

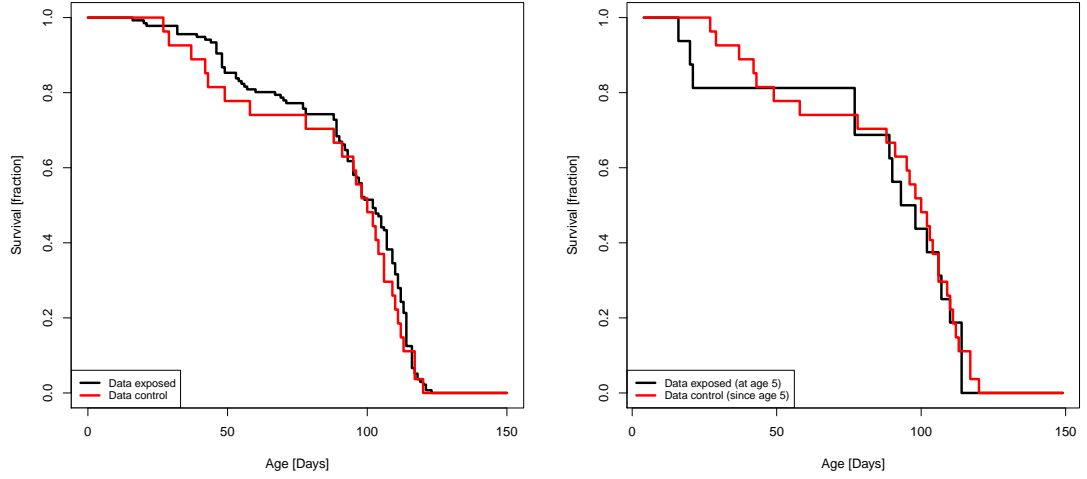
Comparison of survivals in control and exposed populations and its implications for virulence

Petr Kouba

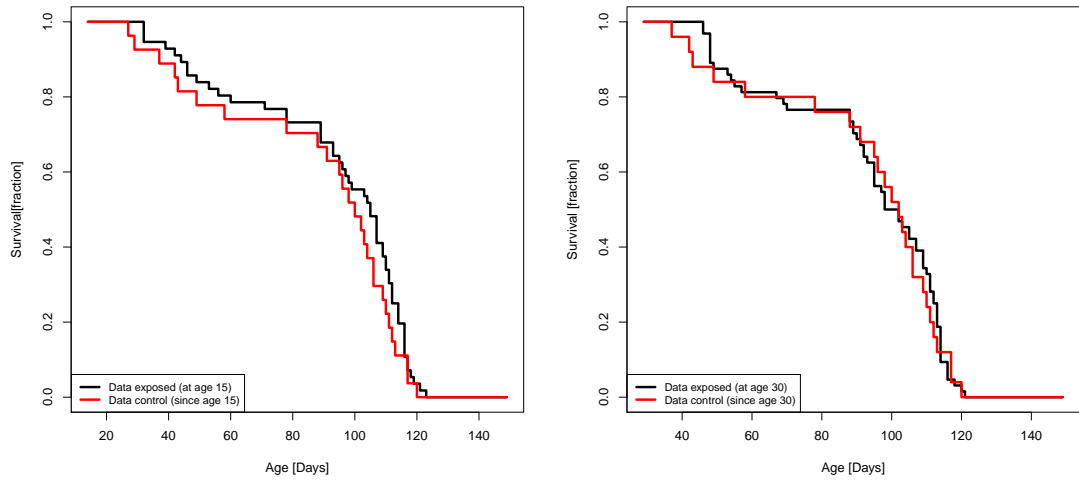
December 10, 2018

This document summarizes the differences between deathrate of exposed and control populations and provides the virulence taken with respect to the mortality of the control population. For the purpose of the whole document by "exposed" population, we mean the population which was exposed to the infection, but did not get infected.

1 Survival in the control population versus survival in the exposed population



(a) Survival of whole exposed [136Daphnias] and (b) Survival of population exposed at age 5 control [27Daphnia] populations, regardless of the [16Daphnias] and control population since age 5 age at infection. [27Daphnias].



(c) Survival of population exposed at age 15 [56Daphnias] and control population since age 15 [64Daphnias] and control population since age 30 [27Daphnias]. (d) Survival of population exposed at age 30 [25Daphnias].

Figure 1: Comparison of survival in exposed and control populations, within compartments based on ages at infection

From the plots in Figure 1, we can see that the survival pattern is very similar in control and exposed populations, except for age at infection 15, where there is a discrepancy. This discrepancy could be however a result of a small sample size in the exposed population for that given age at infection. Lets try to verify the similarity of these survival distributions statistically.

1.1 Hypothesis testing of the sameness of the exposed and control survival distributions

1.1.1 Kolmogorov-Smirnov test

As a first test, we tried the Kolmogorov-Smirnov test. Choosing critical p-value to be $p = 0.05$, we got rejection of the null hypothesis for the group belonging to the age at infection 5 ($p = 0.009$). For ages at infection 15 ($p = 0.057$) and 30 ($p = 0.699$), the null hypothesis (both distributions being drawn from the same distribution) was accepted.

Apperently, the Kolmogorov-Smirnov test might be too harsh on discrete data samples. Therefore we proceeded, by looking for further tests.

Question: What other tests might be convenient for our case?

1.2 Sample sizes for exposed and control populations

If we would assume that the exposed and the control populations follow the same survival distribution (which was confirmend by Kolmogorov-Smirnov test in two age-at-infection-groups out of 3), we could take the deathrate of the control population as the reference deathrate in all three age-at-infection classes and compute the virulence using this natural deathrate. However, the control population only contained 27 individuals, which only exceeds the sample size of population exposed at age 5 (16 individuals). Whereas the populations exposed at age 15 (56 individuals) and at age 30 (64 individuals) both provide us with larger sample.

As for the question regarding the sample sizes you had last time: "Why does the overall sample size (exposed +infected) decrease with age at infection? Does it correspond to the deathrate?"

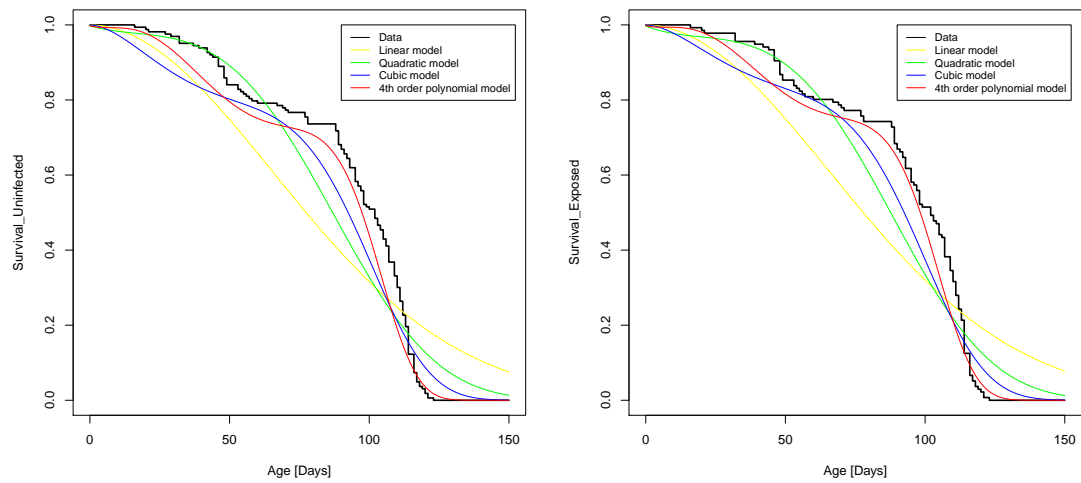
Our samples have the following numbers of individuals: $Infected_{atage5} + Exposed_{atage5} = 113$ $Infected_{atage15} + Exposed_{atage15} = 83$ $Infected_{atage30} + Exposed_{atage30} = 78$

Whereas if we would start with 113, who survived age 5 in each compartment, reflecting the deathrate we would get the sample sizes of 113, 113 and 105 respectively. So I believe this decrease is due to the design of the experiment and I do not see a good reason for it.

Question: If we manage to show that the population exposed at age 5 and the control population follow the same distribution (by using a better suited test), can we just - in order to get the largest sample possible - merge all of the exposed individuals and control into one compartment of uninfecteds and

use this whole compartment for the estimation of the natural deathrate? And if we find difference in exposing the individuals age 5, could we just estimate 2 different natural deathrates?(one for the exposed at age 5 and one common for the rest)

1.3 Fitting survival for "exposed population" and "uninfected population"



(a) Survival of the uninfected population and the corresponding fits (b) Survival of the exposed population and the corresponding fits

Figure 2: Comparison of survival in exposed and uninfected populations

2 Virulences with respect to the chosen reference deathrate

TODO: As soon as we decide on which compartment(s) to use for the estimation of the natural death rate, I will put the updated virulences here

TODO: Examine the statistical significance of the differences in virulence between different ages at infection

What would be the proper way to try to prove the significant difference between virulences for different ages at infection? Could logrank test be applied on the survival distribution models generated by the different virulences?