**Infection Intensity Analysis Methods – version 1\_\_04/27/23**

---

JS: methods on infecting, plating, counting colonies / CFUs.

---

LM: stats

All statistical analyses on the infection intensity data were performed using R statistical software, version 4.2.2 (R-Core-Team, 2022). All code for the infection intensity analysis is available as Sup Data File X.

To identify the effects of warmer temperature, older age, and their interaction on the infection intensity in mosquitoes, we used a zero-inflated, negative binomial model (R package “pscl”) to assess how these factors influenced the number of colony forming units (CFUs) per mosquito at 24 h following infection (Karazsia & van Dulmen, 2008; Loeys, Moerkerke, De Smet, & Buysse, 2012; Wang, Ma, & Wang, 2015; Zeileis, Kleiber, & Jackman, 2008). Some mosquitoes had zero CFUs at 24 h post infection, and these zero-counts were informative to the analysis, indicating the mosquito was able to clear the infection. Thus, the CFU count data was over dispersed, with a variance larger than the mean, and was best fit by a zero-inflated, negative binomial distribution. The zero-inflated component of the model predicts the probability of a mosquito having a zero CFU count value (cleared infection), using a binomial distribution with a logit link. The count component of the model predicts the non-zero number of CFUs per mosquito (active infection), using a negative binomial distribution with a log link.

The final model was determined by stepwise, multidirectional selection from the full model to minimize log-likelihood values and Akaike Information Criterion (AIC). Goodness of fit was also evaluated by comparing the observed and predicted number of zero and non-zero CFU counts. We then used ordinary, non-parametric bootstrapping (R package “boot”) to identify parameter estimates and 95% confidence intervals, as well as incident rate ratios (risk ratios) (for the negative binomial component) and odds ratios (for the zero component) on the multiplicative scale (Canty & Ripley, 2021; Davison & Hinkley, 1997). Additionally, we evaluated the effects of temperature, age, and their interaction by comparing the estimated marginal means (R package “emmeans”) (Lenth, 2022).

**References:**

Canty, A., & Ripley, B. (2021). boot: Bootstrap R (S-Plus) Functions (Version R package version 1.3-28).

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap Methods and Their Applications*. Cambridge: Cambridge University Press.

Karazsia, B. T., & van Dulmen, M. H. M. (2008). Regression Models for Count Data: Illustrations using Longitudinal Predictors of Childhood Injury\*. *Journal of Pediatric Psychology, 33*(10), 1076-1084. doi:10.1093/jpepsy/jsn055

Lenth, R. V. (2022). emmeans: Estimated Marginal Means, aka Least-Squares Means (Version R package version 1.8.3). Retrieved from <https://CRAN.R-project.org/package=emmeans>

Loeys, T., Moerkerke, B., De Smet, O., & Buysse, A. (2012). The analysis of zero-inflated count data: Beyond zero-inflated Poisson regression. *British Journal of Mathematical and Statistical Psychology, 65*(1), 163-180. doi:10.1111/j.2044-8317.2011.02031.x

R-Core-Team. (2022). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>

Wang, Z., Ma, S., & Wang, C.-Y. (2015). Variable selection for zero-inflated and overdispersed data with application to health care demand in Germany. *Biometrical Journal, 57*(5), 867-884. doi:10.1002/bimj.201400143

Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression Models for Count Data in R. *Journal of Statistical Software, 27*(8), 1 - 25. doi:10.18637/jss.v027.i08