Longitudinal analysis of Shigella

0.1 Data description

- Step 1: Load the Dataset

```
df <- read_excel("3.8.2024 Compiled Shigella datav2.xlsx",
                        sheet = "Compiled")
```

- Step 2: Get a Quick Overview of the Data Structure

Table 1: Data summary

Name	df
Number of rows	2374
Number of columns	31
_____	
Column type frequency:	
character	9
numeric	22
_____	
Group variables	None

Variable type: character

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
isotype_name	0	1.00	3	3	0	2	0
dilution_name	0	1.00	6	6	0	1	0
study_name	0	1.00	4	22	0	5	0
sampleID	0	1.00	6	11	0	1017	0
sid	0	1.00	6	11	0	1017	0
unq_id	0	1.00	22	27	0	2374	0
cohort_name	0	1.00	3	11	0	12	0

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
site_name	0	1.00	5	12	0	5	0
treatment	1126	0.53	7	7	0	2	0

Variable type: numeric

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
isotype	0	1.00	1.50	0.50	1.00	1.00	1.50	2.00	2.00	
dilution	0	1.00	1.00	0.00	1.00	1.00	1.00	1.00	1.00	
cohort	0	1.00	8.49	1.66	4.00	9.00	9.00	9.00	15.00	
site	0	1.00	5.33	2.22	1.00	2.50	7.00	7.00	7.00	
age	0	1.00	6.58	8.66	0.08	2.00	3.67	7.00	50.00	
timepoint	0	1.00	969.01	901.92	0.00	0.00	1825.00	1825.00	1825.00	
Actual day	1942	0.18	53.06	62.90	2.00	7.00	30.00	90.00	204.00	
ipab_MFI	0	1.00	43525.84	59594.93	-1.25	818.58	4759.92	86098.81	189355.75	
sf3aospbsa_MFI	0	1.00	3651.27	8563.89	-9.00	183.25	753.71	2956.48	83828.92	
sf2aospbsa_MFI	0	1.00	4681.13	14202.95	-12.00	100.06	489.04	2694.00	214007.67	
sf6ospbsa_MFI	0	1.00	4043.35	9996.03	-7.75	278.25	1036.29	3714.58	166878.50	
sonneiospbsa_MFI	0	1.00	3007.76	12898.10	-14.50	28.08	118.12	887.46	184109.50	
n_ipab_MFI	0	1.00	24971.08	34729.71	-0.63	361.53	1903.47	51093.85	116048.97	
n_sf3aospbsa_MFI	0	1.00	5702.38	12801.36	-12.10	320.15	1273.00	4743.07	112888.86	
n_sf2aospbsa_MFI	0	1.00	5308.11	15272.92	-11.82	135.66	621.84	3295.17	222195.19	
n_sf6ospbsa_MFI	0	1.00	2991.34	6997.96	-6.84	231.92	845.01	2826.64	108831.02	
n_sonneiospbsa_MFI	0	1.00	3747.64	14652.84	-28.75	38.78	183.09	1335.17	219024.17	
ipab_BC	0	1.00	80.03	29.48	17.00	59.00	73.00	91.00	260.00	
sf3aospbsa_BC	0	1.00	103.91	22.52	57.00	88.00	100.00	115.00	226.00	
sf2aospbsa_BC	0	1.00	67.03	18.96	23.00	54.00	63.00	76.00	155.00	
sf6ospbsa_BC	0	1.00	56.18	13.04	27.00	47.00	53.00	63.00	115.00	
sonneiospbsa_BC	0	1.00	58.37	14.60	17.00	49.00	58.00	68.00	114.00	

- Step 3: Identify and Summarize Missing Values (for the entire dataset and the SOSAR study)

```
# A tibble: 31 x 2
   variable      missing_count
   <chr>                 <int>
 1 Actual day             1942
 2 treatment              1126
 3 isotype                   0
 4 isotype_name              0
 5 dilution                  0
 6 dilution_name             0
 7 study_name                0
 8 sampleID                  0
 9 sid                       0
10 unq_id                    0
# i 21 more rows
```

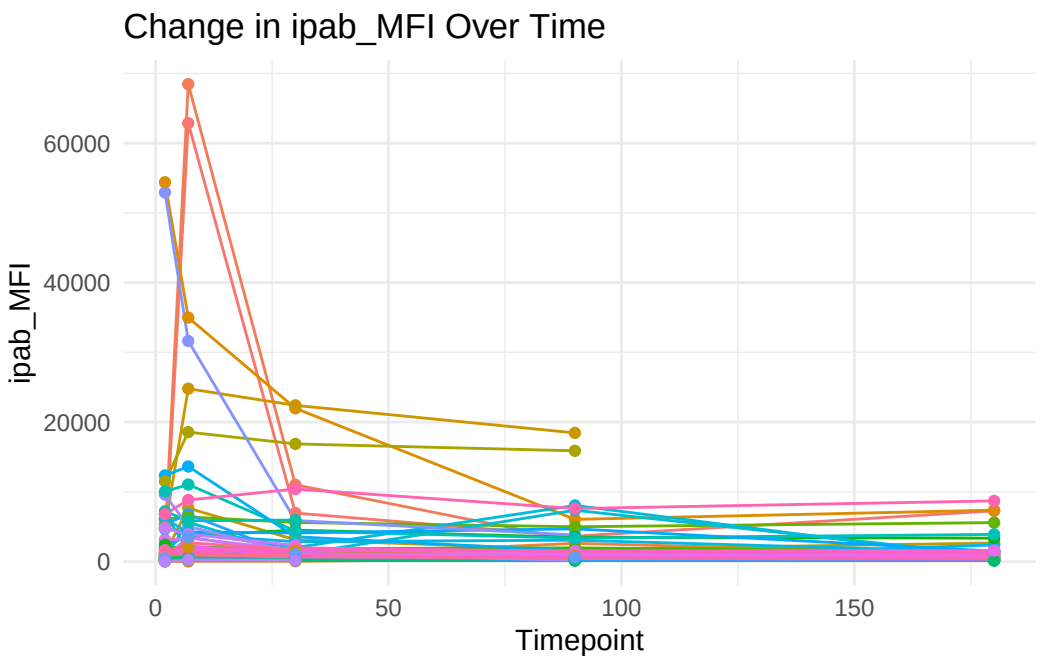
```
# A tibble: 31 x 2
   variable      missing_count
   <chr>                 <int>
 1 treatment               436
 2 Actual day                4
 3 isotype                   0
 4 isotype_name              0
 5 dilution                  0
 6 dilution_name             0
 7 study_name                0
 8 sampleID                  0
 9 sid                       0
10 unq_id                    0
# i 21 more rows
```

0.2 Exploratory Data Analysis

- Step 4: Filter Data for the SOSAR Study and Split Data by Isotype

Verifying the unique sample identification numbers (SIDs) is crucial. There are 48 SIDs in this SOSAR study.

- Step 5: Plot Quantitative Response Change Over Time



0.3 Longitudinal Antibody Decay Modeling

- Step 6: Prepare Data for JAGS Model
- Step 7: Subset Data for Model Testing

```
#subset data for checking
dL_sub <- dL %>%
 filter(index_id %in% sample(unique(index_id), 20))
```

- Step 8: Prepare Data and Priors for JAGS
- Step 9: Run the JAGS Model

```
dump_and_quit <- function() {
  dump.frames(to.file = TRUE)
  q(status = 1)
}
options(error = dump_and_quit)
```

```
[1] "Starting run.jags execution"
```

```
List of 5
 $ smpl.t        : num [1:21, 1:5] 2 2 2 2 2 2 2 2 2 2 ...
  ..- attr(*, "dimnames")=List of 2
  .. ..$ subjects    : chr [1:21] "SOSAR-11001" "SOSAR-12017" "SOSAR-11012" "SOSAR-12005" ..
  .. ..$ visit_number: chr [1:5] "V1" "V2" "V3" "V4" ...
 $ logy          : num [1:21, 1:5, 1:2] -4.61 -4.61 -4.61 -4.61 -4.61 ...
  ..- attr(*, "dimnames")=List of 3
  .. ..$ subjects    : chr [1:21] "SOSAR-11001" "SOSAR-12017" "SOSAR-11012" "SOSAR-12005" ..
  .. ..$ visit_number: chr [1:5] "V1" "V2" "V3" "V4" ...
  .. ..$ antigens    : chr [1:2] "IgG" "IgA"
 $ n_antigen_isos: int 2
 $ nsmpl         : num [1:21] 5 4 5 5 5 5 5 5 5 ...
 $ nsubj         : num 21
 - attr(*, "antigens")= chr [1:2] "IgG" "IgA"
 - attr(*, "n_antigens")= int 2
 - attr(*, "ids")= chr [1:21] "SOSAR-11001" "SOSAR-12017" "SOSAR-11012" "SOSAR-12005" ...
NULL


List of 6
 $ n_params     : num 5
 $ mu.hyp       : num [1:2, 1:5] 1 1 7 7 1 1 -4 -4 -1 -1
 $ prec.hyp     : num [1:2, 1:5, 1:5] 1 1 0 0 0 0 0 0 0 0 ...
 $ omega        : num [1:2, 1:5, 1:5] 1 1 0 0 0 0 0 0 0 0 ...
 $ wishdf       : num [1:2] 20 20
 $ prec.logy.hyp: num [1:2, 1:2] 4 4 1 1
NULL


Calling 4 simulations using the parallel method...
Following the progress of chain 1 (the program will wait for all chains
to finish before continuing):
Welcome to JAGS 4.3.2 on Wed Aug 28 00:47:59 2024
JAGS is free software and comes with ABSOLUTELY NO WARRANTY
Loading module: basemod: ok
Loading module: bugs: ok
. . Reading data file data.txt
. Compiling model graph
   Resolving undeclared variables
   Allocating nodes
Graph information:
   Observed stochastic nodes: 186
   Unobserved stochastic nodes: 54
   Total graph size: 3201
```

```
. Reading parameter file inits1.txt
. Initializing model
. Adapting 100
-------------------------------------------------| 100
+++++++++++++++++++++++++++++++++++++++++++++++++ 100%
Adaptation incomplete.
. Updating 100
-------------------------------------------------| 100
************************************************* 100%
. . . . . . Updating 100
-------------------------------------------------| 100
************************************************* 100%
. . . . Updating 0
. Deleting model
.
All chains have finished

Warning: The adaptation phase of one or more models was not completed in 100
iterations, so the current samples may not be optimal - try increasing the
number of iterations to the "adapt" argument

Simulation complete.  Reading coda files...
Coda files loaded successfully
Finished running the simulation

[1] "run.jags completed successfully"
```

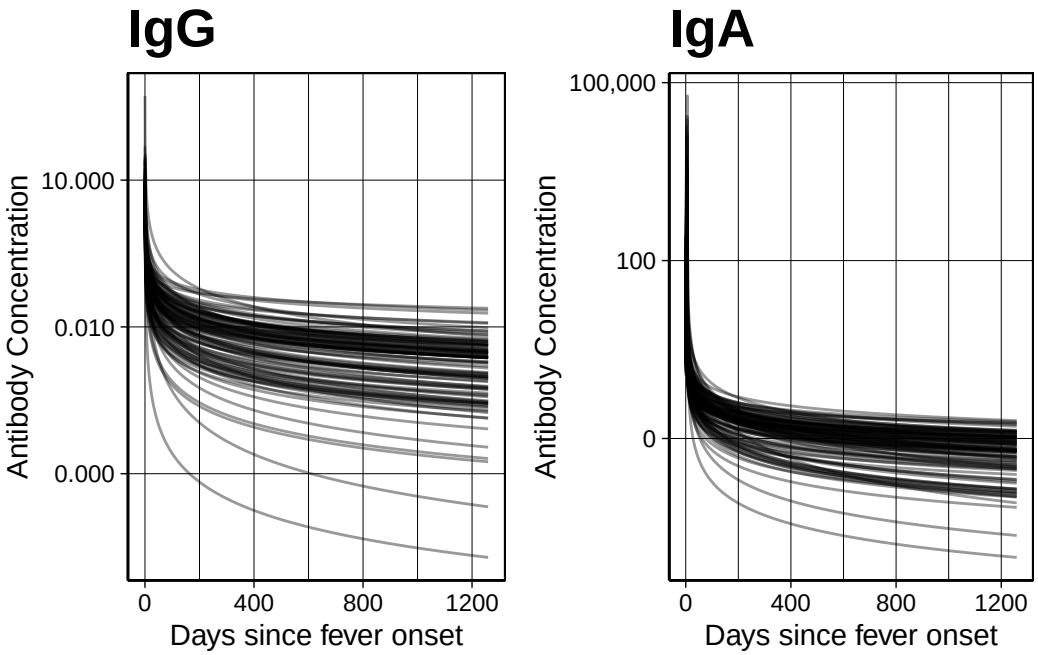
- Step 10: Covert JAGS Output to MCMC List

```r
mcmc_list <- as.mcmc.list(jags.post)

mcmc_df <- ggs(mcmc_list)
```

- Step 11: Prepare Data for Serocalculator Package

- Step 12: Plot Longitudinal Antibody Decay



- Step 13: Define the Antibody Calculation Function

```r
ab <- function(t, y0, y1, t1, alpha, shape) {
  beta <- log(y1/y0) / t1
  if (t <= t1) {
    yt <- y0 * exp(beta * t)
  } else {
    yt <- (y1^(1 - shape) - (1 - shape) * alpha * (t - t1))^(1 / (1 - shape))
  }
  return(yt)
}
```

- Step 14: Generate Time Points for Antibody Levels
- Step 15: Calculate Antibody Levels Over Time

- Step 16: Summarize and Plot Antibody Decay Data