** normalized expression for indicated transcript

ENSG00000144647 normalized expression for indicated transcript  
ENSG00000144677 normalized expression for indicated transcript  
ENSG00000145476 normalized expression for indicated transcript  
ENSG00000145545 normalized expression for indicated transcript  
ENSG00000146243 normalized expression for indicated transcript  
ENSG00000146373 normalized expression for indicated transcript  
ENSG00000147044 normalized expression for indicated transcript  
ENSG00000147180 normalized expression for indicated transcript  
ENSG00000148444 normalized expression for indicated transcript  
ENSG00000148484 normalized expression for indicated transcript  
ENSG00000149131 normalized expression for indicated transcript  
ENSG00000150760 normalized expression for indicated transcript  
ENSG00000150782 normalized expression for indicated transcript  
ENSG00000151135 normalized expression for indicated transcript  
ENSG00000151208 normalized expression for indicated transcript  
ENSG00000151458 normalized expression for indicated transcript  
ENSG00000152409 normalized expression for indicated transcript  
ENSG00000152580 normalized expression for indicated transcript  
ENSG00000152767 normalized expression for indicated transcript  
ENSG00000152778 normalized expression for indicated transcript  
ENSG00000153563 normalized expression for indicated transcript  
ENSG00000154217 normalized expression for indicated transcript  
ENSG00000154743 normalized expression for indicated transcript  
ENSG00000154760 normalized expression for indicated transcript  
ENSG00000154874 normalized expression for indicated transcript  
ENSG00000156381 normalized expression for indicated transcript  
ENSG00000157107 normalized expression for indicated transcript  
ENSG00000157240 normalized expression for indicated transcript  
ENSG00000157873 normalized expression for indicated transcript  
ENSG00000157978 normalized expression for indicated transcript  
ENSG00000158691 normalized expression for indicated transcript  
ENSG00000159339 normalized expression for indicated transcript  
ENSG00000159403 normalized expression for indicated transcript  
ENSG00000159788 normalized expression for indicated transcript  
ENSG00000160685 normalized expression for indicated transcript  
ENSG00000160781 normalized expression for indicated transcript  
ENSG00000161509 normalized expression for indicated transcript

**ENSG00000162433** normalized expression for indicated transcript  
**ENSG00000162614** normalized expression for indicated transcript  
**ENSG00000162676** normalized expression for indicated transcript  
**ENSG00000163412** normalized expression for indicated transcript  
**ENSG00000163702** normalized expression for indicated transcript  
**ENSG00000163814** normalized expression for indicated transcript  
**ENSG00000164086** normalized expression for indicated transcript  
**ENSG00000164172** normalized expression for indicated transcript  
**ENSG00000164442** normalized expression for indicated transcript  
**ENSG00000165272** normalized expression for indicated transcript  
**ENSG00000166165** normalized expression for indicated transcript  
**ENSG00000166435** normalized expression for indicated transcript  
**ENSG00000166987** normalized expression for indicated transcript  
**ENSG00000167291** normalized expression for indicated transcript  
**ENSG00000167565** normalized expression for indicated transcript  
**ENSG00000167851** normalized expression for indicated transcript  
**ENSG00000168026** normalized expression for indicated transcript  
**ENSG00000168209** normalized expression for indicated transcript  
**ENSG00000168502** normalized expression for indicated transcript  
**ENSG00000168939** normalized expression for indicated transcript  
**ENSG00000169203** normalized expression for indicated transcript  
**ENSG00000169247** normalized expression for indicated transcript  
**ENSG00000169504** normalized expression for indicated transcript  
**ENSG00000169860** normalized expression for indicated transcript  
**ENSG00000169991** normalized expression for indicated transcript  
**ENSG00000170035** normalized expression for indicated transcript  
**ENSG00000170180** normalized expression for indicated transcript  
**ENSG00000170456** normalized expression for indicated transcript  
**ENSG00000170522** normalized expression for indicated transcript  
**ENSG00000170909** normalized expression for indicated transcript  
**ENSG00000171121** normalized expression for indicated transcript  
**ENSG00000171222** normalized expression for indicated transcript  
**ENSG00000171476** normalized expression for indicated transcript  
**ENSG00000171813** normalized expression for indicated transcript  
**ENSG00000171962** normalized expression for indicated transcript  
**ENSG00000172197** normalized expression for indicated transcript  
**ENSG00000172236** normalized expression for indicated transcript

**ENSG00000173083** normalized expression for indicated transcript  
**ENSG00000173530** normalized expression for indicated transcript  
**ENSG00000173926** normalized expression for indicated transcript  
**ENSG00000174059** normalized expression for indicated transcript  
**ENSG00000174080** normalized expression for indicated transcript  
**ENSG00000174130** normalized expression for indicated transcript  
**ENSG00000174738** normalized expression for indicated transcript  
**ENSG00000175265** normalized expression for indicated transcript  
**ENSG00000175352** normalized expression for indicated transcript  
**ENSG00000176597** normalized expression for indicated transcript  
**ENSG00000179222** normalized expression for indicated transcript  
**ENSG00000179630** normalized expression for indicated transcript  
**ENSG00000179639** normalized expression for indicated transcript  
**ENSG00000179820** normalized expression for indicated transcript  
**ENSG00000180096** normalized expression for indicated transcript  
**ENSG00000180596** normalized expression for indicated transcript  
**ENSG00000180902** normalized expression for indicated transcript  
**ENSG00000181104** normalized expression for indicated transcript  
**ENSG00000182866** normalized expression for indicated transcript  
**ENSG00000182871** normalized expression for indicated transcript  
**ENSG00000183087** normalized expression for indicated transcript  
**ENSG00000183091** normalized expression for indicated transcript  
**ENSG00000184371** normalized expression for indicated transcript  
**ENSG00000185129** normalized expression for indicated transcript  
**ENSG00000185201** normalized expression for indicated transcript  
**ENSG00000185245** normalized expression for indicated transcript  
**ENSG00000185291** normalized expression for indicated transcript  
**ENSG00000185304** normalized expression for indicated transcript  
**ENSG00000185710** normalized expression for indicated transcript  
**ENSG00000185883** normalized expression for indicated transcript  
**ENSG00000185986** normalized expression for indicated transcript  
**ENSG00000186130** normalized expression for indicated transcript  
**ENSG00000186854** normalized expression for indicated transcript  
**ENSG00000187010** normalized expression for indicated transcript  
**ENSG00000187627** normalized expression for indicated transcript  
**ENSG00000187653** normalized expression for indicated transcript  
**ENSG00000187837** normalized expression for indicated transcript

**ENSG00000188002** normalized expression for indicated transcript  
**ENSG00000188107** normalized expression for indicated transcript  
**ENSG00000188211** normalized expression for indicated transcript  
**ENSG00000188636** normalized expression for indicated transcript  
**ENSG00000188738** normalized expression for indicated transcript  
**ENSG00000188856** normalized expression for indicated transcript  
**ENSG00000189164** normalized expression for indicated transcript  
**ENSG00000189223** normalized expression for indicated transcript  
**ENSG00000196155** normalized expression for indicated transcript  
**ENSG00000196189** normalized expression for indicated transcript  
**ENSG00000196415** normalized expression for indicated transcript  
**ENSG00000196565** normalized expression for indicated transcript  
**ENSG00000197081** normalized expression for indicated transcript  
**ENSG00000197121** normalized expression for indicated transcript  
**ENSG00000197253** normalized expression for indicated transcript  
**ENSG00000197256** normalized expression for indicated transcript  
**ENSG00000197321** normalized expression for indicated transcript  
**ENSG00000197561** normalized expression for indicated transcript  
**ENSG00000197728** normalized expression for indicated transcript  
**ENSG00000197860** normalized expression for indicated transcript  
**ENSG00000197937** normalized expression for indicated transcript  
**ENSG00000197951** normalized expression for indicated transcript  
**ENSG00000198743** normalized expression for indicated transcript  
**ENSG00000198838** normalized expression for indicated transcript  
**ENSG00000199347** normalized expression for indicated transcript  
**ENSG00000200243** normalized expression for indicated transcript  
**ENSG00000201801** normalized expression for indicated transcript  
**ENSG00000203872** normalized expression for indicated transcript  
**ENSG00000204172** normalized expression for indicated transcript  
**ENSG00000205571** normalized expression for indicated transcript  
**ENSG00000205593** normalized expression for indicated transcript  
**ENSG00000208772** normalized expression for indicated transcript  
**ENSG00000213085** normalized expression for indicated transcript  
**ENSG00000213261** normalized expression for indicated transcript  
**ENSG00000213626** normalized expression for indicated transcript  
**ENSG00000213722** normalized expression for indicated transcript  
**ENSG00000213906** normalized expression for indicated transcript

ENSG00000213967 normalized expression for indicated transcript  
ENSG00000214016 normalized expression for indicated transcript  
ENSG00000214425 normalized expression for indicated transcript  
ENSG00000216316 normalized expression for indicated transcript  
ENSG00000220008 normalized expression for indicated transcript  
ENSG00000223345 normalized expression for indicated transcript  
ENSG00000224080 normalized expression for indicated transcript  
ENSG00000225138 normalized expression for indicated transcript  
ENSG00000226471 normalized expression for indicated transcript  
ENSG00000227097 normalized expression for indicated transcript  
ENSG00000227191 normalized expression for indicated transcript  
ENSG00000227615 normalized expression for indicated transcript  
ENSG00000228049 normalized expression for indicated transcript  
ENSG00000229153 normalized expression for indicated transcript  
ENSG00000230076 normalized expression for indicated transcript  
ENSG00000231160 normalized expression for indicated transcript  
ENSG00000231721 normalized expression for indicated transcript  
ENSG00000233927 normalized expression for indicated transcript  
ENSG00000233974 normalized expression for indicated transcript  
ENSG00000234883 normalized expression for indicated transcript  
ENSG00000236876 normalized expression for indicated transcript  
ENSG00000237298 normalized expression for indicated transcript  
ENSG00000237892 normalized expression for indicated transcript  
ENSG00000238160 normalized expression for indicated transcript  
ENSG00000239437 normalized expression for indicated transcript  
ENSG00000241399 normalized expression for indicated transcript  
ENSG00000241489 normalized expression for indicated transcript  
ENSG00000241529 normalized expression for indicated transcript  
ENSG00000244405 normalized expression for indicated transcript  
ENSG00000247627 normalized expression for indicated transcript  
ENSG00000249592 normalized expression for indicated transcript  
ENSG00000250116 normalized expression for indicated transcript  
ENSG00000250251 normalized expression for indicated transcript  
ENSG00000251079 normalized expression for indicated transcript  
ENSG00000253210 normalized expression for indicated transcript  
ENSG00000253276 normalized expression for indicated transcript  
ENSG00000254415 normalized expression for indicated transcript

**ENSG00000259276** normalized expression for indicated transcript  
**ENSG00000260727** normalized expression for indicated transcript  
**ENSG00000261377** normalized expression for indicated transcript  
**ENSG00000264885** normalized expression for indicated transcript  
**ENSG00000264895** normalized expression for indicated transcript  
**ENSG00000267136** normalized expression for indicated transcript  
**ENSG00000267551** normalized expression for indicated transcript  
**ENSG00000267702** normalized expression for indicated transcript  
**ENSG00000268001** normalized expression for indicated transcript  
**ENSG00000268573** normalized expression for indicated transcript  
**ENSG00000270554** normalized expression for indicated transcript  
**ENSG00000270562** normalized expression for indicated transcript  
**ENSG00000271646** normalized expression for indicated transcript  
**ENSG00000273018** normalized expression for indicated transcript  
**ENSG00000273033** normalized expression for indicated transcript

#### Source

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11068580/>

---

amltrain

*AML training data*

---

#### Description

Duration of complete response for 306 cytogenetically normal AML patients and a subset of 320 transcript expression from RNA-sequencing.

#### Usage

amltrain

#### Format

A data frame with 306 rows (subjects) and 322 columns:

**crryr** duration of complete response in years

**relapse.death** censoring indicator: 1 = relapsed or died; 0 = alive at last follow-up

**ENSG00000001561** normalized expression for indicated transcript

**ENSG00000005249** normalized expression for indicated transcript

**ENSG00000006757** normalized expression for indicated transcript

**ENSG00000007062** normalized expression for indicated transcript

**ENSG0000007968** normalized expression for indicated transcript  
**ENSG0000008283** normalized expression for indicated transcript  
**ENSG0000008405** normalized expression for indicated transcript  
**ENSG0000008441** normalized expression for indicated transcript  
**ENSG0000010295** normalized expression for indicated transcript  
**ENSG0000011028** normalized expression for indicated transcript  
**ENSG0000011198** normalized expression for indicated transcript  
**ENSG0000012779** normalized expression for indicated transcript  
**ENSG0000012817** normalized expression for indicated transcript  
**ENSG0000013306** normalized expression for indicated transcript  
**ENSG0000013725** normalized expression for indicated transcript  
**ENSG0000018189** normalized expression for indicated transcript  
**ENSG0000022267** normalized expression for indicated transcript  
**ENSG0000023171** normalized expression for indicated transcript  
**ENSG0000023909** normalized expression for indicated transcript  
**ENSG0000029639** normalized expression for indicated transcript  
**ENSG0000047634** normalized expression for indicated transcript  
**ENSG0000049192** normalized expression for indicated transcript  
**ENSG0000053524** normalized expression for indicated transcript  
**ENSG0000058056** normalized expression for indicated transcript  
**ENSG0000060138** normalized expression for indicated transcript  
**ENSG0000061918** normalized expression for indicated transcript  
**ENSG0000065809** normalized expression for indicated transcript  
**ENSG0000065923** normalized expression for indicated transcript  
**ENSG0000068489** normalized expression for indicated transcript  
**ENSG0000069020** normalized expression for indicated transcript  
**ENSG0000070404** normalized expression for indicated transcript  
**ENSG0000071894** normalized expression for indicated transcript  
**ENSG0000072422** normalized expression for indicated transcript  
**ENSG0000073605** normalized expression for indicated transcript  
**ENSG0000076555** normalized expression for indicated transcript  
**ENSG0000080823** normalized expression for indicated transcript  
**ENSG0000089723** normalized expression for indicated transcript  
**ENSG0000090382** normalized expression for indicated transcript  
**ENSG0000090975** normalized expression for indicated transcript  
**ENSG0000100068** normalized expression for indicated transcript  
**ENSG0000100077** normalized expression for indicated transcript

**ENSG00000100299** normalized expression for indicated transcript  
**ENSG00000100376** normalized expression for indicated transcript  
**ENSG00000100418** normalized expression for indicated transcript  
**ENSG00000100448** normalized expression for indicated transcript  
**ENSG00000100596** normalized expression for indicated transcript  
**ENSG00000100916** normalized expression for indicated transcript  
**ENSG00000102409** normalized expression for indicated transcript  
**ENSG00000102760** normalized expression for indicated transcript  
**ENSG00000104689** normalized expression for indicated transcript  
**ENSG00000104946** normalized expression for indicated transcript  
**ENSG00000105518** normalized expression for indicated transcript  
**ENSG00000105808** normalized expression for indicated transcript  
**ENSG00000106367** normalized expression for indicated transcript  
**ENSG00000106526** normalized expression for indicated transcript  
**ENSG00000106546** normalized expression for indicated transcript  
**ENSG00000106780** normalized expression for indicated transcript  
**ENSG00000107104** normalized expression for indicated transcript  
**ENSG00000107742** normalized expression for indicated transcript  
**ENSG00000107798** normalized expression for indicated transcript  
**ENSG00000107816** normalized expression for indicated transcript  
**ENSG00000107957** normalized expression for indicated transcript  
**ENSG00000109674** normalized expression for indicated transcript  
**ENSG00000110076** normalized expression for indicated transcript  
**ENSG00000110237** normalized expression for indicated transcript  
**ENSG00000110492** normalized expression for indicated transcript  
**ENSG00000110799** normalized expression for indicated transcript  
**ENSG00000111275** normalized expression for indicated transcript  
**ENSG00000112773** normalized expression for indicated transcript  
**ENSG00000113504** normalized expression for indicated transcript  
**ENSG00000114268** normalized expression for indicated transcript  
**ENSG00000114737** normalized expression for indicated transcript  
**ENSG00000115183** normalized expression for indicated transcript  
**ENSG00000115414** normalized expression for indicated transcript  
**ENSG00000115457** normalized expression for indicated transcript  
**ENSG00000115525** normalized expression for indicated transcript  
**ENSG00000116574** normalized expression for indicated transcript  
**ENSG00000117480** normalized expression for indicated transcript

**ENSG00000119280** normalized expression for indicated transcript  
**ENSG00000120594** normalized expression for indicated transcript  
**ENSG00000120675** normalized expression for indicated transcript  
**ENSG00000120832** normalized expression for indicated transcript  
**ENSG00000120913** normalized expression for indicated transcript  
**ENSG00000121005** normalized expression for indicated transcript  
**ENSG00000121039** normalized expression for indicated transcript  
**ENSG00000121274** normalized expression for indicated transcript  
**ENSG00000123080** normalized expression for indicated transcript  
**ENSG00000123836** normalized expression for indicated transcript  
**ENSG00000124019** normalized expression for indicated transcript  
**ENSG00000124882** normalized expression for indicated transcript  
**ENSG00000126822** normalized expression for indicated transcript  
**ENSG00000127152** normalized expression for indicated transcript  
**ENSG00000129824** normalized expression for indicated transcript  
**ENSG00000130702** normalized expression for indicated transcript  
**ENSG00000131188** normalized expression for indicated transcript  
**ENSG00000131370** normalized expression for indicated transcript  
**ENSG00000132122** normalized expression for indicated transcript  
**ENSG00000132530** normalized expression for indicated transcript  
**ENSG00000132819** normalized expression for indicated transcript  
**ENSG00000132849** normalized expression for indicated transcript  
**ENSG00000133401** normalized expression for indicated transcript  
**ENSG00000133619** normalized expression for indicated transcript  
**ENSG00000134531** normalized expression for indicated transcript  
**ENSG00000134897** normalized expression for indicated transcript  
**ENSG00000135074** normalized expression for indicated transcript  
**ENSG00000135245** normalized expression for indicated transcript  
**ENSG00000135272** normalized expression for indicated transcript  
**ENSG00000135362** normalized expression for indicated transcript  
**ENSG00000135363** normalized expression for indicated transcript  
**ENSG00000135916** normalized expression for indicated transcript  
**ENSG00000136026** normalized expression for indicated transcript  
**ENSG00000136193** normalized expression for indicated transcript  
**ENSG00000136231** normalized expression for indicated transcript  
**ENSG00000136997** normalized expression for indicated transcript  
**ENSG00000137193** normalized expression for indicated transcript

ENSG00000137198 normalized expression for indicated transcript  
ENSG00000138722 normalized expression for indicated transcript  
ENSG00000139318 normalized expression for indicated transcript  
ENSG00000140287 normalized expression for indicated transcript  
ENSG00000144036 normalized expression for indicated transcript  
ENSG00000144647 normalized expression for indicated transcript  
ENSG00000144677 normalized expression for indicated transcript  
ENSG00000145476 normalized expression for indicated transcript  
ENSG00000145545 normalized expression for indicated transcript  
ENSG00000146243 normalized expression for indicated transcript  
ENSG00000146373 normalized expression for indicated transcript  
ENSG00000147044 normalized expression for indicated transcript  
ENSG00000147180 normalized expression for indicated transcript  
ENSG00000148444 normalized expression for indicated transcript  
ENSG00000148484 normalized expression for indicated transcript  
ENSG00000149131 normalized expression for indicated transcript  
ENSG00000150760 normalized expression for indicated transcript  
ENSG00000150782 normalized expression for indicated transcript  
ENSG00000151135 normalized expression for indicated transcript  
ENSG00000151208 normalized expression for indicated transcript  
ENSG00000151458 normalized expression for indicated transcript  
ENSG00000152409 normalized expression for indicated transcript  
ENSG00000152580 normalized expression for indicated transcript  
ENSG00000152767 normalized expression for indicated transcript  
ENSG00000152778 normalized expression for indicated transcript  
ENSG00000153563 normalized expression for indicated transcript  
ENSG00000154217 normalized expression for indicated transcript  
ENSG00000154743 normalized expression for indicated transcript  
ENSG00000154760 normalized expression for indicated transcript  
ENSG00000154874 normalized expression for indicated transcript  
ENSG00000156381 normalized expression for indicated transcript  
ENSG00000157107 normalized expression for indicated transcript  
ENSG00000157240 normalized expression for indicated transcript  
ENSG00000157873 normalized expression for indicated transcript  
ENSG00000157978 normalized expression for indicated transcript  
ENSG00000158691 normalized expression for indicated transcript  
ENSG00000159339 normalized expression for indicated transcript

**ENSG00000159403** normalized expression for indicated transcript  
**ENSG00000159788** normalized expression for indicated transcript  
**ENSG00000160685** normalized expression for indicated transcript  
**ENSG00000160781** normalized expression for indicated transcript  
**ENSG00000161509** normalized expression for indicated transcript  
**ENSG00000162433** normalized expression for indicated transcript  
**ENSG00000162614** normalized expression for indicated transcript  
**ENSG00000162676** normalized expression for indicated transcript  
**ENSG00000163412** normalized expression for indicated transcript  
**ENSG00000163702** normalized expression for indicated transcript  
**ENSG00000163814** normalized expression for indicated transcript  
**ENSG00000164086** normalized expression for indicated transcript  
**ENSG00000164172** normalized expression for indicated transcript  
**ENSG00000164442** normalized expression for indicated transcript  
**ENSG00000165272** normalized expression for indicated transcript  
**ENSG00000166165** normalized expression for indicated transcript  
**ENSG00000166435** normalized expression for indicated transcript  
**ENSG00000166987** normalized expression for indicated transcript  
**ENSG00000167291** normalized expression for indicated transcript  
**ENSG00000167565** normalized expression for indicated transcript  
**ENSG00000167851** normalized expression for indicated transcript  
**ENSG00000168026** normalized expression for indicated transcript  
**ENSG00000168209** normalized expression for indicated transcript  
**ENSG00000168502** normalized expression for indicated transcript  
**ENSG00000168939** normalized expression for indicated transcript  
**ENSG00000169203** normalized expression for indicated transcript  
**ENSG00000169247** normalized expression for indicated transcript  
**ENSG00000169504** normalized expression for indicated transcript  
**ENSG00000169860** normalized expression for indicated transcript  
**ENSG00000169991** normalized expression for indicated transcript  
**ENSG00000170035** normalized expression for indicated transcript  
**ENSG00000170180** normalized expression for indicated transcript  
**ENSG00000170456** normalized expression for indicated transcript  
**ENSG00000170522** normalized expression for indicated transcript  
**ENSG00000170909** normalized expression for indicated transcript  
**ENSG00000171121** normalized expression for indicated transcript  
**ENSG00000171222** normalized expression for indicated transcript

ENSG00000171476 normalized expression for indicated transcript  
ENSG00000171813 normalized expression for indicated transcript  
ENSG00000171962 normalized expression for indicated transcript  
ENSG00000172197 normalized expression for indicated transcript  
ENSG00000172236 normalized expression for indicated transcript  
ENSG00000173083 normalized expression for indicated transcript  
ENSG00000173530 normalized expression for indicated transcript  
ENSG00000173926 normalized expression for indicated transcript  
ENSG00000174059 normalized expression for indicated transcript  
ENSG00000174080 normalized expression for indicated transcript  
ENSG00000174130 normalized expression for indicated transcript  
ENSG00000174738 normalized expression for indicated transcript  
ENSG00000175265 normalized expression for indicated transcript  
ENSG00000175352 normalized expression for indicated transcript  
ENSG00000176597 normalized expression for indicated transcript  
ENSG00000179222 normalized expression for indicated transcript  
ENSG00000179630 normalized expression for indicated transcript  
ENSG00000179639 normalized expression for indicated transcript  
ENSG00000179820 normalized expression for indicated transcript  
ENSG00000180096 normalized expression for indicated transcript  
ENSG00000180596 normalized expression for indicated transcript  
ENSG00000180902 normalized expression for indicated transcript  
ENSG00000181104 normalized expression for indicated transcript  
ENSG00000182866 normalized expression for indicated transcript  
ENSG00000182871 normalized expression for indicated transcript  
ENSG00000183087 normalized expression for indicated transcript  
ENSG00000183091 normalized expression for indicated transcript  
ENSG00000184371 normalized expression for indicated transcript  
ENSG00000185129 normalized expression for indicated transcript  
ENSG00000185201 normalized expression for indicated transcript  
ENSG00000185245 normalized expression for indicated transcript  
ENSG00000185291 normalized expression for indicated transcript  
ENSG00000185304 normalized expression for indicated transcript  
ENSG00000185710 normalized expression for indicated transcript  
ENSG00000185883 normalized expression for indicated transcript  
ENSG00000185986 normalized expression for indicated transcript  
ENSG00000186130 normalized expression for indicated transcript

**ENSG00000186854** normalized expression for indicated transcript  
**ENSG00000187010** normalized expression for indicated transcript  
**ENSG00000187627** normalized expression for indicated transcript  
**ENSG00000187653** normalized expression for indicated transcript  
**ENSG00000187837** normalized expression for indicated transcript  
**ENSG00000188002** normalized expression for indicated transcript  
**ENSG00000188107** normalized expression for indicated transcript  
**ENSG00000188211** normalized expression for indicated transcript  
**ENSG00000188636** normalized expression for indicated transcript  
**ENSG00000188738** normalized expression for indicated transcript  
**ENSG00000188856** normalized expression for indicated transcript  
**ENSG00000189164** normalized expression for indicated transcript  
**ENSG00000189223** normalized expression for indicated transcript  
**ENSG00000196155** normalized expression for indicated transcript  
**ENSG00000196189** normalized expression for indicated transcript  
**ENSG00000196415** normalized expression for indicated transcript  
**ENSG00000196565** normalized expression for indicated transcript  
**ENSG00000197081** normalized expression for indicated transcript  
**ENSG00000197121** normalized expression for indicated transcript  
**ENSG00000197253** normalized expression for indicated transcript  
**ENSG00000197256** normalized expression for indicated transcript  
**ENSG00000197321** normalized expression for indicated transcript  
**ENSG00000197561** normalized expression for indicated transcript  
**ENSG00000197728** normalized expression for indicated transcript  
**ENSG00000197860** normalized expression for indicated transcript  
**ENSG00000197937** normalized expression for indicated transcript  
**ENSG00000197951** normalized expression for indicated transcript  
**ENSG00000198743** normalized expression for indicated transcript  
**ENSG00000198838** normalized expression for indicated transcript  
**ENSG00000199347** normalized expression for indicated transcript  
**ENSG00000200243** normalized expression for indicated transcript  
**ENSG00000201801** normalized expression for indicated transcript  
**ENSG00000203872** normalized expression for indicated transcript  
**ENSG00000204172** normalized expression for indicated transcript  
**ENSG00000205571** normalized expression for indicated transcript  
**ENSG00000205593** normalized expression for indicated transcript  
**ENSG00000208772** normalized expression for indicated transcript

ENSG00000213085 normalized expression for indicated transcript  
ENSG00000213261 normalized expression for indicated transcript  
ENSG00000213626 normalized expression for indicated transcript  
ENSG00000213722 normalized expression for indicated transcript  
ENSG00000213906 normalized expression for indicated transcript  
ENSG00000213967 normalized expression for indicated transcript  
ENSG00000214016 normalized expression for indicated transcript  
ENSG00000214425 normalized expression for indicated transcript  
ENSG00000216316 normalized expression for indicated transcript  
ENSG00000220008 normalized expression for indicated transcript  
ENSG00000223345 normalized expression for indicated transcript  
ENSG00000224080 normalized expression for indicated transcript  
ENSG00000225138 normalized expression for indicated transcript  
ENSG00000226471 normalized expression for indicated transcript  
ENSG00000227097 normalized expression for indicated transcript  
ENSG00000227191 normalized expression for indicated transcript  
ENSG00000227615 normalized expression for indicated transcript  
ENSG00000228049 normalized expression for indicated transcript  
ENSG00000229153 normalized expression for indicated transcript  
ENSG00000230076 normalized expression for indicated transcript  
ENSG00000231160 normalized expression for indicated transcript  
ENSG00000231721 normalized expression for indicated transcript  
ENSG00000233927 normalized expression for indicated transcript  
ENSG00000233974 normalized expression for indicated transcript  
ENSG00000234883 normalized expression for indicated transcript  
ENSG00000236876 normalized expression for indicated transcript  
ENSG00000237298 normalized expression for indicated transcript  
ENSG00000237892 normalized expression for indicated transcript  
ENSG00000238160 normalized expression for indicated transcript  
ENSG00000239437 normalized expression for indicated transcript  
ENSG00000241399 normalized expression for indicated transcript  
ENSG00000241489 normalized expression for indicated transcript  
ENSG00000241529 normalized expression for indicated transcript  
ENSG00000244405 normalized expression for indicated transcript  
ENSG00000247627 normalized expression for indicated transcript  
ENSG00000249592 normalized expression for indicated transcript  
ENSG00000250116 normalized expression for indicated transcript

**ENSG00000250251** normalized expression for indicated transcript  
**ENSG00000251079** normalized expression for indicated transcript  
**ENSG00000253210** normalized expression for indicated transcript  
**ENSG00000253276** normalized expression for indicated transcript  
**ENSG00000254415** normalized expression for indicated transcript  
**ENSG00000259276** normalized expression for indicated transcript  
**ENSG00000260727** normalized expression for indicated transcript  
**ENSG00000261377** normalized expression for indicated transcript  
**ENSG00000264885** normalized expression for indicated transcript  
**ENSG00000264895** normalized expression for indicated transcript  
**ENSG00000267136** normalized expression for indicated transcript  
**ENSG00000267551** normalized expression for indicated transcript  
**ENSG00000267702** normalized expression for indicated transcript  
**ENSG00000268001** normalized expression for indicated transcript  
**ENSG00000268573** normalized expression for indicated transcript  
**ENSG00000270554** normalized expression for indicated transcript  
**ENSG00000270562** normalized expression for indicated transcript  
**ENSG00000271646** normalized expression for indicated transcript  
**ENSG00000273018** normalized expression for indicated transcript  
**ENSG00000273033** normalized expression for indicated transcript

### Source

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11068580/>

---

AUC

*AUC for cure prediction using mean score imputation*

---

### Description

This function calculates the AUC for cure prediction using the mean score imputation (MSI) method proposed by Asano et al.

### Usage

```
AUC(object, newdata, cure_cutoff = 5, model.select = "AIC")
```

**Arguments**

<code>object</code>	a <code>mixturecure</code> object resulting from <code>curegmifs</code> , <code>cureem</code> , <code>cv_curegmifs</code> , <code>cv_cureem</code> .
<code>newdata</code>	an optional <code>data.frame</code> that minimally includes the incidence and/or latency variables to use for predicting the response. If omitted, the training data are used.
<code>cure_cutoff</code>	cutoff value for cure, used to produce a proxy for the unobserved cure status; default is 5.
<code>model.select</code>	for models fit using <code>curegmifs</code> or <code>cureem</code> any step along the solution path can be selected. The default is <code>model.select = "AIC"</code> which calculates the predicted values using the coefficients from the model having the lowest AIC. Other options are <code>model.select = "mAIC"</code> for the modified AIC, <code>model.select = "cAIC"</code> for the corrected AIC, <code>model.select = "BIC"</code> , <code>model.select = "mBIC"</code> for the modified BIC, <code>model.select = "EBIC"</code> for the extended BIC, <code>model.select = "logLik"</code> for the step that maximizes the log-likelihood, or any numeric value from the solution path. This option has no effect for objects fit using <code>cv_curegmifs</code> or <code>cv_cureem</code> .

**Value**

Returns the AUC value for cure prediction using the mean score imputation (MSI) method.

**References**

Asano, J., Hirakawa, H., Hamada, C. (2014) Assessing the prediction accuracy of cure in the Cox proportional hazards cure model: an application to breast cancer data. *Pharmaceutical Statistics*, **13**:357–363.

**See Also**

[concordance\\_mcm](#)

**Examples**

```
library(survival)
set.seed(1234)
temp <- generate_cure_data(N = 100, J = 10, nTrue = 10, A = 1.8)
training <- temp$Training
testing <- temp$Testing
fit <- curegmifs(Surv(Time, Censor) ~ .,
                 data = training, x.latency = training,
                 model = "weibull", thresh = 1e-4, maxit = 2000,
                 epsilon = 0.01, verbose = FALSE)
AUC(fit)
AUC(fit, newdata = testing)
```

---

coef.mixturecure      *Extract model coefficients from a fitted mixture cure object*

---

### Description

coef.mixturecure is a generic function which extracts the model coefficients from a fitted mixture cure model object fit using curegmifs, cureem, cv\_curegmifs, or cv\_cureem.

### Usage

```
## S3 method for class 'mixturecure'
coef(object, model.select = "AIC", ...)
```

### Arguments

object	a mixturecure object resulting from curegmifs, cureem, cv_curegmifs, or cv_cureem.
model.select	for models fit using curegmifs or cureem any step along the solution path can be selected. The default is model.select = "AIC" which calculates the predicted values using the coefficients from the model having the lowest AIC. Other options are model.select = "mAIC" for the modified AIC, model.select = "cAIC" for the corrected AIC, model.select = "BIC", model.select = "mBIC" for the modified BIC, model.select = "EBIC" for the extended BIC, model.select = "logLik" for the step that maximizes the log-likelihood, or any numeric value from the solution path. This option has no effect for objects fit using cv_curegmifs or cv_cureem.
...	other arguments.

### Value

a list of estimated parameters extracted from the model object using the model selection criterion

### See Also

[curegmifs](#), [cureem](#), [summary.mixturecure](#), [plot.mixturecure](#), [predict.mixturecure](#)

### Examples

```
library(survival)
set.seed(1234)
temp <- generate_cure_data(N = 100, J = 10, nTrue = 10, A = 1.8)
training <- temp$Training
fit <- curegmifs(Surv(Time, Censor) ~ .,
                 data = training, x.latency = training,
                 model = "weibull", thresh = 1e-4, maxit = 2000, epsilon = 0.01,
                 verbose = FALSE)
coef(fit)
```

---

concordance_mcm	<i>C-statistic for mixture cure models</i>
-----------------	--

---

### Description

This function calculates the C-statistic using the cure status weighting (CSW) method proposed by Asano and Hirakawa.

### Usage

```
concordance_mcm(object, newdata, cure_cutoff = 5, model.select = "AIC")
```

### Arguments

object	a mixturecure object resulting from curegmifs, cureem, cv_curegmifs, cv_cureem.
newdata	an optional data.frame that minimally includes the incidence and/or latency variables to use for predicting the response. If omitted, the training data are used.
cure_cutoff	cutoff value for cure, used to produce a proxy for the unobserved cure status; default is 5.
model.select	for models fit using curegmifs or cureem any step along the solution path can be selected. The default is model.select = "AIC" which calculates the predicted values using the coefficients from the model having the lowest AIC. Other options are model.select = "mAIC" for the modified AIC, model.select = "cAIC" for the corrected AIC, model.select = "BIC", model.select = "mBIC" for the modified BIC, model.select = "EBIC" for the extended BIC, model.select = "logLik" for the step that maximizes the log-likelihood, or any numeric value from the solution path. This option has no effect for objects fit using cv_curegmifs or cv_cureem.

### Value

value of C-statistic for the cure models.

### References

Asano, J. and Hirakawa, H. (2017) Assessing the prediction accuracy of a cure model for censored survival data with long-term survivors: Application to breast cancer data. *Journal of Biopharmaceutical Statistics*, **27**:6, 918–932.

### See Also

[AUC](#)

**Examples**

```

library(survival)
set.seed(1234)
temp <- generate_cure_data(N = 100, J = 10, nTrue = 10, A = 1.8)
training <- temp$Training
testing <- temp$Testing
fit <- curegmifs(Surv(Time, Censor) ~ .,
                 data = training, x.latency = training,
                 model = "weibull", thresh = 1e-4, maxit = 2000,
                 epsilon = 0.01, verbose = FALSE)
concordance_mcm(fit)
concordance_mcm(fit, newdata = testing)

```

cureem

*Fit penalized mixture cure model using the E-M algorithm***Description**

Fits a penalized parametric and semi-parametric mixture cure model (MCM) using the E-M algorithm with user-specified penalty parameters. The lasso (L1), MCP, and SCAD penalty is supported for the Cox MCM while only lasso is currently supported for parametric MCMs.

**Usage**

```

cureem(
  formula,
  data,
  subset,
  x.latency = NULL,
  model = "cox",
  penalty = "lasso",
  penalty.factor.inc = NULL,
  penalty.factor.lat = NULL,
  thresh = 0.001,
  scale = TRUE,
  maxit = NULL,
  inits = NULL,
  lambda.inc = 0.1,
  lambda.lat = 0.1,
  gamma.inc = 3,
  gamma.lat = 3,
  ...
)

```

**Arguments**

**formula** an object of class "formula" (or one that can be coerced to that class): a symbolic description of the model to be fitted. The response must be a survival

	object as returned by the <code>Surv</code> function while the variables on the right side of the formula are the covariates that are included in the incidence portion of the model.
<code>data</code>	a <code>data.frame</code> in which to interpret the variables named in the formula or in the subset argument.
<code>subset</code>	an optional expression indicating which subset of observations to be used in the fitting process, either a numeric or factor variable should be used in subset, not a character variable. All observations are included by default.
<code>x.latency</code>	specifies the variables to be included in the latency portion of the model and can be either a matrix of predictors, a model formula with the right hand side specifying the latency variables, or the same <code>data.frame</code> passed to the <code>data</code> parameter. Note that when using the model formula syntax for <code>x.latency</code> it cannot handle <code>x.latency = ~ ..</code>
<code>model</code>	type of regression model to use for the latency portion of mixture cure model. Can be "cox", "weibull", or "exponential" (default is "cox").
<code>penalty</code>	type of penalty function. Can be "lasso", "MCP", or "SCAD" (default is "lasso").
<code>penalty.factor.inc</code>	vector of binary indicators representing the penalty to apply to each incidence coefficient: 0 implies no shrinkage and 1 implies shrinkage. If not supplied, 1 is applied to all incidence variables.
<code>penalty.factor.lat</code>	vector of binary indicators representing the penalty to apply to each latency coefficient: 0 implies no shrinkage and 1 implies shrinkage. If not supplied, 1 is applied to all latency variables.
<code>thresh</code>	small numeric value. The iterative process stops when the differences between successive expected penalized complete-data log-likelihoods for both incidence and latency components are less than this specified level of tolerance (default is $10^{-3}$ ).
<code>scale</code>	logical, if TRUE the predictors are centered and scaled.
<code>maxit</code>	integer specifying the maximum number of passes over the data for each lambda. If not specified, 100 is applied when <code>penalty = "lasso"</code> and 1000 is applied when <code>penalty = "MCP"</code> or <code>penalty = "SCAD"</code> .
<code>inits</code>	an optional list specifying the initial value for the incidence intercept ( <code>itct</code> ), a numeric vector for the unpenalized incidence coefficients ( <code>b_u</code> ), and a numeric vector for unpenalized latency coefficients ( <code>beta_u</code> ). For parametric models, it should also include a numeric value for the rate parameter ( <code>lambda</code> ) when <code>model = "weibull"</code> or <code>model = "exponential"</code> , and a numeric value for the shape parameter ( <code>alpha</code> ) when <code>model = "weibull"</code> . When <code>model = "cox"</code> , it should also include a numeric vector for the latency survival probabilities $S_u(t_i w_i)$ for $i=1,\dots,N$ ( <code>survprob</code> ). Penalized coefficients are initialized to zero. If <code>inits</code> is not specified or improperly specified, initialization is automatically provided by the function.
<code>lambda.inc</code>	numeric value for the penalization parameter $\lambda$ for variables in the incidence portion of the model.

lambda.lat	numeric value for the penalization parameter $\lambda$ for variables in the latency portion of the model.
gamma.inc	numeric value for the penalization parameter $\gamma$ for variables in the incidence portion of the model when penalty = "MCP" or penalty = "SCAD" (default is 3).
gamma.lat	numeric value for the penalization parameter $\gamma$ for variables in the latency portion of the model when penalty = "MCP" or penalty = "SCAD" (default is 3).
...	additional arguments.

### Value

b_path	Matrix representing the solution path of the coefficients in the incidence portion of the model. Row is step and column is variable.
beta_path	Matrix representing the solution path of lthe coefficients in the latency portion of the model. Row is step and column is variable.
b0_path	Vector representing the solution path of the intercept in the incidence portion of the model.
logLik.inc	Vector representing the expected penalized complete-data log-likelihood for the incidence portion of the model for each step in the solution path.
logLik.lat	Vector representing the expected penalized complete-data log-likelihood for the latency portion of the model for each step in the solution path.
x.incidence	Matrix representing the design matrix of the incidence predictors.
x.latency	Matrix representing the design matrix of the latency predictors.
y	Vector representing the survival object response as returned by the Surv function
model	Character string indicating the type of regression model used for the latency portion of mixture cure model ("weibull" or "exponential").
scale	Logical value indicating whether the predictors were centered and scaled.
method	Character string indicating the EM algoritm was used in fitting the mixture cure model.
rate_path	Vector representing the solution path of the rate parameter for the Weibull or exponential density in the latency portion of the model.
alpha_path	Vector representing the solution path of the shape parameter for the Weibull density in the latency portion of the model.
call	the matched call.

### References

Archer, K. J., Fu, H., Mrozek, K., Nicolet, D., Mims, A. S., Uy, G. L., Stock, W., Byrd, J. C., Hiddemann, W., Braess, J., Spiekermann, K., Metzeler, K. H., Herold, T., Eisfeld, A.-K. (2024) Identifying long-term survivors and those at higher or lower risk of relapse among patients with cytogenetically normal acute myeloid leukemia using a high-dimensional mixture cure model. *Journal of Hematology & Oncology*, **17**:28.

**See Also**[cv\\_cureem](#)**Examples**

```

library(survival)
set.seed(1234)
temp <- generate_cure_data(N = 80, J = 100, nTrue = 10, A = 1.8)
training <- temp$Training
fit <- cureem(Surv(Time, Censor) ~ ., data = training, x.latency = training,
              model = "cox", penalty = "lasso",
              lambda.inc = 0.1, lambda.lat = 0.1, gamma.inc = 6, gamma.lat = 10)

```

curegmifs

---

*Fit penalized parametric mixture cure model using the GMIFS algorithm*

---

**Description**

Fits a penalized Weibull or exponential mixture cure model using the generalized monotone incremental forward stagewise (GMIFS) algorithm and yields solution paths for parameters in the incidence and latency portions of the model.

**Usage**

```

curegmifs(
  formula,
  data,
  subset,
  x.latency = NULL,
  model = "weibull",
  penalty.factor.inc = NULL,
  penalty.factor.lat = NULL,
  epsilon = 0.001,
  thresh = 1e-05,
  scale = TRUE,
  maxit = 10000,
  inits = NULL,
  verbose = TRUE,
  ...
)

```

**Arguments**

**formula** an object of class "formula" (or one that can be coerced to that class): a symbolic description of the model to be fitted. The response must be a survival object as returned by the `Surv` function while the variables on the right side of the formula are the covariates that are included in the incidence portion of the model.

<code>data</code>	a <code>data.frame</code> in which to interpret the variables named in the <code>formula</code> or in the <code>subset</code> argument.
<code>subset</code>	an optional expression indicating which subset of observations to be used in the fitting process, either a numeric or factor variable should be used in <code>subset</code> , not a character variable. All observations are included by default.
<code>x.latency</code>	specifies the variables to be included in the latency portion of the model and can be either a matrix of predictors, a model formula with the right hand side specifying the latency variables, or the same <code>data.frame</code> passed to the <code>data</code> parameter. Note that when using the model formula syntax for <code>x.latency</code> it cannot handle <code>x.latency = ~ ..</code>
<code>model</code>	type of regression model to use for the latency portion of mixture cure model. Can be "weibull" or "exponential"; default is "weibull".
<code>penalty.factor.inc</code>	vector of binary indicators representing the penalty to apply to each incidence coefficient: 0 implies no shrinkage and 1 implies shrinkage. If not supplied, 1 is applied to all incidence variables.
<code>penalty.factor.lat</code>	vector of binary indicators representing the penalty to apply to each latency coefficient: 0 implies no shrinkage and 1 implies shrinkage. If not supplied, 1 is applied to all latency variables.
<code>epsilon</code>	small numeric value reflecting the incremental value used to update a coefficient at a given step (default is 0.001).
<code>thresh</code>	small numeric value. The iterative process stops when the differences between successive expected penalized complete-data log-likelihoods for both incidence and latency components are less than this specified level of tolerance (default is $10^{-5}$ ).
<code>scale</code>	logical, if TRUE the predictors are centered and scaled.
<code>maxit</code>	integer specifying the maximum number of steps to run in the iterative algorithm (default is $10^4$ ).
<code>inits</code>	an optional list specifying the initial value for the incidence intercept ( <code>itct</code> ), a numeric vector for the unpenalized incidence coefficients ( <code>b_u</code> ), and a numeric vector for unpenalized latency coefficients ( <code>beta_u</code> ), a numeric value for the rate parameter ( <code>lambda</code> ), and a numeric value for the shape parameter ( <code>alpha</code> ) when <code>model = "weibull"</code> . If not supplied or improperly supplied, initialization is automatically provided by the function.
<code>verbose</code>	logical, if TRUE running information is printed to the console (default is FALSE).
<code>...</code>	additional arguments.

### Value

<code>b_path</code>	Matrix representing the solution path of the coefficients in the incidence portion of the model. Row is step and column is variable.
<code>beta_path</code>	Matrix representing the solution path of the coefficients in the latency portion of the model. Row is step and column is variable.

b0_path	Vector representing the solution path of the intercept in the incidence portion of the model.
rate_path	Vector representing the solution path of the rate parameter for the Weibull or exponential density in the latency portion of the model.
logLik	Vector representing the log-likelihood for each step in the solution path.
x.incidence	Matrix representing the design matrix of the incidence predictors.
x.latency	Matrix representing the design matrix of the latency predictors.
y	Vector representing the survival object response as returned by the Surv function
model	Character string indicating the type of regression model used for the latency portion of mixture cure model ("weibull" or "exponential").
scale	Logical value indicating whether the predictors were centered and scaled.
alpha_path	Vector representing the solution path of the shape parameter for the Weibull density in the latency portion of the model.
call	the matched call.

## References

Fu, H., Nicolet, D., Mrozek, K., Stone, R. M., Eisfeld, A. K., Byrd, J. C., Archer, K. J. (2022) Controlled variable selection in Weibull mixture cure models for high-dimensional data. *Statistics in Medicine*, **41**(22), 4340–4366.

## See Also

[cv\\_curegmifs](#)

## Examples

```
library(survival)
set.seed(1234)
temp <- generate_cure_data(N = 100, J = 10, nTrue = 10, A = 1.8)
training <- temp$Training

fit <- curegmifs(Surv(Time, Censor) ~ .,
  data = training, x.latency = training,
  model = "weibull", thresh = 1e-4, maxit = 2000, epsilon = 0.01,
  verbose = FALSE)
```

---

cure\_estimate

*Estimate cured fraction*

---

## Description

Estimates the cured fraction using a Kaplan-Meier fitted object.

**Usage**

```
cure_estimate(object)
```

**Arguments**

object            a survfit object.

**Value**

estimated proportion of cured observations

**See Also**

[survfit](#), [sufficient\\_fu\\_test](#), [nonzerocure\\_test](#)

**Examples**

```
library(survival)
set.seed(1234)
temp <- generate_cure_data(N = 100, J = 10, nTrue = 10, A = 1.8)
training <- temp$Training
km.fit <- survfit(Surv(Time, Censor) ~ 1, data = training)
cure_estimate(km.fit)
```

---

cv\_cureem

*Fit penalized mixture cure model using the E-M algorithm with cross-validation for parameter tuning*

---

**Description**

Fits a penalized parametric and semi-parametric mixture cure model (MCM) using the E-M algorithm with with k-fold cross-validation for parameter tuning. The lasso (L1), MCP and SCAD penalty are supported for the Cox MCM while only lasso is currently supported for parametric MCMs. When FDR controlled variable selection is used, the model-X knockoffs method is applied and indices of selected variables are returned.

**Usage**

```
cv_cureem(
  formula,
  data,
  subset,
  x.latency = NULL,
  model = "cox",
  penalty = "lasso",
  penalty.factor.inc = NULL,
  penalty.factor.lat = NULL,
  fdr.control = FALSE,
```

```

    fdr = 0.2,
    grid.tuning = FALSE,
    thresh = 0.001,
    scale = TRUE,
    maxit = NULL,
    inits = NULL,
    lambda.inc.list = NULL,
    lambda.lat.list = NULL,
    nlambda.inc = NULL,
    nlambda.lat = NULL,
    gamma.inc = 3,
    gamma.lat = 3,
    lambda.min.ratio.inc = 0.1,
    lambda.min.ratio.lat = 0.1,
    n_folds = 5,
    measure.inc = "c",
    one.se = FALSE,
    cure_cutoff = 5,
    parallel = FALSE,
    seed = NULL,
    verbose = TRUE,
    ...
)

```

### Arguments

formula	an object of class "formula" (or one that can be coerced to that class): a symbolic description of the model to be fitted. The response must be a survival object as returned by the Surv function while the variables on the right side of the formula are the covariates that are included in the incidence portion of the model.
data	a data.frame in which to interpret the variables named in the formula or in the subset argument.
subset	an optional expression indicating which subset of observations to be used in the fitting process, either a numeric or factor variable should be used in subset, not a character variable. All observations are included by default.
x.latency	specifies the variables to be included in the latency portion of the model and can be either a matrix of predictors, a model formula with the right hand side specifying the latency variables, or the same data.frame passed to the data parameter. Note that when using the model formula syntax for x.latency it cannot handle x.latency = ~ ..
model	type of regression model to use for the latency portion of mixture cure model. Can be "cox", "weibull", or "exponential" (default is "cox").
penalty	type of penalty function. Can be "lasso", "MCP", or "SCAD" (default is "lasso").
penalty.factor.inc	vector of binary indicators representing the penalty to apply to each incidence coefficient: 0 implies no shrinkage and 1 implies shrinkage. If not supplied, 1 is applied to all incidence variables.

<code>penalty.factor.lat</code>	vector of binary indicators representing the penalty to apply to each latency coefficient: 0 implies no shrinkage and 1 implies shrinkage. If not supplied, 1 is applied to all latency variables.
<code>fdr.control</code>	logical, if TRUE, model-X knockoffs are used for FDR-controlled variable selection and indices of selected variables are returned (default is FALSE).
<code>fdr</code>	numeric value in (0, 1) range specifying the target FDR level to use for variable selection when <code>fdr.control=TRUE</code> (default is 0.2).
<code>grid.tuning</code>	logical, if TRUE a 2-D grid tuning approach is used to select the optimal pair of $\lambda_b$ and $\lambda_\beta$ penalty parameters for the incidence and latency portions of the model, respectively. Otherwise the $\lambda_b$ and $\lambda_\beta$ are selected from a 1-D sequence and are equal to one another (default is FALSE).
<code>thresh</code>	small numeric value. The iterative process stops when the differences between successive expected penalized complete-data log-likelihoods for both incidence and latency components are less than this specified level of tolerance (default is $10^{-3}$ ).
<code>scale</code>	logical, if TRUE the predictors are centered and scaled.
<code>maxit</code>	maximum number of passes over the data for each lambda. If not specified, 100 is applied when <code>penalty = "lasso"</code> and 1000 is applied when <code>penalty = "MCP"</code> or <code>penalty = "SCAD"</code> .
<code>inits</code>	an optional list specifying the initial value for the incidence intercept ( <code>itct</code> ), a numeric vector for the unpenalized incidence coefficients ( <code>b_u</code> ), and a numeric vector for unpenalized latency coefficients ( <code>beta_u</code> ). For parametric models, it should also include a numeric value for the rate parameter ( <code>lambda</code> ) when <code>model = "weibull"</code> or <code>model = "exponential"</code> , and a numeric value for the shape parameter ( <code>alpha</code> ) when <code>model = "weibull"</code> . When <code>model = "cox"</code> , it should also include a numeric vector for the latency survival probabilities $S_u(t_i w_i)$ for $i=1,\dots,N$ ( <code>survprob</code> ). Penalized coefficients are initialized to zero. If <code>inits</code> is not specified or improperly specified, initialization is automatically provided by the function.
<code>lambda.inc.list</code>	a numeric vector used to search for the optimal $\lambda_b$ tuning parameter. If not supplied, the function computes a $\lambda_b$ sequence based on <code>nlambda.inc</code> and <code>lambda.min.ratio.inc</code> . If <code>grid.tuning=FALSE</code> , the same sequence should be used for both $\lambda_b$ and $\lambda_\beta$ .
<code>lambda.lat.list</code>	a numeric vector used to search for the optimal $\lambda_\beta$ tuning parameter. If not supplied, the function computes a $\lambda_\beta$ sequence based on <code>nlambda.lat</code> and <code>lambda.min.ratio.lat</code> . If <code>grid.tuning=FALSE</code> , the same sequence should be used for both $\lambda_b$ and $\lambda_\beta$ .
<code>nlambda.inc</code>	an integer specifying the number of values to search for the optimal $\lambda_b$ tuning parameter; default is 10 if <code>grid.tuning=TRUE</code> and 50 otherwise.
<code>nlambda.lat</code>	an integer specifying the number of values to search for the optimal $\lambda_\beta$ tuning parameter; default is 10 if <code>grid.tuning=TRUE</code> and 50 otherwise.
<code>gamma.inc</code>	numeric value for the penalization parameter $\gamma$ for variables in the incidence portion of the model when <code>penalty = "MCP"</code> or <code>penalty = "SCAD"</code> (default is 3).

<code>gamma.lat</code>	numeric value for the penalization parameter $\gamma$ for variables in the latency portion of the model when <code>penalty = "MCP"</code> or <code>penalty = "SCAD"</code> (default is 3).
<code>lambda.min.ratio.inc</code>	numeric value in (0,1) representing the smallest value for $\lambda_b$ as a fraction of <code>lambda.max.inc</code> , the data-derived entry value at which essentially all penalized variables in the incidence portion of the model have a coefficient estimate of 0 (default is 0.1).
<code>lambda.min.ratio.lat</code>	numeric value in (0,1) representing the smallest value for $\lambda_\beta$ as a fraction of <code>lambda.max.lat</code> , the data-derived entry value at which essentially all penalized variables in the latency portion of the model have a coefficient estimate of 0 (default is 0.1).
<code>n_folds</code>	an integer specifying the number of folds for the k-fold cross-valuation procedure (default is 5).
<code>measure.inc</code>	character string specifying the evaluation criterion used in selecting the optimal $\lambda_b$ . Can be "c" or "auc"; default is "c". If <code>measure.inc="c"</code> , the C-statistic using the cure status weighting (CSW) method proposed by Asano and Hirakawa (2017) is used to select both $\lambda_b$ and $\lambda_\beta$ . If <code>measure.inc="auc"</code> , the AUC for cure prediction using the mean score imputation (MSI) method proposed by Asano et al. (2014) is used to select $\lambda_b$ while the C-statistic with CSW is used for $\lambda_\beta$ .
<code>one.se</code>	logical, if TRUE then the one standard error rule is applied for selecting the optimal parameters. The one standard error rule selects the most parsimonious model having evaluation criterion no more than one standard error worse than that of the best evaluation criterion (default is FALSE).
<code>cure_cutoff</code>	numeric value representing the cutoff time value that represents subjects not experiencing the event by this time are cured. This value is used to produce a proxy for the unobserved cure status when calculating C-statistic and AUC (default is 5 representing 5 years). Users should be careful to note the time scale of their data and adjust this according to the time scale and clinical application.
<code>parallel</code>	logical. If TRUE, parallel processing is performed for K-fold CV using <code>foreach</code> and the <b>doParallel</b> package is required.
<code>seed</code>	optional integer representing the random seed. Setting the random seed fosters reproducibility of the results.
<code>verbose</code>	logical, if TRUE running information is printed to the console (default is FALSE).
<code>...</code>	additional arguments.

### Value

<code>b0</code>	Estimated intercept for the incidence portion of the model.
<code>b</code>	Estimated coefficients for the incidence portion of the model.
<code>beta</code>	Estimated coefficients for the latency portion of the model.
<code>alpha</code>	Estimated shape parameter if the Weibull model is fit.
<code>rate</code>	Estimated rate parameter if the Weibull or exponential model is fit.

logLik.inc	Expected penalized complete-data log-likelihood for the incidence portion of the model.
logLik.lat	Expected penalized complete-data log-likelihood for the latency portion of the model.
selected.lambda.inc	Value of $\lambda_b$ selected using cross-validation. NULL when <code>fdr.control</code> is TRUE.
selected.lambda.lat	Value of $\lambda_\beta$ selected using cross-validation. NULL when <code>fdr.control</code> is TRUE.
max.c	Maximum C-statistic achieved.
max.auc	Maximum AUC for cure prediction achieved; only output when <code>measure.inc="auc"</code> .
selected.index.inc	Indices of selected variables for the incidence portion of the model when <code>fdr.control=TRUE</code> . If no variables are selected, <code>int(0)</code> will be returned.
selected.index.lat	Indices of selected variables for the latency portion of the model when <code>fdr.control=TRUE</code> . If no variables are selected, <code>int(0)</code> will be returned.
call	the matched call.

## References

Archer, K. J., Fu, H., Mrozek, K., Nicolet, D., Mims, A. S., Uy, G. L., Stock, W., Byrd, J. C., Hiddemann, W., Braess, J., Spiekermann, K., Metzeler, K. H., Herold, T., Eisfeld, A.-K. (2024) Identifying long-term survivors and those at higher or lower risk of relapse among patients with cytogenetically normal acute myeloid leukemia using a high-dimensional mixture cure model. *Journal of Hematology & Oncology*, **17**:28.

## See Also

[cureem](#)

## Examples

```
library(survival)
set.seed(1234)
temp <- generate_cure_data(N = 200, J = 25, nTrue = 5, A = 1.8)
training <- temp$Training
fit.cv <- cv_cureem(Surv(Time, Censor) ~ ., data = training,
  x.latency = training, fdr.control = FALSE,
  grid.tuning = FALSE, nlambda.inc = 10, nlambda.lat = 10,
  n_folds = 2, seed = 23, verbose = TRUE)
fit.cv.fdr <- cv_cureem(Surv(Time, Censor) ~ ., data = training,
  x.latency = training, model = "weibull", penalty = "lasso",
  fdr.control = TRUE, grid.tuning = FALSE, nlambda.inc = 10,
  nlambda.lat = 10, n_folds = 2, seed = 23, verbose = TRUE)
```

---

cv_curegmifs	<i>Fit a penalized parametric mixture cure model using the GMIFS algorithm with cross-validation for model selection</i>
--------------	--

---

### Description

Fits a penalized Weibull or exponential mixture cure model using the generalized monotone incremental forward stagewise (GMIFS) algorithm with k-fold cross-validation to select the optimal iteration step along the solution path. When FDR controlled variable selection is used, the model-X knockoffs method is applied and indices of selected variables are returned.

### Usage

```
cv_curegmifs(
  formula,
  data,
  subset,
  x.latency = NULL,
  model = "weibull",
  penalty.factor.inc = NULL,
  penalty.factor.lat = NULL,
  fdr.control = FALSE,
  fdr = 0.2,
  epsilon = 0.001,
  thresh = 1e-05,
  scale = TRUE,
  maxit = 10000,
  inits = NULL,
  n_folds = 5,
  measure.inc = "c",
  one.se = FALSE,
  cure_cutoff = 5,
  parallel = FALSE,
  seed = NULL,
  verbose = TRUE,
  ...
)
```

### Arguments

formula	an object of class "formula" (or one that can be coerced to that class): a symbolic description of the model to be fitted. The response must be a survival object as returned by the Surv function while the variables on the right side of the formula are the covariates that are included in the incidence portion of the model.
data	a data.frame in which to interpret the variables named in the formula or in the subset argument.

subset	an optional expression indicating which subset of observations to be used in the fitting process, either a numeric or factor variable should be used in subset, not a character variable. All observations are included by default.
x.latency	specifies the variables to be included in the latency portion of the model and can be either a matrix of predictors, a model formula with the right hand side specifying the latency variables, or the same data.frame passed to the data parameter. Note that when using the model formula syntax for x.latency it cannot handle <code>x.latency = ~ ..</code>
model	type of regression model to use for the latency portion of mixture cure model. Can be "weibull" or "exponential"; default is "weibull".
penalty.factor.inc	vector of binary indicators representing the penalty to apply to each incidence coefficient: 0 implies no shrinkage and 1 implies shrinkage. If not supplied, 1 is applied to all incidence variables.
penalty.factor.lat	vector of binary indicators representing the penalty to apply to each latency coefficient: 0 implies no shrinkage and 1 implies shrinkage. If not supplied, 1 is applied to all latency variables.
fdr.control	logical, if TRUE, model-X knockoffs are used for FDR-controlled variable selection and indices of selected variables are returned (default is FALSE).
fdr	numeric value in (0, 1) range specifying the target FDR level to use for variable selection when fdr.control=TRUE (default is 0.2).
epsilon	small numeric value reflecting incremental value used to update a coefficient at a given step (default is 0.001).
thresh	small numeric value. The iterative process stops when the differences between successive expected penalized complete-data log-likelihoods for both incidence and latency components are less than this specified level of tolerance (default is $10^{-5}$ ).
scale	logical, if TRUE the predictors are centered and scaled.
maxit	integer specifying the maximum number of steps to run in the iterative algorithm (default is $10^4$ ).
inits	an optional list specifying the initial value for the incidence intercept ( <code>itct</code> ), a numeric vector for the unpenalized incidence coefficients ( <code>b_u</code> ), and a numeric vector for unpenalized latency coefficients ( <code>beta_u</code> ), a numeric value for the rate parameter ( <code>lambda</code> ), and a numeric value for the shape parameter ( <code>alpha</code> ) when <code>model = "weibull"</code> . If not supplied or improperly supplied, initialization is automatically provided by the function.
n_folds	an integer specifying the number of folds for the k-fold cross-validation procedure (default is 5).
measure.inc	character string specifying the evaluation criterion used in selecting the optimal $\lambda_b$ . Can be "c" or "auc"; default is "c". If <code>measure.inc="c"</code> , the C-statistic using the cure status weighting (CSW) method proposed by Asano and Hirakawa (2017) is used to select both $\lambda_b$ and $\lambda_\beta$ . If <code>measure.inc="auc"</code> , the AUC for cure prediction using the mean score imputation (MSI) method proposed by Asano et al. (2014) is used to select $\lambda_b$ while the C-statistic with CSW is used for $\lambda_\beta$ .

one.se	logical, if TRUE then the one standard error rule is applied for selecting the optimal parameters. The one standard error rule selects the most parsimonious model having evaluation criterion no more than one standard error worse than that of the best evaluation criterion (default is FALSE).
cure_cutoff	numeric value representing the cutoff time value that represents subjects not experiencing the event by this time are cured. This value is used to produce a proxy for the unobserved cure status when calculating C-statistic and AUC (default is 5 representing 5 years). Users should be careful to note the time scale of their data and adjust this according to the time scale and clinical application.
parallel	logical. If TRUE, parallel processing is performed for K-fold CV using foreach and the <b>doParallel</b> package is required.
seed	optional integer representing the random seed. Setting the random seed fosters reproducibility of the results.
verbose	logical, if TRUE running information is printed to the console (default is FALSE).
...	additional arguments.

### Value

b0	Estimated intercept for the incidence portion of the model.
b	Estimated coefficients for the incidence portion of the model.
beta	Estimated coefficients for the latency portion of the model.
alpha	Estimated shape parameter if the Weibull model is fit.
rate	Estimated rate parameter if the Weibull or exponential model is fit.
logLik	Log-likelihood value.
selected.step.inc	Iteration step selected for the incidence portion of the model using cross-validation. NULL when <code>fdr.control</code> is TRUE.
selected.step.lat	Iteration step selected for the latency portion of the model using cross-validation. NULL when <code>fdr.control</code> is TRUE.
max.c	Maximum C-statistic achieved
max.auc	Maximum AUC for cure prediction achieved; only output when <code>measure.inc="auc"</code> .
selected.index.inc	Indices of selected variables for the incidence portion of the model when <code>fdr.control=TRUE</code> . If none selected, <code>int(0)</code> will be returned.
selected.index.lat	Indices of selected variables for the latency portion of the model when <code>fdr.control=TRUE</code> . If none selected, <code>int(0)</code> will be returned.
call	the matched call.

### References

Fu, H., Nicolet, D., Mrozek, K., Stone, R. M., Eisfeld, A. K., Byrd, J. C., Archer, K. J. (2022) Controlled variable selection in Weibull mixture cure models for high-dimensional data. *Statistics in Medicine*, **41**(22), 4340–4366.

**See Also**[curegmifs](#)[curegmifs](#)**Examples**

```
library(survival)
set.seed(123)
temp <- generate_cure_data(N = 100, J = 15, nTrue = 3, A = 1.8, rho = 0.2)
training <- temp$Training

fit.cv <- cv_curegmifs(Surv(Time, Censor) ~ ., data = training,
                      x.latency = training, fdr.control = FALSE,
                      maxit = 450, epsilon = 0.01,
                      n_folds = 2, seed = 23, verbose = TRUE)
```

---

generate\_cure\_data      *Simulate data under a mixture cure model*

---

**Description**

Simulate data under a mixture cure model

**Usage**

```
generate_cure_data(
  N = 400,
  J = 500,
  nonp = 2,
  train.prop = 3/4,
  nTrue = 10,
  A = 1,
  rho = 0.5,
  itct_mean = 0.5,
  cens_ub = 20,
  alpha = 1,
  lambda = 2,
  same_signs = FALSE,
  model = "weibull"
)
```

**Arguments**

N                    an integer denoting the total sample size.

J                    an integer denoting the number of penalized predictors which is the same for both the incidence and latency portions of the model.

nonp	an integer less than J denoting the number of unpenalized predictors (which is the same for both the incidence and latency portions of the model).
train.prop	a numeric value in 0, 1 representing the fraction of N to be used in forming the Training dataset.
nTrue	an integer denoting the number of variables truly associated with the outcome (i.e., the number of covariates with nonzero parameter values) among the penalized predictors.
A	a numeric value denoting the effect size which is the same for both the incidence and latency portions of the model.
rho	a numeric value in 0, 1 representing the correlation between adjacent covariates in the same block. See details below.
itct_mean	a numeric value representing the expectation of the incidence intercept which controls the cure rate.
cens_ub	a numeric value representing the upper bound on the censoring time distribution which follows a uniform distribution on 0, cens_ub.
alpha	a numeric value representing the shape parameter in the Weibull density.
lambda	a numeric value representing the rate parameter in the Weibull density.
same_signs	logical, if TRUE the incidence and latency coefficients have the same signs.
model	type of regression model to use for the latency portion of mixture cure model. Can be "weibull", "GG", "Gompertz", "nonparametric", or "GG_baseline".

### Value

Training	Training data.frame which includes Time, Censor, and covariates.
Testing	Testing data.frame which includes Time, Censor, and covariates.
parameters	A list including: the indices of true incidence signals (nonzero_b), indices of true latency signals (nonzero_beta), unpenalized incidence parameter values (b_u), unpenalized latency parameter values (beta_u), parameter values for the true incidence signals among penalized covariates (b_p_nz), parameter values for the true latency signals among penalized covariates (beta_p_nz), parameter value for the incidence intercept (itct)

### Examples

```
library(survival)
set.seed(1234)
data <- generate_cure_data(N = 200, J = 50, nTrue = 10, A = 1.8, rho = 0.2)
training <- data$Training
testing <- data$Testing
fit <- cureem(Surv(Time, Censor) ~ ., data = training,
             x.latency = training, model = "cox", penalty = "lasso",
             lambda.inc = 0.05, lambda.lat = 0.05,
             gamma.inc = 6, gamma.lat = 10)
```

---

nonzerocure\_test      *Non-parametric test for a non-zero cured fraction*

---

### Description

Tests the null hypothesis that the proportion of observations susceptible to the event = 1 against the alternative that the proportion of observations susceptible to the event is < 1. If the null hypothesis is rejected, there is a significant cured fraction.

### Usage

```
nonzerocure_test(object, Reps = 1000, seed = NULL, plot = FALSE, B = NULL)
```

### Arguments

object	a survfit object.
Reps	number of simulations on which to base the p-value (default = 1000).
seed	optional random seed.
plot	logical. If TRUE a histogram of the estimated susceptible proportions over all simulations is produced.
B	optional. If specified the maximum observed time for the uniform distribution for generating the censoring times. If not specified, an exponential model is used for generating the censoring times (default).

### Value

proportion_susceptible	estimated proportion of susceptibles
proportion_cured	estimated proportion of those cured
p.value	p-value testing the null hypothesis that the proportion of susceptibles = 1 (cured fraction = 0) against the alternative that the proportion of susceptibles < 1 (non-zero cured fraction)
time_95_percent_of_events	estimated time at which 95% of events should have occurred

### References

Maller, R. A. and Zhou, X. (1996) *Survival Analysis with Long-Term Survivors*. John Wiley & Sons.

### See Also

[survfit](#), [cure\\_estimate](#), [sufficient\\_fu\\_test](#)

**Examples**

```

library(survival)
set.seed(1234)
temp <- generate_cure_data(N = 100, J = 10, nTrue = 10, A = 1.8)
training <- temp$Training
km.fit <- survfit(Surv(Time, Censor) ~ 1, data = training)
nonzerocure_test(km.fit)

```

---

plot.mixturecure      *Plot fitted mixture cure model*

---

**Description**

This function plots either the coefficient path, the AIC, the cAIC, the BIC, or the log-likelihood for a fitted curegmifs or cureem object. This function produces a lollipop plot of the coefficient estimates for a fitted cv\_curegmifs or cv\_cureem object.

**Usage**

```

## S3 method for class 'mixturecure'
plot(x, type = "trace", xlab = NULL, ylab = NULL, main = NULL, ...)

```

**Arguments**

x	a mixturecure object resulting from curegmifs or cureem, cv_curegmifs or cv_cureem.
type	default is "trace" which plots the coefficient path for the fitted object. Also available are "AIC", "cAIC", "mAIC", "BIC", "mBIC", "EBIC", and "logLik". This option has no effect for objects fit using cv_curegmifs or cv_cureem.
xlab	a default x-axis label will be used which can be changed by specifying a user-defined x-axis label.
ylab	a default y-axis label will be used which can be changed by specifying a user-defined y-axis label.
main	a default main title will be used which can be changed by specifying a user-defined main title. This option is not used for cv_curegmifs or cv_cureem fitted objects.
...	other arguments.

**Value**

this function has no returned value but is called for its side effects

**See Also**

[curegmifs](#), [cureem](#), [coef.mixturecure](#), [summary.mixturecure](#), [predict.mixturecure](#)

**Examples**

```

library(survival)
set.seed(1234)
temp <- generate_cure_data(N = 100, J = 10, nTrue = 10, A = 1.8)
training <- temp$Training
fit <- curegmifs(Surv(Time, Censor) ~ .,
                 data = training, x.latency = training,
                 model = "weibull", thresh = 1e-4, maxit = 2000,
                 epsilon = 0.01, verbose = FALSE)

plot(fit)

```

---

predict.mixturecure    *Predicted probabilities for susceptibles, linear predictor for latency, and risk class for latency for mixture cure fit*

---

**Description**

This function returns a list that includes the predicted probabilities for susceptibles as well as the linear predictor for the latency distribution and a dichotomous risk for latency for a curegmifs, cureem, cv\_curegmifs or cv\_cureem fitted object.

**Usage**

```

## S3 method for class 'mixturecure'
predict(object, newdata, model.select = "AIC", ...)

```

**Arguments**

object	a mixturecure object resulting from curegmifs, cureem, cv_curegmifs, cv_cureem.
newdata	an optional data.frame that minimally includes the incidence and/or latency variables to use for predicting the response. If omitted, the training data are used.
model.select	for models fit using curegmifs or cureem any step along the solution path can be selected. The default is model.select = "AIC" which calculates the predicted values using the coefficients from the model having the lowest AIC. Other options are model.select = "mAIC" for the modified AIC, model.select = "cAIC" for the corrected AIC, model.select = "BIC", model.select = "mBIC" for the modified BIC, model.select = "EBIC" for the extended BIC, model.select = "logLik" for the step that maximizes the log-likelihood, or any numeric value from the solution path. This option has no effect for objects fit using cv_curegmifs or cv_cureem.
...	other arguments

**Value**

- `p.uncured` a vector of probabilities from the incidence portion of the fitted model representing the P(uncured).
- `linear.latency` a vector for the linear predictor from the latency portion of the model.
- `latency.risk` a dichotomous class representing low (below the median) versus high risk for the latency portion of the model.

**See Also**

[curegmifs](#), [cureem](#), [coef.mixturecure](#), [summary.mixturecure](#), [plot.mixturecure](#)

**Examples**

```
library(survival)
set.seed(1234)
temp <- generate_cure_data(N = 100, J = 10, nTrue = 10, A = 1.8)
training <- temp$Training
fit <- curegmifs(Surv(Time, Censor) ~ .,
                 data = training, x.latency = training,
                 model = "weibull", thresh = 1e-4, maxit = 2000,
                 epsilon = 0.01, verbose = FALSE)
predict.train <- predict(fit)
names(predict.train)
testing <- temp$Testing
predict.test <- predict(fit, newdata = testing)
```

---

`print.mixturecure` *Print the contents of a mixture cure fitted object*

---

**Description**

This function prints the names of the list objects from a `curegmifs`, `cureem`, `cv_cureem`, or `cv_curegmifs` fitted model.

**Usage**

```
## S3 method for class 'mixturecure'
print(x, ...)
```

**Arguments**

- `x` a `mixturecure` object resulting from `curegmifs`, `cureem`, `cv_cureem`, or `cv_curegmifs`.
- `...` other arguments.

**Value**

names of the objects in a `mixturecure` object fit using `cureem`, `curegmifs`, `cv_cureem`, or `cv_curegmifs`.

**Note**

The contents of an `mixturecure` fitted object differ depending upon whether the EM (`cureem`) or GMIFS (`curegmifs`) algorithm is used for model fitting. Also, the output differs depending upon whether `x.latency` is specified in the model (i.e., variables are included in the latency portion of the model fit) or only terms on the right hand side of the equation are included (i.e., variables are included in the incidence portion of the model).

**See Also**

[curegmifs](#), [cureem](#), [coef.mixturecure](#), [summary.mixturecure](#), [plot.mixturecure](#), [predict.mixturecure](#)

**Examples**

```
library(survival)
set.seed(1234)
temp <- generate_cure_data(N = 100, J = 10, nTrue = 10, A = 1.8)
training <- temp$Training
fit <- curegmifs(Surv(Time, Censor) ~ .,
                 data = training, x.latency = training,
                 model = "weibull", thresh = 1e-4, maxit = 2000,
                 epsilon = 0.01, verbose = FALSE)

print(fit)
```

---

sufficient\_fu\_test      *Test for sufficient follow-up*

---

**Description**

Tests for sufficient follow-up using a Kaplan-Meier fitted object.

**Usage**

```
sufficient_fu_test(object)
```

**Arguments**

`object`            a `survfit` object.

**Value**

<code>p.value</code>	p-value from testing the null hypothesis that there was not sufficient follow-up against the alternative that there was sufficient follow-up
<code>Nn</code>	total number of events that occurred at time $> \text{pmax}(0, 2 * (\text{last observed event time}) - (\text{last observed time}))$ and $<$ the last observed event time
<code>N</code>	number of observations in the dataset

**References**

Maller, R. A. and Zhou, X. (1996) *Survival Analysis with Long-Term Survivors*. John Wiley & Sons.

**See Also**

[survfit](#), [cure\\_estimate](#), [nonzerocure\\_test](#)

**Examples**

```
library(survival)
set.seed(1234)
temp <- generate_cure_data(N = 100, J = 10, nTrue = 10, A = 1.8)
training <- temp$Training
km.fit <- survfit(Surv(Time, Censor) ~ 1, data = training)
sufficient_fu_test(km.fit)
```

---

summary.mixturecure     *Summarize a Fitted Mixture Cure Object.*

---

**Description**

summary method for a mixturecure object fit using curegmifs, cureem, cv\_curegmifs, or cv\_cureem.

**Usage**

```
## S3 method for class 'mixturecure'
summary(object, ...)
```

**Arguments**

object	a mixturecure object resulting from curegmifs, cureem, cv_curegmifs, or cv_cureem.
...	other arguments.

**Value**

prints the following items extracted from the object fit using curegmifs or cureem: the step and value that maximizes the log-likelihood; the step and value that minimizes the AIC, modified AIC (mAIC), corrected AIC (cAIC), BIC, modified BIC (mBIC), and extended BIC (EBIC). Returns log-likelihood, AIC, and BIC if the object was fit using cv\_curegmifs or cv\_cureem at the optimal cross-validated values if no FDR control; the number of non-zero incidence and latency variables is returned when cross-validation is used together with FDR control.

**See Also**

[curegmifs](#), [cureem](#), [coef.mixturecure](#), [plot.mixturecure](#), [predict.mixturecure](#)

**Examples**

```
library(survival)
set.seed(1234)
temp <- generate_cure_data(N = 100, J = 10, nTrue = 10, A = 1.8)
training <- temp$Training
fit <- curegmifs(Surv(Time, Censor) ~ .,
                 data = training, x.latency = training,
                 model = "weibull", thresh = 1e-4, maxit = 2000,
                 epsilon = 0.01, verbose = FALSE)

summary(fit)
```

# Index

- \* **datasets**
    - amltest, 2
    - amltrain, 11
  - \* **htest**
    - nonzerocure\_test, 40
    - sufficient\_fu\_test, 44
  - \* **methods**
    - coef.mixturecure, 22
    - plot.mixturecure, 41
    - predict.mixturecure, 42
    - print.mixturecure, 43
    - summary.mixturecure, 45
  - \* **models**
    - cureem, 24
    - curegmifs, 27
    - cv\_cureem, 30
    - cv\_curegmifs, 35
  - \* **regression**
    - cureem, 24
    - curegmifs, 27
    - cv\_cureem, 30
    - cv\_curegmifs, 35
  - \* **univar**
    - AUC, 20
    - concordance\_mcm, 23
    - cure\_estimate, 29
- amltest, 2  
amltrain, 11  
AUC, 20, 23
- coef.mixturecure, 22, 41, 43–45  
concordance\_mcm, 21, 23  
cure\_estimate, 29, 40, 45  
cureem, 22, 24, 34, 41, 43–45  
curegmifs, 22, 27, 38, 41, 43–45  
cv\_cureem, 27, 30  
cv\_curegmifs, 29, 35
- generate\_cure\_data, 38
- nonzerocure\_test, 30, 40, 45
- plot.mixturecure, 22, 41, 43–45  
predict.mixturecure, 22, 41, 42, 44, 45  
print.mixturecure, 43
- sufficient\_fu\_test, 30, 40, 44  
summary.mixturecure, 22, 41, 43, 44, 45  
survfit, 30, 40, 45