MATH552 Group Project

Group 12

35427962 & 35762970

November 12, 2020

Congye Wang (35427962) and Kieran Molloy (35762970) have contributed equally to the whole project and agree to share the mark equally

The Influence of Multiple Variables on High-Risk Neonatal Survivors Based on the Quasi-Binomial Distribution GLM Model

Wang, Congye

c.wang35@lancaster.ac.uk

Molloy, Kieran

k.molloy@lancaster.ac.uk

November 12, 2020

Abstract

Insert Abstract Here

1 Introduction

High rates of infant morbidity arising from low birth weight impose a significant burden on families, in addition to health, education ans social services, as shown in Boyle et al. 1983's landmark journal which found that for infants of $< 1500g$ admitted to neonatal intensive care units, the long-term costs outweighed the measurable economic benefits. As reported in Salt 2001 and described further in Petrou et al. 2001, low birth weight is increasing in industrialised nations. In the UK, it increased from 0.8% to 1.2% from 1983 to 1998, similar trends are noticed in other industrialised nations. Additionally, advances in medicine and medical practices such as increased use of assisted ventilation during delivery and various disinfectant methods have led to increasing premature births surviving and with that, low birth weights increasing. Trends such as these are only likely to increase, with medical advances, leading to an increased requirement of specialist services for high risk - low birth weight infants requiring further support.

The purpose of this report is to quantify the impact of Community Neonatal Services (CNS) on mother and child satisfaction compared to centres without CNS. In Langley et al. 2002 service was measured by a number of responses such as:

- Length of stay in hospital from birth to initial discharge

- Number of hospital readmissions required by these infants

- Support and equality of care provided to the family in the first year of life after discharge

- Overall costs of treatment in the first year of life

Their report focused on objectives 1 and 2, as objective 3 is a well studied problem, and objective 4 is difficult to measure with insufficient data.
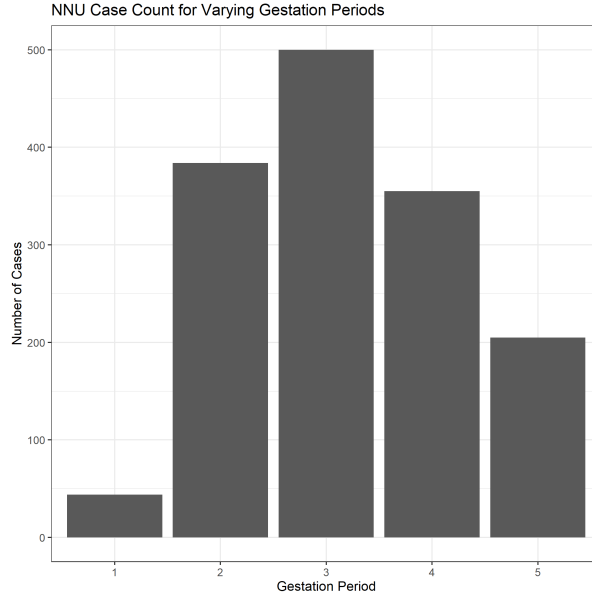

Figure 1: Gestation

Investigating the impact of on high risk neonatal survivors within the first year of life, based on data-sets derived from Langley et al. 2002. In Langley et al. 2002, a total of 166 units (76%) responded, and 120 (55%) were willing to participate in research, and it stated that there were no consistent guidelines / definitions for what a CNS was. Henceforth, some guidelines were created to categorise units:

1. Home visits offering nursing care and specialist advice for a minimum of five days a week for the first month after discharge from the Neonatal Unit (NNU)

2. Ad-hoc advice to families until the infant was at least 12 months old

3. A named nurse providing the link between primary and secondary health care services.

From the initial data set, an random set of 1488 samples was extracted, considering a response variable (readmission) for 10 covariates (see table 1 for 9 of the covariates, with total length of stay in hospital during initial admission not in this table). This data was collected from an 18 NNU subsample of Langley et al. 2002, with surveys sent to every mother who has visited within the previous 3 years. There was a total of 1488 responses, shown in Table 1

	NNU w/ CNS	NNU w/o CNS	Total	
n	702	786	1488	
Birth weight (g)	1649.42	1743.20		
Gestation				
- >26	19	25		
- 26-29	191	193		
- 30-32	261	239		
- 33-36	209	146		
- >36	106	99		
Gender				
- Female	353	355		
- Male	433	347		
Mother Employment				
-Yes	694	631		
- No	92	71		
Father Employment				
- Yes	355	364		
- No	431	338		
Age Mother Finish Edu				
- <16	24	83		
- 16-17	455	416		
- 18-20	224	152		
- >20	83	81		
Accommodation				
- Yes	552	534		
- No	234	168		
NNU Size				
- Large	552	505		
- Small	234	197		

Table 1: Analysis of provided dataset for NNU's with and without CNS

2 Methods

To analyse the data, the R packages provided by 'tidyverse' are used throughout. Before any statistical methods are created, the data is factorised, to allow for consistency. Upon this data some exploratory analysis is performed to see if there is any obvious relationships that can be seen from the visualisations of continuous variables versus factor variables. Violin plots composed upon scatter plots are used to depict the true distributions behind them. See Figure 2 for a comparison of gestation period with birthweight, it is clear that there is positive correlation between the factors, in addition, readmission to the NNU correlates with a lower birthweight respectively to each gestation period.

By creating an initial generalised linear model, based on the binomial distribution presented in Equation 1, which considers all variable interactions and transformations. Thus the 'los' variable, total length of stay in hospital in log(days) during the initial admission, is added into a secondary model to determine if these are identical. Which would prove how impactful the secondary post-birth response variable is
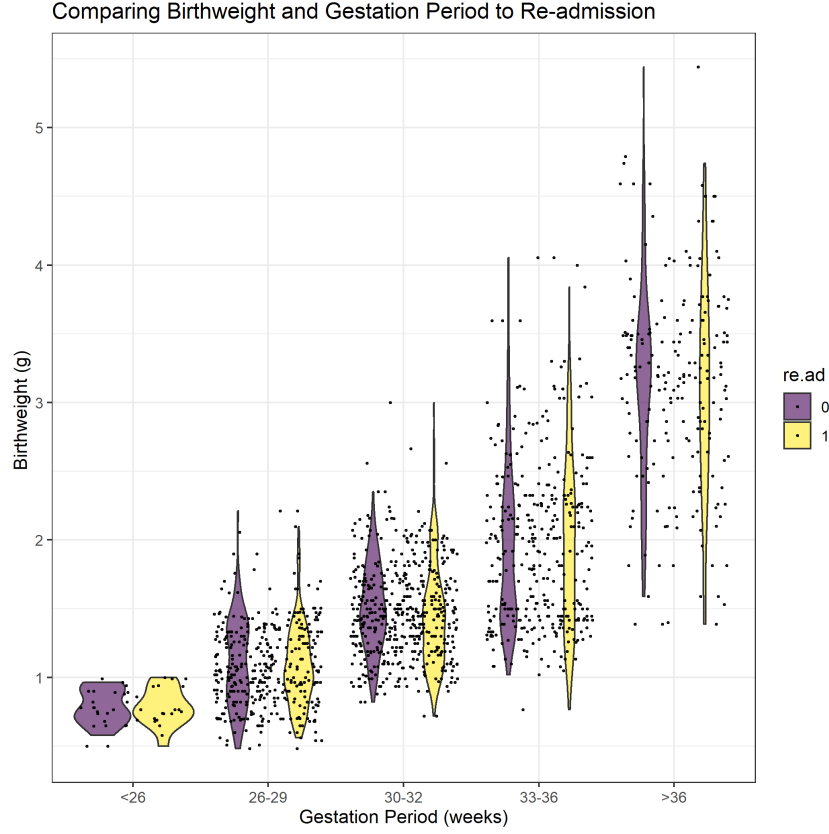
Figure 2: Gestation vs Birthweight for Re-admission. Purple - No readmission, Yellow - readmission

on the overall model.

$$f(x, \pi) = \binom{n}{k} \pi^x (1 - \pi)^{(n-x)} \qquad\qquad x = 0, 1, 2, \ldots, n. \tag{1}$$

$$\tag{2}$$

Using the binomial distribution (from Eq.1), and factoring the quantitative variables to form into dummy

matrices, as in Equation 3 and 4. Where I is the identity matrix.

$$X_{\text{gest}} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{bmatrix} \tag{3}$$

$$\Rightarrow X_{\text{gest}} = X_{1488 \times 5} = I \tag{4}$$

After preparing the data for analysis, a regression model is run on the indepenent variables, excluding

're.ad' as the readmission variable because it is the response variable and also excluding 'los' as it too is

a post-birth response variable, and consider the interaction between variables After data preparation is

complete, we first run a regression on all the independent variables except los and consider the interaction

(a) Birthweight and Gender vs Re-admission

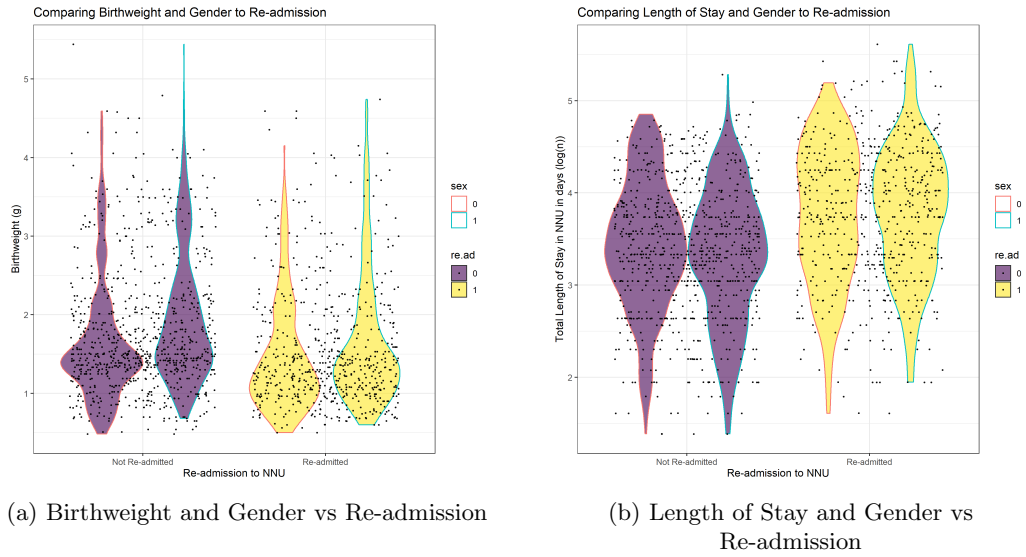(b) Length of Stay and Gender vs Re-admission

Figure 3: Birthweight and Length of Stay with Gender vs Re-admission. Outline : Blue - Male, Pink - Female

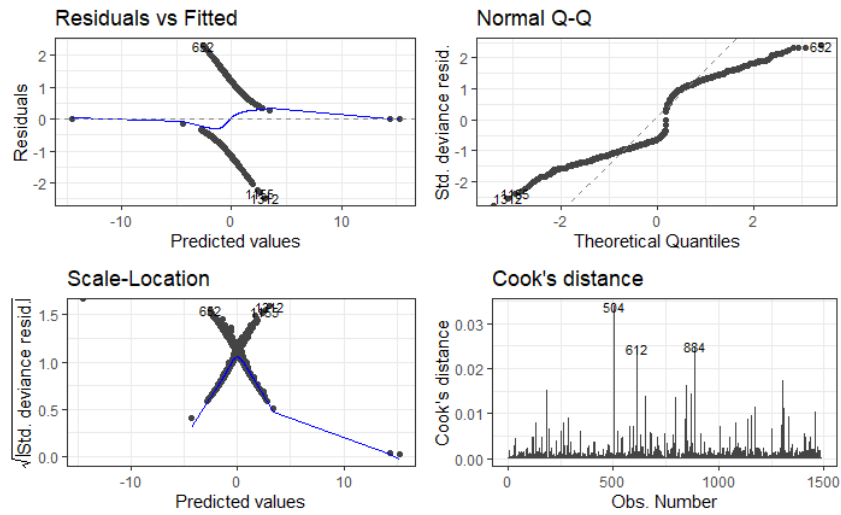Figure 4: Preliminary GLM Model with 9 Covariates

effects of all variables as follow:

3 Results

Insert Results

4 Conclusions

Insert Conclusions

References

Boyle, Michael H., Torrance, George W., Sinclair, John C., and Horwood, Sargent P. (1983). "Economic Evaluation of Neonatal Intensive Care of Very-Low-Birth-Weight Infants". In: *New England Journal of Medicine* 308.22, pp. 1330–1337. DOI: 10.1056/NEJM198306023082206. eprint: https://doi.org/10.1056/NEJM198306023082206. URL: https://doi.org/10.1056/NEJM198306023082206.

Langley, D, Hollis, S, Friede, T, MacGregor, D, and Gatrell, A (2002). "Impact of community neonatal services: a multicentre survey". In: *Archives of Disease in Childhood-Fetal and Neonatal Edition* 87.3, F204–F208.

Petrou, S, Sach, T, and Davidson, L (2001). "The long-term costs of preterm birth and low birth weight: results of a systematic review". In: *Child: Care, Health and Development* 27.2, pp. 97–115. DOI: https://doi.org/10.1046/j.1365-2214.2001.00203.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1365-2214.2001.00203.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2214.2001.00203.x.

Salt, Alison (2001). "Birth Counts: Statistics of Pregnancy and Childbirth. Volumes I and II". In: *Developmental Medicine &; Child Neurology* 43.6, pp. 428–432. DOI: 10.1017/S0012162201260794.