

Three (Groups of) Blind Mice. Familial Clusters of Cataract Development in Irradiated Mice

Alyssa Allsop and Amira Burns
Department of Statistics, Colorado State University

June 22, 2022

Abstract

As astronauts travel farther into space, outside of the earth's protective magnetic field, they are exposed to radiation that could have serious effects. One of those side effects is cataract development. It has also been thought that there might be a genetic predisposition to develop cataracts that is passed down in families. The goal of this study was to use mice in families exposed to radiation to determine if there is cataract clustering in families. A mixed effects model was created to account for the family effect on cataract development. The model included terms for sex, treatment group, sex by treatment group interaction, and a random intercept for family. Family was shown to be an important source of variation in predicting cataract development. Males had big differences in cataract development among treatment groups whereas females had only small differences. Males were also more likely to develop severe cataracts than females in all three treatment groups. Other risks involved in radiation exposure could cause death before cataracts have time to develop. This complicates the inferences made about cataract development from this dataset; however, the conclusions indicate the necessity of further research to better understand how humans might react to radiation as they leave the earth's magnetic field.

Keywords: Hierarchical modeling, Logistic Regression, Bayesian analysis

1 Introduction

Little is known about the effects of high atomic number and energy (HZE) radiation, a main component of space radiation to which exposure is unavoidable beyond Earth’s magnetic field. In contrast, extensive research shows exposure to high doses of gamma radiation leads to acute radiation sickness. Effects include damage to the blood forming system, GI system, immune system, increased risk for cancer, cardiovascular disease, neurodegenerative disease, death, and cataracts. Adverse effects from radiation may also be attributable to other factors, including genetics; several different cancers have been observed to cluster in mice families (Chernyavskiy et al. 2017). Accounting for potential family clustering allows for thorough examination of the primary research questions.

Is there a genetic susceptibility to cataracts caused by radiation? Accounting for potential genetic susceptibility, is there a difference in cataract presentation between HZE radiation, gamma radiation and unirradiation? 1820 individual mice from 48 unique families, were randomly assigned with equal family weight to each of three treatment groups. The mice were bred over several generations to create a genetically heterogeneous sample to better represent biological diversity within the human population. The treatments are HZE irradiation, gamma irradiation, and non-irradiated control. The HZE group is irradiated with either silicon or iron nuclei HZE ions, which are considered as a single treatment group. The second group is subjected to ^{137}Cs gamma irradiation. Mice in both irradiated groups are exposed to radiation at 7-12 weeks of age. The third group is unirradiated control. All mice are monitored until 800 days of age –effectively a survival study. Mice are checked weekly for symptoms of cataracts, cataract risk factors, and other symptoms of radiation exposure such as tumors and carcinomas. Resources do not allow for weekly measurement of every mouse; previous measurements are carried forward if a mouse is not assessed on a particular week.

This analysis uses a simplified version of the data set - a snapshot of the 1169 mice that are alive at 552 days. This cutoff is chosen because it is the median survival time for the group with the shortest median survival. There are 47 unique families with $n = 396$ in the HZE group, $n = 277$ in the gamma radiation group, and $n = 496$ in the unirradiated control group. There are up to two generations of mice pups from each family in the dataset, but

generation is not distinguished in the simplified data. Family size ranges from $n = 11$ to $n = 48$, with median family size of 24.

The response is Merriam-Focht cataract score (Merriam & Focht, 1957), an ordinal categorical variable corresponding to radiation-associated ocular changes in the eye. A score of 0 is associated with a completely clear lens, while a score of 5 is associated with a completely occluded lens. This dataset contains cataract score levels = $[1, 2, 3, 4]$. A score ≥ 2 indicates presence of cataracts, and the small sample sizes for score > 2 across all treatment

Treatment	Cataract Score				Total
	1	2	3	4	
Unirradiated	438	45	11	2	496
Gamma	214	53	6	4	277
HZE	281	107	6	2	396

Table 1: Counts of score by treatment group

groups raise concerns about making inference on an ordinal analysis; consequently, the response is converted to binary *Cataracts* with score = 1 converted to 0, and score ≥ 2 converted to 1. The main experimental factor is *Treatment*, a categorical variable with three levels = [HZE radiation, gamma radiation, non-irradiated control]. A single random effect, genetic *Family*, is a categorical factor with 47 levels. Both Treatment and Family are central to the experimental design and will serve as the basis for all models considered. Additional covariates under consideration were Sex, coat color, weight in grams, body condition score (BCS) age in days, and presence three cancers: myeloid leukemia, harderian tumors, and PreT lymphoma. The experimental design prescribed the inclusion of Treatment and Family in the analysis.

2 Summary Statistics

Exploratory data analysis focused on assessing binary Cataracts status in terms of Treatment. Tables and graphs were produced to explore (wc) any potential associations between covariates and the response, and correlations between covariates and either Treatment or the response. Further examination of potential covariates show distinct patterns of cataracts status between male and female mice. Figure 1 highlights the differences between males and females across all treatment groups.

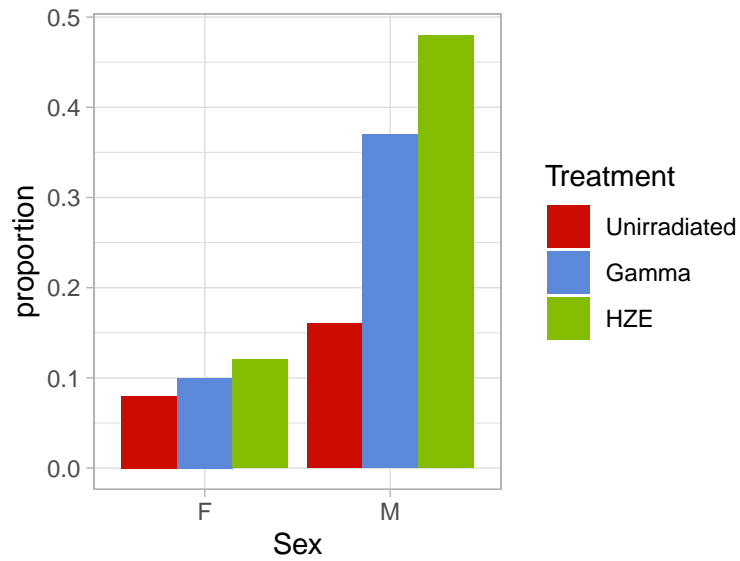


Figure 1: Sample proportions with cataracts by sex, treatment group

Figure 2 allows us to visualize any patterns between families within other predictors of interest. Different patterns are displayed across families, but distinct clustering is not apparent.

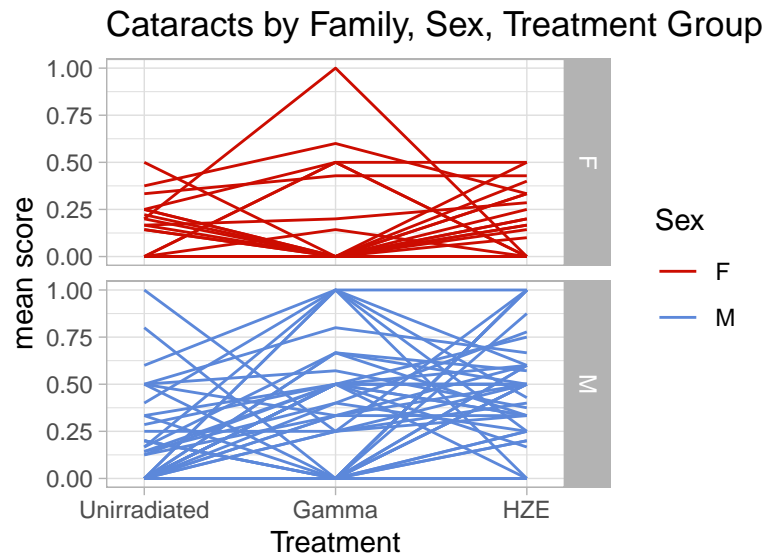


Figure 2: Average family cataract score by sex, treatment group

3 Statistical Methods

Hierarchical logistic regression was selected as the primary analytic method, due to the binary response and random effect of interest to the research question. Figure 1 indicated potential interaction between Sex and Treatment group; therefore, an interaction term was included in the model. Various models were fit and assessed using Aikike's Information Criterion (AIC); the full list of models fit can be viewed in the **APPENDIX - add cross ref!**. The final model, Model (1), can be written as:

$$\begin{aligned}
 Y_{ij} &\sim \text{Bernoulli}(p) \\
 \log\left(\frac{p}{1-p}\right) &= \beta_0 + \beta_1 * \text{Gamma}_i + \beta_2 * \text{HZE}_i + \beta_3 * M_i + \\
 &\quad \beta_4 * \text{Gamma}_i * M_i + \beta_5 * \text{HZE}_i * M_i + [v_j + \epsilon_{ij}] \\
 i &= 1, \dots, 1169 \text{ mice} \\
 j &= 1, \dots, 47 \text{ families}
 \end{aligned} \tag{1}$$

where M_i is an indicator for males, v_j is the variance of the random intercept for Family j , and ϵ_{ij} is the residual variance.

A complementary Bayesian model with non-informative priors was fit to examine the distribution of the estimates from the final model, as well as test its robustness to alternative approaches. The Bayes model is parameterized as:

$$\begin{aligned}
Y_{ij} &\sim \text{Bernoulli}(p) \\
\log\left(\frac{p}{1-p}\right) &= \beta_0 * \text{Control}_i * F_i + \beta_1 * \text{Gamma}_i * F_i + \beta_2 * \text{HZE}_i * F_i + \\
&\quad \beta_3 * \text{Control}_i * M_i + \beta_4 * \text{Gamma} * M + \beta_5 * \text{HZE} * M + \\
&\quad v_j + \epsilon_{ij} \\
i &= 1, \dots, 1169 \text{ mice} \\
j &= 1, \dots, 47 \text{ families, and} \\
v_j &\sim N(0, \tau) \\
&\text{with non-informative priors:} \\
\beta_0 &\sim N(0, 0.001) \\
\beta_1 &\sim N(0, 0.001) \\
\beta_2 &\sim N(0, 0.001) \\
\beta_3 &\sim N(0, 0.001) \\
\beta_4 &\sim N(0, 0.001) \\
\beta_5 &\sim N(0, 0.001) \\
v_i &\sim N(0, \sigma^2) \\
\tau &\sim \text{Gamma}(0.001, 0.001) \text{ where } \sigma^2 = 1/\tau
\end{aligned} \tag{2}$$

Model (2), was run through a Gibbs sampler Markov Chain Monte Carlo (MCMC) algorithm in the (Plummer 2022) package to approximate the posterior distributions of the estimates of all fixed effects and the variance of the random effect. 3 chains of 60000 total iterations with a 10000-iteration burn-in period; starting values were obtained from the estimates generated by Model (1), with random noise added to avoid false convergence.

Estimates for the final analysis:

4 Limitations and Alternatives

Presence						
GLMM Est	MCMC Mean	MCMC Median	MCMC Mode	MCMC SD	HPD Lower	HPD Upper
-2.70	-2.73	-2.72	-2.68	0.27	-3.25	-2.20
0.21	0.20	0.20	0.19	0.38	-0.53	0.95
0.50	0.51	0.51	0.49	0.33	-0.14	1.15
0.89	0.90	0.90	0.90	0.30	0.31	1.50

com-
pet-
ing
risks
in
the
ex-
per-
i-
men-

tal design obscure the presence of family clusters of cataracts. It is reasonable to assume that individuals in the irradiated Treatment groups were more likely to die due to more acute effects of irradiation before developing cataracts, compared to individuals in the control group. Simplifying the dataset by instituting a cutoff age attempts to address this limitation, but the issue persists even after excluding individuals deceased before Age = 552.

5 Results

The final probability for developing cataracts can be written as:

$$p_i = 0.063 + 0.55 * Gamma_i + 0.623 * HZE_i + 0.708 * Male_i + 0.726 * Gamma_i * Male_i + 0.770 * HZE_i * Male_i \quad (3)$$

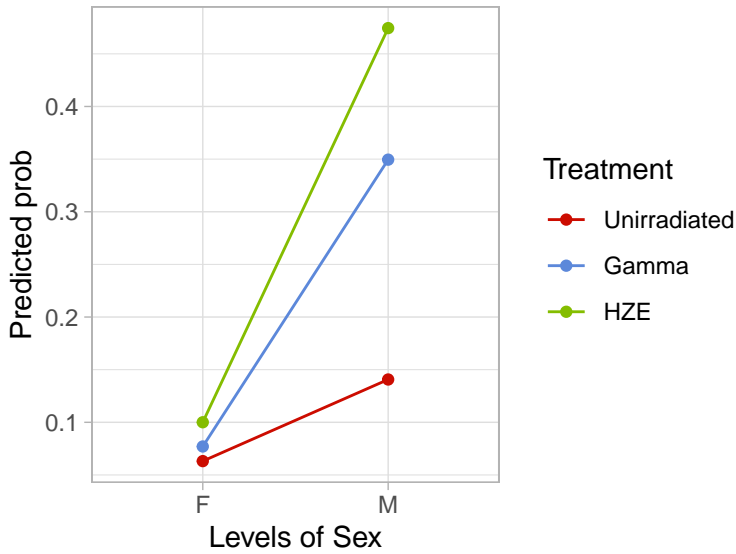


Figure 3: Predicted probability of cataracts by sex, treatment group

Figure 3 shows the probability of developing cataracts for each combination of sex by treatment group. Differences between both sex and treatment group are clearly visible, notably the differences between treatment group across gender.

Assessing relative risk for cataracts between groups:

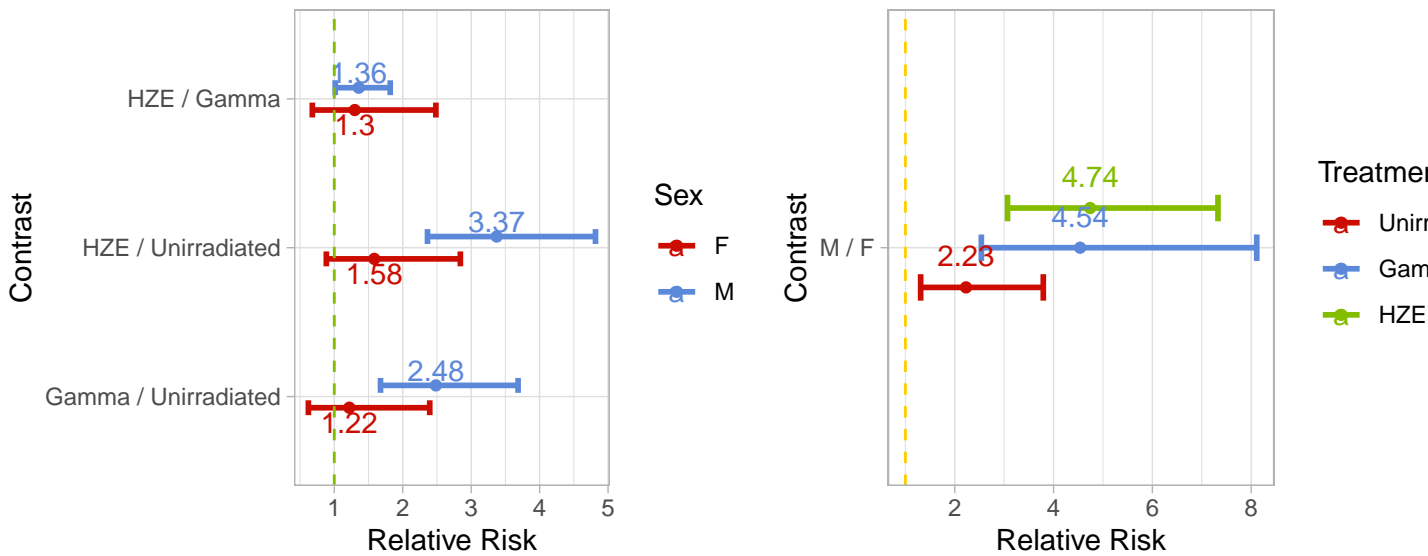


Figure ?? shows the relative risk of cataract development for all groups in Model (1).

add interpretation

Visualizing random effect by family:

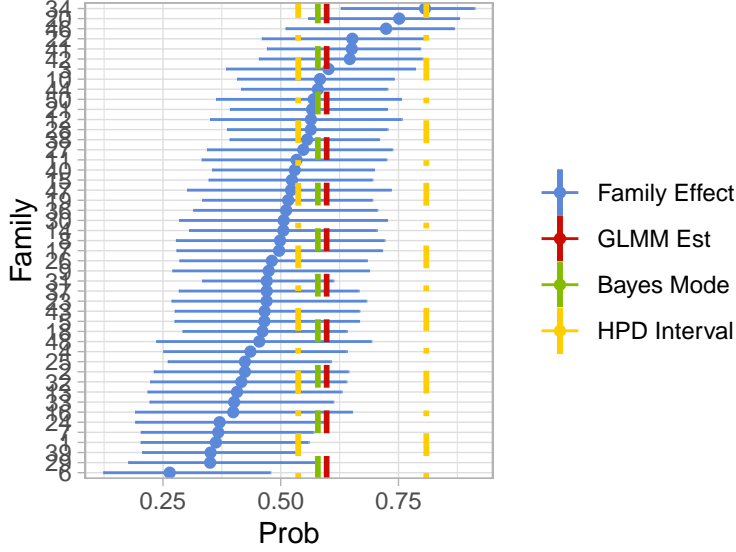


Figure @ref(fig:re_plot) visualizes the estimated probability of developing cataracts for each family in the dataset, along with their confidence intervals. The vertical lines plot the GLMM estimated variance from Model (1) and the MCMC mode and HPD interval for σ^2 from Model (2). Many of the probabilities shown are close to 0.5, with small numbers of families at the top and bottom of the graph that indicate some respective genetic predisposition or resistance to cataract development.

6 Conclusions

Turn this into a nice paragraph or two; add quantitative info from results to back these up, or keep it general?

Females in general face less probability of developing cataracts than males, across all treatment groups. We could argue that the effect of treatment group is not of clinical significance for female mice in this dataset. As you can see, the probability of developing cataracts is quite low for all 3 treatment groups. Males, on the other hand, have a higher probability of developing cataracts than females across all treatment groups. For males, the differences of development of cataracts across treatment groups is of clinical importance.

7 Author’s Statements

Both authors contributed to statistical analysis and writing for this report. Both authors read and approved the final report. Both authors contributed equally to introduction, exploratory data analysis, summary statistics, and results. Alyssa Allsop served as the primary statistician on fitting Frequentist models; Amira Burns served as the primary statistician on fitting the Bayesian model.

8 Appendix

Model	AIC
Full Mixed Model	1017.98
Final Model	1038.9
Fixed Effects Model	1040.3
Mixed Model with no Interaction	1023.19
Base Model	1123.57

Table 3: Model selection via AIC comparison

Sample citations:

(Bates et al. 2022)

(Plummer 2022) (Lenth 2022)

References

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2022), *lme4: Linear Mixed-Effects Models using Eigen and S4*. R package version 1.1-29.

URL: <https://github.com/lme4/lme4/>

Chernyavskiy, P., Edmondson, E. F., Weil, M. M. & Little, M. P. (2017), ‘High-energy particle beam and gamma radiation exposure, familial relatedness and cancer in mice’,

British journal of cancer **117**(1), 41–50.

URL: <https://www.nature.com/articles/bjc2017141>

Lenth, R. V. (2022), *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.7.4-1.

URL: <https://github.com/rvlenth/emmeans>

Plummer, M. (2022), *rjags: Bayesian Graphical Models using MCMC*. R package version 4-13.

URL: <https://mcmc-jags.sourceforge.io>