SWISS TPH Assignment

Ekerette Emmanuel Udoh

2023-02-20

ABSTRACT

Background

Although optimal exclusive breastfeeding for 6 months is recommended, challenges in provision may lead to reduced infant nutrition and hemoglobin levels causing malnutrition. The objective of this study was to investigate the relationship between exclusive breastfeeding duration and child hemoglobin levels.

Methods

This was an analysis of the SWIS cross-sectional survey data of 20 mother-child pair. Interviews with mothers and infant clinical characteristics were obtained at approximately 6 months postpartum. Mother and child demographic, and child hemoglobin variables were measured. Descriptive statistics were used to present all variables, while bivariate and multivariate linear regression were employed to model the association between maternal and child characteristics with hemoglobin levels.

Result

Average infant age was 5.81 (SD: 0.24) months, and 10/19 (53%) were female. Exclusive breastfeeding duration was 3.04 (SD: 2.09), and all the infants (19/19) received high quality dietary meals. Mean infant hemoglobin level was 89.21 (SD:17.26). Exclusive breastfeeding duration was not a predictor of infant hemoglobin levels, although high exclusive breastfeeding levels indicated increasing hemoglobin levels mainly among girls. All maternal and child demographic variables were not significant predictors of hemoglobin levels.

Conclusion

Exclusive breastfeeding is not a significant predictor of infant hemoglobin levels. Nonetheless, promoting exclusive breastfeeding and providing high quality dietary meals to infants should remain a public health priority as they have important health benefits for both the mother and child.

Word count: 216

Analysis plan

1. Descriptive Statistics:

- Child and maternal demographic and clinical characteristics will be summarized using descriptive statistics.
- Frequencies and percentages will be calculated for categorical variables such as gender, age, and education level.
- Measures of central tendency such as mean and median will be calculated for continuous variables such as hemoglobin levels and months of exclusive breastfeeding.

2. Linear Regression Analysis:

- A linear regression model will be fitted to examine the relationship between months of exclusive breastfeeding and hemoglobin levels.
- Bivariate linear regression analysis will be conducted to test the association between other predictor variables and hemoglobin levels.
- Multivariate linear regression analysis will be used to control for the independent effects of predictor variables of the breastfeeding levels.

3. Statistical Significance:

- The threshold for statistical significance will be set at $p<0.05$.
- The statistical significance of all associations will be assessed using this threshold.

4. Reporting Results:

- Tables and graphs will be used to present the results visually.

The following codes show the packages used and how data was cleaned for analysis. This document with the codes for the task and the analysis done can be found on GIthub with the link.

```
# Packages used for data wrangling and analysis
library(readxl)
library(janitor)
library(tidyverse)
library(gtsummary)
```

```
# recode education
mat_data <- mat_data %>%
mutate (education_fac =
          as.factor(recode(education,
          "1" = "College or higher",
          "2" = "High school graduate",
          "3" = "Less than high school")))

# recode diet
mat_data <- mat_data %>%
mutate (diet_fac =
          as.factor(recode(diet, "1" = "High", "2" = "Middle", "3" = "Low")))

# recode ses
mat_data <- mat_data %>%
mutate (ses_fac =
          as.factor(recode(ses, "1" = "Low", "2" = "Middle", "3" = "High")))



# recode sex
mat_data <- mat_data %>%
mutate (sex_fac =
          as.factor(recode(sex, "m" = "Male", "f" = "Female")))

# Reclass age
mat_data <- mat_data %>%
mutate (age = as.numeric(age))


# Create dummy varaibles for linear reg models

mat_data <- mat_data %>%
mutate (education_dummy = ifelse(education_fac == 'College or higher', 1, 0))

mat_data <- mat_data %>%
mutate (ses_fac_dummy = ifelse(ses_fac == 'High', 1, 0))

mat_data <- mat_data %>%
mutate (ses_fac_dummy = ifelse(ses_fac == 'High', 1, 0))
```

Codes and Results

Demographic characteristics of mother and child.

```
demo_char <- mat_data %>%
            select(age, sex_fac,    education_fac, ses_fac)

tbl_summary(demo_char,
            digits = list(age ~ 2),
            statistic = list(age ~ "{mean} ({sd})"),
            type = list(age ~ "continuous"),
    missing ="no",
      label = c(age ~ "Infant age, months",
              sex_fac ~ "Gender assigned at birth",
              education_fac ~ "Mothers highest education",
              ses_fac ~ "Household secio-economic status"))%>%
      italicize_levels() %>%
    bold_labels() %>%
     as_gt() %>%
   gt::tab_header("Table 1. Maternal and child demographic Characteristics")
```

Table 1. Maternal and child demographic Characteristics

Characteristic	**N = 20**[1]
Infant age, months	5.81 (0.24)
Gender assigned at birth	
Female	10 (53%)
Male	9 (47%)
Mothers highest education	
College or higher	7 (39%)
High school graduate	7 (39%)
Less than high school	4 (22%)
Household secio-economic status	
High	6 (32%)
Low	6 (32%)
Middle	7 (37%)

[1]Mean (SD); n (%)

Child nutritional and clinical characteristics

```r
nut_char <- mat_data %>%
            select(ebf, diet_fac, hemo)

tbl_summary(nut_char,
    digits = list(ebf ~ 2, hemo ~ 2),
    statistic = list(ebf ~ "{mean} ({sd})",
            hemo ~ "{mean} ({sd})"),
            type = list(diet_fac ~ "categorical",
            hemo ~ "continuous",
            ebf ~ "continuous"),
    missing ="no",
    label = c(ebf ~ "Exclusive breastfeeding duration, months",
              diet_fac ~ "Dietary quality",
              hemo ~ "Hemoglobin levels (g/L)")) %>%
     italicize_levels() %>%
    bold_labels() %>%
       as_gt() %>%
  gt::tab_header("Table 2. Child nutritional and clinical characteristics")
```

Table 2. Child nutritional and clinical characteristics

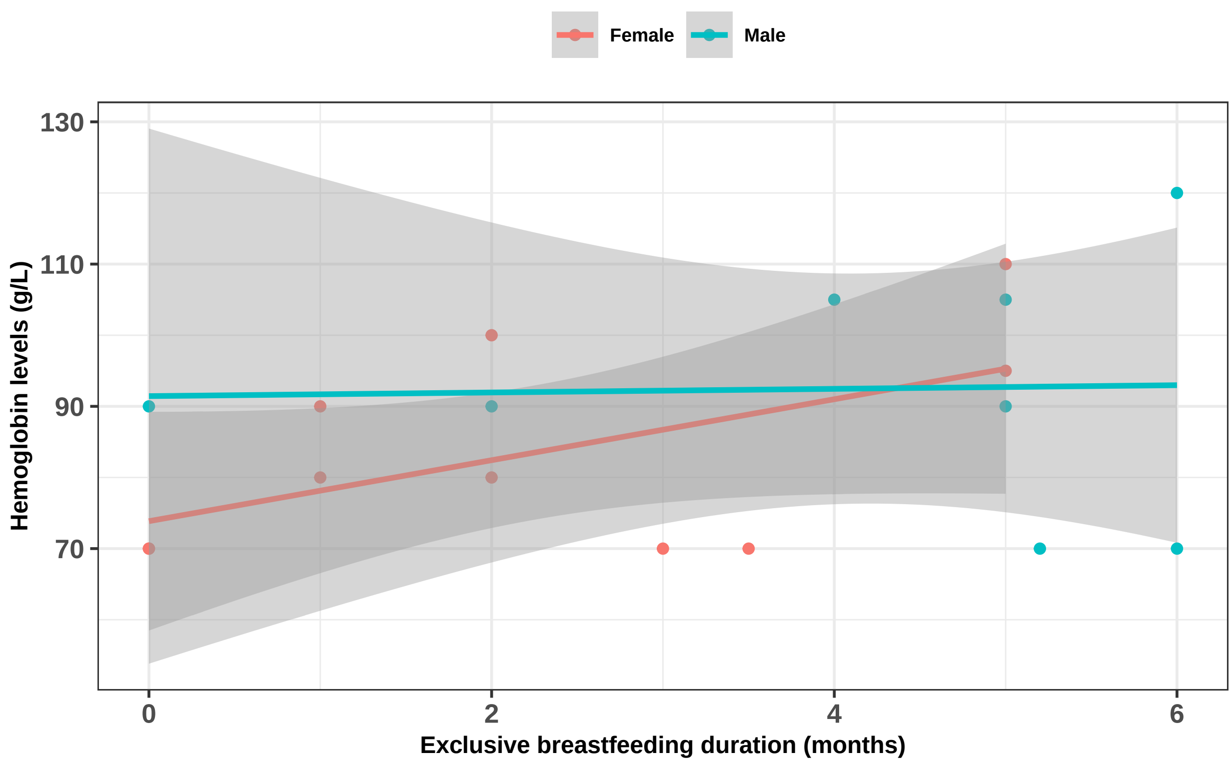
Characteristic	**N = 20**[1]
Exclusive breastfeeding duration, months	3.04 (2.09)
Dietary quality	
High	19 (100%)
Hemoglobin levels (g/L)	89.21 (17.26)

[1]Mean (SD); n (%)

Association between exclusive breastfeeding and hemoglobin levels

```
mat_data <- na.omit(mat_data)
# Plot linear regression



# Determine p-value
p_value <- round(cor.test(mat_data$ebf, mat_data$hemo)$p.value , 2)

# Determine R2
correlation <- cor(mat_data$ebf, mat_data$ebf)

r_squared <- correlation^2


ggplot(data = mat_data, aes(x = ebf, y = hemo, colour = sex_fac)) +
  geom_point()+
    geom_smooth(method = "lm")+
    ggtitle("Fig. 1. Linear reltionship between exclusive breasfeeding duration and
infant hemoglobin levels")+
   xlab("Exclusive breastfeeding duration (months)") + ylab("Hemoglobin levels (g/L)")+
  theme_bw()+
  theme(axis.text = element_text(size = 10),
        text = element_text(size = 9, face = "bold"),
        legend.title = element_blank(),
        legend.position = "top")
```

Fig. 1. Linear reltionship between exclusive breasfeeding duration and infant hemoglobin levels



Bivariate and multivariate linear models for relationship with hemoglobin levels

```
# Univariate model

hemo_model_uni <- mat_data %>%
  select(ebf,age, sex_fac, education_fac, ses_fac, hemo) %>%
tbl_uvregression(
    method = lm,
    y = hemo,
    hide_n = TRUE,
    label =  c(ebf ~ "Exclusive breastfeeding duration, months",
              age ~ "Infant age, months",
              sex_fac ~ "Gender assigned at birth",
              education_fac ~ "Mothers highest education",
              ses_fac ~ "Household secio-economic status"))%>%
   add_global_p()

#  multivariate model

hemo_model_multi <- lm(hemo ~ ebf + age + sex_fac + education_fac + ses_fac,
                       data = mat_data)

hemo_modelmulti_tab <- tbl_regression(hemo_model_multi,
    label =  c(ebf ~ "Exclusive breastfeeding duration, months",
              age ~ "Infant age, months",
              sex_fac ~ "Gender assigned at birth",
              education_fac ~ "Mothers highest education",
              ses_fac ~ "Household secio-economic status"))

# Merged the univariate models and multivariate model in a single table
tbl_merge(
    tbls = list(hemo_model_uni, hemo_modelmulti_tab),
    tab_spanner = c("**Unadjusted model**", "**Adjusted model**")) %>%
      bold_labels() %>%
    italicize_levels() %>%
      as_gt() %>%
  gt::tab_header("Table 3. Predictors of hemoglobin levels")
```

Table 3. Predictors of hemoglobin levels

	Unadjusted model			**Adjusted model**		
Characteristic	**Beta**	**95% CI**[1]	**p-value**	**Beta**	**95% CI**[1]	**p-value**
Exclusive breastfeeding duration, months	2.8	-0.92, 6.4	0.13	3.6	-3.1, 10	0.3
Infant age, months	9.2	-27, 46	0.6	-1.1	-60, 57	>0.9
Gender assigned at birth			0.2			
Female	—	—		—	—	
Male	9.0	-6.9, 25		5.3	-20, 30	0.6
Mothers highest education			0.5			
College or higher	—	—		—	—	
High school graduate	-10	-29, 8.6		-2.8	-26, 21	0.8
Less than high school	-3.4	-25, 18		7.8	-26, 41	0.6

Household secio-economic status			0.7			
High	—	—		—	—	
Low	-3.7	-25, 18		12	-22, 46	0.5
Middle	-8.4	-29, 12		-3.0	-27, 21	0.8

[1] CI = Confidence Interval

Reflection on the assignment

Working on this assignment was interesting and it was an important learning experience, especially having to refresh my knowledge on a few of the analysis I have not done recently in my current work. The assignment required me to demonstrate analytical techniques to write codes in R to carry out data analysis. The assignment requires the knowledge of statistical concepts and skills in cleaning and transforming the data, performing exploratory data analysis, and conducting statistical tests.

While working on the assignment, I encountered a few challenges, but none that could not be overcome given the level of data wrangling and analysis required for this task. One of the common challenges I faced was recalling certain codes correctly, which required me to reference the package documentation. Additionally, I had to ensure that the analysis plan and results accurately reflected the data set and addressed the research objective. I had no challenge writing the abstract because I have the knowledge and skills to do this from the experience I have gained writing many scientific papers, and from the training I have attended on scientific writing.

Overall, this exercise provided a valuable opportunity to showcase my analytical capabilities and problem-solving skills. While it required significant effort and attention to detail, the exercise was a worthwhile exercise that allowed me to apply my skills in a practical setting.