**Figure 1.1 Writing**

# **Results**

# **3.1: *APOE* variants modulate outcome of mouse SARS-CoV-2 MA10 infection**

To access the impact of APOE variation on SARS-CoV-2 infection on mice, we used pre-processed data provided from the paper to perform Kaplan-Meier estimate for the survival function [citation]. The goal was to estimate a population survival curve from a sample. To assess the survival rates of the SARS-CoV-2 MA10-infected *APOE*-knock-in mice stratified by age, sex, and *APOE* genotype, the Kaplan-Meier analysis approach was used to analyze the data. The Kaplan-Meier analysis allows estimation of survival over time, even when data points were omitted (for instance due to animal death in the study) or are studied for different lengths of time.

The statistical significance *P* values of Figures 1a and 1b were computed by log-rank tests, similarly to the original paper’s method. The log rank test allows us to test the null hypothesis of no difference in survival between two or more independent groups. The test compares the entire survival experience between groups and can be thought of as a test of whether the survival curves are identical (overlapping) or not.

Figure 1a demonstrates the survival of combined male and female SARS-CoV-2 MA10-infected *APOE*-knock-in mice stratified by age with the cutoff at 30 weeks, with the p-values of 0.00012.

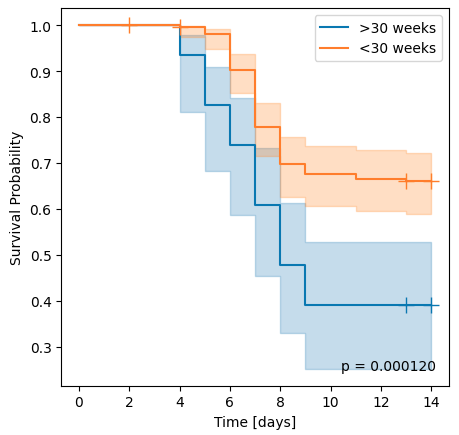


Figure 1a. Survival rate of COVID-19 infected mice by age

Figure 1b demonstrates the survival SARS-CoV-2 MA10-infected *APOE*-knock-in mice stratified by gender, male and female with p-value of 2.597e-13.

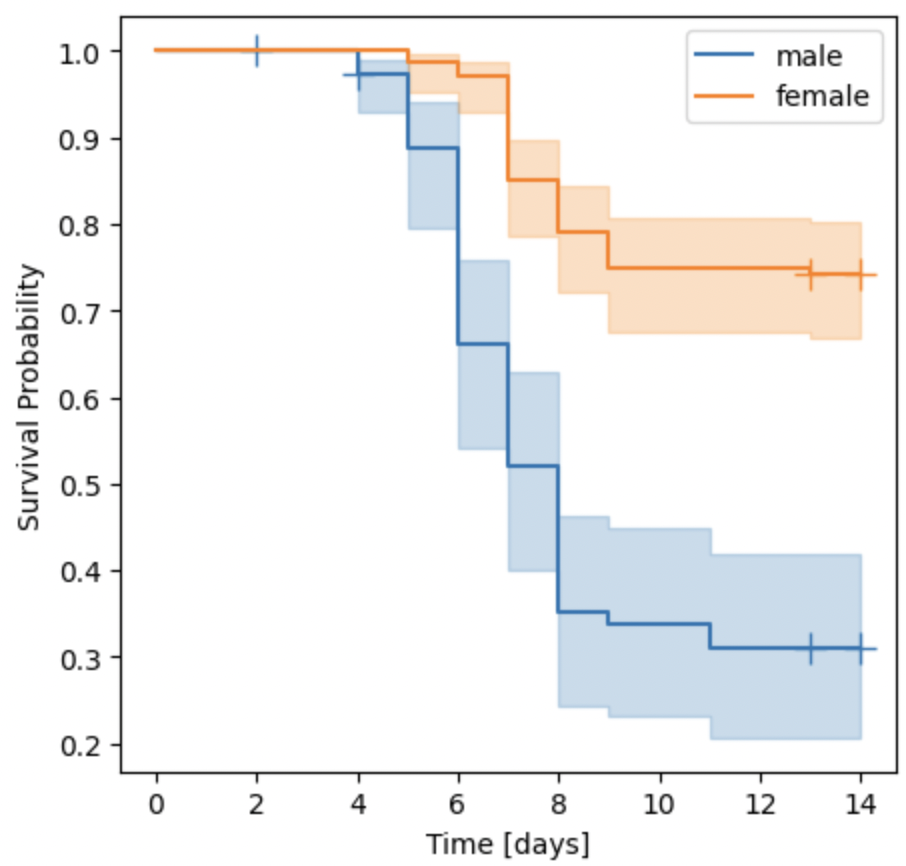


Figure 1b. Survival rate stratified by gender

The statistical significance *P* values of Figures 1c, 1d, and 1e were computed by Cox proportional hazards model, to allow for computing of multiple variables including the three genotypes including *APOE2*, *APOE3*, and *APOE4.*The Cox proportional hazards model approach allows us to evaluate simultaneously the effect of several factors on survival rate. Simply put, it allows us to examine how specific factors influence the rate of a particular event happening (e.g., genotype) at a particular point in time. This rate is commonly referred as the hazard rate. Predictor variables (or factors) are usually termed covariates in the survival-analysis literature.

ADDED: The drawback of Log-rank test is that it does not analyze other independent variables affecting the survival time. However, Cox proportional hazards (CPH) model is a semiparametric model which can analyzes multiple independent variables for estimating differences between the survival curves. Independent variables can include the variable of interest (e.g. treatments) and other potential confounders (e.g. age of the patients). For this reason we decided to use the CPH method for the following figures.

Figure 1c demonstrates the survival of combined male and female SARS-CoV-2 MA10-infected *APOE*-knock-in mice stratified by genotype including APOE2, APOE3, and APOE4 with p-value of 3.1e-5.

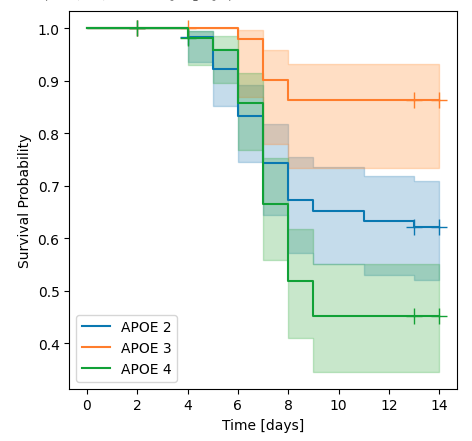


Figure 1c. Survival rate stratified by genotype

Figure 1d demonstrates the survival of male SARS-CoV-2 MA10-infected *APOE*-knock-in mice stratified by genotype including APOE2, APOE3, and APOE4, with p-value of 0.000355

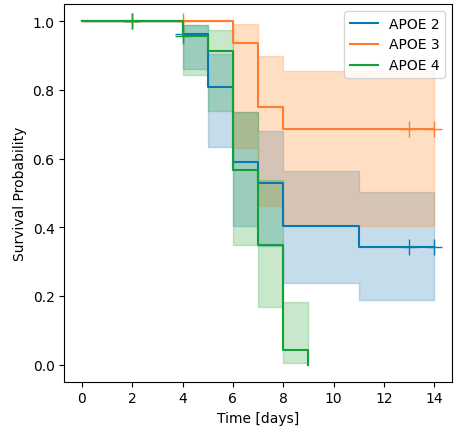


Figure 1d. Survival rate of male subjects stratified by genotype

Figure 1m demonstrates the survival of female SARS-CoV-2 MA10-infected *APOE*-knock-in mice stratified by genotype including APOE2, APOE3, and APOE4, with p-value of 0.0021

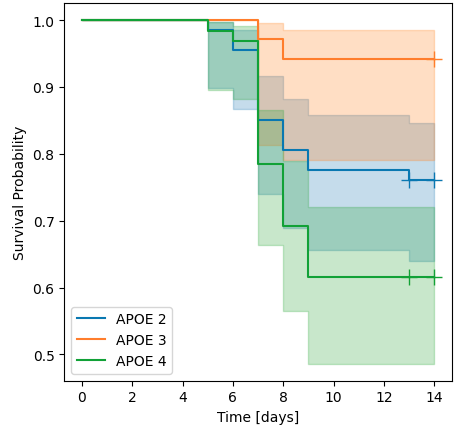


Figure 1e. Survival rate of female subjects stratified by genotype

ADDED: Additionally, while Cox proportional hazards model and log rank test methods are usually considered very similar, our reasoning to use the Cox proportional hazards model was because the Cox proportional hazards test allows the risk of death to vary within the model, whereas the log-rank test assumes it to be the same throughout. Because there were animal deaths in this study, we believed it would be a more appropriate approach.

ignore

(*P* values according to Kruskal–Wallis test (e,j), two-sided *t*-tests (g,l), log-rank test (h,m) and Cox proportional hazard models (i,n); note that f and k show group averages

## **The Kaplan-Meier parametric models is one method of modeling survival rate.**

“The **Cox proportional hazards test allows the risk of death to vary within the model**, whereas the log-rank test assumes it to be the same throughout.”<https://www.sciencedirect.com/topics/nursing-and-health-professions/log-rank-test>

<https://www.reneshbedre.com/blog/survival-analysis.html>

<https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_survival/BS704_Survival5.html#:~:text=The%20log%20rank%20test%20is,identical%20(overlapping)%20or%20not>.

(1) unlike the Cox model, log-rank does not generalize to a Bayesian framework, (2) the log-rank test only works for mutually exclusive categories and does not handle a continuous exposure variable, (3) log-rank does not allow for general covariate adjustment. Since the log-rank test is a special case of the Cox model, it does not have fewer assumptions or more power.

<https://stats.stackexchange.com/questions/486806/the-logrank-test-statistic-is-equivalent-to-the-score-of-a-cox-regression-is-th>

<https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_survival/BS704_Survival5.html#:~:text=The%20log%20rank%20test%20is,identical%20(overlapping)%20or%20not>.

<https://www.reneshbedre.com/blog/survival-analysis.html>

<http://www.sthda.com/english/wiki/cox-proportional-hazards-model#:~:text=The%20purpose%20of%20the%20model,referred%20as%20the%20hazard%20rate>.