And here's the R code that produces the chart.

```
library(tidyverse)
library(cowplot)
library(ggtext)
library(magick)

# Get data from Doutre et al.

df <- tribble(
  ~condition, ~awareness, ~incidence,
  "Congenital Cytomegalovirus (CMV)", 6.7, 6000,
  "Congenital Toxoplasmosis", 8.53, 400,
  "Congenital Rubella Syndrome", 13.27, 3,
  "Beta Strep (Group B Strep)", 16.91, 380,
  "Parvovirus B19 (Fifth Disease)", 19.63, 1045,
  "Fetal Alcohol Syndrome", 61.04, 1200,
  "Spina Bifida", 64.54, 1500,
  "Sudden Infant Death Syndrome (SIDS)", 78.7, 1500,
  "Down Syndrome", 85.44, 6000,
  "Congenital HIV/AIDS", 86.33, 30
)

# Get National CMV logo

logo <- image_read("https://github.com/seth-dobson/cmv-charts/blob/master/CMV-Full-Tagline-Logo_Transparent.png?raw=true")

# Create chart

p <- df %>%
  ggplot(aes(x = reorder(condition, desc(awareness)), y = awareness)) +
  geom_col(fill = "#28C1DB") +
  geom_point(
    aes(x = condition, y = incidence / 70),
    size = 4,
    pch = 21,
    fill = "#FB791A"
  ) +
  scale_y_continuous(
    sec.axis = sec_axis(
      ~ . * 70,
      name = "Number of Children Born with the Condition Each Year (Dots)",
      labels = scales::comma_format()
    )
  ) +
  coord_flip() +
  labs(
    x = "",
    y = "Percentage of Women Who Have Heard of the Condition (Bars)",
    title = "Awareness vs Incidence of Congenital Conditions",
    caption = "Based on US data from Doutré SM *et al.* (2016) Losing Ground:
Awareness of Congenital Cytomegalovirus
    in the United States. *Journal of Early Hearing Detection and Intervention*
1:39-48. Chart by Artful Analytics,
    LLC (@_sethdobson).
```

```
For more information, visit nationalcmv.org."
  ) +
  theme_bw() +
  theme(
    plot.title = element_text(face = "bold", hjust = .5),
    plot.caption = element_textbox_simple(size = 6, margin = margin(10, 0, 0,
0)),
    axis.text = element_text(color = "black"),
    axis.title = element_text(size = 10)
  ) +
  background_grid(major = "none") +
  annotate(
    geom = "text",
    label = "Number of children\nborn with CMV",
    x = 7.8,
    y = 75,
    color = "#FB791A",
    size = 3
  ) +
  annotate(
    geom = "curve",
    x = 8.5,
    y = 75,
    xend = 10,
    yend = 84,
    curvature = -.3,
    arrow = arrow(length = unit(2, "mm")),
    color = "#FB791A"
  ) +
  annotate(
    geom = "text",
    label = "% of women who have\nheard of CMV",
    x = 7.8,
    y = 30,
    color = "#28C1DB",
    size = 3
  ) +
  annotate(
    geom = "curve",
    x = 8.5,
    y = 30,
    xend = 10,
    yend = 7,
    curvature = .20,
    arrow = arrow(length = unit(2, "mm")),
    color = "#28C1DB"
  )

# Combine chart with logo

ggdraw() +
  draw_plot(p) +
  draw_image(
    logo,
    x = .075,
    y = .1,
    scale = .2,
```

```
    hjust = .5,
    vjust = .5
  )
```

A few things to note about the code above:

- The secondary x-axis is actually coded as a secondary y-axis since you have to use `coord_flip()` to get the categorical variable on the y-axis when using `geom_col()`.
- The `sec_axis()` function is used in conjuction with the `sec.axis` option within `scale_y_continuous()`. In order to align the two y-axes, I multiplied the secondary axis by 70 within `sec_axis()` and divided incidence by 70 within the aesthetics of `geom_point()`. I arrived at the number 70 by trial and error. Not sure why this works, but it does.