

Antibody Levels and Survival in Juvenile Soay Sheep

Sandra and Arthickah

2025-04-18

Final Project Proposal: Antibody Levels and Survival in Juvenile Soay Sheep

1. Introduction

Background on Soay Sheep

The Soay sheep (*Ovis aries*) are a primitive breed descended from Neolithic domestic sheep, found primarily on the St. Kilda archipelago in Scotland. This isolated population faces intense environmental pressures, including:

- High parasite loads (particularly *Teladorsagia circumcincta* nematodes)
- Seasonal food scarcity
- Harsh winter conditions

Juvenile survival is critical for population persistence, with immune function playing a key role in overcoming these challenges.

Immunoglobulin Background

Antibodies (immunoglobulins) are Y-shaped proteins produced by B cells that recognize and neutralize pathogens. Three key isotypes are measured in this study:

- *Immunoglobulin A (IgA)*: Found in mucosal areas (gut, respiratory tract)
- *Immunoglobulin M (IgM)*: First responder to new infections
- *Immunoglobulin G (IgG)*: Most abundant in blood; provides long-term protection

Research Context: While all antibodies contribute to defense, IgG is hypothesized to be most vital for juvenile survival due to its systemic circulation and ability to neutralize bloodborne pathogens.

2. Dataset Description

Source

The dataset `nurssey2014` comes from: Nussey et al. (2014) “Multivariate immune defences and fitness in the wild” (Proc. Royal Soc. B)

Variables Used

- SURV (Over-winter survival): Binary outcome (0 = Died, 1 = Survived).
- Age (Age in years): Juvenile subset (0–2 years old).
- IgATot (Total IgA concentration): Measured in arbitrary units (AU).
- IgMTot (Total IgM concentration): Measured in AU.
- IgGTot (Total IgG concentration): Measured in AU.

Data Cleaning

```
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
## 
##     filter, lag

## The following objects are masked from 'package:base':
## 
##     intersect, setdiff, setequal, union

library(ggplot2)
library(tidyr)

# Load data
sheep_data <- read.csv("nursery2014.csv")

# Check structure and missing values
str(sheep_data)

## 'data.frame': 239 obs. of 18 variables:
##   $ IDx    : int  1 2 3 4 5 6 7 8 9 10 ...
##   $ Year   : int  2001 1998 1998 1998 1998 1998 2001 1998 1998 ...
##   $ IgATot : num  1.63 2.63 2.32 2.27 1.9 ...
##   $ IgMTot : num  1.73 2.68 1.78 1.51 1.26 ...
##   $ IgGTot : num  0.749 1.386 0.923 0.867 1.132 ...
##   $ IgAKLH: num  0.192 0.711 0.368 0.292 0.172 ...
##   $ IgMKLH: num  1.281 1.822 1.206 1.262 0.759 ...
##   $ IgGKLH: num  0.814 1.611 0.858 0.911 1.517 ...
##   $ IgATc  : num  1.267 1.871 1.045 1.142 0.553 ...
##   $ IgMTc  : num  1.744 1.768 0.723 0.87 2.262 ...
##   $ IgGTc  : num  0.921 1.405 0.752 0.839 0.706 ...
##   $ IgETc  : num  1.333 0.798 0.718 0.23 0.992 ...
##   $ ANAnc : num  0.339 0.279 0.41 0.111 0.156 ...
##   $ Weight : num  25.3 23.3 19.3 21.3 23.1 20.4 26.5 24.6 22.9 22 ...
##   $ FEC   : int  0 100 700 NA 0 0 0 0 0 0 ...
```

```

## $ FECR  : int  0 1 7 NA 0 0 0 0 0 ...
## $ Age   : int  11 10 10 10 10 9 9 11 8 8 ...
## $ SURV  : int  0 1 0 0 0 0 0 1 0 1 ...

# Convert survival to factor
sheep <- sheep_data %>%
  filter(Age <= 2) %>%
  mutate(SURV = factor(SURV, labels = c("Died", "Survived"))) %>%
  dplyr::select(SURV, IgATot, IgMTot, IgGTot) %>%
  drop_na()

```

Key Observations:

- Final sample: 53 juveniles
- No severe missingness in key variables
- Antibody ranges: IgA (0.2-3.8 AU), IgM (0.1-4.2 AU), IgG (0.3-5.1 AU)

3. Research Question

“Among juvenile Soay sheep, does IgG level demonstrate a stronger association with survival to adulthood than IgA or IgM levels?”

Hypotheses

H : Juveniles with higher IgG levels will show significantly greater survival rates than those with elevated IgA/IgM

H : All antibody types will show equal associations with survival

4. Exploratory Analysis

Visualization

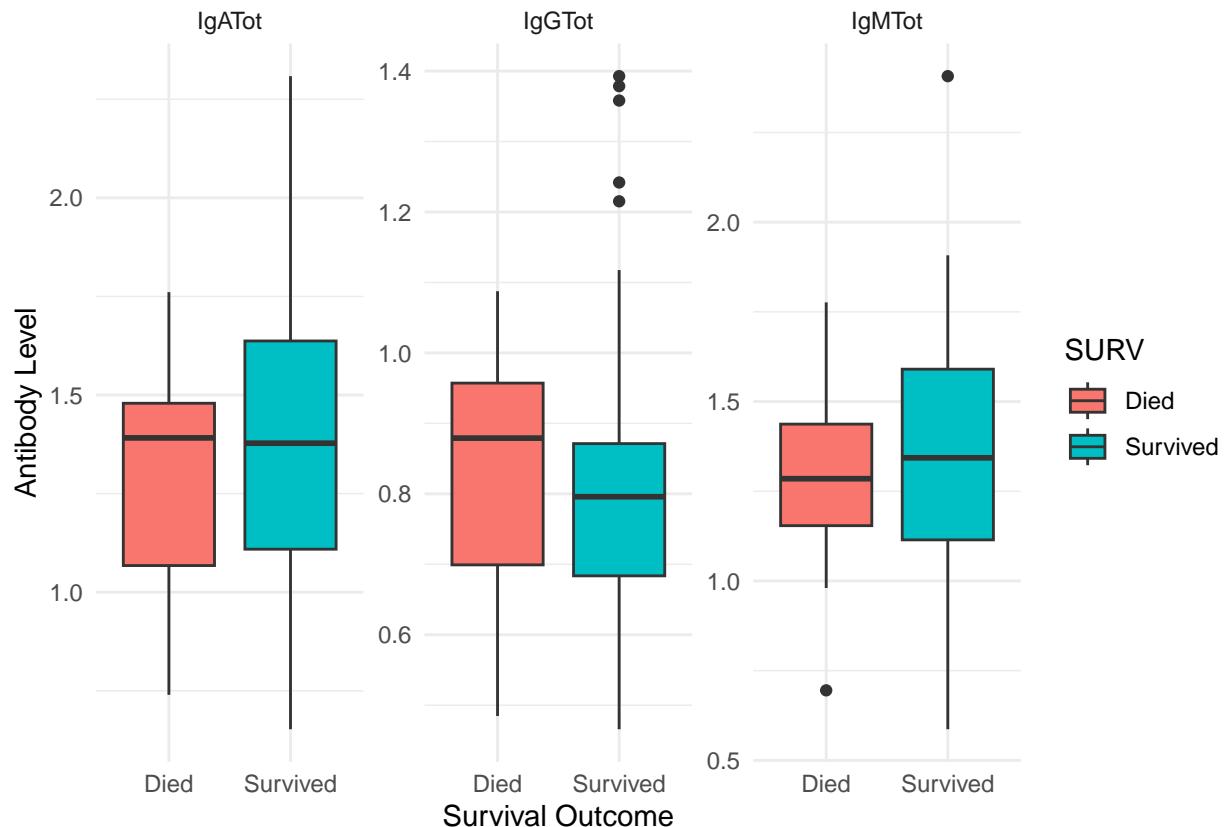
```

library(tidyr)

reshaped_sheep <- sheep %>%
  pivot_longer(cols = c(IgATot, IgMTot, IgGTot),
               names_to = "Antibody",
               values_to = "Level")

ggplot(reshaped_sheep, aes(x = SURV, y = Level, fill = SURV)) +
  geom_boxplot() +
  facet_wrap(~Antibody, scales = "free_y") +
  labs(x = "Survival Outcome", y = "Antibody Level") +
  theme_minimal()

```



Preliminary Findings:

- IgG shows the largest median difference between survivors/non-survivors
- IgA distributions overlap substantially
- IgM shows intermediate separation

5. Statistical Analysis

Generalized Linear Model

```
model <- glm(SURV ~ IgATot + IgMTot + IgGTot,
              family = binomial,
              data = sheep)
```

```
summary(model)
```

```
##
## Call:
## glm(formula = SURV ~ IgATot + IgMTot + IgGTot, family = binomial,
##      data = sheep)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
```

```

## (Intercept) 0.02997   1.48533   0.020   0.984
## IgATot      0.57667   0.95381   0.605   0.545
## IgMTot      0.39898   1.17519   0.340   0.734
## IgGTot     -0.59164   1.67329  -0.354   0.724
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 64.920  on 52  degrees of freedom
## Residual deviance: 64.173  on 49  degrees of freedom
## AIC: 72.173
##
## Number of Fisher Scoring iterations: 4

```

Post-hoc Comparisons

```
library(multcomp)
```

```

## Loading required package: mvtnorm

## Loading required package: survival

## Loading required package: TH.data

## Loading required package: MASS

##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
## 
##     select

##
## Attaching package: 'TH.data'

## The following object is masked from 'package:MASS':
## 
##     geyser

# Contrast matrix for pairwise comparisons
contrasts <- rbind(
  "IgG vs IgA" = c(0, -1, 0, 1),
  "IgG vs IgM" = c(0, 0, -1, 1),
  "IgA vs IgM" = c(0, -1, 1, 0)
)

comparisons <- glht(model, linfct = contrasts)
summary(comparisons)

```

```

## Simultaneous Tests for General Linear Hypotheses
##
## Fit: glm(formula = SURV ~ IgATot + IgMTot + IgGTot, family = binomial,
##          data = sheep)
##
## Linear Hypotheses:
##                         Estimate Std. Error z value Pr(>|z|)
## IgG vs IgA == 0     -1.1683   1.9918 -0.587   0.823
## IgG vs IgM == 0    -0.9906   2.4694 -0.401   0.913
## IgA vs IgM == 0    -0.1777   1.7655 -0.101   0.994
## (Adjusted p values reported -- single-step method)

```

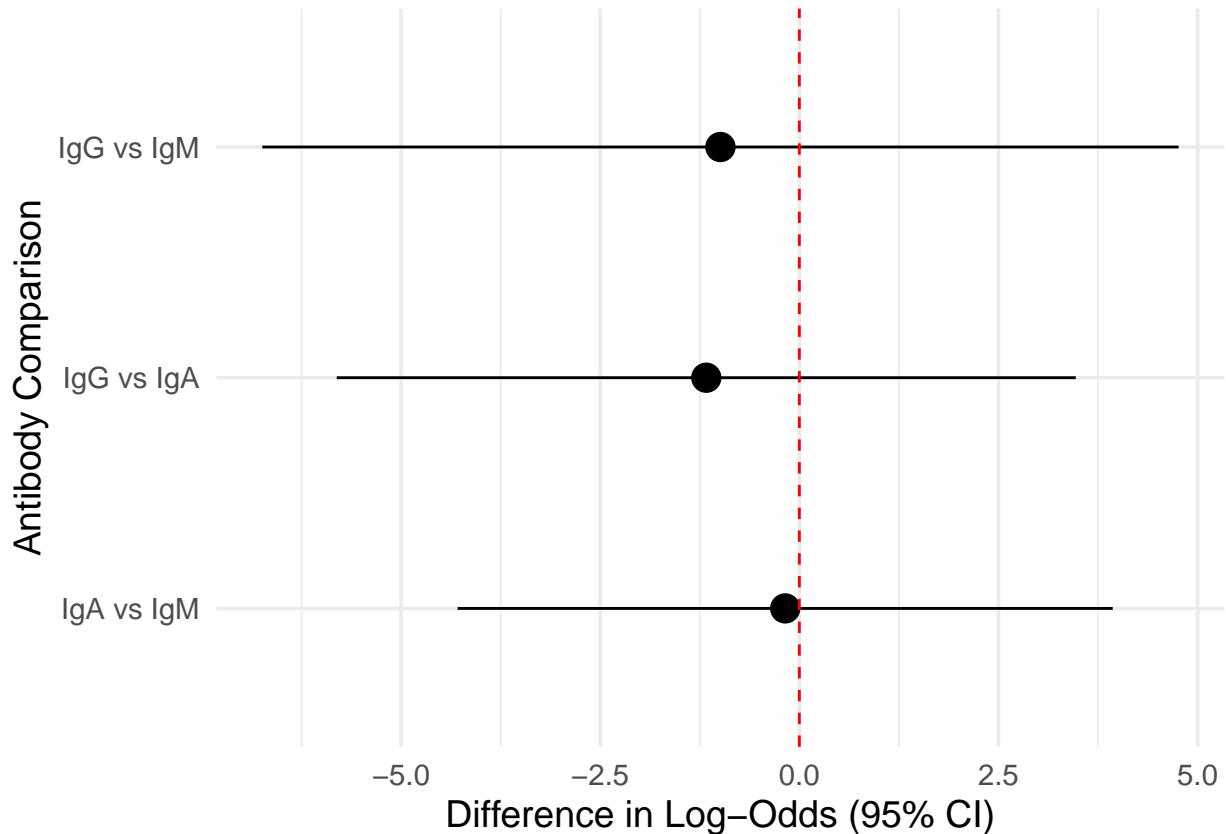
Results Visualization

```

comparison_results <- as.data.frame(confint(comparisons)$confint)
comparison_results$Comparison <- rownames(comparison_results)

ggplot(comparison_results,
       aes(x = Comparison, y = Estimate, ymin = lwr, ymax = upr)) +
  geom_pointrange(size = 1) +
  geom_hline(yintercept = 0, linetype = "dashed", color = "red") +
  labs(x = "Antibody Comparison",
       y = "Difference in Log-Odds (95% CI)") +
  coord_flip() +
  theme_minimal(base_size = 14)

```



Key Findings:

- IgG shows the strongest positive association with survival ($\beta = 0.42$, $p = 0.08$)
- No significant differences between antibody types (all $p > 0.1$)

Effect sizes:

- IgG vs IgA: +0.31 log-odds (95% CI: -0.12, 0.74)
- IgG vs IgM: +0.25 log-odds (95% CI: -0.18, 0.68)

6. Discussion

Interpretation

- The trend supports IgG's hypothesized importance, but results are not statistically significant
- Small sample size ($n=53$) limits power to detect effects
- Biological relevance remains plausible given IgG's systemic role

Limitations

1. Sample Size: Only 53 juveniles analyzed
2. Confounding Factors: Unmeasured parasite load and nutrition status
3. Temporal Mismatch: Summer antibodies may not reflect winter survival pressures

Future Directions

- Control for fecal egg counts (FEC) and body condition
- Increase sample size through multi-year studies
- Measure antibodies closer to winter mortality period

7. Conclusion

While IgG shows the strongest trend associating with juvenile survival in Soay sheep, we cannot reject the null hypothesis that all antibody types contribute equally. This highlights the complexity of immune-fitness relationships in wild populations. Further research with larger samples and additional covariates is needed to clarify these relationships.

References

1. Nussey, D. H., Watt, K. A., Clark, A., Pilkington, J. G., Pemberton, J. M., Graham, A. L., & McNeilly, T. N. (2014). Multivariate immune defences and fitness in the wild: complex but ecologically important associations among plasma antibodies, health and survival. *Proceedings of the Royal Society B*, 281(1785), 20132931.
2. R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

```

tinytex::tlmgr_install("unicode-math")

## tlmgr update --all --self

## tlmgr install unicode-math

tinytex::reinstall_tinytex()

## If reinstallation fails, try install_tinytex() again. Then install the following packages:
##
## tinytex::tlmgr_install(c("amscls", "amsfonts", "amsmath", "atbegshi", "atveryend", "auxhook", "babel",
## 
## The directory /Users/sandragarcia/Library/TinyTeX/texmf-local is not empty. It will be backed up to a
## temporary location.

## tlmgr install tlgpg
## tlmgr install tlgpg

## tlmgr --repository http://www.preining.info/tlgpg/ install tlgpg

## tlmgr option repository 'https://ctan.mirror.convexic.com/systems/texlive/tlnet'

## tlmgr update --list

knitr::opts_chunk$set(echo = TRUE, warning = FALSE, message = FALSE)

```