

# Package ‘chest’

December 16, 2019

**Type** Package

**Title** Change-in-Estimate Approach to Assess Confounding Effects

**Version** 0.2.0

**Description** The 'chest' package applies the change-in-effect estimate method for assessing confounding effects in medical and epidemiological research. It starts with a crude model including only the outcome and exposure variables. At each of the subsequent steps, one variable which creates the largest change among the remaining variables is selected. This process is repeated until all variables have been entered into the model (Wang Z (2007) <doi:10.1177/1536867X0700700203>). Currently, the 'chest' package has functions for linear regression, logistic regression, Cox proportional hazards model and conditional logistic regression.

**License** GPL-2

**Encoding** UTF-8

**LazyData** true

**Imports** tidyverse, survival, grid, speedglm, forestplot, broom, tibble, MASS, scales, tidymodels

**Depends** R (>= 2.10)

**RoxygenNote** 7.0.2

**Suggests** spelling, knitr, rmarkdown

**VignetteBuilder** knitr

**Language** en-US

**NeedsCompilation** no

**Author** Zhiqiang Wang [aut, cre]

**Maintainer** Zhiqiang Wang <zhiqiang.wang@menzies.edu.au>

**Repository** CRAN

**Date/Publication** 2019-12-16 15:00:02 UTC

## R topics documented:

chest . . . . .	2
chest_clogit . . . . .	3
chest_cox . . . . .	4
chest_glm . . . . .	6
chest_lm . . . . .	8
chest_speedglm . . . . .	9
diab_df . . . . .	11

<b>Index</b>	<b>13</b>
--------------	-----------

---

chest	<i>Change-in-Estimate Approach to Assess Confounding Effects</i>
-------	--

---

### Description

In clinical trials and epidemiological studies, the association between an exposure and the outcome of interest in a study can be estimated by *regression coefficients*, *odds ratios* or *hazard ratios* depending on the nature of study designs and outcome measurements. We use a general term *effect estimate* here for any of those measurements in this document. Based on those measurements, we determine if a treatment is effective (or detrimental) or a factor is a risk factor. Imbalanced distributions of other factors could bias the effect estimates, called *confounding*. One way to assess the confounding effect of a factor is to examine the difference in effect estimates between models with and without a specific factor. 'chest' allows users quickly calculate the changes when potential confounding factors are sequentially added to the model in a stepwise fashion. At each step, one variable which creates the largest change (%) of the effect estimate among the remaining variables is added to the model. 'chest' returns a graph and a data frame (table) with effect estimates (95% CI) and change (%) values. The package currently has 5 main functions: 'chest\_lm' for linear regression, 'chest\_glm' for logistic regression and Poisson regression, 'chest\_speedglm' using 'speedlm' as a faster alternative of 'chest\_glm', 'chest\_clogit' for matched logistic regression, and 'chest\_cox' for Cox Proportional Hazards Models.

### References

Zhiqiang Wang (2007) <<https://doi.org/10.1177/1536867X0700700203>>

### Examples

```
? chest_speedglm
? chest_glm
? chest_cox
? chest_clogit
? chest_lm
```

---

chest_clogit	<i>Assessing confounding effects using conditional logistic regression models</i>
--------------	---

---

## Description

'chest\_clogit' is used to fit many *Conditional Logistic Regression* models to assess confounding effects.

## Usage

```
chest_clogit(
  crude,
  xlist,
  data,
  var_lab = "Variables",
  est_lab = "OR (95% CI)",
  change_lab = "Change, %",
  digits = "%.2f",
  digits_change = "%.1f",
  method = "exact",
  na_omit = FALSE,
  hrz_l_lines = gpar(col = "#444444"),
  text_size = 1,
  xlab_size = 1,
  xtick_size = 1,
  xlab = NULL,
  plus = " + ",
  indicate = FALSE,
  ...
)
```

## Arguments

crude	An object of <i>formula</i> for initial model, generally crude model. However, any other variables can also be included here as the initial model.
xlist	A <i>vector</i> of characters with all variable names of potential confounders.
data	<i>Data frame</i> .
var_lab	<i>Character</i> string for the column name of variables in the graph.
est_lab	<i>Character</i> string for the column name of effect estimates in the graph.
change_lab	<i>Character</i> string for the column name of "Change" in the graph
digits	Set the display format for number in the graph other than the "Change" column. Default: "%.2f"
digits_change	Set the format for the "Change" column. Default: "%.1f"
method	See 'clogit', default is the "exact" method.

na_omit	Remove all missing values.
hrzl_lines	A logic to include or remove the horizontal line in the graph.
text_size	Set text size.
xlab_size	Set x label size.
xtick_size	Set x tick label size.
xlab	Add x label.
plus	Change the + sign before variable names.
indicate	indicate the calculation progress.
...	Further optional arguments for forestplot.

### Value

A table with effect estimates and their changes at all steps.

### See Also

chest

'clogit' in **'survival'**

**'forestplot'**

### Examples

```
vlist <- c("Age", "Sex", "Married", "Cancer", "CVD", "Education", "Income")

chest_clogit(crude = "Endpoint ~ Diabetes + strata(mid)",
             xlist = vlist, data = diab_df)
```

---

chest_cox	<i>Assessing confounding effects using Cox Proportional Hazards regression models</i>
-----------	---

---

### Description

'chest\_cox' is used to assess confounding effects using Proportional Hazards Regression Model ('coxph' from **'survival'** package). It presents the effect estimates (such as hazard ratios) for the association between exposure and outcome variables by adding other variables (potential confounders) to the model sequentially. The order of variables to be added is based on the magnitudes of the changes in effect estimates.

**Usage**

```

chest_cox(
  crude,
  xlist,
  data,
  var_lab = "Variables",
  est_lab = "HR (95% CI)",
  change_lab = "Change, %",
  digits = "%.2f",
  digits_change = "%.1f",
  na_omit = FALSE,
  hrzl_lines = gpar(col = "#444444"),
  text_size = 1,
  xlab_size = 1,
  xtick_size = 1,
  xlab = NULL,
  plus = " + ",
  indicate = FALSE,
  ...
)

```

**Arguments**

<code>crude</code>	An object of <i>formula</i> for initial model, generally crude model. However, any other variables can also be included here as the initial model.
<code>xlist</code>	A <i>vector</i> of characters with variable names of potential confounders.
<code>data</code>	<i>Data frame</i> .
<code>var_lab</code>	<i>Character</i> string for the column name of variables in the graph.
<code>est_lab</code>	<i>Character</i> string for the column name of effect estimates.
<code>change_lab</code>	<i>Character</i> string for the column name of "Changes".
<code>digits</code>	Set the display format for number in the graph other than the " <i>Change</i> " column. Default: "%.2f"
<code>digits_change</code>	Set the format for the " <i>Change</i> " column. Default: "%.1f"
<code>na_omit</code>	Remove all missing values.
<code>hrzl_lines</code>	A <i>logic</i> to include or remove horizontal line.
<code>text_size</code>	Set text size.
<code>xlab_size</code>	Set x label size.
<code>xtick_size</code>	Set x tick label size.
<code>xlab</code>	Add x label.
<code>plus</code>	Change the + sign before variable names.
<code>indicate</code>	indicate the progress.
<code>...</code>	Further optional arguments for forestplot.

**Value**

A table with effect estimates and their changes at all steps.

**See Also**

'forestplot'

'survival'

**Examples**

```
vlist <- c("Age", "Sex", "Married", "Cancer", "CVD", "Education", "Income")
chest_cox(crude = "Surv(t0, t1, Endpoint) ~ Diabetes", xlist = vlist, data = diab_df)
```

---

chest_glm	<i>Assessing confounding effects using Generalized Linear regression models</i>
-----------	---

---

**Description**

Please note: There is a faster option: 'chest\_speedglm'. 'chest\_glm' is used to assess confounding effects using Generalized Linear Models, such as logistic regression and Poisson regression with 'glm'. It presents *odds ratios* or *rate ratios* for the association between exposure and outcome variables by adding other variables (potential confounders) to the model sequentially. The order of variables to be added is based on the magnitudes of the changes in effect estimates.

**Usage**

```
chest_glm(
  crude,
  xlist,
  data,
  var_lab = "Variables",
  est_lab = "OR (95% CI)",
  change_lab = "Change, %",
  digits = "%.2f",
  digits_change = "%.1f",
  family = "binomial",
  method = "glm.fit",
  maxit = 25,
  epsilon = 1e-08,
  na_omit = FALSE,
  hrzl_lines = gpar(col = "#444444"),
  text_size = 1,
  xlab_size = 1,
```

```

    xtick_size = 1,
    xlab = NULL,
    indicate = FALSE,
    plus = " + ",
    ...
)

```

## Arguments

crude	An object of <i>formula</i> for initial model, generally crude model. However, any other variables can also be included here as the initial model.
xlist	A <i>vector</i> of characters with all variable names of potential confounders.
data	<i>Data frame</i> .
var_lab	<i>Character</i> string for the column name of variables in the graph.
est_lab	<i>Character</i> string for the column name of effect estimates.
change_lab	<i>Character</i> string for the column name of "Changes".
digits	Set the display format for number in the graph other than the "Change" column. Default: "%.2f"
digits_change	Set the format for the "Change" column. Default: "%.1f"
family	Description of the error distribution. Default is "binomial".
method	Method to detect for singularity.
maxit	See 'glm'.
epsilon	See 'glm'.
na_omit	Remove all missing values.
hrzl_lines	A <i>logic</i> to include or remove horizontal line.
text_size	Set text size.
xlab_size	Set x label size.
xtick_size	Set x tick label size.
xlab	Add x label.
indicate	indicate progress
plus	Change the + sign before variable names.
...	Further optional arguments for forestplot.

## Value

A table with effect estimates and their changes at all steps.

## See Also

'forestplot'

'glm' 'stats'

for effects estimates from all possible models.

## Examples

```
chest_glm(crude = "Endpoint ~ Diabetes", xlist = c("Age", "Sex", "Married"),
na_omit=TRUE, data = diab_df)
```

---

chest\_lm

*Assessing confounding effects using Linear regression models*

---

## Description

'chest\_lm' is used to assess confounding effects using Linear Regression Models. It presents linear regression coefficients as effect estimates and changes when other variables are added sequentially to the model.

## Usage

```
chest_lm(
  crude,
  xlist,
  data,
  method = "qr",
  var_lab = "Variables",
  est_lab = "Coeff. (95% CI)",
  change_lab = "Change, %",
  digits = "%.2f",
  digits_change = "%.1f",
  na_omit = FALSE,
  hrzl_lines = gpar(col = "#444444"),
  text_size = 1,
  xlab_size = 1,
  xtick_size = 1,
  xlab = NULL,
  indicate = FALSE,
  plus = " + ",
  ...
)
```

## Arguments

crude	An object of <i>formula</i> for initial model, generally crude model. However, any other variables can also be included here as the initial model.
xlist	A <i>vector</i> of characters with all variable names of potential confounders.
data	<i>Data frame</i> .
method	The method to be used; see ' <b>lm</b> '.
var_lab	<i>Character</i> string for the column name of variables in the graph.

est_lab	Character string for the column name of effect estimates in the graph.
change_lab	Character string for the column name of "Change" in the graph
digits	Set the display format for number in the graph other than the "Change" column. Default: "%.2f"
digits_change	Set the format for the "Change" column. Default: "%.1f"
na_omit	Remove all missing values.
hrzl_lines	A <i>logic</i> to to include or remove horizontal line.
text_size	Set text size.
xlab_size	Set x label size.
xtick_size	Set x tick label size.
xlab	Add x label.
indicate	indicate progress
plus	Change the + sign before variable names.
...	Further optional arguments for forestplot.

**Value**

A table with effect estimates and their changes at all steps.

**See Also**

'forestplot'

'lm' of 'stats'

**Examples**

```
vlist<-c("Age", "Sex", "Married", "Cancer", "CVD", "Education", "Income")
chest_lm(crude = "BMI ~ Diabetes", xlist = vlist, data = diab_df, na_omit = TRUE)
```

---

chest_speedglm	<i>Assessing confounding effects using Generalized Linear regression models</i>
----------------	---

---

**Description**

This is a faster option to 'chest\_glm'. It presents the effect estimates (such as *odds ratios*) for the association between exposure and outcome variables by adding other variables (potential confounders) to the model sequentially. The order of variables to be added is based on the magnitudes of the changes in effect estimates. It returns a graph and a table with effect estimates, 95% confidence intervals and changes (%) at different steps.

**Usage**

```
chest_speedglm(
  crude,
  xlist,
  data,
  var_lab = "Variables",
  est_lab = "OR (95% CI)",
  change_lab = "Change, %",
  family = binomial(),
  digits = "%.2f",
  digits_change = "%.1f",
  method = c("eigen", "Cholesky", "qr"),
  maxit = 25,
  epsilon = 1e-08,
  na_omit = FALSE,
  hrzl_lines = gpar(col = "#444444"),
  text_size = 1,
  xlab_size = 1,
  xtick_size = 1,
  xlab = NULL,
  plus = " + ",
  indicate = FALSE,
  ...
)
```

**Arguments**

crude	An object of <i>formula</i> for initial model, generally crude model. However, any other variables can also be included here as the initial model.
xlist	A <i>vector</i> of characters with all variable names of potential confounders.
data	<i>Data frame</i> .
var_lab	<i>Character</i> string for the column name of <i>variables</i> in the graph.
est_lab	<i>Character</i> string for the column name of effect estimates in the graph.
change_lab	<i>Character</i> string for the column name of "Change" in the graph
family	Description of the error distribution. Default is "binomial".
digits	Set the display format for number in the graph other than the "Change" column. Default: "%.2f"
digits_change	Set the format for the "Change" column. Default: "%.1f"
method	Method to detect for singularity.
maxit	See 'glm'.
epsilon	See 'glm'.
na_omit	Remove all missing values.
hrzl_lines	A logic to include or remove horizontal line.
text_size	Set text size.

xlab_size	Set x label size.
xtick_size	Set x tick label size.
xlab	Add x label.
plus	Change the + sign before variable names.
indicate	indicate progress
...	Further optional arguments for forestplot.

**Value**

A table with effect estimates and their changes at all steps.

**See Also**

'forestplot'  
'speedglm'  
'glm' of 'stats'

**Examples**

```
vlist<-c("Age", "Sex", "Married", "Cancer", "CVD", "Education", "Income")

chest_speedglm(crude = "Endpoint ~ Diabetes",
               xlist = vlist, na_omit=TRUE, data = diab_df)

# adding derived terms such as an interaction between Age and Sex, and age squared:

library(tidyverse)
diab_df <- diab_df %>%
mutate(Age_Sex = Age*Sex, Age2 = Age^2)

vlist<-c("Age", "Sex", "Age2", "Age_Sex", "Married", "Cancer", "CVD", "Education", "Income")

chest_speedglm(crude = "Endpoint ~ Diabetes", xlist = vlist, na_omit=TRUE, data = diab_df)
```

---

diab\_df

*Example data: Health outcomes of 2372 adults with and without diabetes*

---

**Description**

A data frame with 2372 rows and 14 variables with diabetes status `diabetes` and mortality status `endpoint`. For the purpose of demonstrate, assume that we are interested in the association between diabetes and endpoint. Other variables are considered as possible confounders. The purposes of this dataset is to illustrate those functions in **chest** package only. Therefore, we assume it is a cohort design for Cox Proportional Hazard regression, and a case-control design for logistic regression.

**Usage**

diab\_df

**Format**

A data frame with 2372 rows and 14 variables:

**Diabetes** diabetes status 1: with diabetes 0: without diabetes

**Endpoint** mortality status 1: reached end point, and 0: survived

**Age** Age, in years

**Sex** sex, 1: male, 2: Female

**BMI** Body mass index

**Married** marital status 1: married, 0: not

**Smoke** smoking status 1: smoker, 0: non-smoker

**CVD** cardiovascular disease 1: yes 0: no

**Cancer** cancer 1: yes, 0: no

**Education** education 1: high, 0: low

**Income** income 1: high, 0: low

**t0** time (age) at the start of the follow-up

**t1** time (age) at the end of the follow-up

**mid** matched set id, for conditional logistic regression

# Index

## \*Topic **datasets**

diab\_df, [11](#)

chest, [2](#)

chest\_clogit, [3](#)

chest\_cox, [4](#)

chest\_glm, [6](#)

chest\_lm, [8](#)

chest\_speedglm, [9](#)

diab\_df, [11](#)