Plotting Cumulative Incidence Curves

Zachary McCaw

2024-11-02

Abstract

This vignette illustrates plotting cumulative incidence for one or more strata.

```r
suppressPackageStartupMessages({
  library(CICs)
  library(dplyr)
  library(ggplot2)
})
```

Cumulative incidence of the event of interest

Built-in functions are available for plotting the cumulative incidence curves (CICs) of the event of interest for 1 or 2 treatment arms.

One-sample setting

Case of a single group or treatment arm.

Simulate data

```r
set.seed(100)
n <- 1e2

# Single arm.
data <- CICs::GenData(
  n = n,
  event_rate = 1.0,
  death_rate = 0.25,
  censor_rate = 0.25,
  tau = 4
)
```

Plotting

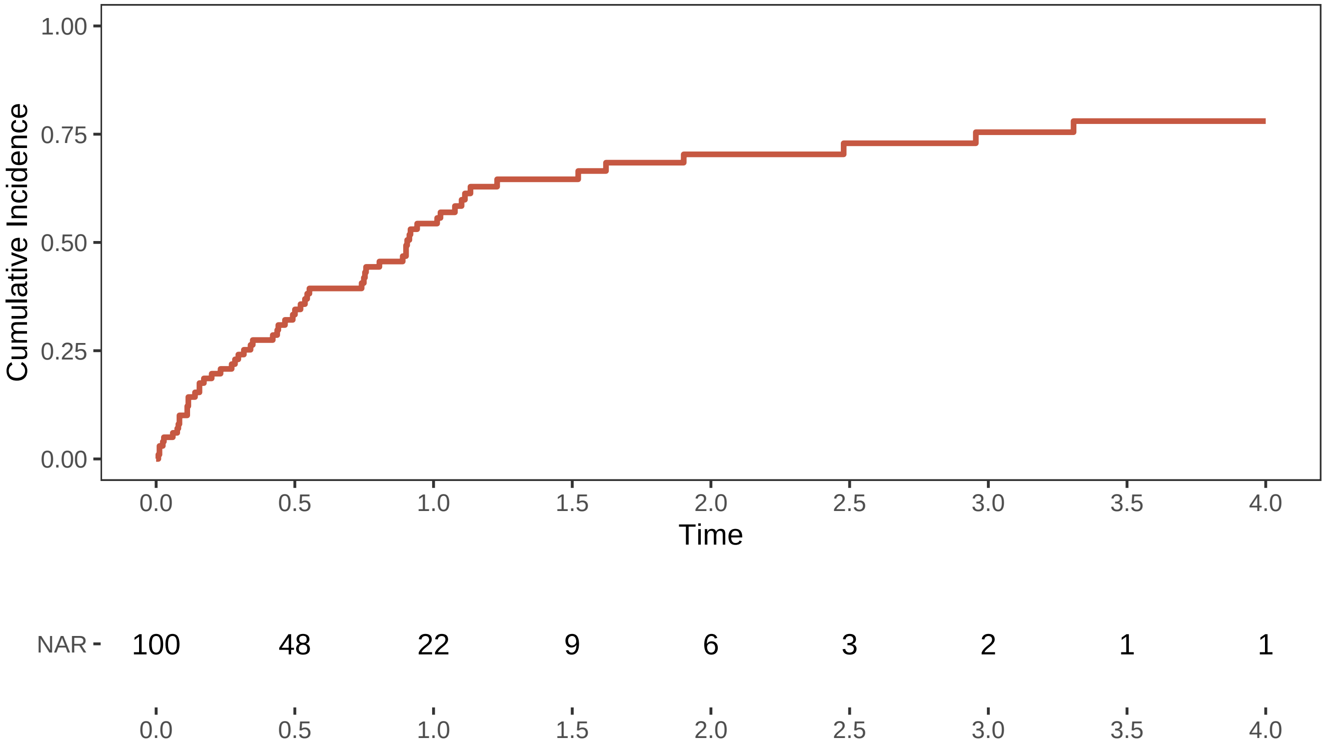
```r
# Set the time points at which to calculate NARs.
# Pass these to both PlotCICs and PlotNARs for consistent labeling.
x_breaks <- seq(from = 0, to = 4, by = 0.5)
```

```r
x_labs <- sprintf("%.1f", x_breaks)

# Plot the cumulative incidence curves.
q_cic <- CICs::PlotOneSampleCIC(
  data = data,
  x_breaks = x_breaks,
  x_labs = x_labs
)

# Plot the numbers at risk.
q_nar <- CICs::PlotOneSampleNARs(
  data = data,
  x_breaks = x_breaks,
  x_labs = x_labs
)

# Final plot.
q <- cowplot::plot_grid(
  plotlist = list(q_cic, q_nar),
  align = "v",
  axis = "l",
  ncol = 1,
  rel_heights = c(3, 1)
)
show(q)
```



Two-sample setting

Case of two treatment arms.

Simulate data Data are simulated for two treatment arms, a reference arm (`arm == 0`) and a treatment arm (`arm == 1`). Note that the plotting functions assume the arms are labeled 0 and 1.

```r
set.seed(101)
n <- 1e2

# Reference arm.
df0 <- CICs::GenData(
  n = n,
  event_rate = 1.0,
  death_rate = 0.25,
  censor_rate = 0.25,
  tau = 4
)
df0$arm <- 0

# Treatment arm.
df1 <- CICs::GenData(
  n = n,
  event_rate = 0.5,
  death_rate = 0.25,
  censor_rate = 0.25,
  tau = 4
)
df1$arm <- 1

data <- rbind(df0, df1)
```

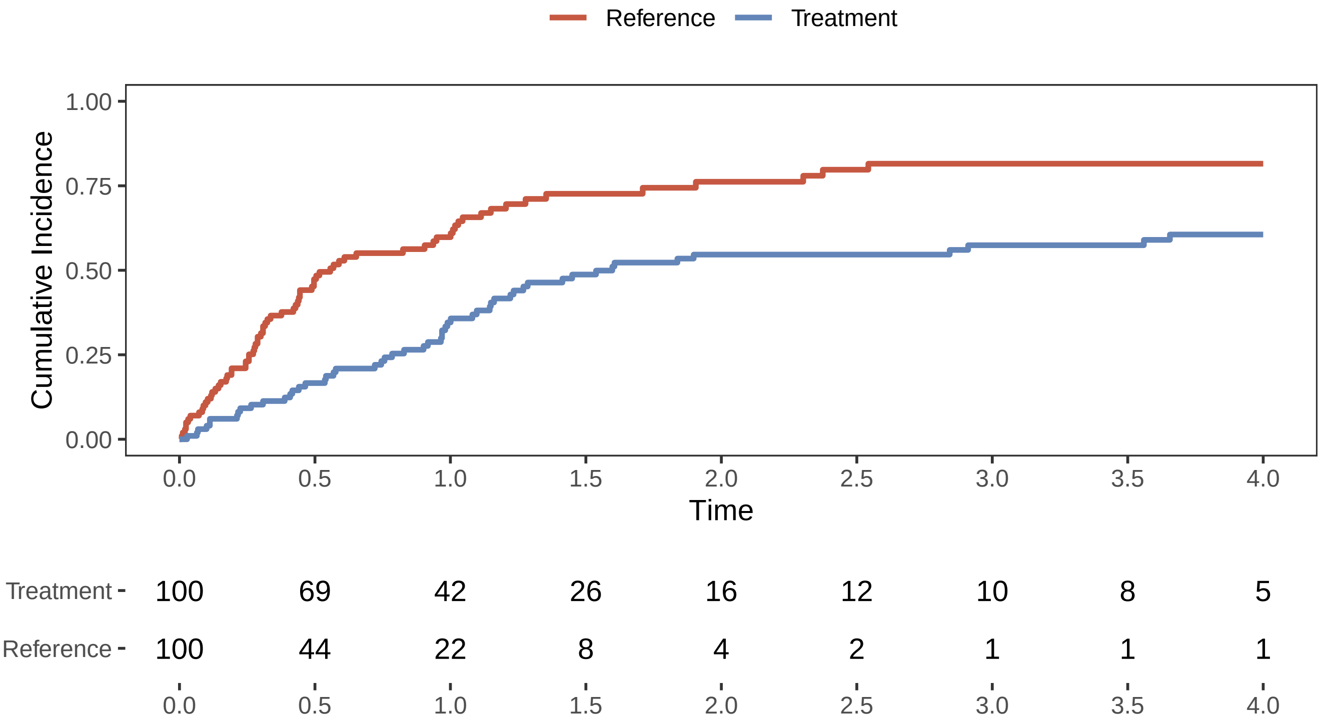
Plotting The built-in functions provide only the cumulative incidence curve (CIC) for the event of interest (i.e. the event with `status == 1`). To plot both the event of interest and the competing risk, see a subsequent example.

```r
# Set the time points at which to calculate NARs.
# Pass these to both PlotCICs and PlotNARs for consistent labeling.
x_breaks <- seq(from = 0, to = 4, by = 0.5)
x_labs <- sprintf("%.1f", x_breaks)

# Plot the cumulative incidence curves.
q_cic <- CICs::PlotCICs(
  data = data,
  color_labs = c("Reference", "Treatment"),
  x_breaks = x_breaks,
  x_labs = x_labs
)

# Plot the numbers at risk.
q_nar <- CICs::PlotNARs(
  data = data,
  x_breaks = x_breaks,
  x_labs = x_labs,
  y_labs = c("Reference", "Treatment")
)

# Final plot.
q <- cowplot::plot_grid(
  plotlist = list(q_cic, q_nar),
  align = "v",
  axis = "l",
  ncol = 1,
  rel_heights = c(3, 1)
)
show(q)
```



Stratified setting

Although built-in functions are not available for the stratified setting, the following code can readily be adapted for plotting CICs with NARs for any number of strata.

Simulate data Data are simulated for 3 strata with increasing event rates.

```
set.seed(102)
n <- 1e2
n_strata <- 3

# Simulate data for each stratum.
data <- lapply(seq_len(n_strata), function(i) {
  df <- CICs::GenData(
    n = n,
    event_rate = i / n_strata,
    death_rate = 0.25,
    censor_rate = 0.25,
    tau = 4
  )
  df$stratum <- i
  return(df)
})
data <- do.call(rbind, data)
```

Prepare plotting frames First calculate the cumulative incidence curves and numbers at risk for each stratum. Below, these are stored as a *list of functions*.

```
# Calculate CICs.
strata <- sort(unique(data$stratum))
cics <- lapply(strata, function(n){
  cic <- CICs::CICurve(data %>% dplyr::filter(stratum == n))
  return(cic)
})
names(cics) <- strata

# Calculate NARs.
nars <- lapply(strata, function(n){
  nar <- CICs::NARCurve(data %>% dplyr::filter(stratum == n))
  return(nar)
})
names(nars) <- strata
```

Prepare the plotting frame for the cumulative incidence curves.

```
# Set a dense grid of points at which to evaluate the CICs.
eval_points <- seq(from = 0, to = 4, length.out = 1000)

df_cic <- lapply(strata, function(n) {
  cic_fn <- cics[[n]]
  out <- data.frame(
    time = eval_points,
    cic = cic_fn(eval_points),
    stratum = n
  )
  return(out)
})
df_cic <- do.call(rbind, df_cic)

# Ensure stratum is a factor for plotting.
df_cic$stratum <- factor(
  x = df_cic$stratum,
  levels = sort(unique(df_cic$stratum))
)
```

Prepare the plotting frame for the numbers at risk.

```
# Choose the points at which to evaluate the NARs.
# Also set the X-axis labels.
x_breaks <- seq(from = 0, to = 4, by = 0.5)
```

```r
x_labs <- sprintf("%.1f", x_breaks)

df_nar <- lapply(strata, function(n) {
  nar_fn <- nars[[n]]
  out <- data.frame(
    time = x_breaks,
    nar = nar_fn(x_breaks),
    stratum = n
  )
  return(out)
})
df_nar <- do.call(rbind, df_nar)

# Ensure stratum is a factor for plotting.
df_nar$stratum <- factor(
  x = df_nar$stratum,
  levels = sort(unique(df_nar$stratum))
)
```

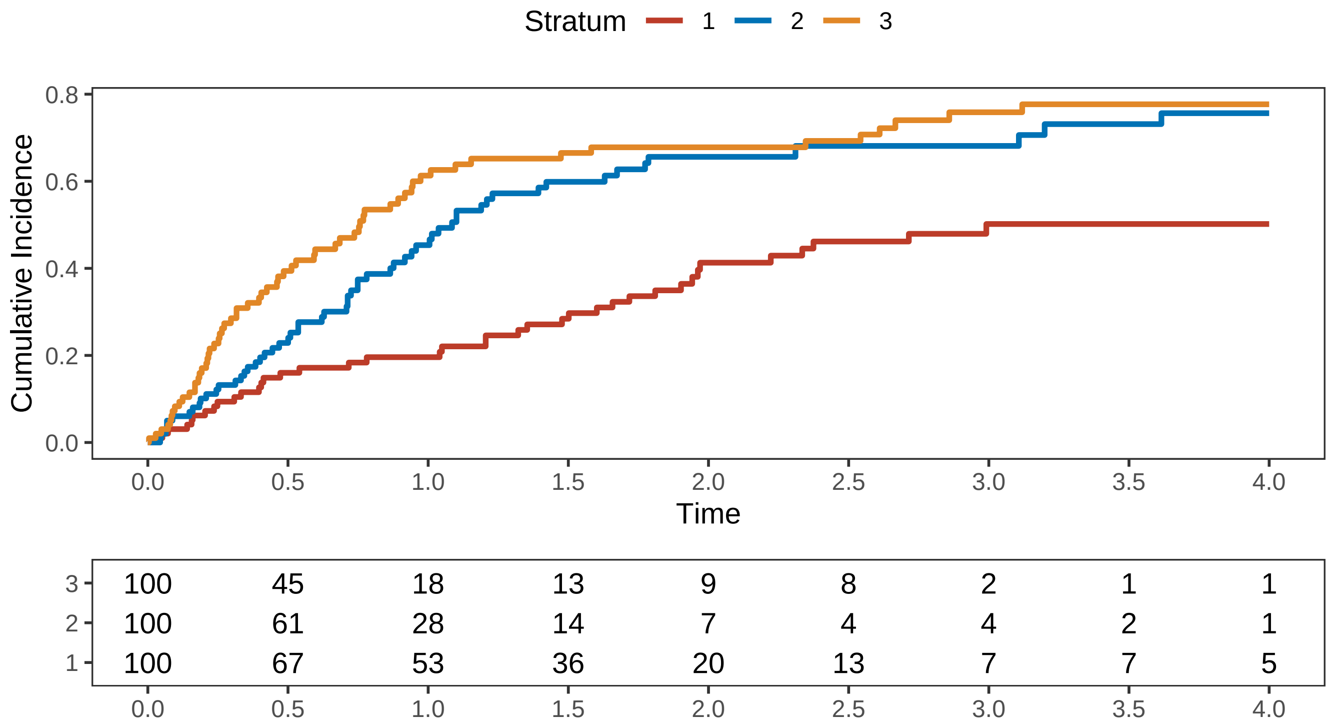
```r
# Common plotting options.
gg_opts <- theme_bw() +
  theme(
    panel.grid.major = ggplot2::element_blank(),
    panel.grid.minor = ggplot2::element_blank(),
    legend.position = "top"
  )

# Plot the cumulative incidence curve.
q_cic <- ggplot(data = df_cic) +
  gg_opts +
  geom_step(
    aes(x = time, y = cic, color = stratum),
    linewidth = 1
  ) +
  scale_x_continuous(
    name = "Time",
    breaks = x_breaks,
    labels = x_labs
  ) +
  scale_y_continuous(
    name = "Cumulative Incidence"
  ) +
  ggsci::scale_color_nejm(
    name = "Stratum"
  )

# Plot the number at risk.
q_nar <- ggplot(data = df_nar) +
  gg_opts +
  geom_text(
    aes(x = time, y = stratum, label = nar)
  ) +
  scale_x_continuous(
    name = NULL,
    breaks = x_breaks,
    labels = x_labs
  ) +
  scale_y_discrete(
    name = NULL
  )

# Final plot.
q <- cowplot::plot_grid(
  plotlist = list(q_cic, q_nar),
  align = "v",
  axis = "l",
  ncol = 1,
  rel_heights = c(3, 1)
)
show(q)
```

Plotting


Stratum
Cumulative Incidence
Time