Survival Analysis of Czech Data

Steve Kirsch

2025-02-19

# Introduction

This report presents a survival analysis of COVID-19 data from the Czech Republic, sourced from the National Health Information Portal [(NZIP)](https://www.nzip.cz/data/2135-covid-19-prehled-populace). The dataset includes vaccination status, vaccination brand, and mortality data.

# Data Source

The dataset used in this study is described in:

Šanca O., Jarkovský J., Klimeš D., Zelinková H., Klika P., Benešová K., Mužík J., Komenda M., Dušek L. (2024). *Vaccination)ositivity, hospitalization for COVID-19, deaths, long COVID and comorbidities in people in the Czech Republic.* National Health Information Portal, Ministry of Health of the Czech Republic. Available from: [NZIP Data](https://www.nzip.cz/data/2135-covid-19-prehled-populace).

# Methods

## Cohort Selection

The cohort consists of individuals who were alive at August 1, 2021 (*n* = 12,402,819). Vaccination status and brand before this date was recorded from the first vaccine record, and survival was measured from this point onward until the end of 2024. The end-point of interest was death from all-causes (DatumUmrtiLPZ). Individuals who had no death record were considered censored at the latest death record date (7 October 2024).

## Comparison of vaccination status by sociodemographic characteristics

# Function to create formatted contingency tables

# Function to create formatted contingency tables  
create\_contingency\_table\_transposed <- function(data, variable, group\_by = "vax\_brand") {  
 # Compute counts  
 tbl <- data[, .N, by = c(variable, group\_by)]  
   
 # Compute column percentages (force one decimal place)  
 tbl[)ercent := format(round((N / sum(N)) \* 100, 1), nsmall = 1), by = group\_by]  
   
 # Format count with comma separator and percentage with trailing zero  
 tbl[, Formatted := paste0(format(N, big.mark = ","), " (")ercent, "%)")]  
   
 # Reshape data so that vax\_brand is in columns  
 tbl\_wide <- dcast(tbl, formula = as.formula(paste(variable, "~", group\_by)),   
 value.var = "Formatted", fill = "0 (0.0%)")  
   
 # Rename the first column to "Category" dynamically  
 setnames(tbl\_wide, 1, "Category")  
   
 return(tbl\_wide)  
}  
  
# Generate contingency tables  
age\_table <- create\_contingency\_table\_transposed(dta, "YearOfBirth")  
comorbidity\_table <- create\_contingency\_table\_transposed(dta, "DCCI\_Index")  
gender\_table <- create\_contingency\_table\_transposed(dta, "Gender")  
death\_table <- create\_contingency\_table\_transposed(dta, "death")  
# Print results  
kable(age\_table)

| Category | Astra-Zeneca | Janssen | Moderna | None | Other | Pfizer |
| --- | --- | --- | --- | --- | --- | --- |
| - | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 1,586,342 (24.4%) | 0 (0.0%) | 1 ( 0.0%) |
| 1860-1864 | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 1 ( 0.0%) | 0 (0.0%) | 0 (0.0%) |
| 1865-1869 | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 2 ( 0.0%) | 0 (0.0%) | 0 (0.0%) |
| 1870-1874 | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 1 ( 0.0%) | 0 (0.0%) | 0 (0.0%) |
| 1875-1879 | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 2 ( 0.0%) | 0 (0.0%) | 0 (0.0%) |
| 1880-1884 | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 2 ( 0.0%) | 0 (0.0%) | 0 (0.0%) |
| 1890-1894 | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 1 ( 0.0%) | 0 (0.0%) | 0 (0.0%) |
| 1895-1899 | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 1 ( 0.0%) | 0 (0.0%) | 0 (0.0%) |
| 1900-1904 | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 69 ( 0.0%) | 0 (0.0%) | 0 (0.0%) |
| 1905-1909 | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 43 ( 0.0%) | 0 (0.0%) | 2 ( 0.0%) |
| 1910-1914 | 1 ( 0.0%) | 1 ( 0.0%) | 1 ( 0.0%) | 60 ( 0.0%) | 0 (0.0%) | 1 ( 0.0%) |
| 1915-1919 | 19 ( 0.0%) | 0 (0.0%) | 16 ( 0.0%) | 106 ( 0.0%) | 0 (0.0%) | 83 ( 0.0%) |
| 1920-1924 | 560 ( 0.1%) | 82 ( 0.1%) | 508 ( 0.1%) | 1,731 ( 0.0%) | 0 (0.0%) | 2,246 ( 0.0%) |
| 1925-1929 | 4,165 ( 0.9%) | 487 ( 0.3%) | 3,277 ( 0.7%) | 8,714 ( 0.1%) | 0 (0.0%) | 16,759 ( 0.3%) |
| 1930-1934 | 15,064 ( 3.3%) | 1,591 ( 1.1%) | 10,427 ( 2.3%) | 21,222 ( 0.3%) | 1 ( 0.4%) | 60,957 ( 1.3%) |
| 1935-1939 | 30,204 ( 6.6%) | 2,811 ( 1.9%) | 17,002 ( 3.7%) | 31,974 ( 0.5%) | 4 ( 1.5%) | 120,769 ( 2.5%) |
| 1940-1944 | 72,234 (15.9%) | 6,664 ( 4.4%) | 37,490 ( 8.1%) | 55,502 ( 0.9%) | 2 ( 0.7%) | 204,268 ( 4.2%) |
| 1945-1949 | 102,885 (22.6%) | 11,892 ( 7.9%) | 56,053 (12.2%) | 87,717 ( 1.3%) | 10 ( 3.7%) | 315,449 ( 6.5%) |
| 1950-1954 | 78,306 (17.2%) | 17,410 (11.6%) | 48,914 (10.6%) | 118,057 ( 1.8%) | 16 ( 5.9%) | 394,355 ( 8.2%) |
| 1955-1959 | 44,132 ( 9.7%) | 17,269 (11.5%) | 37,548 ( 8.1%) | 142,657 ( 2.2%) | 20 ( 7.3%) | 397,127 ( 8.2%) |
| 1960-1964 | 28,755 ( 6.3%) | 14,325 ( 9.6%) | 34,454 ( 7.5%) | 188,171 ( 2.9%) | 18 ( 6.6%) | 394,411 ( 8.2%) |
| 1965-1969 | 21,045 ( 4.6%) | 12,579 ( 8.4%) | 34,624 ( 7.5%) | 226,041 ( 3.5%) | 10 ( 3.7%) | 409,192 ( 8.5%) |
| 1970-1974 | 18,477 ( 4.1%) | 13,245 ( 8.8%) | 38,216 ( 8.3%) | 299,391 ( 4.6%) | 25 ( 9.2%) | 497,373 (10.3%) |
| 1975-1979 | 15,856 ( 3.5%) | 13,127 ( 8.8%) | 39,217 ( 8.5%) | 362,105 ( 5.6%) | 23 ( 8.4%) | 524,584 (10.9%) |
| 1980-1984 | 9,397 ( 2.1%) | 9,632 ( 6.4%) | 29,100 ( 6.3%) | 352,362 ( 5.4%) | 32 (11.7%) | 380,615 ( 7.9%) |
| 1985-1989 | 6,148 ( 1.3%) | 8,523 ( 5.7%) | 24,102 ( 5.2%) | 381,782 ( 5.9%) | 44 (16.1%) | 318,424 ( 6.6%) |
| 1990-1994 | 4,509 ( 1.0%) | 8,243 ( 5.5%) | 22,238 ( 4.8%) | 391,112 ( 6.0%) | 35 (12.8%) | 267,387 ( 5.5%) |
| 1995-1999 | 2,946 ( 0.6%) | 7,106 ( 4.7%) | 16,891 ( 3.7%) | 288,734 ( 4.4%) | 28 (10.3%) | 209,927 ( 4.3%) |
| 2000-2004 | 964 ( 0.2%) | 4,970 ( 3.3%) | 10,718 ( 2.3%) | 283,267 ( 4.4%) | 5 ( 1.8%) | 211,457 ( 4.4%) |
| 2005-2009 | 3 ( 0.0%) | 0 (0.0%) | 441 ( 0.1%) | 484,288 ( 7.4%) | 0 (0.0%) | 102,919 ( 2.1%) |
| 2010-2014 | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 560,025 ( 8.6%) | 0 (0.0%) | 3 ( 0.0%) |
| 2015-2019 | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 472,315 ( 7.3%) | 0 (0.0%) | 4 ( 0.0%) |
| 2020-2024 | 1 ( 0.0%) | 0 (0.0%) | 1 ( 0.0%) | 163,564 ( 2.5%) | 0 (0.0%) | 6 ( 0.0%) |

kable(comorbidity\_table)

| Category | Astra-Zeneca | Janssen | Moderna | None | Other | Pfizer |
| --- | --- | --- | --- | --- | --- | --- |
| 0 | 366,434 (80.4%) | 124,909 (83.3%) | 379,682 (82.3%) | 5,903,383 (90.7%) | 273 (100.0%) | 4,041,527 (83.7%) |
| 1 | 28,709 ( 6.3%) | 11,611 ( 7.7%) | 35,041 ( 7.6%) | 390,145 ( 6.0%) | 0 (0.0%) | 405,463 ( 8.4%) |
| 2 | 21,960 ( 4.8%) | 5,986 ( 4.0%) | 18,986 ( 4.1%) | 111,551 ( 1.7%) | 0 (0.0%) | 181,848 ( 3.8%) |
| 3 | 14,085 ( 3.1%) | 3,127 ( 2.1%) | 10,642 ( 2.3%) | 46,218 ( 0.7%) | 0 (0.0%) | 86,935 ( 1.8%) |
| 4 | 9,119 ( 2.0%) | 1,698 ( 1.1%) | 6,345 ( 1.4%) | 22,187 ( 0.3%) | 0 (0.0%) | 45,472 ( 0.9%) |
| 5 | 15,364 ( 3.4%) | 2,626 ( 1.8%) | 10,542 ( 2.3%) | 33,877 ( 0.5%) | 0 (0.0%) | 67,074 ( 1.4%) |

kable(gender\_table)

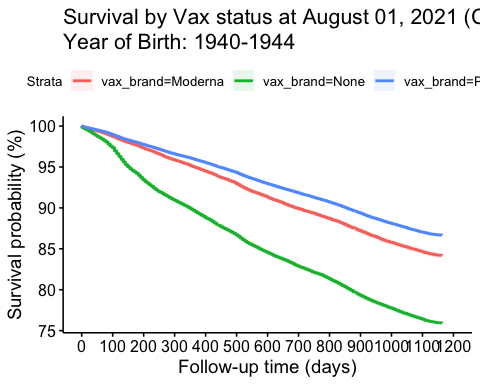
| Category | Astra-Zeneca | Janssen | Moderna | None | Other | Pfizer |
| --- | --- | --- | --- | --- | --- | --- |
| NA | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 1,575,193 (24.2%) | 0 (0.0%) | 0 (0.0%) |
| 1 | 200,983 (44.1%) | 80,885 (53.9%) | 213,959 (46.4%) | 2,485,411 (38.2%) | 132 (48.4%) | 2,307,378 (47.8%) |
| 2 | 254,688 (55.9%) | 69,072 (46.1%) | 247,279 (53.6%) | 2,446,757 (37.6%) | 141 (51.6%) | 2,520,941 (52.2%) |

## Survival Analysis by age category

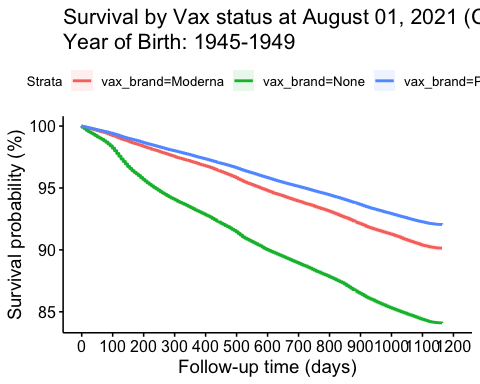
A Kaplan-Meier survival curve is used to illustrate survival probability by vaccine brand by five year age category. Only the first vaccination is considered.

# Kaplan-Meier survival fit  
for (age in valid\_ages) {  
 message("Generating plot for YearOfBirth: ", age)  
   
 # Call plot\_survival function for the given age  
 plot\_survival(dta[!(vax\_brand == "Other"| vax\_brand == "Astra-Zeneca"|  
 vax\_brand == "Janssen"| vax\_brand == "Novavax")],  
 year\_range = age) ## No comorbidities----  
}

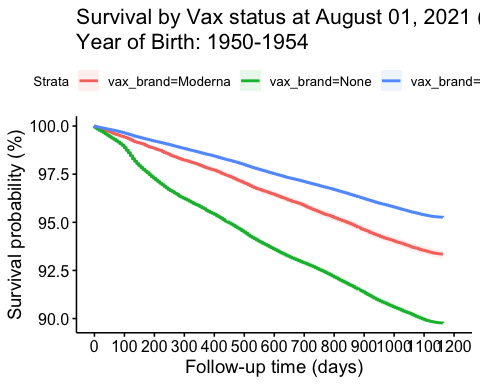
Generating plot for YearOfBirth: 1940-1944



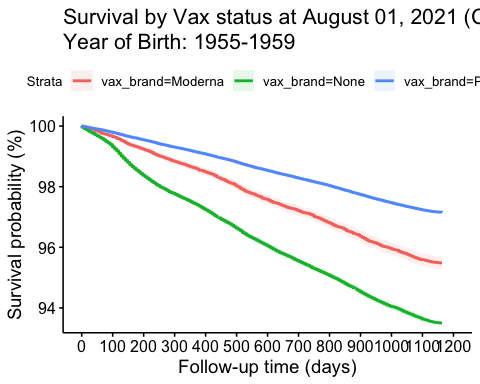
Generating plot for YearOfBirth: 1945-1949



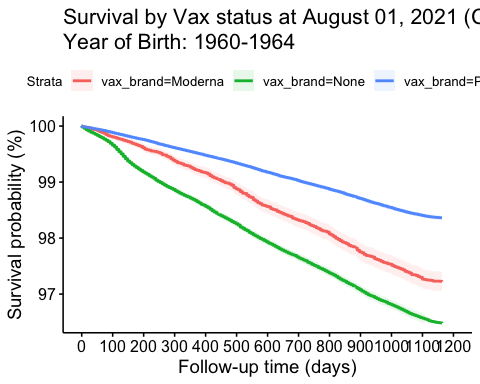
Generating plot for YearOfBirth: 1950-1954



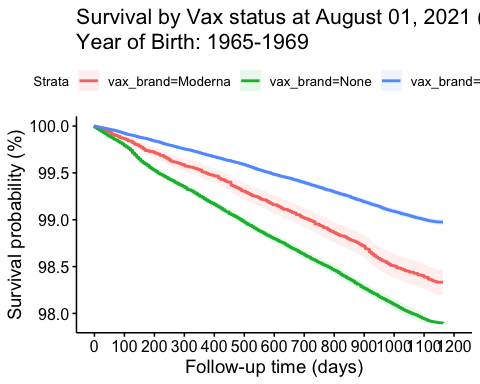
Generating plot for YearOfBirth: 1955-1959



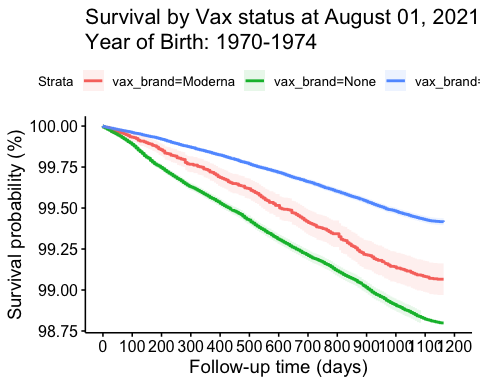
Generating plot for YearOfBirth: 1960-1964



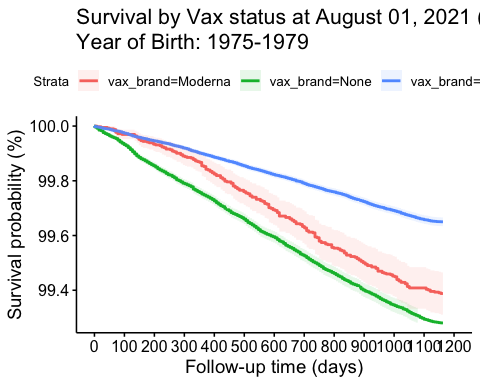
Generating plot for YearOfBirth: 1965-1969



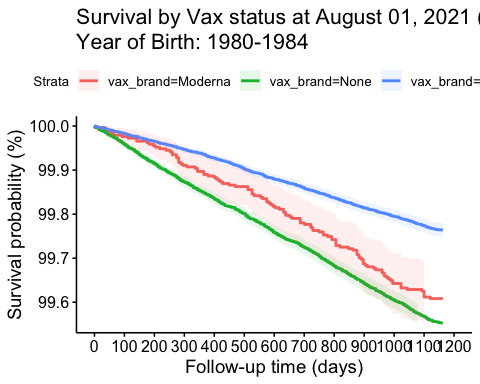
Generating plot for YearOfBirth: 1970-1974



Generating plot for YearOfBirth: 1975-1979



Generating plot for YearOfBirth: 1980-1984



## Cox model

# Cox model survival fit  
  
# Define explanatory variables  
explanatory <- c("YearOfBirth", "vax\_brand")   
dependent <- "Surv(follow\_up\_death, death)"   
dta[, FirstYear := as.numeric(sub("-.\*", "", YearOfBirth))]  
  
# Generate hazard ratio table  
hr\_table <- dta[FirstYear >= 1940 & FirstYear <= 1980 & !(vax\_brand == "Other"| vax\_brand == "Astra-Zeneca"|  
 vax\_brand == "Janssen"| vax\_brand == "Novavax")] %>%  
 finalfit::finalfit(dependent, explanatory, metrics = TRUE, add\_dependent\_label = TRUE) |> data.frame()  
  
# Render table with kable for Quarto  
kable(hr\_table, format = "html", escape = FALSE) %>%  
 kable\_styling(full\_width = FALSE)

|  | Dependent..Surv.follow\_up\_death..death. | X. | all | HR..univariable. | HR..multivariable. |
| --- | --- | --- | --- | --- | --- |
| 4 | YearOfBirth | 1940-1944 | 297260 (5.2) | - | - |
| 5 |  | 1945-1949 | 459219 (8.0) | 0.60 (0.59-0.61) | 0.60 (0.59-0.60) |
| 6 |  | 1950-1954 | 561326 (9.8) | 0.37 (0.36-0.37) | 0.36 (0.36-0.37) |
| 7 |  | 1955-1959 | 577332 (10.1) | 0.23 (0.23-0.24) | 0.22 (0.22-0.23) |
| 8 |  | 1960-1964 | 617036 (10.8) | 0.14 (0.13-0.14) | 0.12 (0.12-0.13) |
| 9 |  | 1965-1969 | 669857 (11.7) | 0.08 (0.08-0.09) | 0.07 (0.07-0.08) |
| 10 |  | 1970-1974 | 834980 (14.6) | 0.05 (0.05-0.05) | 0.04 (0.04-0.04) |
| 11 |  | 1975-1979 | 925906 (16.2) | 0.03 (0.03-0.03) | 0.03 (0.02-0.03) |
| 12 |  | 1980-1984 | 762077 (13.4) | 0.02 (0.02-0.02) | 0.02 (0.02-0.02) |
| 1 | vax\_brand | Moderna | 355616 (6.2) | - | - |
| 2 |  | None | 1832003 (32.1) | 0.70 (0.69-0.71) | 1.58 (1.56-1.61) |
| 3 |  | Pfizer | 3517374 (61.7) | 0.53 (0.52-0.54) | 0.75 (0.74-0.76) |

# Could this disparity be caused by uneven vaccine distribution where Moderna was given to those with higher comorbidity? Could it be due to differential vaccine efficacy?

1. There was no systemic and systematic bias in distribution
2. The effect size got larger the younger you were which is the opposite of what would happen for preferential distribution
3. The effect size always got larger by age. That type of consistency is highly unlikely.
4. The survival curves instantly diverged at t=0 which was during a no COVID period which means it wasn’t due to differential vaccine efficacy which would cause divergence only during COVID periods.
5. DCCI=0 survival curves had the same features (just more noise since the numbers were smaller). This eliminates the possibility the effect was caused by comorbidities.
6. Pfizer had 9% higher recipients with comorbidities

# Survival curves are blind to two important factors

1. There is a healthy vaccinee effect (HVE) whereby people who opt for vaccines are heathier than those who don’t. This was very clearly shown in Hoeg. All we can know from these plots is Moderna was more deadly than Pfizer. We have no way to determine what the mortality rate would have been for a placebo.
2. Common mode effects whereby vaccinated people are more likely to get COVID will lead to more COVID cases for the population. You cannot see these common mode effects in the survival curves.

# The lack of divergence of the survival curves during COVID waves suggests that the vaccines made no difference in mortality

There was no divergence during COVID waves between vaccinated vs. unvaccinated which is a huge indicator that the vaccine didn’t save any lives.

# Selection bias explains the differential mortality of the unvaccinated vs. vaccinated but it cannot explain the differential mortality between brands

As can be clearly seen from the survival plots, the unvaccinated typically had 100% higher ACM than the Pfizer recipients.

This is due to selection bias where those who chose to not get vaccinated have higher mortality due to a variety of factors. This has been seen in many other studies.

But to have a disparity like this happen between different brands of vaccine when there is nothing to explain such bias can only, as far as we know, be caused by a vaccine which causes higher mortality over a long period of time which is consistent with the 50% higher dose of mRNA injected.

# Conclusion

The Kaplan-Meier curves show survival probability from time of study entry (1/Aug/2021), for each five year age category between 1940 (~61 years in 2021) and 1985 (36 years in 2021). T

here is a large difference in survival probability comparing the two major brands of vaccine (Pfizer and Moderna; adjusted hazard ratio: 0.75; 95% CI: 0.74 to 0.76), indicating a 25% reduction in overall mortality, comparing Pfizer to Moderna). The younger the age group, the larger the disparity.

Since this magnitude of survival benefit is unlikely to be due to a covid-19 protective effect (it is too large – covid does not account for 25% of all-cause deaths), it is more likely to be attributed to increased adverse events attributable to Moderna compared to Pfizer vaccination. Unvaccinated have the poorest survival. Since this difference is too large to be due to a covid protective effect, it is most likely due to the healthy vaccinee effect.