Package 'DOVE3'

September 24, 2022

Title Durability of Vaccine Efficacy 3

Version 1.0

Description Implements maximum likelihood methods for assessing potentially time-varying effectiveness of vaccination and prior infection against potentially recurrent SARS-CoV-2 infection under staggered enrollment and time-varying community transmission, allowing crossover of placebo volunteers to the vaccine arm.

License GPL-2

Encoding UTF-8

Suggests rmarkdown, knitr

VignetteBuilder utils, knitr, rmarkdown

RoxygenNote 7.2.1

Imports Rcpp (>= 1.0.7), methods, stats, tidyverse, ggplot2

LinkingTo Rcpp, RcppArmadillo

Depends R (>= 3.5.0)

Collate 'CoxReg.R'
'RcppExports.R'
'VEplot.R'
'outcome.R'
'exposure.R'
'dove3.R'
'exampleData.R'
'processOutput.R'

R topics documented:

dove3 *Effectiveness of Vaccination and Prior Infection*

Description

Estimates the potentially waning long-term effectiveness of vaccination and prior infection against COVID-19 outcomes in observational studies. Effects of all exposures (i.e., vaccination and prior infection) are estimated simultaneously under a single Cox regression model, allowing the outcome of interest to be a recurrent event.

Usage

```
dove3(
  formula,
  data,
  vaccine_infection_interaction = FALSE,
  vaccine_knots = NULL,
  vaccine_uninfected_knots = NULL,
  vaccine_infected_knots = NULL,
  prior_infection_knots = NULL,
  related_vaccine_types = NULL,
  last_piece_constant = FALSE,
  reinfection_cutoff = 14,
  plots = TRUE
)
```

Arguments

`formula` A formula object, with the response on the left-hand side of a '~' operator, and the covariates and exposure() function on the right-hand side. The response contains the outcome information and must be specified through function 'outcome()'. The exposure() function must be used to specify the information of each exposure. See ?exposure, ?outcome, and Details for further information.

`data` A data.frame object. The data.frame in which to interpret the variable names in formula. Must contain the subject ID, the entry and censoring times, the event time, the vaccination time and vaccine type, the infection time and infection type (determined by the dominant variant), and the covariates. See Details.

`vaccine_infection_interaction` A logical object. If TRUE, interaction term between vaccination and prior infection is included in the model. That is, vaccine effects are allowed to be different between previously uninfected subjects and previously infected subjects. If FALSE (default), there is no interaction term in the model, i.e., average vaccine effects among all subjects are estimated, regardless of the prior infection status.

`vaccine_knots` A list object or NULL. The kth element specifies the knots of the piecewise linear function for the log rate/hazard ratio of the kth vaccine type. If NULL, knots will be placed at every month by default. This input is ignored when vaccine_infection_interaction = TRUE.

`vaccine_uninfected_knots` A list object or NULL. The kth element specifies the knots of the piecewise linear function for the log rate/hazard ratio of the kth vaccine type given no prior

infection before vaccination. If NULL, knots will be placed at every month by default. This input is ignored when vaccine_infection_interaction = FALSE.

`vaccine_infected_knots`
: A list object or NULL. The kth element specifies the knots of the piecewise linear function for the log rate/hazard ratio of the kth vaccine type given at least one prior infection before vaccination. If NULL, knots will be placed at every other month by default. This input is ignored when vaccine_infection_interaction = FALSE.

`prior_infection_knots`
: A list object or NULL. The kth element specifies the knots of the piecewise linear function for the log rate/hazard ratio of the kth infection type. The first knot should be placed at reinfection_cutoff. If NULL, a default set of knots will be used, with the first knot placed at reinfection_cutoff and the subsequent knots placed at every three months.

`related_vaccine_types`
: A list object or NULL. Each element of the list takes the form c(i, j) and imposes a constraint that the slope of the first piece of the piecewise linear function is the same between the ith and jth vaccine types. This input is intended to account for the fact that the 1-dose and 2-dose regimens of an mRNA vaccine should have the same effectiveness from the receipt of the first dose until the receipt of the second dose. If NULL (default), there is no constraint on the piecewise linear functions for different vaccine types during the estimation.

`last_piece_constant`
: A logical object. If FALSE (default), effectiveness is allowed to vary after the last knot. If TRUE, effectiveness is assumed to be constant after the last knot.

`reinfection_cutoff`
: A positive scalar object. Positive results at two time points separated by a gap larger than this number (in days) are considered as two different infections.

`plots`
: A logical object. If TRUE (default), plots of the estimated effectiveness of each vaccine type and prior infection type against the clinical outcome of interest and their 95% confidence intervals will be automatically generated. If FALSE, plots will not be generated.

Details

The log intensity or hazard ratio for an exposure is a piecewise linear function of time since the exposure. Specific constraint on the first and last piece is allowed. In addition, vaccine effects can be evaluated among previously uninfected and infected subjects separately.

The information required for an analysis is

Subject ID: Number that identifies subjects.

Entry Time: Calendar time when the subject enters the risk set.

Event Time: Calendar time when the subject experiences the clinical outcome of interest (e.g., SARS-CoV-2 Infection, hospitalization, or death caused by infection), with NA, Inf, or an arbitrary value greater than the censoring time if the subject does not experience an event.

Censoring Time: Calendar time when the subject moves out of the risk set.

Vaccination Time: Calendar time when vaccination takes place, with NA, Inf, or an arbitrary value greater than the study end time if the subject is not vaccinated during the study period.

Vaccine Type: Categorical variable indicating which vaccine type the subject receives. Must be an integer between 1 and the total number of vaccine types, with NA or an arbitrary value within this range if the subject is not vaccinated during the study period.

Infection Time: Calendar time when the subject is infected, with NA, Inf, or an arbitrary value that is larger than the study end time if the subject is not infected during the study period.

Infection Type: Categorical variable indicating which dominant variant the infection is associated with. Must be an integer between 1 and the total number of variants under investigation, with NA or an arbitrary value within this range if the subject is not infected during the study period.

Covariates: Baseline covariates (e.g., age, gender, ethnicity).

Note that a subject can have multiple rows and each row corresponds to an infection. The event_time is the time of the clinical outcome of interest caused by the infection in the same row, with NA, Inf, or an arbitrary value greater than the censor_time if that infection does not cause the outcome. If a subject never experiences the clinical outcome of interest during the entire study period, this subject will have only one record, with the event_time being NA, Inf, or an arbitrary value greater than the censor_time, and with the infection_time being NA, Inf, or any arbitrary value.

All the time variables are measured from the same time origin and are specified in units of days. For each individual, the entry_time and censor_time must satisfy entry_time $\leq$ censor_time. If entry_time > censor_time, the case will be removed from the analysis and a message will be generated. The definitions of entry_time and censor_time depend on specific analyses. For example, entry_time is set to be 0 for all subjects if effectiveness of primary vaccine series is of interest, and is set to be the completion time of the primary vaccine series if effectiveness of boosters is of interest. Thus, the entry_time and censor_time are not necessarily the beginning and end of the follow-up period for a subject, and it is possible that event_time is smaller than entry_time or larger than censor_time.

All the categorical variables (i.e., subject_id, vaccine_type, infection_type) must be labelled as 1, 2, . . ., K, where K is the total number of categories.

The general structure of the formula input is

```
outcome(subject_id, entry_time, event_time, censor_time) ~ covariates +
  exposure(vaccine_time, vaccine_type, infection_time, infection_type)
```

The response variable contains the information of event time and event status for each infection and must be specified through function 'outcome()'. All the variables related to vaccination and infection are specified through function 'exposure()'. The covariates can be either numerical or categorical. If categorical covariates are provided, all other categories are compared to the first category.

Value

An list object with elements

`covariates`	A matrix containing the estimated (log) hazard ratio of each covariate, together with the estimated standard error, the 95% confidence interval, and the two-sided p-value for testing no covariate effect.
`effectiveness`	A list of matrices, one for each type of exposure (vaccination comes first and prior infection next, both in order of type). Each matrix contains the daily effectiveness estimates in reducing the rate or hazard of the clinical outcome of interest, together with the standard errors and the 95% confidence intervals.
`plots`	A list of plot objects returned by ggplot(), one for each type of exposure (vaccination comes first and prior infection next, both in order of type).

References

Lin D, Gu Y, Xu Y, et al. Association of Primary and Booster Vaccination and Prior Infection With SARS-CoV-2 Infection and Severe COVID-19 Outcomes. JAMA. Published online September 26, 2022. doi:10.1001/jama.2022.17876

Examples

```
data(exampleData)

# specify the knots for each exposure
vaccine.knots = list("vac.type1" = c(30),
                     "vac.type2" = c(30,60),
                     "vac.type3" = c(30,60))
prior.infection.knots = list("inf.type1" = c(14, 120),
                             "inf.type2" = c(14,120))

# Fit the simple model without interaction or related vaccine types
fit1 <- dove3(formula = outcome(subject.id, entry.time, event.time, censor.time) ~
                age + gender + priority +
                exposure(Vtime, Vtype, infection.time, infection.type),
      data = exampleData,
      vaccine_knots = vaccine.knots,
      prior_infection_knots = prior.infection.knots)

# Specify the knots for vaccination without and with prior infection
vaccine.uninfected.knots = list("vac.noinf.type1" = c(30),
                                "vac.noinf.type2" = c(30,60),
                                "vac.noinf.type3" = c(30,60))

vaccine.infected.knots = list("vac.noinf.type1" = c(30),
                                "vac.noinf.type2" = c(60),
                                "vac.noinf.type3" = c(60))

# Fit the model with interaction between vaccination and prior infection status,
# and impose a constraint on the first pieces of the first two vaccine types.
fit2 <- dove3(formula = outcome(subject.id, entry.time, event.time, censor.time) ~
                age + gender + priority +
                exposure(Vtime, Vtype, infection.time, infection.type),
      data = exampleData,
      vaccine_infection_interaction = TRUE,
      vaccine_uninfected_knots = vaccine.uninfected.knots,
      vaccine_infected_knots = vaccine.infected.knots,
      prior_infection_knots = prior.infection.knots,
      related_vaccine_types = list(c(1,2)))
```

---

exampleData *Toy Dataset For Illustration*

Description

This data set is provided for the purposes of illustrating the use of the software. It is a simulated data set that mimics the surveillance data of vaccination and infection.

Usage

```
data(exampleData)
```

Format

exampleData is a data.frame containing 8,000 records. The data.frame contains 11 columns,

subject.id The subject ID of each record.

event.time The event (infection, hospitalization or death) time in days. Infection is treated as recurrent event.

censor.time The censoring time for each subject in days.

entry.time The entry time in days.

Vtime The time of vaccination/booster in days.

Vtype The type of vaccination/booster.

infection.time The time of infection in days. If the outcome is infection, infection.time is the same as event.time.

infection.type The type of infection.

age A categorical variable of age (<18, 18-34, 35-49, 50-64, >=65).

gender A binary indicator of gender (1 for male and 0 for female).

priority A composite baseline risk score taking values 1-3.

---

exposure *Specify Exposure Variables*

Description

This function is used in the model statement of dove3() to specify the vaccination time, vaccine type, infection time, and infection type.

Usage

```
exposure(vaccine_time, vaccine_type, infection_time, infection_type)
```

Arguments

`vaccine_time`	The variable for the time when the subject is vaccinated, with NA, Inf, or an arbitrary value that is larger than the study end time if the subject is never vaccinated during the study period.
`vaccine_type`	The variable for the vaccine type the subject receives, must be an integer between 1 and the total number of vaccine types, with NA or an arbitrary value within this range if the subject is never vaccinated during the study period.
`infection_time`	The variable for the time when the subject is infected, with NA, Inf, or an arbitrary value that is larger than the study end time if the subject is never infected during the study period.
`infection_type`	The variable for the dominant variant at the time when the subject is infected, must be an integer between 1 and the total number of variants under investigation, with NA or an arbitrary value within this range if the subject is never infected during the study period.

Value

This function is intended to be used only in the model statement of dove3(). The result, a matrix, is used internally.

`outcome` *Specify Outcome Variables*

Description

This function is used in the model statement of dove3() to specify the subject id, entry time, event time, and censoring time.

Usage

```
outcome(subject_id, entry_time, event_time, censor_time)
```

Arguments

`subject_id`	The variable for the subject ID that is used to identify records from the same subject, must be an integer between 1 and the total number of subjects.
`entry_time`	The variable for the time when the subject enters the risk set. For example, entry_time is set to be 0 for all subjects if effectiveness of primary vaccine series is of interest, and is set to be the time of primary vaccine series if effectiveness of boosters is of interest.
`event_time`	The variable for the time when the subject has an event.
`censor_time`	The variable for the time when the subject is censored. For example, subjects are censored when they receive boosters if effectiveness of primary vaccine series is of interest. Likewise, subjects are censored when they receive second boosters if effectiveness of first boosters is of interest.

Value

This function is intended to be used only in the model statement of dove3(). The result, a matrix, is used internally.

Index