

Examining Serum Levels of Zinc and Copper in Individuals with Chronic Problematic Tinnitus

Reference

This is an original study conducted by Valerie Ingalls, as well as Dr. Ishan Bhatt, PhD, member and PI, respectively, of the Audiogenomics Lab, part of the Department of Communication Sciences and Disorders at the University of Iowa. Data for the study are taken from the National Health and Nutrition Examination Survey (NHANES), which has its data published publically available at their website.

Introduction

Tinnitus, or the phantom perception of sound in absence of external stimulus, which is commonly but not exclusively experienced as ringing in the ears, affects an estimated 10% of US adults each year. This high prevalence makes it a very common disorder. The rate increases with age, with nearly 5 million total individuals in their 50's experiencing tinnitus during a given year. There are numerous differences between tinnitus phenotypes that individuals experience. Tinnitus may manifest only in the short term, which is known as acute tinnitus, or it may be experienced for multiple years, which is known as chronic tinnitus. Even among chronic tinnitus patients, some people perceive tinnitus consistently, while for others it is only present or discernible intermittently. A significant portion of individuals with tinnitus report that it presents at least a moderate problem in their life, and over 1 million U.S. adults report it is a big problem. As a common disorder with a notable negative impact on those who have it, the importance of treatment for tinnitus is clear. Those with chronic, problematic tinnitus represent the population of greatest clinical interest.

While noise exposure is known to be the primary causal trigger for the manifestation of tinnitus, the exact biological origins of the disorder yet remain unclear. One leading theory is that tinnitus is connected to damage within the cochlea, the primary structure of the inner ear that is responsible for transduction of sound into an electrical signal before it passed to the brain. Cochlear synaptopathy due to glutamatergic excitotoxicity is one known mechanism of cochlear damage that is triggered by noise exposure. Overstimulation at synapses within the cochlea has demonstrated to lead to large-scale loss of synapses in animal models, and while the story appears to be more complicated in humans, it remains an avenue worth exploring further. Zinc and copper are elements able to neutralize the superoxide radical, thereby preventing some amount of long-term cochlear damage. It is theorized that these metals may have a protective effect against noise-induced cochlear damage, and therefore tinnitus. It follows that their deficiency may result in susceptibility to the same. This provides one potential explanation for individual differences in tinnitus susceptibility.

Previous studies examining this hypothesis have showed some small promise, but they have largely been underpowered and their results lacking in statistical significance. Further, several previous studies that attempted to use zinc for treating tinnitus neglected to examine pre-existing differences in zinc levels and rates of deficiency between individuals with and without tinnitus. Using epidemiological data, we can easily rectify the sample size issue. With this project, we aim to complete a large-scale epidemiological analysis of differences in the levels of zinc and copper in individuals with and without tinnitus, as well as the rates of deficiency, in order to determine whether there are population-level variations that could explain tinnitus susceptibility.

Example Figure

As this is an original study, I do not have a figure to replicate. Additionally, the previously mentioned studies did not use the type of visualization that I am planning to use, because their sample sizes were too small

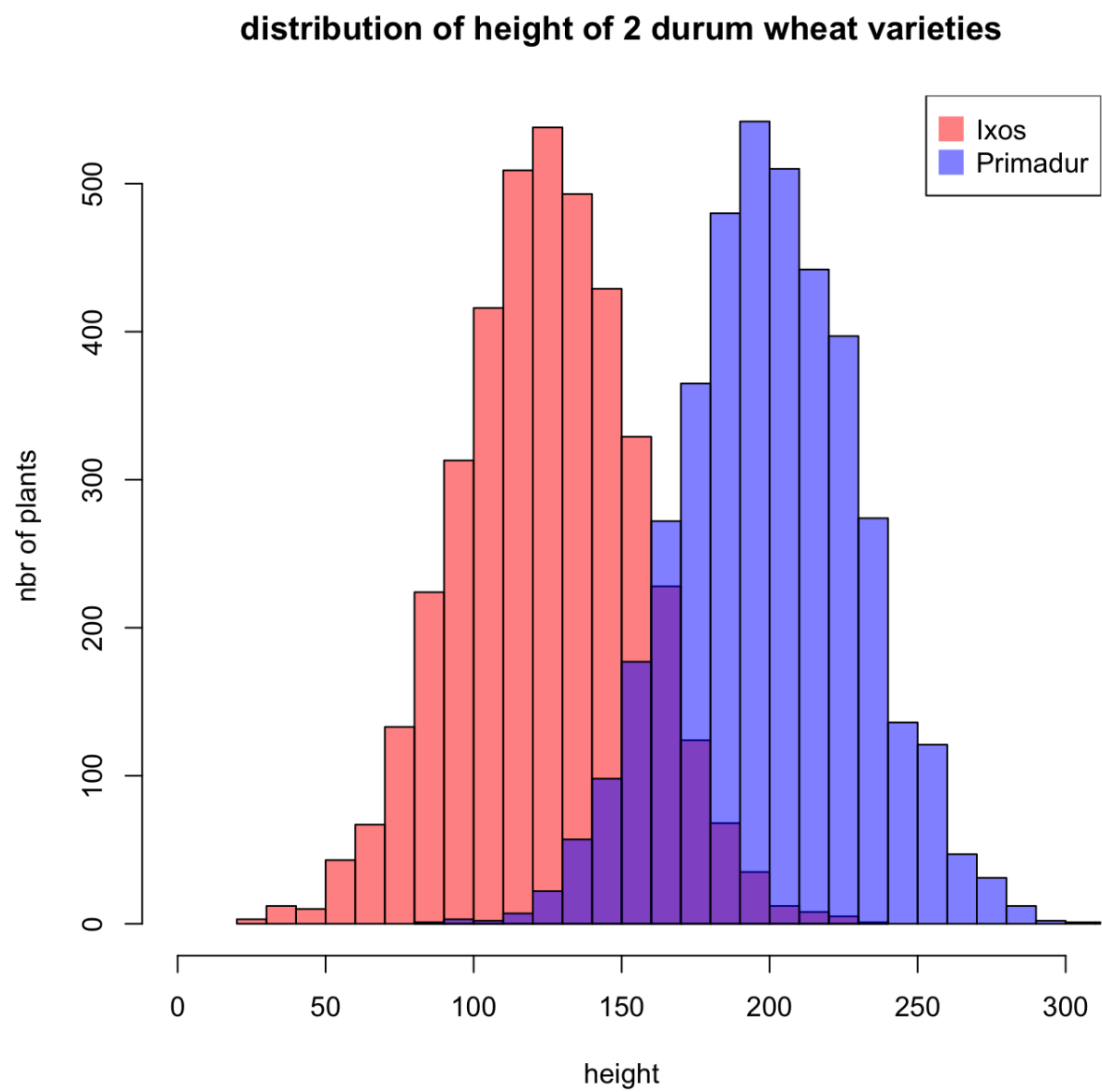


Figure 1: An example histogram

for a histogram to be an effective visualization. As such, I have pulled a somewhat similar example and will describe the differences that I plan for my final version.

The goal is to create a plot with multiple distributions on the same axis. I will be making two figures with the same underlying design, the only difference being one figure will be copper and the other for zinc. Due to differences in normative serum levels of copper and zinc, it is impractical to measure both metals on the same figure. On the x-axis for each figure will be level of serum copper (or zinc), while the y-axis will be density (preferable to a count y-axis in this case because our weighted sample size is in the millions). One of the distributions will be for the no tinnitus control group, one distribution will be for individuals with chronic problematic tinnitus, while a third distribution will be for individuals with other tinnitus phenotypes (either acute or unproblematic). I will also add a vertical line at the clinical cutoff value for deficiency in those metals. In this way, I will visualize the differences in zinc and/or copper in those with chronic problematic tinnitus compared to the control groups of those with no tinnitus and those with clinically non-significant tinnitus, while simultaneously showing the proportion of individuals in each population who have a deficiency in that metal.

In the case where there are minimal differences between groups, we will also generate box plots for each metal in order to easily visualize that similarity.

Materials and Methods

This is an epidemiological study using data from the National Health and Nutrition Examination Survey (NHANES).

NHANES Data Collection

The NHANES is a nationally representative health survey that is unique in that it makes use of both questionnaires and physical examinations to gather a robust set of health data. It began several decades ago with a series of multi-year surveys, the NHANES I, II, and III. The modern NHANES, known as the Continuous NHANES, is composed of a series of survey cycles, each lasting for two years. Individual cycles are constructed to be fully representative of the non-incarcerated United States population. Researchers are encouraged to combine multiple cycles in order to increase the robustness of their dataset. The NHANES utilizes a complex stratified survey design in order to increase its representativeness. Subpopulations of particular public health interest are first identified (eg. non-white Hispanic individuals). The full population is broken up into sampling strata based on location, often but not always by county. These strata are assigned a weight that increases their likelihood of being randomly selected for sampling based on the proportion of their population that includes the targeted subpopulation(s). Some strata are weighted so heavily as to be guaranteed for selection. Once these weights are assigned, weighted randomized selection of the strata occurs. Within each stratum, households and finally individuals are randomly selected. Weights and strata information are then included in the NHANES data so that it is possible to accurately extrapolate results to the population level.

Data Selection and Processing

Data from the NHANES are stored in .XPT format and are accessible via the website. In order to preserve metadata and avoid data corruption in the present study, we acquired data via CLI using the `wget` command in a bash shell. Two cycles, 2011-2012 and 2015-2016 were selected for this study, as these cycles are only ones that have full audiological datasets available alongside serum metal levels.

NHANES stores data separately by cycle and by subject matter. In the case of this study, we used data from a total of eight different files, four from each cycle. All data wrangling and analysis was performed in RStudio

(RStudio 2023.03.0+386 “Cherry Blossom” Release (3c53477afb13ab959aeb5b34df1f10c237b256c3, 2023-03-09) for Windows Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) RStudio/2023.03.0+386 Chrome/108.0.5359.179 Electron/22.0.3 Safari/537.36) under R version 4.2.1. In addition to base R functions, we used several packages for this work: `foreign`, `tidyverse`, `bestNormalize`, `ggplot2`, `moments`, and `survey`.

The first step in our data wrangling process was to combine the eight initial data sheets into one master data frame that could then be properly worked with. Many variables of interest were obtained directly from the raw data, while others needed to be derived or recoded. Ethnicity refactored into two groups (White and Non-White/Hispanic), a method commonly used in tinnitus studies that is generally true to prevalence. We derived a factor variable with 3 levels for tinnitus type: no tinnitus, chronic bothersome tinnitus, and other tinnitus. “Other tinnitus” includes those individuals who have experienced tinnitus within the last 12 months, but either **(a)** have not experienced it for at least a total of 1 year **(b)** do not report their tinnitus causing any problem in their life or **(c)** only experience tinnitus as a temporary response to exposure to loud noise. Deficiency was also derived for both zinc and copper levels, using pre-defined clinical cutoff values. Zinc cutoff values controlled for sex as well as whether or not the individual had fasted before the blood draw. No such nuanced thresholds were found for copper, so a global cutoff value was used instead. Finally, the sample weights are halved for each individual, as is the standard NHANES protocol when combining two cycles together. For this study, we used the laboratory sub sample weights as our serum metal measurements come from this sub sample.

Lab Data

Serum metal levels were obtained as part of the laboratory examination portion of the NHANES. Approximately one third of sampled individuals participate in this portion of the survey. A portion of these individuals are instructed to fast before their blood sample is taken. For details on how the samples are obtained, see the full lab methods provided on the NHANES website [here](#).

Analysis

The NHANES provides survey design details that allow us to extrapolate our results to the entire US population. In order to make use of the complex survey design in our analysis, we need to use the `survey` package in R. This package contains functions for regression modeling, t tests, chisq tests, and various forms of graphing, along with other analysis tools. With that package, we can implement the survey design in both our histogram figure as well as any statistical analyses.

For full comparison of tinnitus groups, we will perform several types of statistical analyses. First, in order to examine stratification within our sample population, we will perform chi square goodness of fit tests, examining differences in sex and ethnicity across the tinnitus types. Secondly, we will perform two multiple linear regression analyses, one with serum zinc level as the response variable, and one with serum copper level as the response variable, with sex, age, ethnicity, speech-in-noise perception, and tinnitus type as the predictor variables. This will allow us to examine the effect of tinnitus type on serum metal levels while controlling for these other demographic variables. SIN perception is included as a related phenotype of some tangential interest. Finally, we will perform two logistic regression analyses with the same predictor variables, while the response variables will be deficiency in the metals, as defined by pre-established clinical values. This allows us to examine not just the discrete quantity of serum zinc and copper, but also the rates of deficiency across the different tinnitus groups.

Figure Production

Figures for this study were generated using a mixture of the `ggplot2` and `survey` packages in RMarkdown, with minor aesthetic changes performed in MS Paint 3D. For our density plots, the primary density plots are created with the `geom_density()` function. `geom_vline()` was also used to show the deficiency cutoff point. The `fill` aesthetic performs the grouping by tinnitus types, while the `weight` aesthetic allows us to implement the weighting from the complex survey design into these plots. Further aesthetics specify details that make the graph more visually appealing. A code snippet for one of the graphs is included below for easy reference:

```
ggplot(data = df) +  
  geom_density(mapping = aes(x = zinc, weight = weight, fill = tinnitus_type), alpha = .5) +  
  geom_vline(xintercept = 61, color = "blue", linetype = 2) +  
  labs(x = "Serum Zinc Level (\u03BCg/dL)", y = "Density",  
       title = "Distribution of Serum Zinc Levels by Tinnitus Type", fill = "Tinnitus Type")
```

Boxplots were created using the `svyboxplot()` function from the `survey` package rather than `ggplot2` in order to fully implement the survey weighting in these plots. This function includes an argument for the survey design object that we created for use in our statistical analyses. Sample code is provided here for those plots as well:

```
svyboxplot(normedZinc~tinnitus_type, design = tinnitus_design,  
xlab = "Tinnitus Type", ylab = expression("Serum Zinc Level ( " * sqrt(mu *g/dL) * ")"),  
main = "Serum Zinc Levels by Tinnitus Type", all.outliers = TRUE, col = "#00beff", boxwex = .75)
```

Results

Sample

A total of 2841 individuals (age range: 20-69, mean: 43.88) with complete data were collected for our study sample. Of these individuals, 175 reported chronic bothersome tinnitus, 228 reported other tinnitus, and 2438 reported no tinnitus, for an unweighted overall tinnitus prevalence of 14.2%. 1404 individuals were male, while 1437 individuals were female. 885 individuals reported their ethnicity as White, while 1956 individuals reported a different ethnicity. This imbalance is expected from the targeted oversampling of the NHANES and is accounted for in the implementation of the survey design.

Statistical Analyses

Population Stratification Chi square goodness of fit tests with Rao & Scott adjustment were performed to examine stratification of tinnitus types across sex and ethnicity. No significant differences were found between males and females ($F = 0.22$, $p = 0.8$). Differences were found between White and Non-White/Hispanic individuals ($F = 12.26$, $p < 0.001$), which was anticipated.

Skew and Data Transformation Preliminary analysis of the data revealed right skew in the distributions for both serum zinc levels and serum copper levels, primarily due to the presence of several extreme observations. Therefore, we opted to perform a transformation on these values in order to pull in extreme observations and to improve our data's conformation to the normality assumption for our regression analyses. In order to select a transformation, we made use of the `bestNormalize()` function from the `bestNormalize`

package. This tool runs a series of transformations and provides a skewness value for the results of each. This makes it straightforward to make an informed decision for the choice of transformation. Through this method, we selected a square root transformation for serum zinc levels, and a Yeo-Johnson transformation with $\lambda = -0.519$ for serum copper levels. The Yeo-Johnson transformation is a power transformation and a modified version of the Box-Cox transformation.

Zinc Multiple linear regression analysis with square root serum zinc as the outcome variable revealed correlations with sex ($p < 0.001$) and ethnicity ($p < 0.01$). No significant relationship was found between square root serum zinc and tinnitus type or speech in noise perception ($p > 0.05$).

Logistic regression with zinc deficiency as the outcome showed a significant correlation between zinc deficiency and speech in noise perception ($B = -0.15$, $p = 0.02$), but no relationship with tinnitus type.

Copper Multiple regression with Yeo-Johnson transformed copper as the outcome variable showed significant correlations with ethnicity ($p = 0.015$) and sex ($p < 0.001$) but not with SIN perception ($p = .18$) or tinnitus type.

Logistic regression with copper deficiency as the outcome variable showed a significant correlation with age ($p < 0.01$). Additionally, individuals with other forms of tinnitus showed significantly lower rates of copper deficiency than individuals with chronic bothersome tinnitus ($B = -16.57$, $p < 0.001$).

Figures

We generated a total of four figures for this study. The goal was to visualize the distributions of the levels of serum zinc and copper across different tinnitus groups. In order to accomplish this, we started with density plots, with a different color representing each group and a vertical line indicating a cutoff value for deficiency in that metal. These plots are similar to the example that is included earlier in this report. One notable difference is that we opted to use a density plot rather than a histogram. This is because our variable of interest, serum metal level, is truly continuous, and a binned histogram would not fully capture that level of nuance in the data. While a small bin width could show a similar shape, it is both simpler and more true to the data to use a density plot. These density plots capture the lack of difference between the distributions of our groups, but they do so with some ambiguity. Therefore, we opted to also include box and whisker plots. For these plots, we opted to use the transformed values for zinc and copper that were also used for all of our statistical analyses. These figures clearly show that, with the exception of some outliers, the distributions of zinc and copper are practically identical across tinnitus groups. This agrees with our statistical analyses.

Discussion and Conclusions

Our analyses show no evidence for a correlation between serum zinc or serum copper level and tinnitus phenotype. However, we did find a few interesting correlations with deficiency. Copper specifically showed a clear difference in rate of deficiency between types of tinnitus, although not between chronic tinnitus and no tinnitus. This perhaps suggests that copper deficiency does not play a role in if tinnitus manifests, but that it may be related to how tinnitus manifests. Zinc deficiency showed correlation with speech-in-noise perceptual difficulties, but not tinnitus type. Notably, however, the direction of correlation indicates that those with SIN perceptual difficulties were the ones with lower deficiency rates. This runs directly counter to our hypothesis regarding a potential protective effect of zinc on the cochlea.

Overall, it appears that serum levels of copper and zinc have little relationship with tinnitus prevalence in the US adult population. Further study is warranted to examine the effects that copper may have on

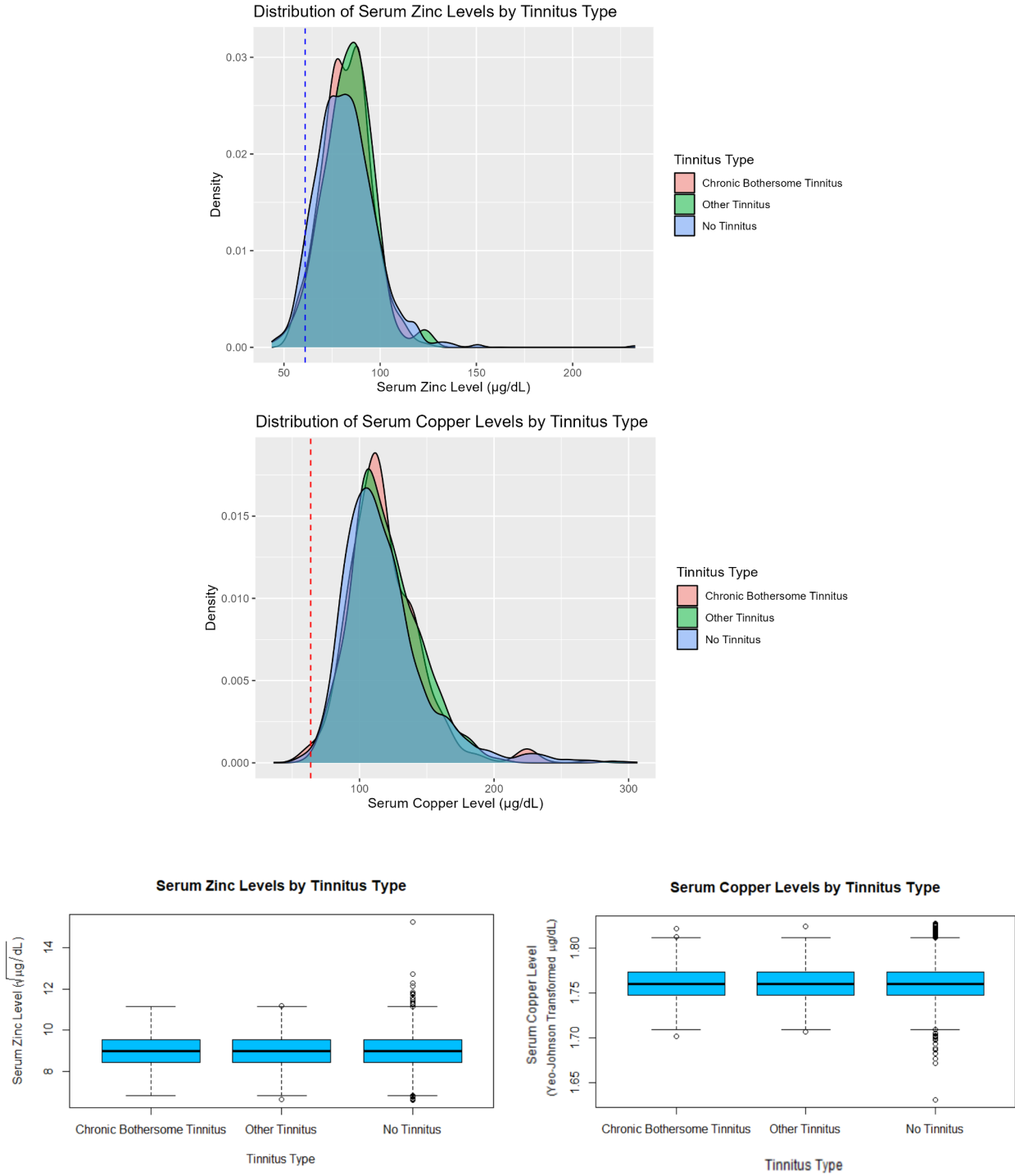


Figure 2: Density plots and box plots for the distributions of serum zinc and copper. Density plots contain a vertical line indicating a threshold for clinical deficiency. Box plots used transformed values for zinc and copper. Both sets of graphs show very little difference in serum metal levels between the different tinnitus groups.

tinnitus presentation, with our findings indicating a potential modulation of severity or duration that can be achieved through avoiding deficiency.

Reflection

One of the chief challenges that I discovered during this project concerning reproducibility was in making my code clear for others! Originally, the code I had written was very “one and done”, and meant only for my eyes. Throughout the course of this project, I spent a lot of time going back into my code and providing more and more detailed comments, scrubbing out obsolete parts leftover from previous iterations of the project, and so on. I still think there are probably some improvements that could be made, but the result is pretty good. Most of all, it kept me honest, because I felt obligated to make edits that might not even affect the final product, just because I knew it would be clearer. Two elements of reproducibility that I actually really liked were using GitHub and having a clear folder directory. I work on two different computers, so Git was a great tool for version syncing across devices. Having a clean folder structure was also very convenient any time I needed to grab a file, for example in my R analysis. Knowing my working directory well, and it being easy to navigate made that a cinch. Normally I find getting the file path right to be a pain in R but I had no issues because everything was set up well.

Another challenge that I faced was not allowing the scope of my project to expand too much. I added another tinnitus group partway through on the project, which created changes in my statistical analyses, which created more work, etc. I also spent multiple hours trying to get multinomial logistic regression to work, only to eventually decide there was no good way to do it within the complex survey design. ChatGPT also hallucinated many solutions to that problem. Definitely a frustrating moment! However, I’m overall quite pleased with how my project turned out. I found an interesting if small result, and even a null result feels valuable for this research question. I think that my graphs turned out pretty well, and making clean, commented code, while a pain, is something that I know is valuable not just to you guys but to my future self! After all, I wouldn’t have had to make so many edits if my code had been good and clear the first time around.

Finally, I think the biggest challenge for this project was this very README file! Honestly it’s a pain to do text-editing on a markdown file compared to a word processor, and there are some elements of it that I still don’t completely get. For example, I can confidently say that the figures are not going to land where I want them to be, and the size may be questionable. Sorry about that! However, it’s finished, and that was the goal here!