

# Package ‘incidence’

August 24, 2018

**Type** Package

**Title** Compute, Handle, Plot and Model Incidence of Dated Events

**Version** 1.4.1

**Description** Provides functions and classes to compute, handle and visualise incidence from dated events for a defined time interval. Dates can be provided in various standard formats. The class 'incidence' is used to store computed incidence and can be easily manipulated, subsetted, and plotted. In addition, log-linear models can be fitted to 'incidence' objects using 'fit'. This package is part of the RECON (<<http://www.repidemicsconsortium.org/>>) toolkit for outbreak analysis.

**Encoding** UTF-8

**License** MIT + file LICENSE

**URL** <http://www.repidemicsconsortium.org/incidence/>

**BugReports** <http://github.com/reconhub/incidence/issues>

**RoxygenNote** 6.1.0

**Imports** ggplot2, ISOweek

**Suggests** magrittr, outbreaks, testthat, vdiffr, covr, knitr, rmarkdown, scales, cowplot

**VignetteBuilder** knitr

**NeedsCompilation** no

**Author** Thibaut Jombart [aut],  
Zhian N. Kamvar [aut, cre],  
Rich FitzJohn [aut],  
Jun Cai [ctb],  
Sangeeta Bhatia [ctb]

**Maintainer** Zhian N. Kamvar <[zkamvar@gmail.com](mailto:zkamvar@gmail.com)>

**Repository** CRAN

**Date/Publication** 2018-08-24 20:32:20 UTC

**R topics documented:**

as.data.frame.incidence . . . . .	2
bootstrap . . . . .	4
cumulate . . . . .	5
estimate_peak . . . . .	6
find_peak . . . . .	8
fit . . . . .	9
get_counts . . . . .	11
get_dates . . . . .	12
get_fit . . . . .	13
get_interval . . . . .	15
incidence . . . . .	16
incidence_pall . . . . .	19
plot.incidence . . . . .	20
pool . . . . .	22
subset.incidence . . . . .	23
<b>Index</b>	<b>25</b>

---

as.data.frame.incidence  
*Conversion of incidence objects*

---

**Description**

These functions convert incidence objects into other classes.

**Usage**

```
## S3 method for class 'incidence'
as.data.frame(x, ..., long = FALSE)

as.incidence(x, ...)

## S3 method for class 'matrix'
as.incidence(x, dates = NULL, interval = NULL,
             isoweeks = TRUE, ...)

## S3 method for class 'data.frame'
as.incidence(x, dates = NULL, interval = NULL,
             isoweeks = TRUE, ...)

## S3 method for class 'numeric'
as.incidence(x, dates = NULL, interval = NULL,
             isoweeks = TRUE, ...)
```

## Arguments

x	An incidence object, or an object to be converted as incidence (see details).
...	Further arguments passed to other functions (no used).
long	A logical indicating if the output data.frame should be 'long', i.e. where a single column containing 'groups' is added in case of data computed on several groups.
dates	A vector of dates, each corresponding to the (inclusive) lower limit of the bins.
interval	An integer indicating the time interval used in the computation of the incidence. If NULL, it will be determined from the first time interval between provided dates. If only one date is provided, it will trigger an error.
isoweeks	A logical indicating whether isoweeks should be used in the case of weekly incidence; defaults to TRUE.

## Details

Conversion to incidence objects should only be done when the original dates are not available. In such case, the argument x should be a matrix corresponding to the \$counts element of an incidence object, i.e. giving counts with time intervals in rows and named groups in columns. In the absence of groups, a single unnamed columns should be given. data.frame and vectors will be coerced to a matrix.

## Author(s)

Thibaut Jombart <thibautjombart@gmail.com>, Rich Fitzjohn

## See Also

the [incidence\(\)](#) function to generate the 'incidence' objects.

## Examples

```
## create fake data
data <- c(0,1,1,2,1,3,4,5,5,5,5,4,4,26,6,7,9)
sex <- sample(c("m","f"), length(data), replace=TRUE)

## get incidence per group (sex)
i <- incidence(data, groups = sex)
i
plot(i)

## convert to data.frame
as.data.frame(i)

## same, 'long format'
as.data.frame(i, long = TRUE)

## conversion from a matrix of counts to an incidence object
i$counts
```

```
new_i <- as.incidence(i$counts, i$dates)
new_i
all.equal(i, new_i)
```

---

bootstrap

*Bootstrap incidence time series*

---

### Description

This function can be used to bootstrap incidence objects. Bootstrapping is done by sampling with replacement the original input dates. See details for more information on how this is implemented.

### Usage

```
bootstrap(x, randomise_groups = FALSE)
```

### Arguments

**x** An incidence object.

**randomise\_groups** A logical indicating whether groups should be randomised as well in the re-sampling procedure; respective group sizes will be preserved, but this can be used to remove any group-specific temporal dynamics. If FALSE (default), data are resampled within groups.

### Details

As original data are not stored in incidence objects, the bootstrapping is achieved by multinomial sampling of date bins weighted by their relative incidence.

### Value

An incidence object.

### Author(s)

Thibaut Jombart <thibautjombart@gmail.com>

### See Also

[incidence::find\\_peak](#) to use estimate peak date using bootstrap

## Examples

```
if (require(outbreaks) && require(ggplot2)) { withAutoprint({
  i <- incidence(fluH7N9_china_2013$date_of_onset)
  i
  plot(i)

  ## one simple bootstrap
  x <- bootstrap(i)
  x
  plot(x)

})}
```

---

cumulate	<i>Compute cumulative 'incidence'</i>
----------	---------------------------------------

---

## Description

cumulate is an S3 generic to compute cumulative numbers, with methods for different types of objects:

## Usage

```
cumulate(x)

## Default S3 method:
cumulate(x)

## S3 method for class 'incidence'
cumulate(x)
```

## Arguments

x                    An incidence object.

## Details

- default method is a wrapper for cumsum
- incidence objects: computes cumulative incidence over time
- projections objects: same, for projections objects, implemented in the similarly named package; see ?cumulate.projections for more information, after loading the package

## Author(s)

Thibaut Jombart <thibautjombart@gmail.com>

**See Also**

The `incidence()` function to generate the 'incidence' objects.

**Examples**

```
dat <- as.integer(c(0,1,2,2,3,5,7))
group <- factor(c(1, 2, 3, 3, 3, 3, 1))
i <- incidence(dat, groups = group)
i
plot(i)

i_cum <- cumulate(i)
i_cum
plot(i_cum)
```

---

estimate\_peak

*Estimate the peak date of an incidence curve using bootstrap*

---

**Description**

This function can be used to estimate the peak of an epidemic curve stored as incidence, using bootstrap. See `incidence::bootstrap` for more information on the resampling.

**Usage**

```
estimate_peak(x, n = 100, alpha = 0.05)
```

**Arguments**

x	An incidence object.
n	The number of bootstrap datasets to be generated; defaults to 100.
alpha	The type 1 error chosen for the confidence interval; defaults to 0.05.

**Details**

Input dates are resampled with replacement to form bootstrapped datasets; the peak is reported for each, resulting in a distribution of peak times. When there are ties for peak incidence, only the first date is reported.

Note that the bootstrapping approach used for estimating the peak time makes the following assumptions:

- the total number of event is known (no uncertainty on total incidence)
- dates with no events (zero incidence) will never be in bootstrapped datasets
- the reporting is assumed to be constant over time, i.e. every case is equally likely to be reported

**Value**

A list containing the following items:

- observed: the peak incidence of the original dataset
- estimated: the mean peak time of the bootstrap datasets
- ci: the confidence interval based on bootstrap datasets
- peaks: the peak times of the bootstrap datasets

**Author(s)**

Thibaut Jombart <thibautjombart@gmail.com>, with inputs on caveats from Michael Höhle.

**See Also**

[incidence::bootstrap](#) for the bootstrapping underlying this approach and [incidence::find\\_peak](#) to find the peak in a single incidence object.

**Examples**

```
if (require(outbreaks) && require(ggplot2)) { withAutoprint({
  i <- incidence(fluH7N9_china_2013$date_of_onset)
  i
  plot(i)

  ## one simple bootstrap
  x <- bootstrap(i)
  x
  plot(x)

  ## find 95% CI for peak time using bootstrap
  peak_data <- estimate_peak(i)
  peak_data
  summary(peak_data$peaks)

  ## show confidence interval
  plot(i) + geom_vline(xintercept = peak_data$ci, col = "red", lty = 2)

  ## show the distribution of bootstrapped peaks
  df <- data.frame(peak = peak_data$peaks)
  plot(i) + geom_density(data = df,
                        aes(x = peak, y = 10 * ..scaled..),
                        alpha = .2, fill = "red", color = "red")
})}
```

---

`find_peak`*Find the peak date of an incidence curve*

---

**Description**

This function can be used to find the peak of an epidemic curve stored as an incidence object.

**Usage**

```
find_peak(x, pool = TRUE)
```

**Arguments**

<code>x</code>	An incidence object.
<code>pool</code>	If TRUE (default), any groups will be pooled before finding a peak. If FALSE, separate peaks will be found for each group.

**Value**

The date of the (first) highest incidence in the data.

**Author(s)**

Thibaut Jombart <thibautjombart@gmail.com>, Zhian N. Kamvar <zkamvar@gmail.com>

**See Also**

[estimate\\_peak\(\)](#) for bootstrap estimates of the peak time

**Examples**

```
if (require(outbreaks) && require(ggplot2)) { withAutoprint({
  i <- incidence(fluH7N9_china_2013$date_of_onset)
  i
  plot(i)

  ## one simple bootstrap
  x <- bootstrap(i)
  x
  plot(x)

  ## find 95% CI for peak time using bootstrap
  find_peak(i)

  ## show confidence interval
  plot(i) + geom_vline(xintercept = find_peak(i), col = "red", lty = 2)
```



```
}}}
```

---

 fit

*Fit exponential models to incidence data*


---

## Description

The function `fit` fits two exponential models to incidence data, of the form:  $\log(y) = r * t + b$  where 'y' is the incidence, 't' is time (in days), 'r' is the growth rate, and 'b' is the origin. The function `fit` will fit one model by default, but will fit two models on either side of a splitting date (typically the peak of the epidemic) if the argument `split` is provided. When groups are present, these are included in the model as main effects and interactions with dates. The function `fit_optim_split()` can be used to find the optimal 'splitting' date, defined as the one for which the best average R2 of the two models is obtained. Plotting can be done using `plot`, or added to an existing incidence plot by the piping-friendly function `add_incidence_fit()`.

## Usage

```
fit(x, split = NULL, level = 0.95, quiet = FALSE)

fit_optim_split(x, window = x$timespan/4, plot = TRUE, quiet = TRUE,
  separate_split = TRUE)

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

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

## Arguments

<code>x</code>	An incidence object, generated by the function <code>incidence()</code> . For the plotting function, an <code>incidence_fit</code> object.
<code>split</code>	An optional time point identifying the separation between the two models. If <code>NULL</code> , a single model is fitted. If provided, two models would be fitted on the time periods on either side of the split.
<code>level</code>	The confidence interval to be used for predictions; defaults to 95%.
<code>quiet</code>	A logical indicating if warnings from <code>fit</code> should be hidden; <code>FALSE</code> by default. Warnings typically indicate some zero incidence, which are removed before performing the log-linear regression.
<code>window</code>	The size, in days, of the time window either side of the split.
<code>plot</code>	A logical indicating whether a plot should be added to the output ( <code>TRUE</code> , default), showing the mean R2 for various splits.

`separate_split` If groups are present, should separate split dates be determined for each group? Defaults to TRUE, in which separate split dates and thus, separate models will be constructed for each group. When FALSE, the split date will be determined from the pooled data and modelled with the groups as main effects and interactions with date.

... currently unused.

## Value

For `fit()`, a list with the class `incidence_fit` (for a single model), or a list containing two `incidence_fit` objects (when fitting two models). `incidence_fit` objects contain:

- `$model`: the fitted linear model
- `$info`: a list containing various information extracted from the model (detailed further)
- `$origin`: the date corresponding to day '0'

The `$info` item is a list containing:

- `r`: the growth rate
- `r.conf`: the confidence interval of 'r'
- `pred`: a `data.frame` containing predictions of the model, including the true dates (`dates`), their numeric version used in the model (`dates.x`), the predicted value (`fit`), and the lower (`lwr`) and upper (`upr`) bounds of the associated confidence interval.
- `doubling`: the predicted doubling time in days; exists only if 'r' is positive
- `doubling.conf`: the confidence interval of the doubling time
- `halving`: the predicted halving time in days; exists only if 'r' is negative
- `halving.conf`: the confidence interval of the halving time

For `fit_optim_split`, a list containing:

- `df`: a `data.frame` of dates that were used in the optimization procedure, and the corresponding average R2 of the resulting models.
- `split`: the optimal splitting date
- `fit`: an `incidence_fit_list` object containing the fit for each split. If the `separate_split = TRUE`, then the `incidence_fit_list` object will contain these splits nested within each group. All of the `incidence_fit` objects can be retrieved with `get_fit()`.
- `plot`: a plot showing the content of `df` (`ggplot2` object)

## Author(s)

Thibaut Jombart <[thibautjombart@gmail.com](mailto:thibautjombart@gmail.com)>, Zhian N. Kamvar <[zkamvar@gmail.com](mailto:zkamvar@gmail.com)>.

## See Also

the `incidence()` function to generate the 'incidence' objects. The `get_fit()` function to flatten `incidence_fit_list` objects to a list of `incidence_fit` objects.

**Examples**

```

if (require(outbreaks)) { withAutoprint({
  dat <- ebola_sim$linelist$date_of_onset

  ## EXAMPLE WITH A SINGLE MODEL

  ## compute weekly incidence
  i.7 <- incidence(dat, interval=7)
  plot(i.7)
  plot(i.7[1:20])

  ## fit a model on the first 20 weeks
  f <- fit(i.7[1:20])
  f
  names(f)
  head(get_info(f, "pred"))

  ## plot model alone (not recommended)
  plot(f)

  ## plot data and model (recommended)
  plot(i.7, fit = f)
  plot(i.7[1:25], fit = f)

  ## piping versions
  if (require(magrittr)) { withAutoprint({
    plot(i.7) %>% add_incidence_fit(f)

    ## EXAMPLE WITH 2 PHASES
    ## specifying the peak manually
    f2 <- fit(i.7, split = as.Date("2014-10-15"))
    f2
    plot(i.7) %>% add_incidence_fit(f2)

    ## finding the best 'peak' date
    f3 <- fit_optim_split(i.7)
    f3
    plot(i.7) %>% add_incidence_fit(f3$fit)
  })}

  })}

```

---

get\_counts

*Get counts from an incidence object*


---

**Description**

Get counts from an incidence object

**Usage**

```
get_counts(x, groups = NULL)

## S3 method for class 'incidence'
get_counts(x, groups = NULL)
```

**Arguments**

x                    an incidence object.

groups                if there are groups, use this to specify a group or groups to subset. Defaults to NULL indicating that all groups are returned.

**Value**

a matrix of counts where each row represents a date bin

**Examples**

```
if (require(outbreaks)) { withAutoprint({
  dat <- ebola_sim$linelist$date_of_onset
  gend <- ebola_sim$linelist$gender
  i <- incidence(dat, interval = "week", groups = gend)

  ## Use with an object and no arguments gives the counts matrix
  head(get_counts(i))

  ## Specifying a position or group name will return a matrix subset to that
  ## group
  head(get_counts(i, 1L))
  head(get_counts(i, "f"))

  ## Specifying multiple groups allows you to rearrange columns
  head(get_counts(i, c("m", "f")))

  ## If you want a vector, you can use drop
  drop(get_counts(i, "f"))
})}
```

---

get\_dates

*Retrieve dates from an incidence object*

---

**Description**

Retrieve dates from an incidence object

**Usage**

```

get_dates(x, ...)

## Default S3 method:
get_dates(x, ...)

## S3 method for class 'incidence'
get_dates(x, position = "left", count_days = FALSE,
  ...)

```

**Arguments**

x	an <a href="#">incidence</a> object
...	Unused
position	One of "left", "center", "middle", or "right" specifying what side of the bin the date should be drawn from.
count_days	If TRUE, the result will be represented as the number of days from the first date.

**Value**

a vector of dates or numerics

**Examples**

```

set.seed(999)
dat <- as.Date(Sys.Date()) + sample(-3:50, 100, replace = TRUE)
x <- incidence(dat, interval = "month")
get_dates(x)
get_dates(x, position = "middle")
set.seed(999)
dat <- as.Date(Sys.Date()) + sample(-3:50, 100, replace = TRUE)
x <- incidence(dat, interval = "month")
get_dates(x)
get_dates(x, "center")
get_dates(x, "right")

# Return dates by number of days from the first date
get_dates(x, count_days = TRUE)
get_dates(incidence(-5:5), count_days = TRUE)

```

---

get\_fit

*Accessors for incidence\_fit objects*


---

**Description**

Accessors for `incidence_fit` objects

**Usage**

```

get_fit(x)

## S3 method for class 'incidence_fit'
get_fit(x)

## S3 method for class 'incidence_fit_list'
get_fit(x)

get_info(x, what = "r", ...)

## S3 method for class 'incidence_fit'
get_info(x, what = "r", ...)

## S3 method for class 'incidence_fit_list'
get_info(x, what = "r", groups = NULL,
         na.rm = TRUE, ...)

```

**Arguments**

x	an incidence_fit or incidence_fit_list object.
what	the name of the item in the "info" element of the incidence_fit object.
...	currently unused.
groups	if what = "pred" and x is an incidence_fit_list object, then this indicates what part of the nesting hierarchy becomes the column named "groups". Defaults to NULL, indicating that no groups column will be added/modified.
na.rm	when TRUE (default), missing values will be excluded from the results.

**Value**

a list of incidence\_fit objects.

**Examples**

```

if (require(outbreaks)) { withAutoprint({

  dat <- ebola_sim$linelist$date_of_onset

  ## EXAMPLE WITH A SINGLE MODEL

  ## compute weekly incidence
  sex <- ebola_sim$linelist$gender
  i.sex <- incidence(dat, interval = 7, group = sex)

  ## Compute the optimal split for each group separately
  fits <- fit_optim_split(i.sex, separate_split = TRUE)

  ## `fits` contains an `incidence_fit_list` object

```

```

fits$fit

## Grab the list of `incidence_fit` objects
get_fit(fits$fit)

## Get the predictions for all groups
get_info(fits$fit, "pred", groups = 1)

## Get the predictions, but set `groups` to "before" and "after"
get_info(fits$fit, "pred", groups = 2)

## Get the reproduction number
get_info(fits$fit, "r")

## Get the doubling confidence interval
get_info(fits$fit, "doubling.conf")

## Get the halving confidence interval
get_info(fits$fit, "halving.conf")
}}

```

---

get\_interval

*Get the interval for an incidence object.*


---

## Description

Get the interval for an incidence object.

## Usage

```

get_interval(x, ...)

## S3 method for class 'incidence'
get_interval(x, integer = TRUE, ...)

```

## Arguments

x	an <a href="#">incidence</a> object.
...	Unused
integer	When TRUE (default), the interval will be converted to an integer vector if it is stored as a character in the incidence object.

## Value

if integer = TRUE: an integer vector, otherwise: the value stored in x\$interval

**Examples**

```
set.seed(999)
dat <- as.Date(Sys.Date()) + sample(-3:50, 100, replace = TRUE)
x <- incidence(dat, interval = "month")
get_interval(x)
get_interval(x, integer = FALSE)
```

---

incidence

---

*Compute incidence of events from a vector of dates.*


---

**Description**

This function computes incidence based on dates of events provided in various formats. A fixed interval, provided as numbers of days, is used to define time intervals. Counts within an interval always include the first date, after which they are labeled, and exclude the second. For instance, intervals labeled as 0, 3, 6, ... mean that the first bin includes days 0, 1 and 2, the second interval includes 3, 4 and 5 etc.

**Usage**

```
incidence(dates, interval = 1L, ...)

## Default S3 method:
incidence(dates, interval = 1L, ...)

## S3 method for class 'Date'
incidence(dates, interval = 1L, standard = TRUE,
  groups = NULL, na_as_group = TRUE, first_date = NULL,
  last_date = NULL, ...)

## S3 method for class 'integer'
incidence(dates, interval = 1L, groups = NULL,
  na_as_group = TRUE, first_date = NULL, last_date = NULL, ...)

## S3 method for class 'numeric'
incidence(dates, interval = 1L, groups = NULL,
  na_as_group = TRUE, first_date = NULL, last_date = NULL, ...)

## S3 method for class 'POSIXt'
incidence(dates, interval = 1L, standard = TRUE,
  groups = NULL, na_as_group = TRUE, first_date = NULL,
  last_date = NULL, ...)

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



**Arguments**

<code>dates</code>	A vector of dates, which can be provided as objects of the class: <code>integer</code> , <code>numeric</code> , <code>Date</code> , <code>POSIXct</code> . Note that decimal numbers will be floored with a warning.
<code>interval</code>	An integer or character indicating the (fixed) size of the time interval used for computing the incidence; defaults to 1 day. This can also be a text string that corresponds to a valid date interval: <code>day</code> , <code>week</code> , <code>month</code> , <code>quarter</code> , or <code>year</code> . See Note.
<code>...</code>	Additional arguments passed to other methods (none are used).
<code>standard</code>	(Only applicable to <code>Date</code> objects) When <code>TRUE</code> (default) and the <code>interval</code> one of <code>"week"</code> , <code>"month"</code> , <code>"quarter"</code> , or <code>"year"</code> , then this will cause the bins for the counts to start at the beginning of the interval (See Note). This is overridden by defining a non- <code>NULL</code> <code>first_date</code> .
<code>groups</code>	An optional factor defining groups of observations for which incidence should be computed separately.
<code>na_as_group</code>	A logical value indicating if missing group ( <code>NA</code> ) should be treated as a separate group.
<code>first_date</code> , <code>last_date</code>	optional first/last dates to be used in the epicurve. When these are <code>NULL</code> (default), the dates from the first/last dates are taken from the observations. If these dates are provided, the observations will be trimmed to the range of <code>[first_date, last_date]</code> .
<code>x</code>	An 'incidence' object.

**Details**

For details about the `incidence` class, see the dedicated vignette:  
`vignette("incidence_class", package = "incidence")`

**Value**

An list with the class `incidence`, which contains the following items:

- **dates**: The dates marking the left side of the bins used for counting events. When ISO week-based weekly incidence is computed, the dates are the first days of corresponding isoweeks.
- **counts**: A matrix of incidence counts, which one column per group (and a single column if no groups were used).
- **timespan**: The length of the period for which incidence is computed, in days.
- **interval**: The bin size, in number of days; e.g. `7` indicates weekly incidence.
- **n**: The total number of cases.
- **isoweeks**: ISO 8601 week format `yyyy-Www`, which is returned only when ISO week-based weekly incidence is computed.

**Note**

If `interval` is a valid character (e.g. "week" or "month"), then the bin will start at the beginning of the interval just before the first observation by default. For example, if the first case was recorded on Wednesday, 2018-05-09:

- "week" : first day of the ISOweek (i.e. Monday, 2018-05-07)
- "month" : first day of the month (i.e. 2018-05-01)
- "quarter" : first day of the quarter (i.e. 2018-04-01)
- "year" : first day of the calendar year (i.e. 2018-01-01)

These default intervals can be overridden in two ways:

1. Specify `standard = FALSE`, which sets the interval to begin at the first observed case.
2. Specify a date in the `first_date` field.

The intervals for "month", "quarter", and "year" will necessarily vary in the number of days they encompass and warnings will be generated when the first date falls outside of a calendar date that is easily represented across the interval.

**Author(s)**

Thibaut Jombart, Rich Fitzjohn, Zhian Kamvar

**See Also**

The main other functions of the package include:

- `incidence::plot.incidence()`: Plot epicurves from an incidence object.
- `incidence::fit()`: Fit log-linear model to computed incidence.
- `incidence::fit_optim_split()`: Find the optimal peak of the epidemic and fits log-linear models on either side of the peak.
- `incidence::subset()`: Handling of incidence objects.
- `incidence::pool()`: Sum incidence over groups.
- `incidence::as.data.frame.incidence()`: Convert an incidence object to a data.frame.

The following vignettes are also available:

- `overview`: Provides an overview of the package's features.
- `customize_plot`: Provides some tips on finer plot customization.
- `incidence_class`: Details the content of the incidence class.

**Examples**

```

## toy example
incidence(c(1, 5, 8, 3, 7, 2, 4, 6, 9, 2))
incidence(c(1, 5, 8, 3, 7, 2, 4, 6, 9, 2), 2)

## example using simulated dataset
if(require(outbreaks)) { withAutoprint({
  onset <- ebola_sim$linelist$date_of_onset

  ## daily incidence
  inc <- incidence(onset)
  inc
  plot(inc)

  ## weekly incidence
  inc.week <- incidence(onset, interval = 7, standard = FALSE)
  inc.week
  plot(inc.week)
  plot(inc.week, border = "white") # with visible border
  inc.isoweek <- incidence(onset, interval = 7, standard = TRUE)
  inc.isoweek
  ## use group information
  sex <- ebola_sim$linelist$gender
  inc.week.gender <- incidence(onset, interval = 7,
                              groups = sex, standard = FALSE)

  inc.week.gender
  head(inc.week.gender$counts)
  plot(inc.week.gender)
  inc.isoweek.gender <- incidence(onset, interval = 7,
                                 groups = sex, standard = TRUE)

  inc.isoweek.gender
}}})

```

---

incidence\_pal1

*Color palettes used in incidence*


---

**Description**

These functions are color palettes used in incidence.

**Usage**

```
incidence_pal1(n)
```

```
incidence_pal1_light(n)
```

```
incidence_pal1_dark(n)
```

**Arguments**

n                    a number of colors

**Author(s)**

Thibaut Jombart <thibautjombart@gmail.com>

**Examples**

```
plot(1:4, cex=8, pch=20, col = incidence_pal1(4),
     main = "palette: incidence_pal1")
plot(1:100, cex=8, pch=20, col = incidence_pal1(100),
     main = "palette: incidence_pal1")
plot(1:100, cex=8, pch=20, col = incidence_pal1_light(100),
     main = "palette: incidence_pal1_light")
plot(1:100, cex=8, pch=20, col = incidence_pal1_dark(100),
     main = "palette: incidence_pal1_dark")
```

---

plot.incidence

*Plot function for incidence objects*

---

**Description**

This function is used to visualise the output of the [incidence\(\)](#) function, using the package `ggplot2`.

**Usage**

```
## S3 method for class 'incidence'
plot(x, ..., fit = NULL, stack = is.null(fit),
     color = "black", border = NA, col_pal = incidence_pal1,
     alpha = 0.7, xlab = "", ylab = NULL,
     labels_iso = !is.null(x$isoweeks), n_breaks = 6)

add_incidence_fit(p, x, col_pal = incidence_pal1)

## S3 method for class 'incidence_fit'
plot(x, ...)

## S3 method for class 'incidence_fit_list'
plot(x, ...)
```

**Arguments**

x	An incidence object, generated by the function <code>incidence()</code> .
...	Further arguments passed to other methods (currently not used).
fit	An 'incidence_fit' object as returned by <code>fit()</code> .
stack	A logical indicating if bars of multiple groups should be stacked, or displayed side-by-side.
color	The color to be used for the filling of the bars; NA for invisible bars; defaults to "black".
border	The color to be used for the borders of the bars; NA for invisible borders; defaults to NA.
col_pal	The color palette to be used for the groups; defaults to <code>incidence_pal1</code> . See <code>incidence_pal1()</code> for other palettes implemented in <code>incidence</code> .
alpha	The alpha level for color transparency, with 1 being fully opaque and 0 fully transparent; defaults to 0.7.
xlab	The label to be used for the x-axis; empty by default.
ylab	The label to be used for the y-axis; by default, a label will be generated automatically according to the time interval used in <code>incidence</code> computation.
labels_iso	a logical value indicating whether labels x axis tick marks are in ISO 8601 week format <code>yyyy-Www</code> when plotting ISO week-based weekly incidence; defaults to be TRUE.
n_breaks	the ideal number of breaks to be used for the x-axis labeling
p	An existing incidence plot.

**Author(s)**

Thibaut Jombart <thibautjombart@gmail.com> Zhian N. Kamvar <zkamvar@gmail.com>

**See Also**

The `incidence()` function to generate the 'incidence' objects.

**Examples**

```
if(require(outbreaks)) { withAutoprint({
  onset <- ebola_sim$linelist$date_of_onset

  ## daily incidence
  inc <- incidence(onset)
  inc
  plot(inc)

  ## weekly incidence
  inc.week <- incidence(onset, interval = 7)
  inc.week
  plot(inc.week) # default to label x axis tick marks with isoweeks
```

```

plot(inc.week, labels_iso = FALSE) # label x axis tick marks with dates
plot(inc.week, border = "white") # with visible border

## use group information
sex <- ebola_sim$linelist$gender
inc.week.gender <- incidence(onset, interval = 7, groups = sex)
plot(inc.week.gender)
plot(inc.week.gender, labels_iso = FALSE)

## adding fit
fit <- fit_optim_split(inc.week.gender)$fit
plot(inc.week.gender, fit = fit)
plot(inc.week.gender, fit = fit, labels_iso = FALSE)
}}

```

---

pool

*Pool 'incidence' across groups*

---

### Description

This function pools incidence across all groups of an incidence object. The resulting `incidence()` object will contains counts summed over all groups present in the input.

### Usage

```
pool(x)
```

### Arguments

x                    An 'incidence' object.

### Author(s)

Thibaut Jombart <thibautjombart@gmail.com>

### See Also

The `incidence()` function to generate the 'incidence' objects.

### Examples

```

dat <- as.integer(c(0,1,2,2,3,5,7))
group <- factor(c(1, 2, 3, 3, 3, 3, 1))
i <- incidence(dat, groups = group)
i
i$counts

## pool all groups
pool(i)

```

```
pool(i)$counts

## pool only groups 1 and 3
pool(i[,c(1,3)])
pool(i[,c(1,3)]$counts
```

---

subset.incidence      *Subsetting 'incidence' objects*

---

### Description

Two functions can be used to subset incidence objects. The function `subset` permits to retain dates within a specified range and, optionally, specific groups. The operator "[" can be used as for matrices, using the syntax `x[i, j]` where 'i' is a subset of dates, and 'j' is a subset of groups.

### Usage

```
## S3 method for class 'incidence'
subset(x, ..., from = min(x$dates),
       to = max(x$dates), groups = TRUE)

## S3 method for class 'incidence'
x[i, j]
```

### Arguments

x	An incidence object, generated by the function <code>incidence()</code> .
...	Further arguments passed to other methods (not used).
from	The starting date; data strictly before this date are discarded.
to	The ending date; data strictly after this date are discarded.
groups	(optional) The groups to retained, indicated as subsets of the columns of <code>x\$counts</code> .
i	a subset of dates to retain
j	a subset of groups to retain

### Author(s)

Thibaut Jombart <thibautjombart@gmail.com>

### See Also

The `incidence()` function to generate the 'incidence' objects.

**Examples**

```
## example using simulated dataset
if(require(outbreaks)) { withAutoprint({
  onset <- ebola_sim$linelist$date_of_onset

  ## weekly incidence
  inc <- incidence(onset, interval = 7)
  inc
  inc[1:10] # first 10 weeks
  plot(inc[1:10])
  inc[-c(11:15)] # remove weeks 11-15
  plot(inc[-c(11:15)])
}}}
```



# Index

## \*Topic **accessors**

get\_dates, [12](#)  
get\_interval, [15](#)  
[.incidence (subset.incidence), [23](#)

add\_incidence\_fit (plot.incidence), [20](#)  
as.data.frame.incidence, [2](#)  
as.incidence (as.data.frame.incidence),  
[2](#)

bootstrap, [4](#)

cumulate, [5](#)

estimate\_peak, [6](#)  
estimate\_peak(), [8](#)

find\_peak, [8](#)  
fit, [9](#)  
fit(), [21](#)  
fit\_optim\_split (fit), [9](#)

get\_counts, [11](#)  
get\_dates, [12](#)  
get\_fit, [13](#)  
get\_fit(), [10](#)  
get\_info (get\_fit), [13](#)  
get\_interval, [15](#)

incidence, [13](#), [15](#), [16](#)  
incidence(), [3](#), [6](#), [9](#), [10](#), [20–23](#)  
incidence::as.data.frame.incidence(),  
[18](#)

incidence::bootstrap, [6](#), [7](#)  
incidence::find\_peak, [4](#), [7](#)  
incidence::fit(), [18](#)  
incidence::fit\_optim\_split(), [18](#)  
incidence::plot.incidence(), [18](#)  
incidence::pool(), [18](#)  
incidence::subset(), [18](#)  
incidence\_pal1, [19](#)

incidence\_pal1(), [21](#)  
incidence\_pal1\_dark (incidence\_pal1), [19](#)  
incidence\_pal1\_light (incidence\_pal1),  
[19](#)

palettes (incidence\_pal1), [19](#)  
plot.incidence, [20](#)  
plot.incidence\_fit (plot.incidence), [20](#)  
plot.incidence\_fit\_list  
(plot.incidence), [20](#)  
pool, [22](#)  
print.incidence (incidence), [16](#)  
print.incidence\_fit (fit), [9](#)  
print.incidence\_fit\_list (fit), [9](#)

subset.incidence, [23](#)