Kaplan-Meier analysis of Czech vaccine and all-cause data with extrapolated survival using linear tangents.

By Epidemiologist who would be fired if name was disclosed

2025-04-09

# 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 covid-19 vaccination status, date of vaccination, vaccination brand, and mortality data.

This analysis uses a ‘predicted’ survival using a straight line derived from the tangent to the survival curve for each brand including the unvaccinated. The tangent is taken from early in the survival (April 12, 2021 to May 31, 2021) trajectory to consider evidence that late patterns of survival in each group differ from early. It is expected that mortality benefits or harms from the covid-19 vaccinations are difficult to compare within groups due to unmeasured confounding. This method assumes that the early survival trajectory represents the ‘true’ survival trajectory of the group and that latter deviation from the tangent reflects the effect of the treatment.

The average deviation between observed survival in the Kaplan-Meier plot compared to the groups corresponding tangent, summarises the effect of the treatment group (unvaccinated [“None”], “Pfizer”, “Moderna” or “Astra-Zeneca”) up until the end of 2022. Values greater than zero indicate poorer than expected survival, whereas negative values indicated better than expected survival. The straight line is considered a rough approximation to expected survival.

Stratified Kaplan-Meier survival plots are presented by five-year age category to control, as best as possible, for age confounding in the cohort.

# 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, positivity, 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 April 05, 2021 (*n* = 12,436,895).

## Outcome

The end-point of interest was survival with the end-point being death from all-causes (DatumUmrtiLPZ).

## Censoring

Individuals who had no death record were considered censored at the latest death record date October 07, 2024.

## Exposure

Vaccination status and brand before April 05, 2021 was recorded from the first vaccine record (OckovaciLatkaKod\_Prvni\_davka and Datum\_Prvni\_davka), and survival was measured from this point onward until October 07, 2024. If individuals had no record of first vaccination before April 05, 2021, they were considered “unvaccinated”.

The following codes were used to assign vaccination brand from the OckovaciLatkaKod\_Prvni\_davka (VaccineCode\_First\_Dose) variable:

* Pfizer: “CO01”, “CO09”, “CO16”, “CO21”, “CO23”
* Moderna: “CO02”, “CO15”, “CO19”
* Astra-Zeneca: “CO03”

Other brands were excluded due to low uptake and low statistical power.

## Survival Analysis by age category

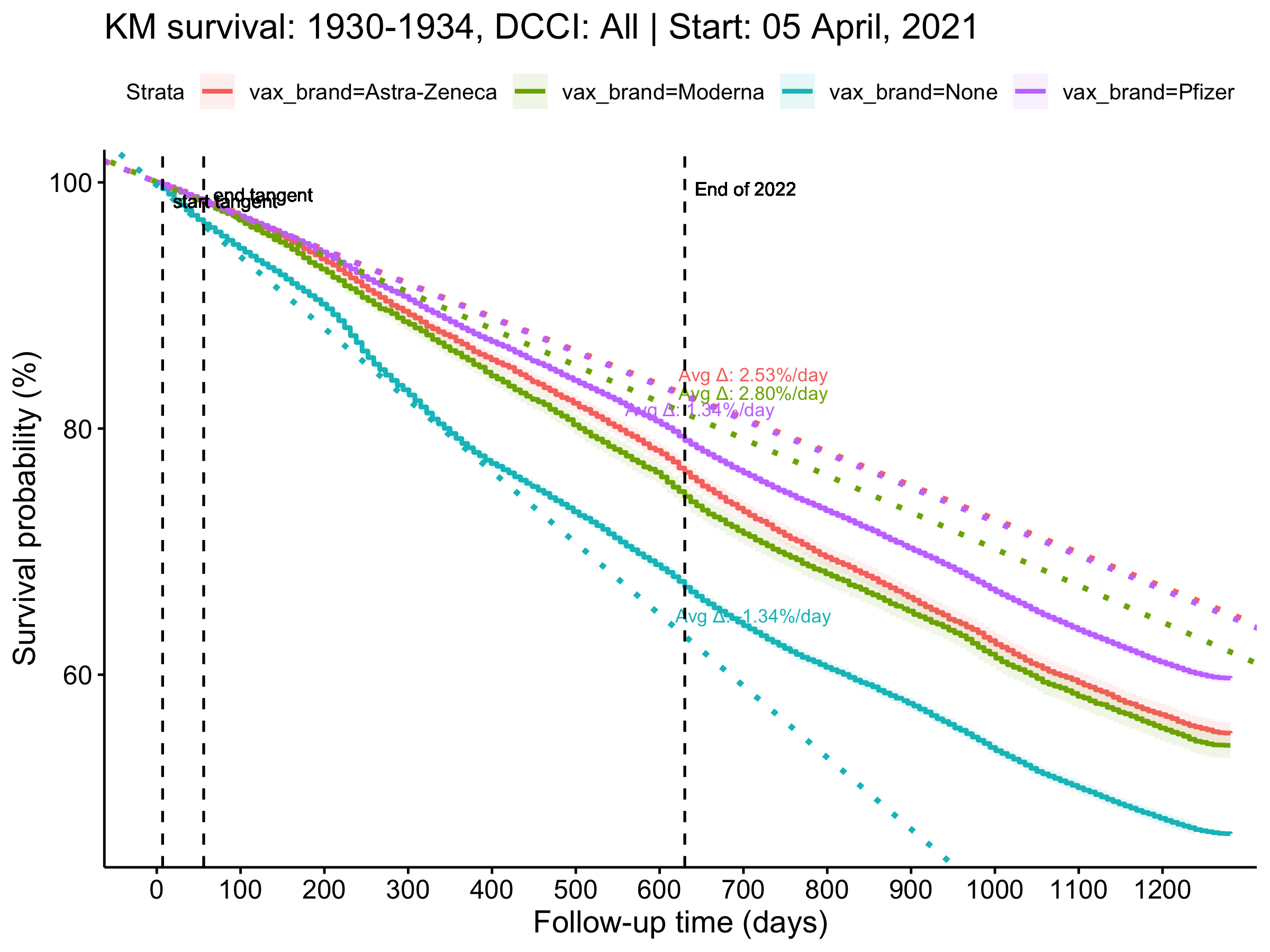
A Kaplan-Meier survival curve is used to illustrate survival probability by vaccine brand by five year age category.

# Kaplan-Meier survival fit  
for (age in valid\_ages) {  
 message("Generating plot for YearOfBirth: ", age)  
   
 # Call plot\_survival function for the given age  
 plot\_km\_with\_tangents(dta,  
 age\_group = age) |> print()## No comorbidities----  
}

Generating plot for YearOfBirth: 1930-1934

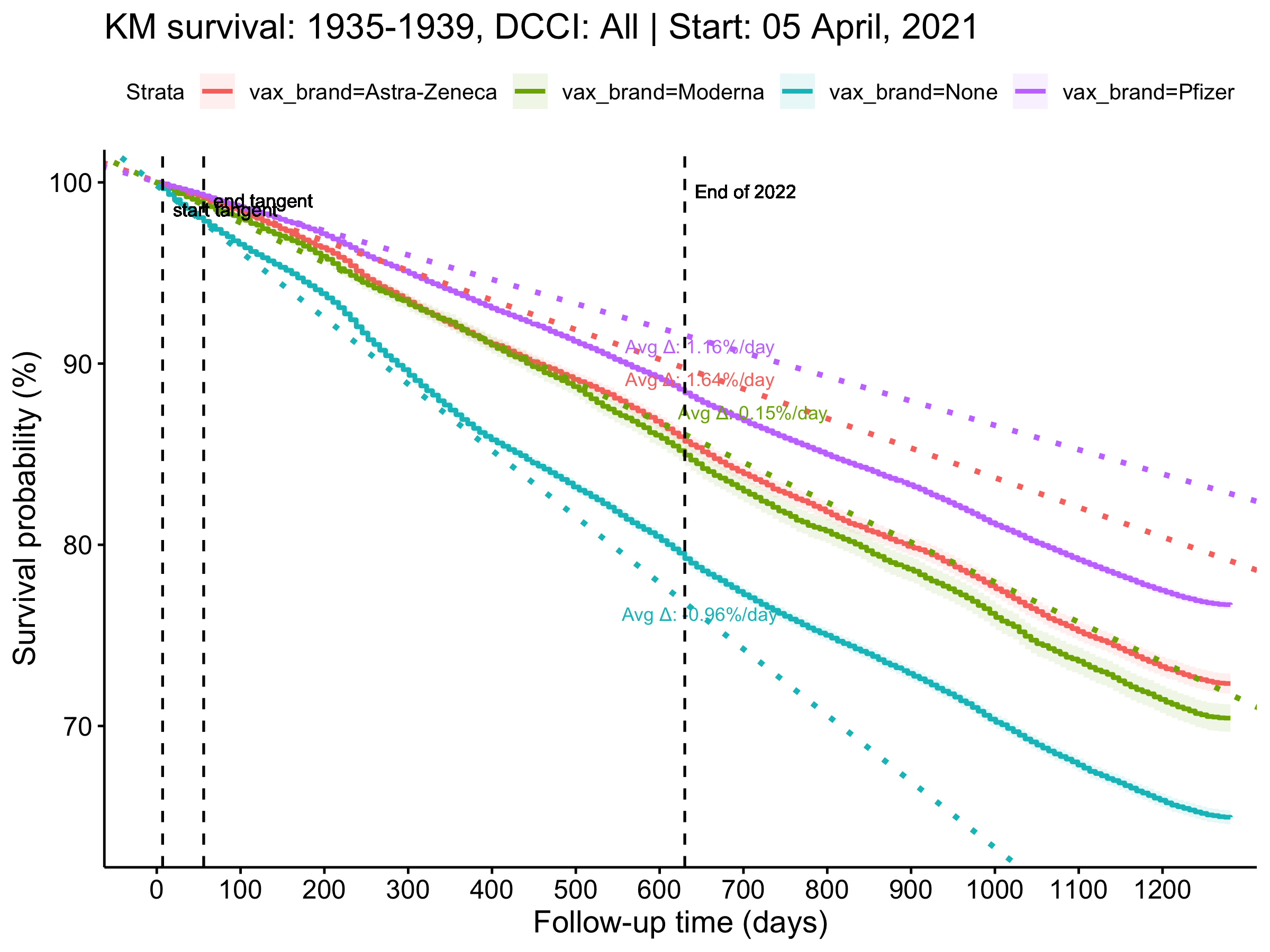
Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.  
ℹ Please use `linewidth` instead.

[1] "vax\_brand=Astra-Zeneca Avg Δ: 2.53%/day"  
[1] "vax\_brand=Moderna Avg Δ: 2.80%/day"  
[1] "vax\_brand=None Avg Δ: -1.34%/day"  
[1] "vax\_brand=Pfizer Avg Δ: 1.34%/day"



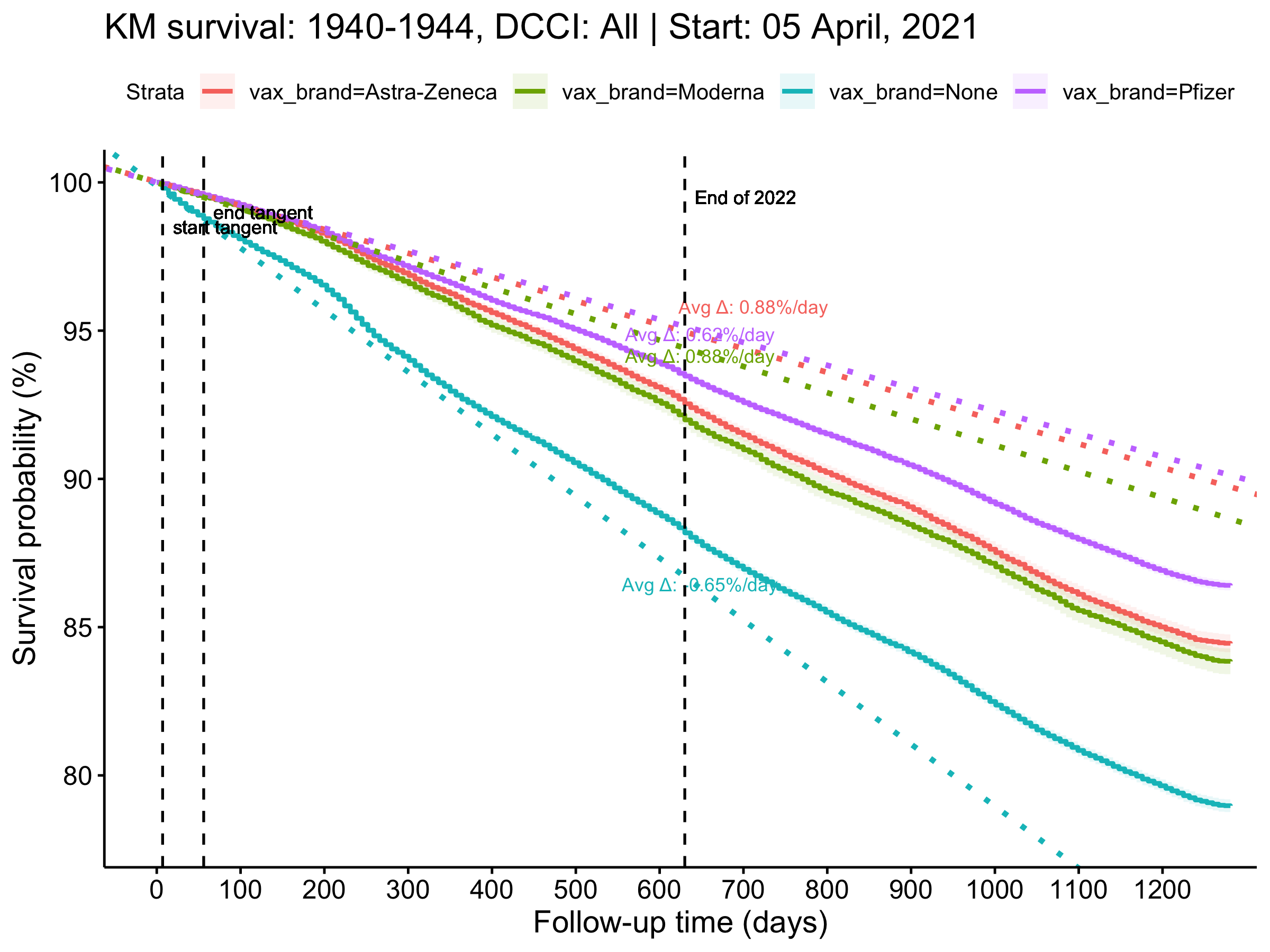
Generating plot for YearOfBirth: 1935-1939

[1] "vax\_brand=Astra-Zeneca Avg Δ: 1.64%/day"  
[1] "vax\_brand=Moderna Avg Δ: 0.15%/day"  
[1] "vax\_brand=None Avg Δ: -0.96%/day"  
[1] "vax\_brand=Pfizer Avg Δ: 1.16%/day"



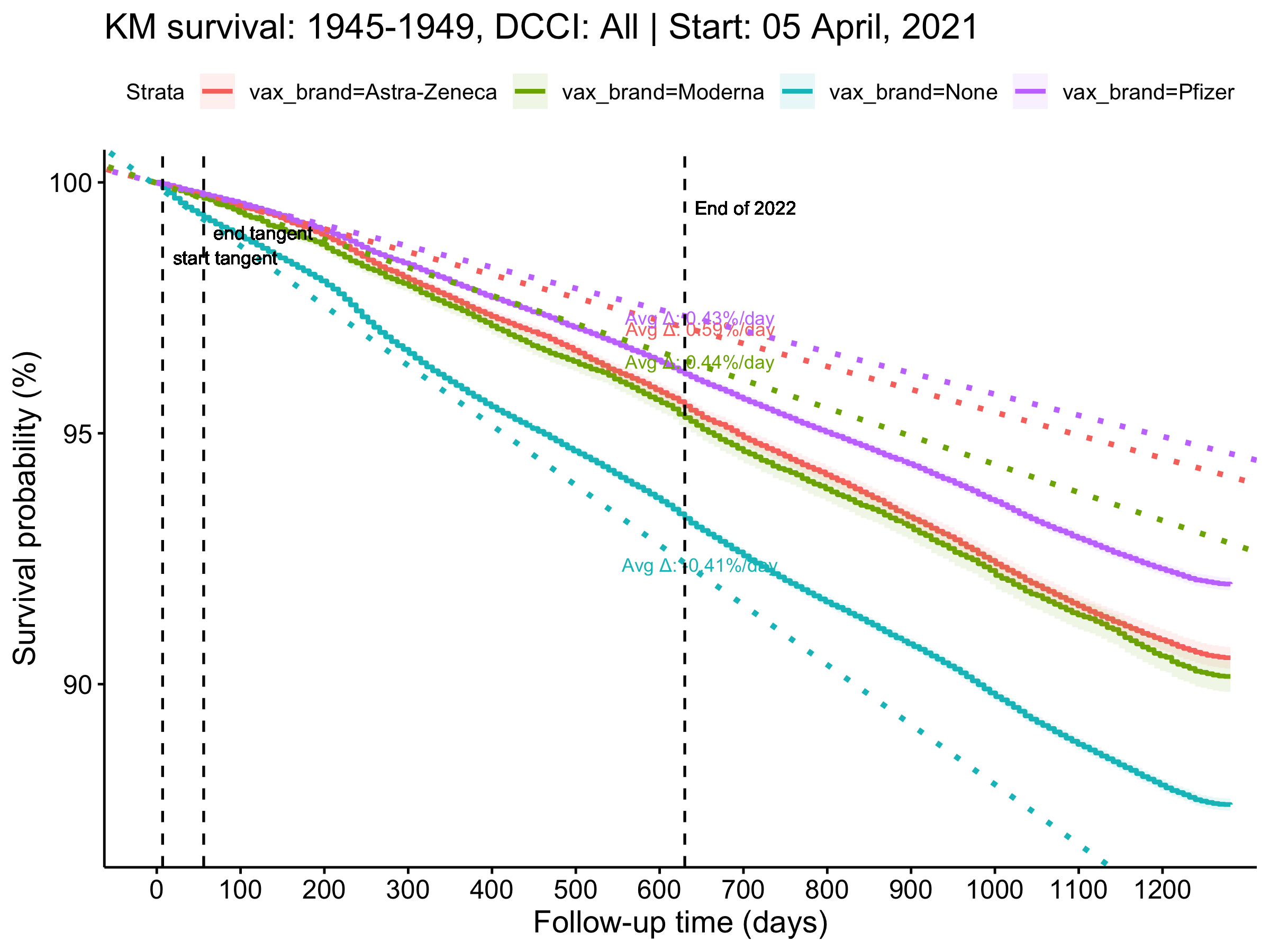
Generating plot for YearOfBirth: 1940-1944

[1] "vax\_brand=Astra-Zeneca Avg Δ: 0.88%/day"  
[1] "vax\_brand=Moderna Avg Δ: 0.88%/day"  
[1] "vax\_brand=None Avg Δ: -0.65%/day"  
[1] "vax\_brand=Pfizer Avg Δ: 0.62%/day"



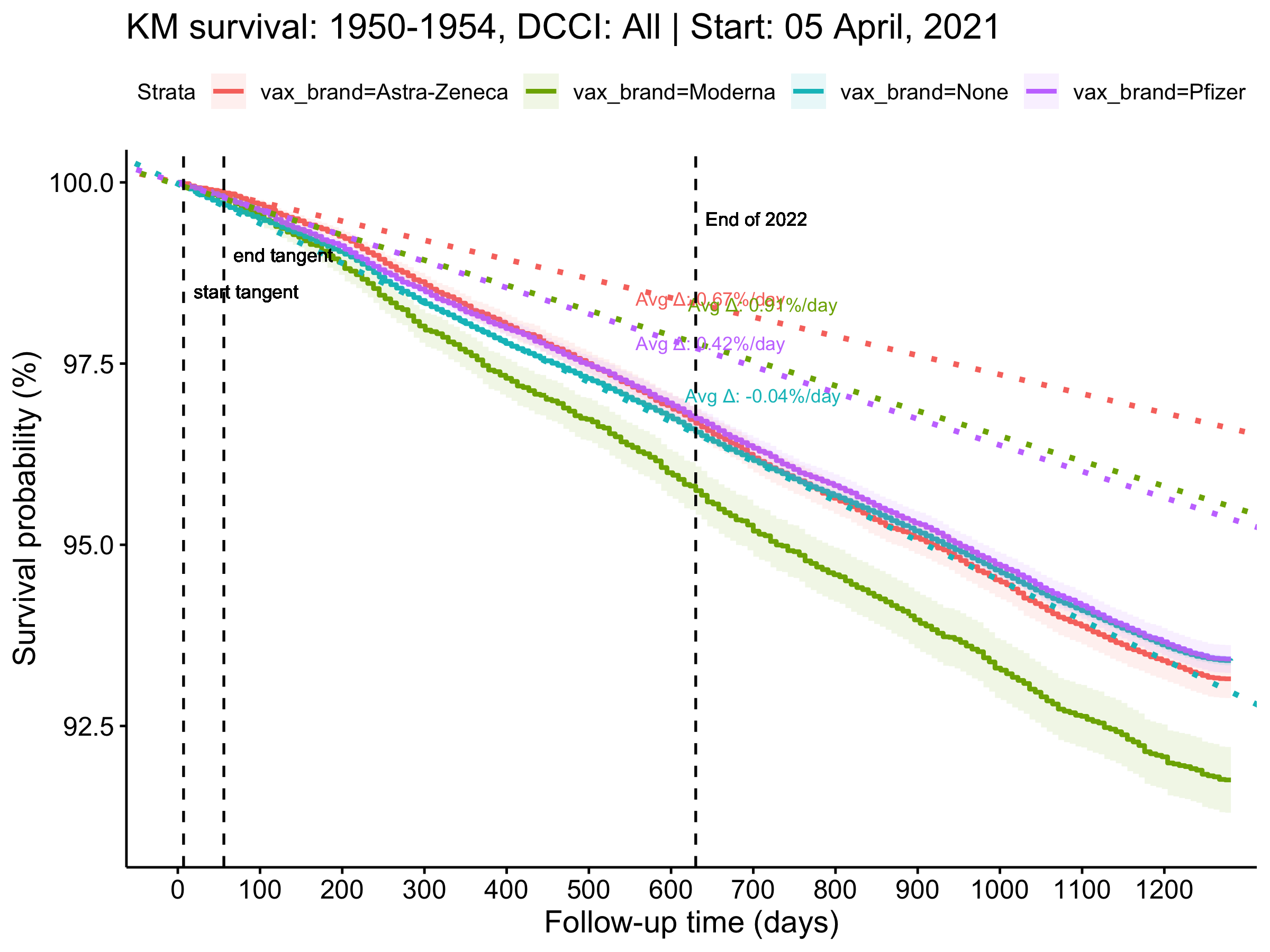
Generating plot for YearOfBirth: 1945-1949

[1] "vax\_brand=Astra-Zeneca Avg Δ: 0.59%/day"  
[1] "vax\_brand=Moderna Avg Δ: 0.44%/day"  
[1] "vax\_brand=None Avg Δ: -0.41%/day"  
[1] "vax\_brand=Pfizer Avg Δ: 0.43%/day"



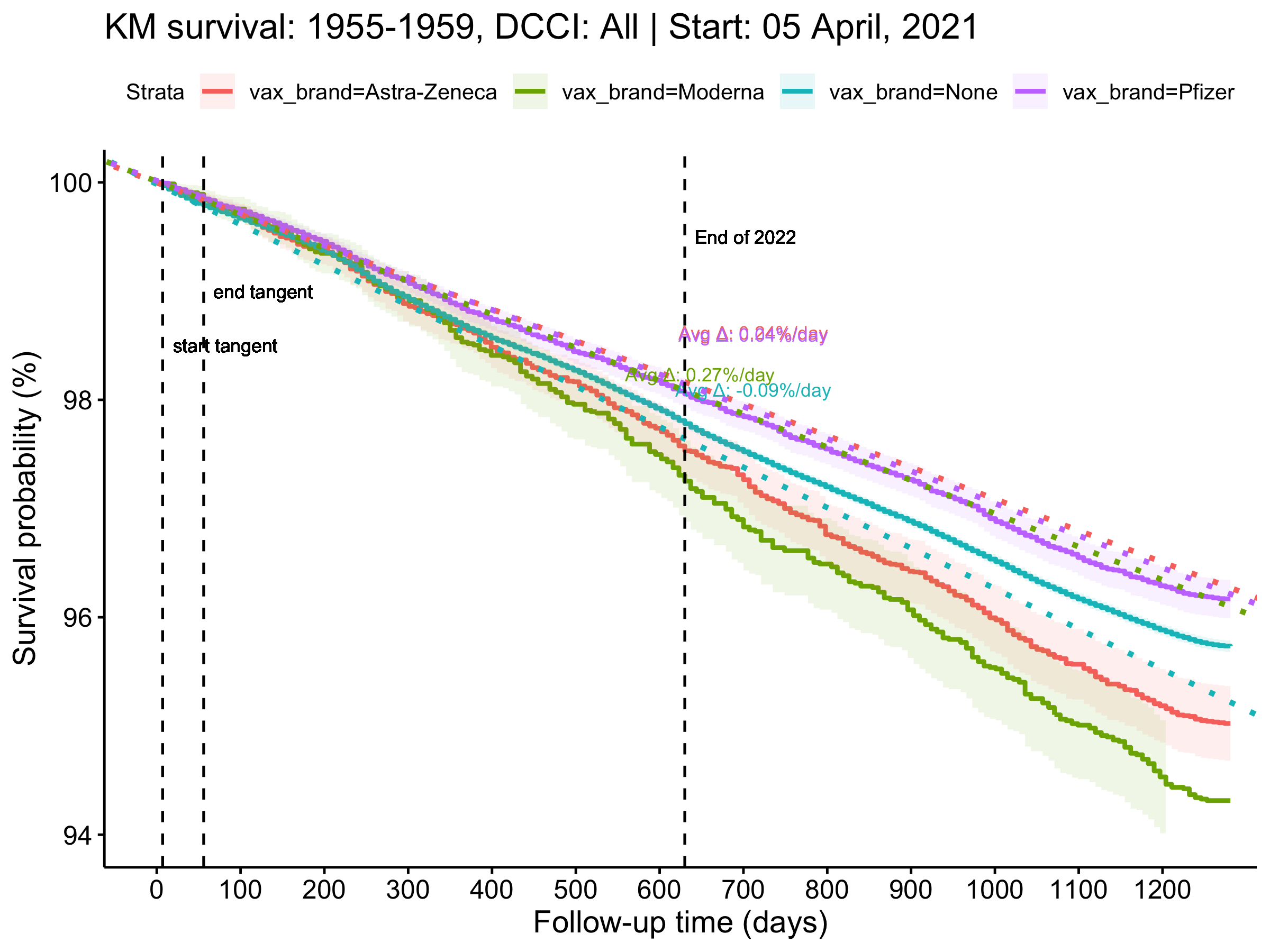
Generating plot for YearOfBirth: 1950-1954

[1] "vax\_brand=Astra-Zeneca Avg Δ: 0.67%/day"  
[1] "vax\_brand=Moderna Avg Δ: 0.91%/day"  
[1] "vax\_brand=None Avg Δ: -0.04%/day"  
[1] "vax\_brand=Pfizer Avg Δ: 0.42%/day"



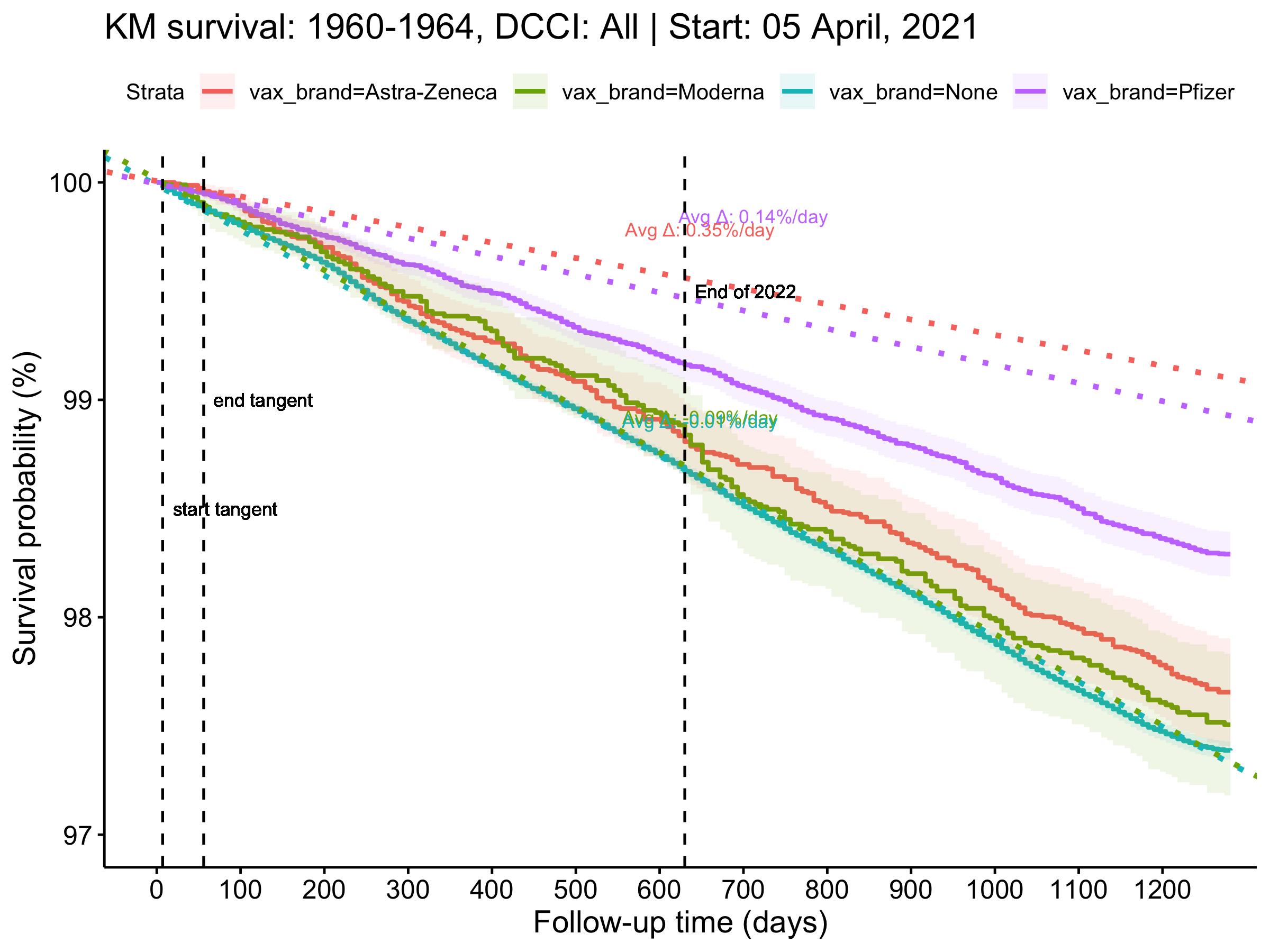
Generating plot for YearOfBirth: 1955-1959

[1] "vax\_brand=Astra-Zeneca Avg Δ: 0.24%/day"  
[1] "vax\_brand=Moderna Avg Δ: 0.27%/day"  
[1] "vax\_brand=None Avg Δ: -0.09%/day"  
[1] "vax\_brand=Pfizer Avg Δ: 0.04%/day"



Generating plot for YearOfBirth: 1960-1964

[1] "vax\_brand=Astra-Zeneca Avg Δ: 0.35%/day"  
[1] "vax\_brand=Moderna Avg Δ: -0.09%/day"  
[1] "vax\_brand=None Avg Δ: -0.01%/day"  
[1] "vax\_brand=Pfizer Avg Δ: 0.14%/day"



Generating plot for YearOfBirth: 1965-1969

[1] "vax\_brand=Astra-Zeneca Avg Δ: -0.11%/day"  
[1] "vax\_brand=Moderna Avg Δ: 0.16%/day"  
[1] "vax\_brand=None Avg Δ: 0.04%/day"  
[1] "vax\_brand=Pfizer Avg Δ: 0.02%/day"

