Evaluation of risk factors for infant hospitalized pneumonia incidence

Lina Gao, Yu-Fen Huang, Feng Li

Summary

Using the data 11.3 in the textbook, we identified four important risk factors associated with the incidence of infant hospitalized pneumonia. They include mom's age, cigarette, number of sibling and breastfeeding. Infants with older mom (age >=23) have lower incidence of hospitalized pneumonia than the ones with young mom (age <23) with the hazard ratio 0.081. Similar results were obtained for breastfeeding and non-breastfeeding babies. It's hazard ratio is 0.4. Smokers have almost two times of probability giving birth to a infant with hospitalized pneumonia than non-smokers. Infants with more siblings (n >1) have ~ 1.9 times of pneumonia incidence than the ones with one or no sibling (n<=1). Although region is not a significant factor if all four regions were considered, the significant difference between north central compared with the others (northeast, south, and west) were observed. The rest factors were not identified as the significant ones in this study. Therefore, an infant with breastfeeding, one or no sibling, older and non-smoking mom has lower risk of hospitalized pneumonia.

Abbreviation

Age_child: age of child had pneumonia
Indic_hos: hospitalized =1, not hospitalized =0
Age_mom: age of mother
Urban: urban environment
Alcohol: alcohol use by mother during pregnancy
Age_in_hos: months indicator for hospitalization
Cigarette: cigarette use by mother
Region: region of the country
Poverty: mother at poverty level
Bweight: normal birthweight
Race_mom: race of mother, white, black or other
Edu_mom: education of mother
N_sibling: number of siblings
M_solidfood: month the child on solid food
M_weaned: month of weaning

Introduction

Pneumonia is an infection in the lungs most often caused by bacteria or viruses. The air sacs in the lungs (called alveoli) fill up with pus and other fluid, which makes it hard for oxygen to reach the bloodstream.

Common bacteria and viruses that may cause pneumonia are:

- Streptococcus pneumoniae
- Mycoplasma pneumonia.
- Group B streptococcus
- Staphylococcus aureus
- Respiratory syncytial virus (RSV).
- Parainfluenza virus
- Influenza virus
- Adenovirus

Pneumonia remains the leading infectious cause of death among children under five, killing approximately 2,400 children a day (1-2). Pneumonia accounted for approximately 16 percent of the 5.6 million under-five deaths, killing around 880,000 children in 2016. Most of its victims were less than 2 years old. Hospitalization due to childhood pneumonia is strongly linked to poverty-related factors such as undernutrition, lack of safe water and sanitation, indoor air pollution, inadequate access to health care and the health conditions of the pregnant mom. An integrative approach is urgently needed to tackle this important public health issue and develop a practical guidelines for pregnant women.

Studies suggest that optimal breastfeeding practices, including exclusive breastfeeding during the first six months of life and continued breastfeeding until 24 months of age, are critical for reducing the burden of pneumonia among infants and young children. The protective effect of human milk against respiratory infection is attributed to its numerous immunobiological components (3-4). Using the data 1.13 from the textbook (5), we will explore the association of breast-feeding with hospitalized infant pneumonia incidence, as well as other potential risks. This data was gathered from 3470 personal interviews conducted by the National Longitudinal Survey of Youth (NLSY, 1995) from 1979 to 1986. Totally 15 variables are collected, which can be divided into three categories: mom-related, infant-related and social environment-related. Among them, the age of the child suffered from pneumonia and hospitalized or not as the variables for time to event survival analysis in this study. The other variables are used as covariates, which include birth weight, mom's race, number of siblings (0-6), age of the mom (14-29 years old), mom's education, region of the country (northeast, north central, south and west), poverty, urban (whether the mother live in an urban environment), alcohol (alcohol consumption by mother during pregnancy), cigarette (mother's cigarette consumption), month of weaned and month of solid food (month the baby begin to eat solid food).

Using all the relevant statistical techniques we learned in life data analysis class, we will try to answer two questions:

•Primary question: Does breastfeeding protest the infant against hospitalized pneumonia in the first year of infants?

•Secondary questions: Do any other factors have close association with infant hospitalized pneumonia incidence? Based on our results, can we provide a preventive guideline for pregnant women to decrease the risk of infant pneumonia incidence?

Methods

Cox model is well-recognized as a powerful statistical technique for assessing simultaneously the effect of several risk factors on the survival analysis. It can provided an estimate of the effect of a variable on the survival curves after adjustment for the other variables. Hence, cox model will be our main technique to explore the association of infant hospitalized pneumonia incidence and 12 potential risk factors.

After identifying the key risk factors, we will further confirm them using nonparametric test (Wilcoxon and logrank test) and Kaplan-Meier curves. They can describe the survival according to one factor under investigation, but ignore the impact of any others. Additionally, because these techniques are useful only when the predictor is categorical, we need define proper dummy variables before analysis.

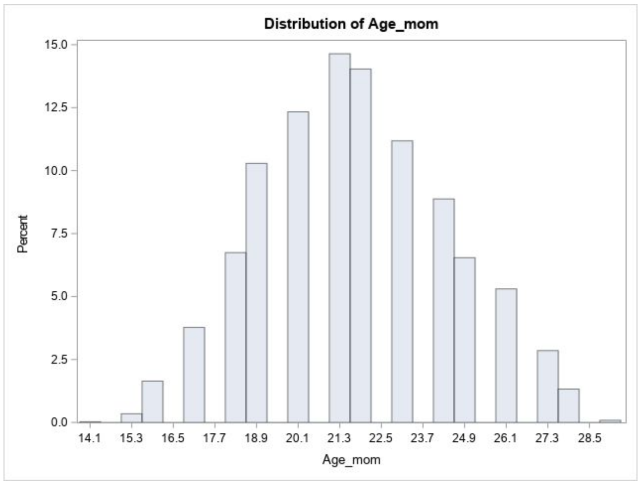
Although the golden age of parametric model in life data analysis is gone. It is still worthwhile to explore its application in this study. Hence, we will choose two dominant techniques, Weilbull and log logistic regression to fit the data.

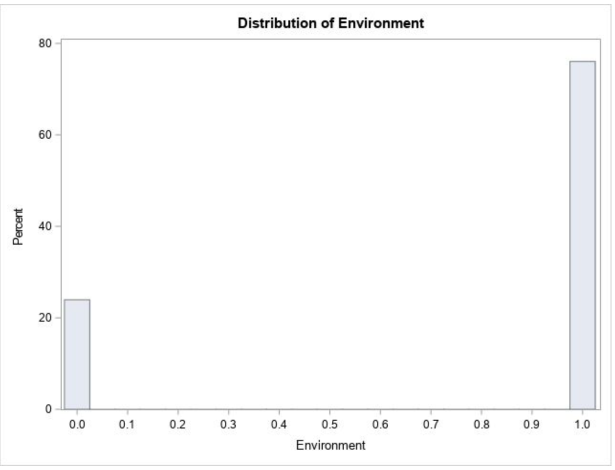
Results

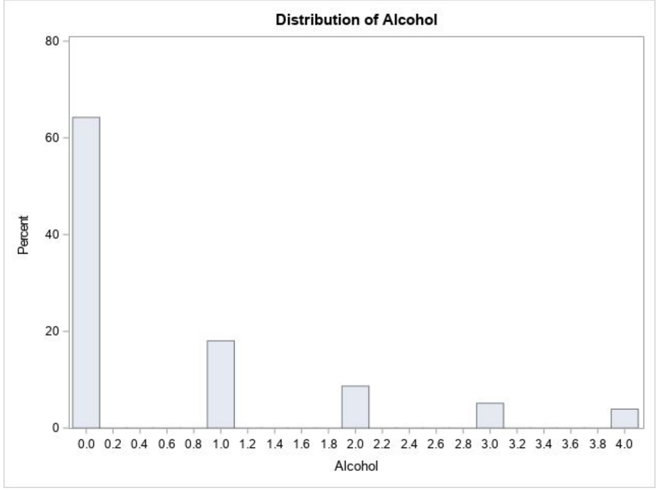
1. Distribution of covariates in this study

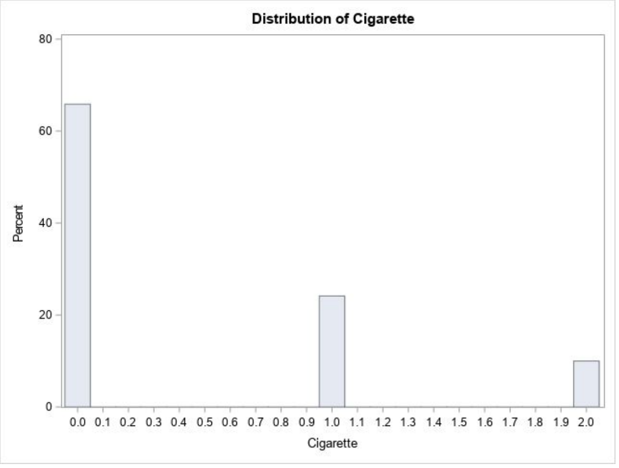
•Age of the mother	(avg 21.6 years, range 14-19 years)
•Education of the mother	(avg 11.4 years, range 0-19 years)

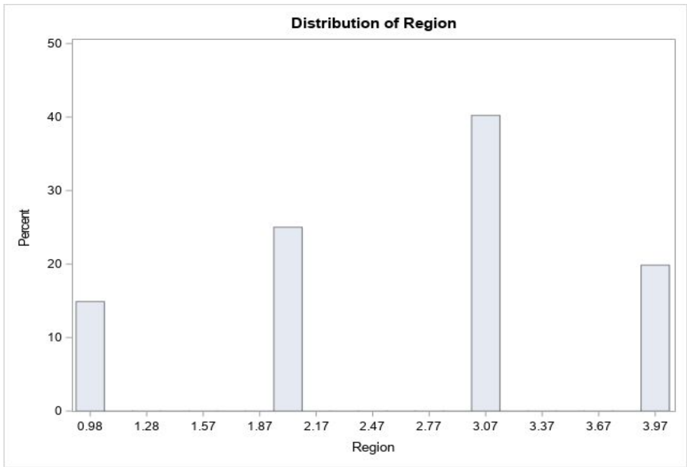
- Poverty status mother (Yes: 92%, No: 8%)
- Urban environment for mother (Yes: 76%, No: 24%)
- Alcohol use mother (Yes: 36%, No: 64%, range 0-4 drinks per month)
- Cigarette use mother (Yes: 34%, No: 66%, range 0-2 pack per day)
- Normal birthweight (>5.5 lbs.) (Yes: 36%, No: 64%)
- Siblings presence (Yes: 48%, No: 52%, range 0-6)
- Month the child was weaned (avg 1.9 months, range 0-28 months)
- Month the child on solid food (avg 1.1 months, range 0-18 months)
- Race of the mother (1=white: 56%, 2=black: 28%, 3=other:16%)
- Region of the country (1=northeast: 15%, 2=north central: 25%, 3=south: 40%, 4=west: 20%)

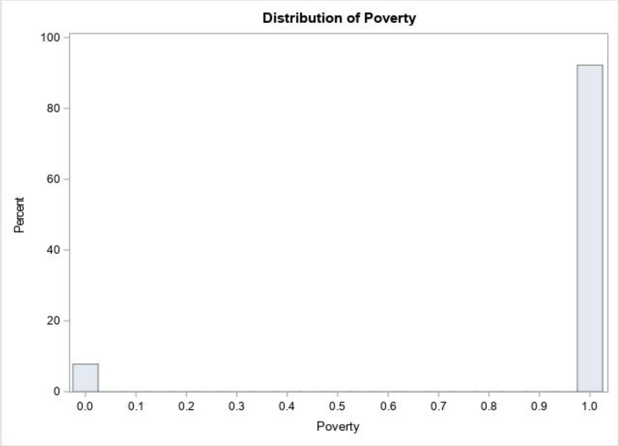


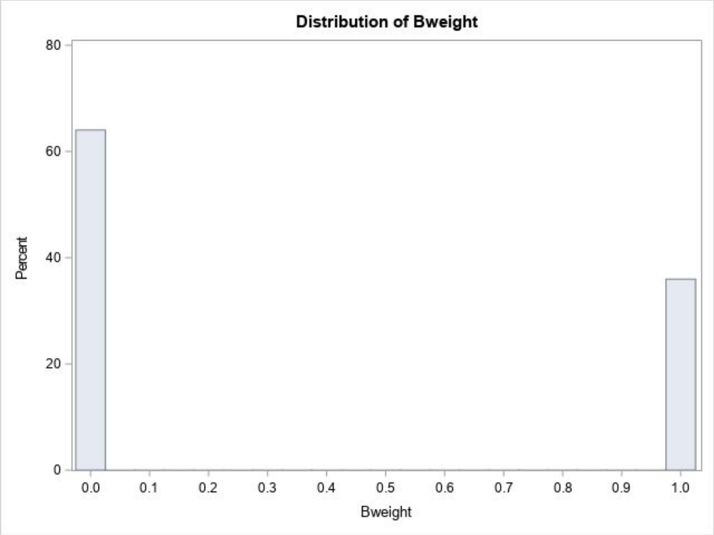


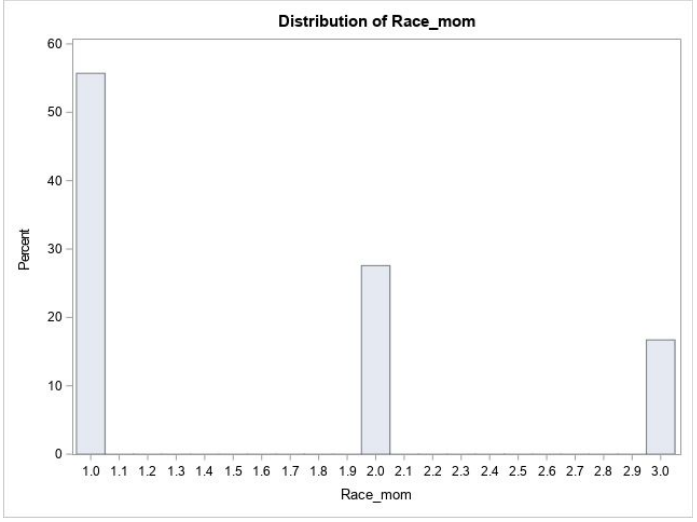


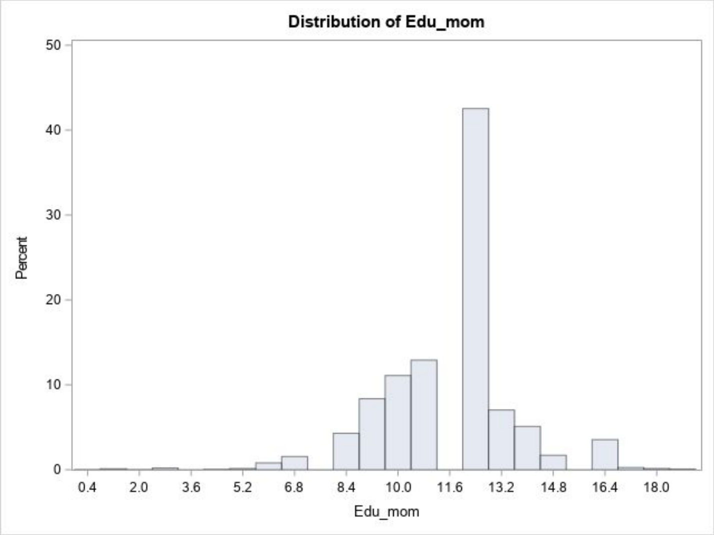


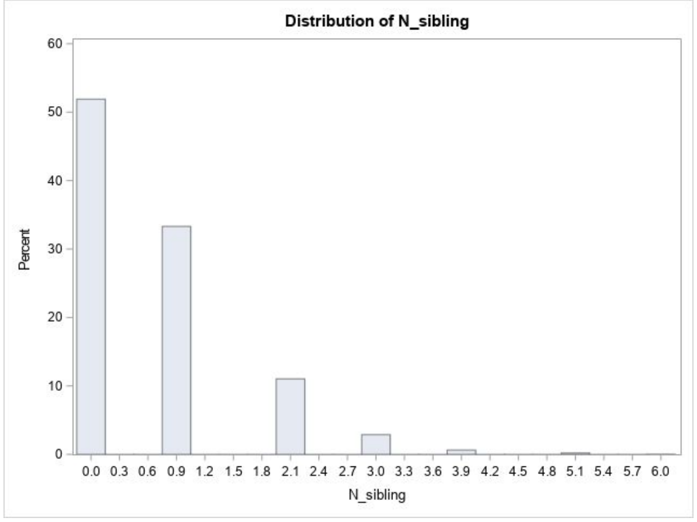

Distribution of Region
Percent
0
10
20
30
40
50
0.98
1.28
1.57
1.87
2.17
2.47
2.77
3.07
3.37
3.67
3.97
Region


Distribution of Poverty
Percent
0
20
40
60
80
100
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Poverty


Distribution of Bweight
Percent
0
20
40
60
80
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Bweight


Distribution of Race_mom
Percent
0
10
20
30
40
50
60
1.0
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
2.0
2.1
2.2
2.3
2.4
2.5
2.6
2.7
2.8
2.9
3.0
Race_mom


Distribution of Edu_mom
Percent
0
10
20
30
40
50
0.4
2.0
3.6
5.2
6.8
8.4
10.0
11.6
13.2
14.8
16.4
18.0
Edu_mom


Distribution of N_sibling
Percent
0
10
20
30
40
50
60
0.0
0.3
0.6
0.9
1.2
1.5
1.8
2.1
2.4
2.7
3.0
3.3
3.6
3.9
4.2
4.5
4.8
5.1
5.4
5.7
6.0
N_sibling

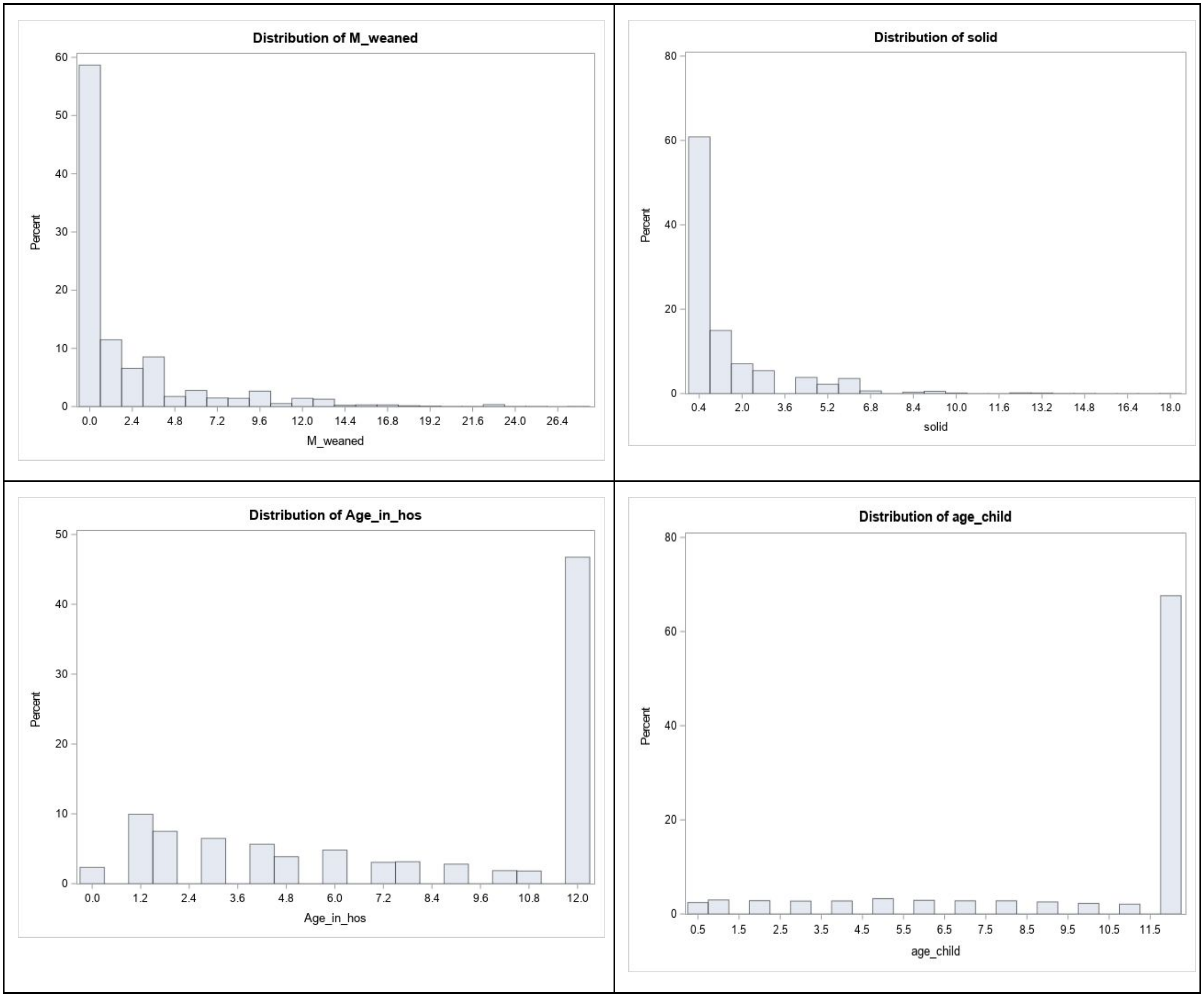


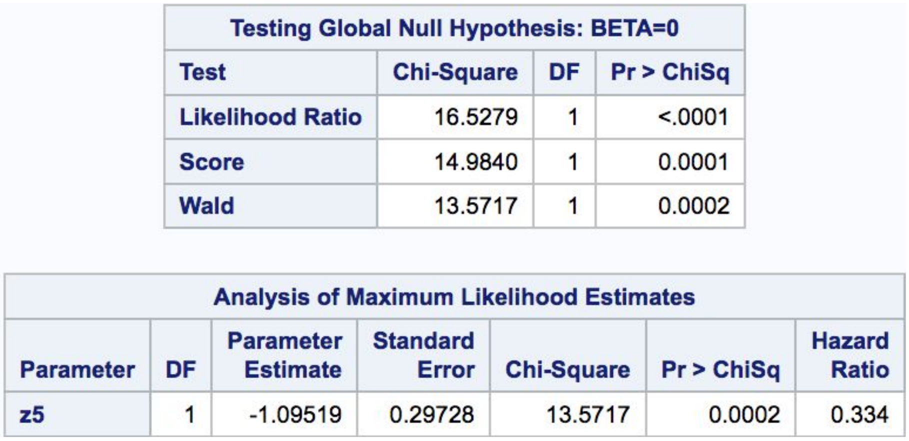
Figure 1. Overall Histogram of distribution of covariates in this study

2. Cox model development

- **Primary question:** Does breastfeeding protest the infant against hospitalized pneumonia in the first year of infants?

  Define the dummy variable Z5=1 if infants were breastfed at birth, 0 otherwise. By testing the hypothesis of beta=0, the results consistently suggested rejection of hypothesis using the likelihood ratio, test score, and Wald tests. The hazard ratio , 0.33, means breastfeeding infant has decrease 67% hospitalized pneumonia incidence of the ones without breastfeeding.

Table 1. Global test results for z5 in the survival function



Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	16.5279	1	<.0001
Score	14.9840	1	0.0001
Wald	13.5717	1	0.0002

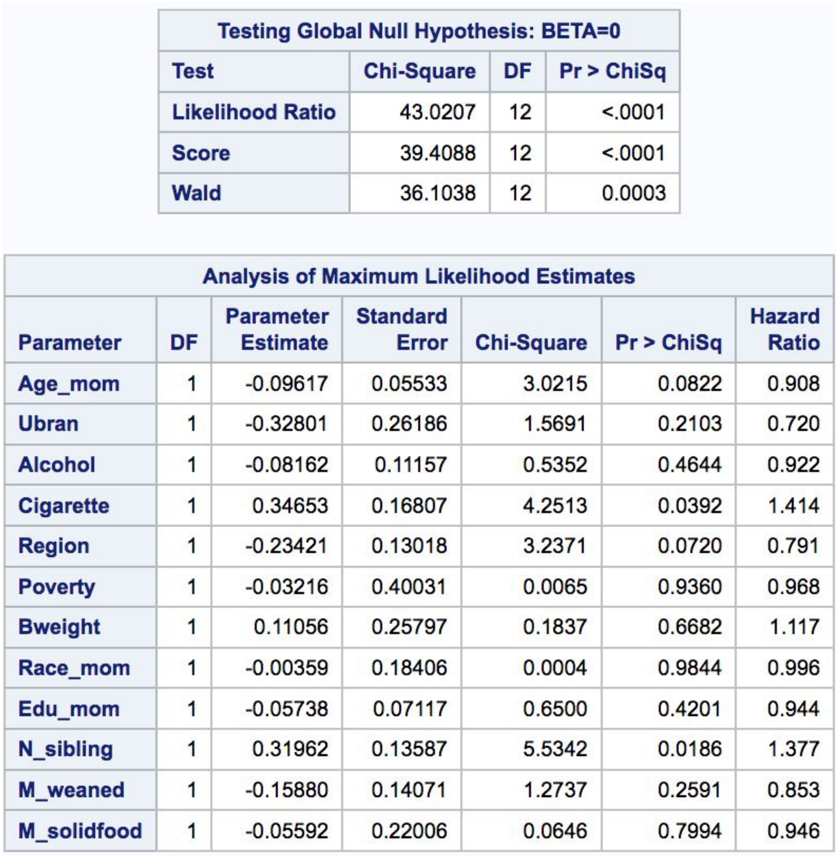
Analysis of Maximum Likelihood Estimates						
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
z5	1	-1.09519	0.29728	13.5717	0.0002	0.334

Table 1. Result of maximum likelihood estimation of breastfeeding variable

Secondary question: Compared with breastfeeding choice, can any other factors have effect on the hospitalization of pneumonia for infants?.

- Cox full model development

Table 2. ANOVA results of Cox full model in this study

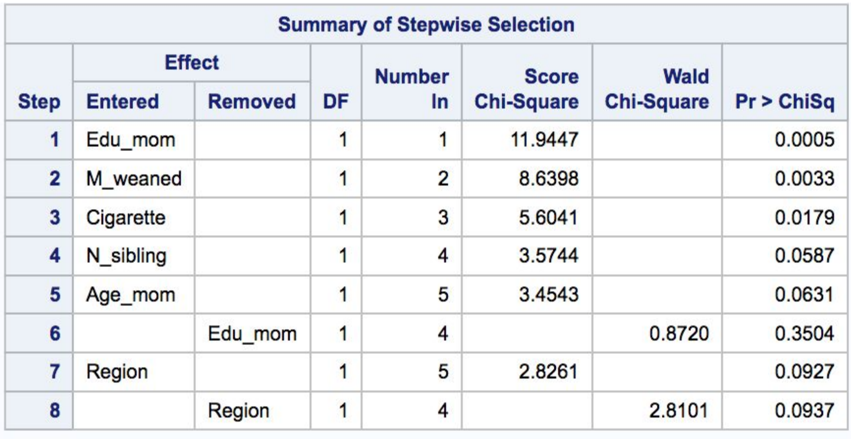


Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	43.0207	12	<.0001
Score	39.4088	12	<.0001
Wald	36.1038	12	0.0003

Analysis of Maximum Likelihood Estimates						
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
Age_mom	1	-0.09617	0.05533	3.0215	0.0822	0.908
Ubran	1	-0.32801	0.26186	1.5691	0.2103	0.720
Alcohol	1	-0.08162	0.11157	0.5352	0.4644	0.922
Cigarette	1	0.34653	0.16807	4.2513	0.0392	1.414
Region	1	-0.23421	0.13018	3.2371	0.0720	0.791
Poverty	1	-0.03216	0.40031	0.0065	0.9360	0.968
Bweight	1	0.11056	0.25797	0.1837	0.6682	1.117
Race_mom	1	-0.00359	0.18406	0.0004	0.9844	0.996
Edu_mom	1	-0.05738	0.07117	0.6500	0.4201	0.944
N_sibling	1	0.31962	0.13587	5.5342	0.0186	1.377
M_weaned	1	-0.15880	0.14071	1.2737	0.2591	0.853
M_solidfood	1	-0.05592	0.22006	0.0646	0.7994	0.946

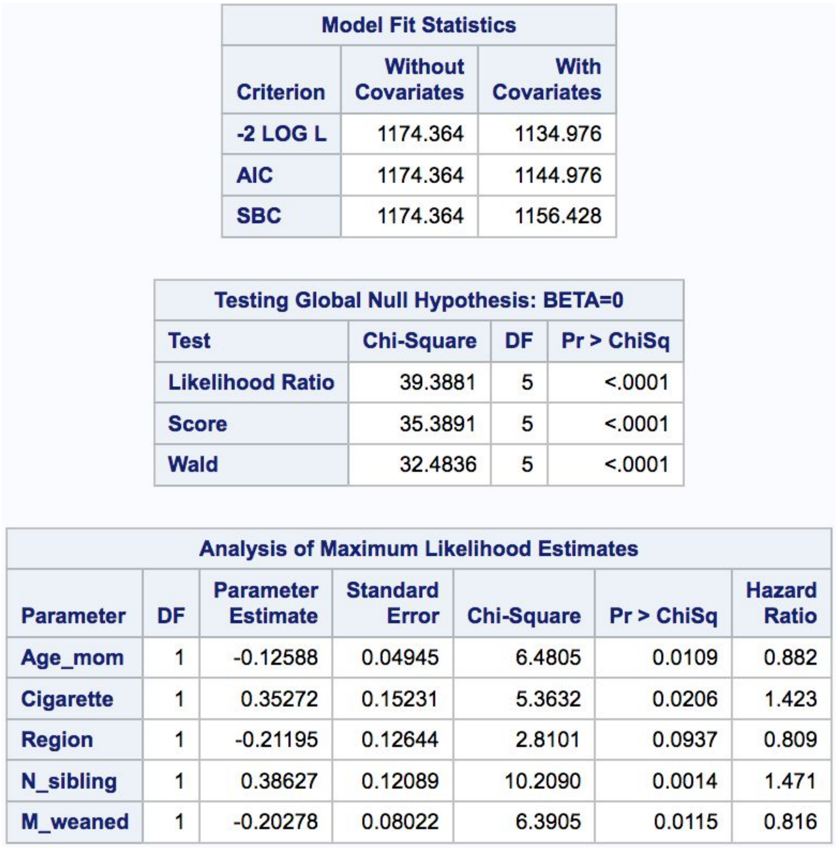
- Best model selection by stepwise and AIC

Table 3. Result of stepwise variable selection



Summary of Stepwise Selection							
	Effect						
Step	Entered	Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
1	Edu_mom		1	1	11.9447		0.0005
2	M_weaned		1	2	8.6398		0.0033
3	Cigarette		1	3	5.6041		0.0179
4	N_sibling		1	4	3.5744		0.0587
5	Age_mom		1	5	3.4543		0.0631
6		Edu_mom	1	4		0.8720	0.3504
7	Region		1	5	2.8261		0.0927
8		Region	1	4		2.8101	0.0937

Table 4. Final model determination by the smallest AIC



Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	1174.364	1134.976
AIC	1174.364	1144.976
SBC	1174.364	1156.428

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	39.3881	5	<.0001
Score	35.3891	5	<.0001
Wald	32.4836	5	<.0001

Analysis of Maximum Likelihood Estimates						
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
Age_mom	1	-0.12588	0.04945	6.4805	0.0109	0.882
Cigarette	1	0.35272	0.15231	5.3632	0.0206	1.423
Region	1	-0.21195	0.12644	2.8101	0.0937	0.809
N_sibling	1	0.38627	0.12089	10.2090	0.0014	1.471
M_weaned	1	-0.20278	0.08022	6.3905	0.0115	0.816

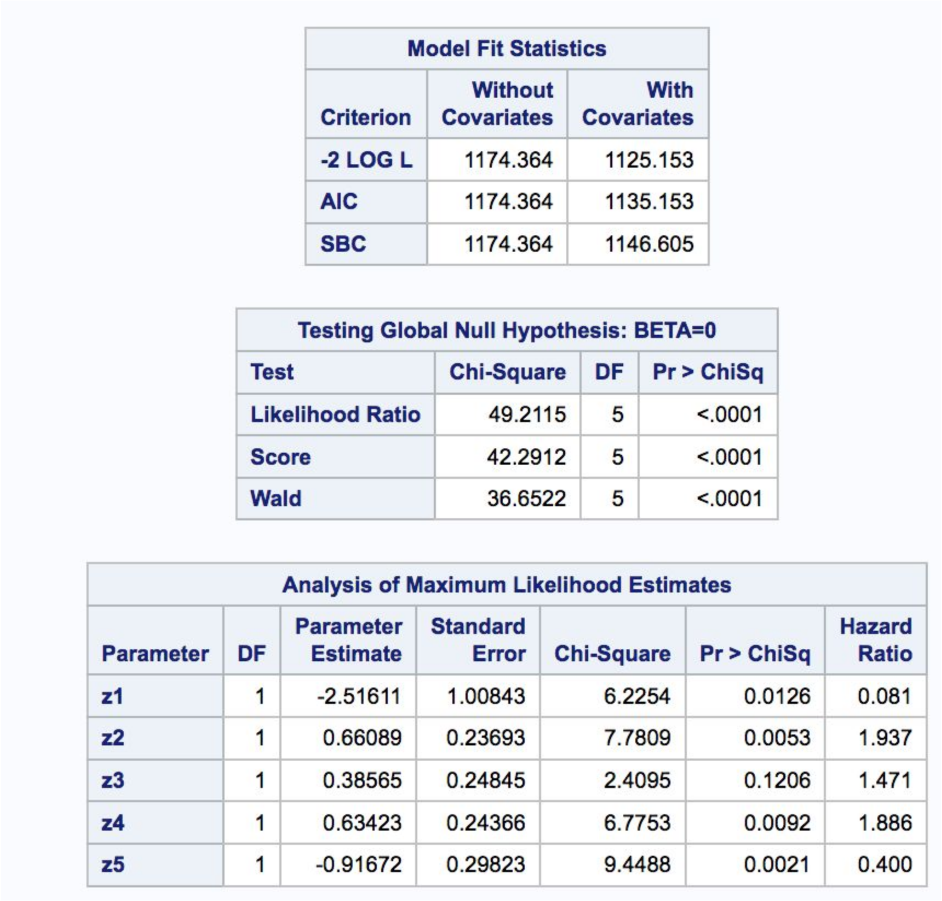
Using Cox model, we selected five predictors as the important risk factors in this study, including age_mom, cigarette, region, N_sibling and M_weaned.

- Data transformation and final model development
  Dummy variable definition for testing key predictors

```
if (Age_mom >= 23)   then z1=1; else z1=0;
if (Cigarette >= 1)   then z2=1; else z2=0;
if (Region = 2)       then z3=1; else z3=0;
if (N_sibling >= 1)   then z4=1; else z4=0;
if (M_weaned >= 1)    then z5=1; else z5=0;
```

The definition of dummy variables uses the median value in each variable considering censored or not censored as well.

Table 5. Final best model fitting results



Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	1174.364	1125.153
AIC	1174.364	1135.153
SBC	1174.364	1146.605

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	49.2115	5	<.0001
Score	42.2912	5	<.0001
Wald	36.6522	5	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
z1	1	-2.51611	1.00843	6.2254	0.0126	0.081
z2	1	0.66089	0.23693	7.7809	0.0053	1.937
z3	1	0.38565	0.24845	2.4095	0.1206	1.471
z4	1	0.63423	0.24366	6.7753	0.0092	1.886
z5	1	-0.91672	0.29823	9.4488	0.0021	0.400

3. nonparametric test

- Mom's age

  Both log-rank test (p=0.022) and Wilcoxon test (p=0.017) shows significant difference between older mom (>=23 years old, z1=1) and younger mom (< 23 years old, z1=0). The K-M survival curves also showed significant difference between young and mature moms.
- Cigarette

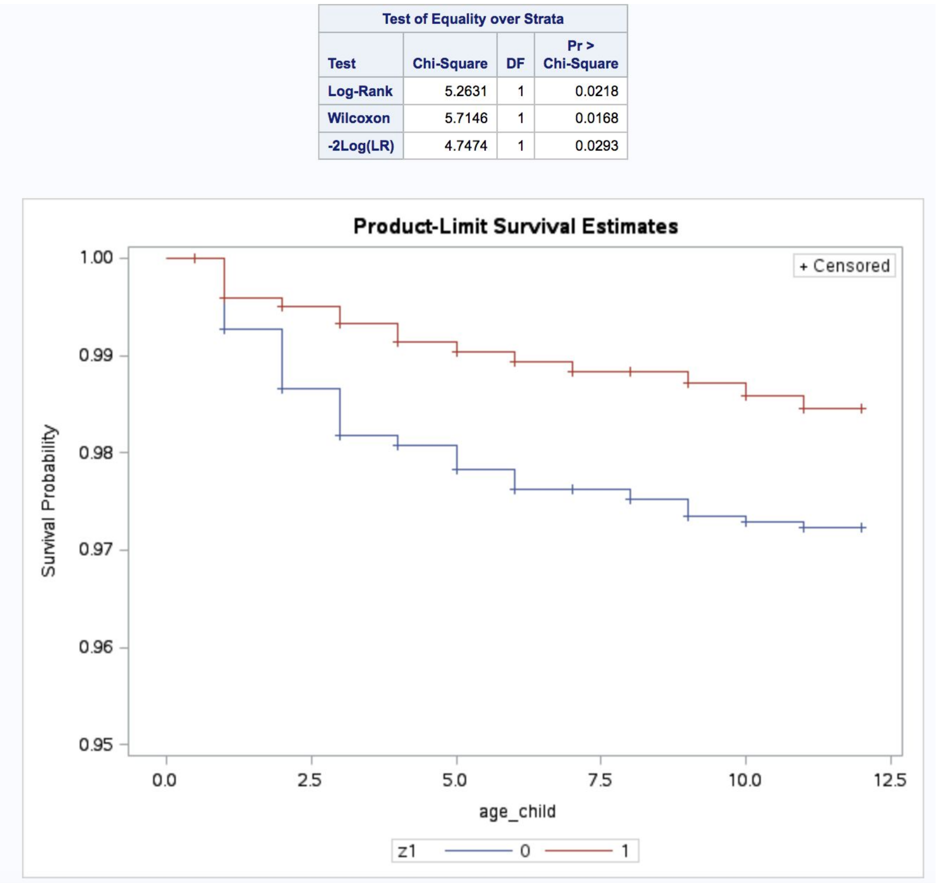
  Both log-rank test (p=0.0005) and Wilcoxon test (p=0.0006) shows significant difference between non cigarette consumer and cigarette consumer. Consistent results also implied in figure 3.

- Region 2 vs other region

  Both log-rank test (p=0.061) and Wilcoxon test (p=0.068) shows marginally significant difference between live in north central and other region of country. Also, figure 4 showed the difference between them.
- Sibling

  Both log-rank test (p=0.0068) and Wilcoxon test (p=0.0061) shows significant difference between number of siblings of infant <= 1 and number of siblings of infant > 1. It was also confirmed by the P-L plot in figure 5.
- Breastfeeding

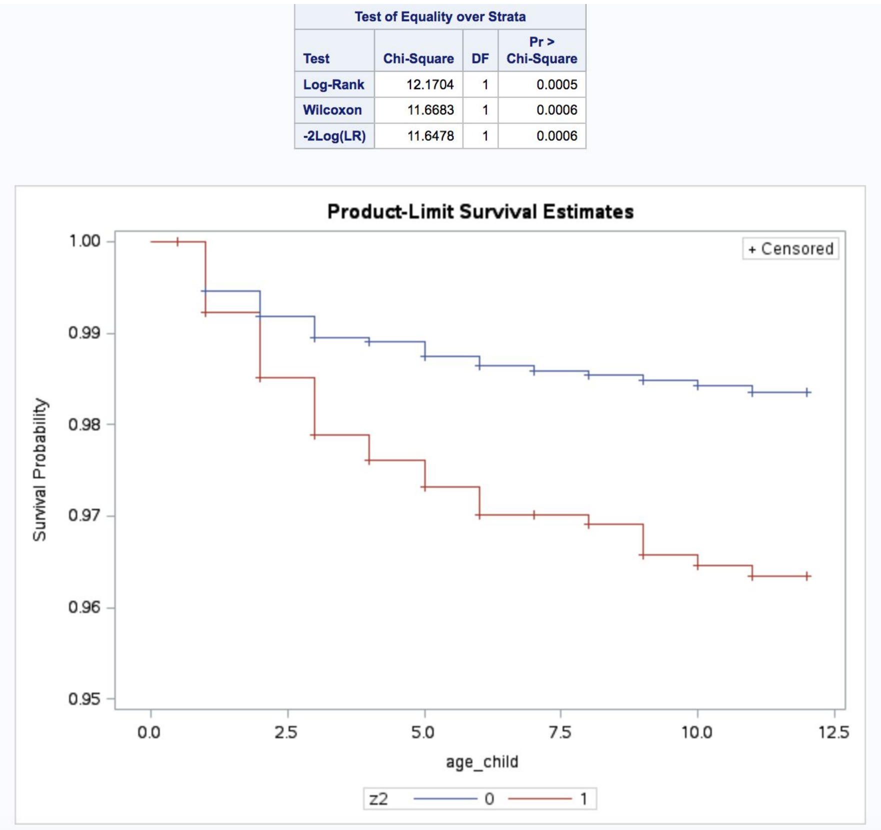
  log-rank test (p=0.0001), Wilcoxon test (p<0.0001) and figure 6 consistently supported the significant difference between breastfeeding and non breastfeeding survival curves.



Test of Equality over Strata

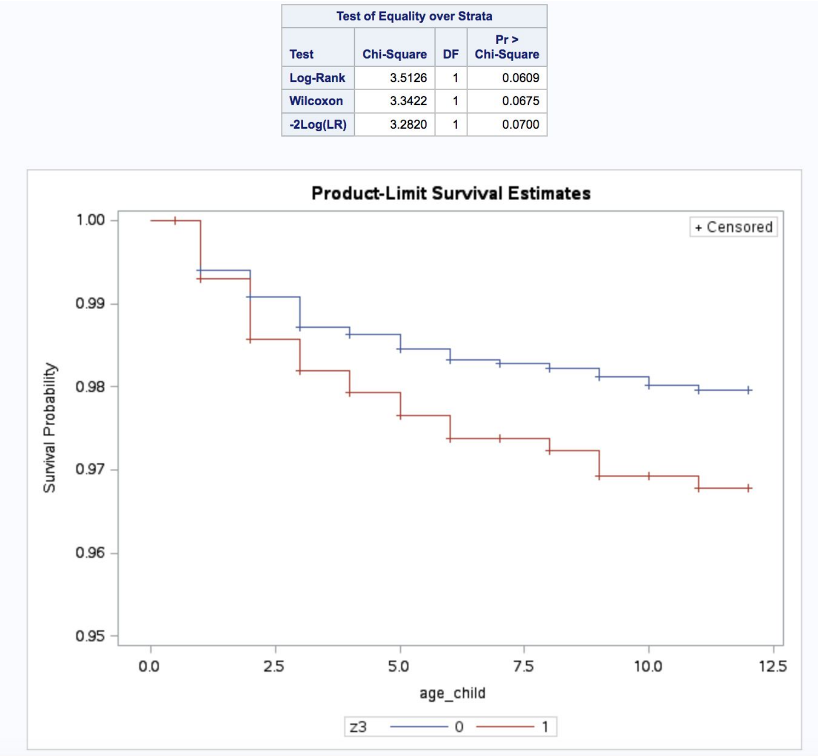
Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	5.2631	1	0.0218
Wilcoxon	5.7146	1	0.0168
-2Log(LR)	4.7474	1	0.0293

Figure 2. Result of PL survival estimation about mom’s age in this study



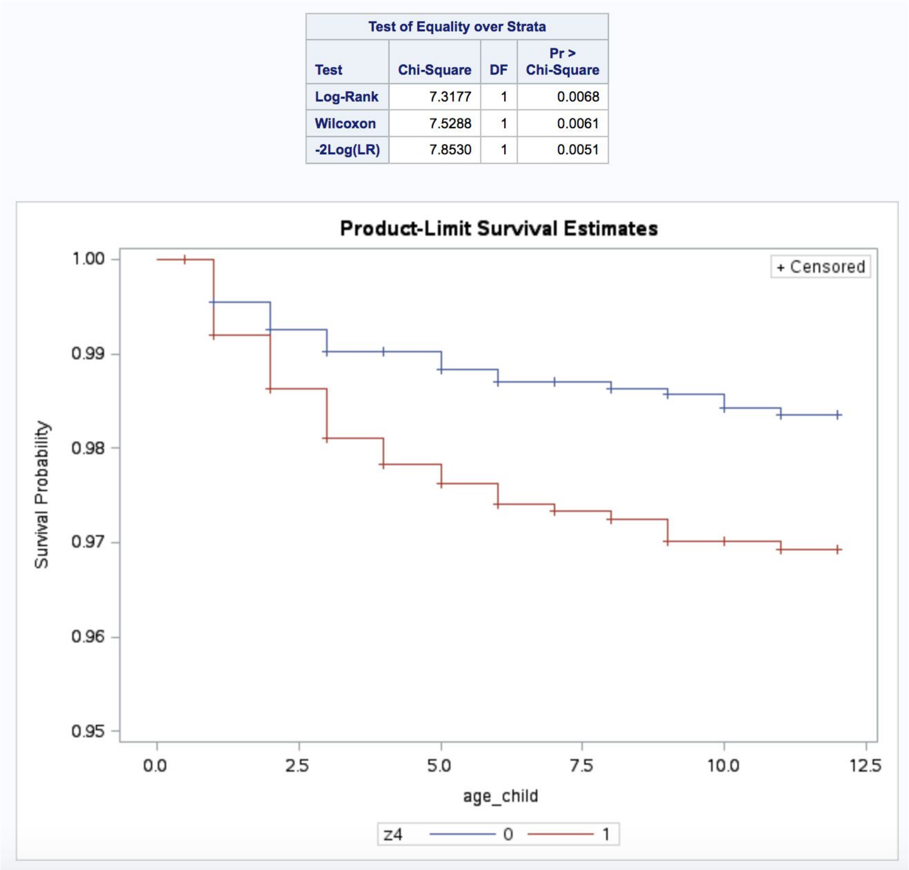
Test of Equality over Strata			
Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	12.1704	1	0.0005
Wilcoxon	11.6683	1	0.0006
-2Log(LR)	11.6478	1	0.0006

Figure 3. Result of PL survival estimation about cigarette in this study



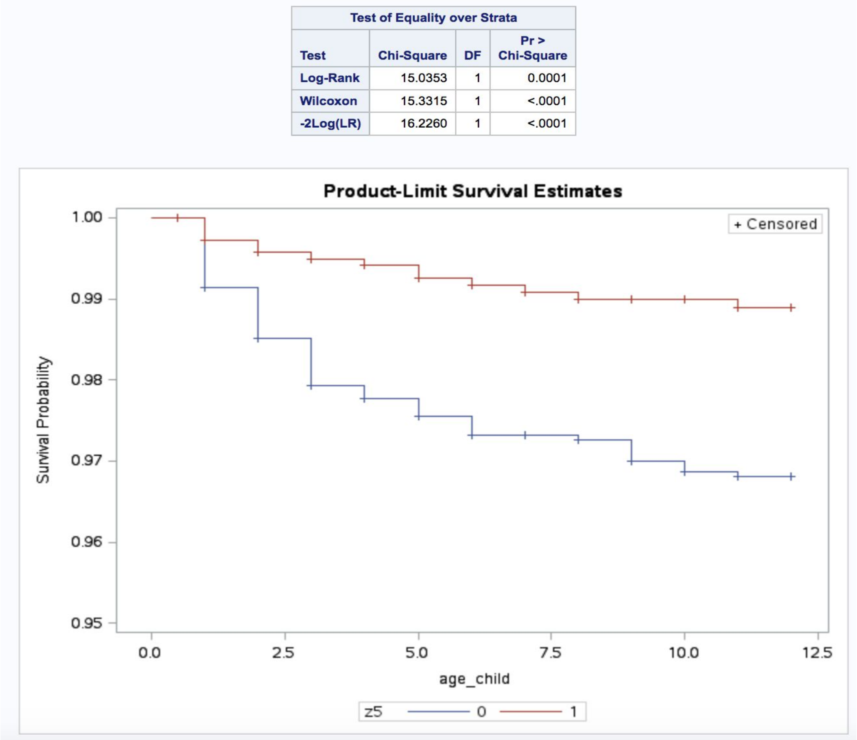
Test of Equality over Strata			
Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	3.5126	1	0.0609
Wilcoxon	3.3422	1	0.0675
-2Log(LR)	3.2820	1	0.0700

Figure 4. Result of PL survival estimation between region 2 and other regions



Test of Equality over Strata			
Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	7.3177	1	0.0068
Wilcoxon	7.5288	1	0.0061
-2Log(LR)	7.8530	1	0.0051

Figure 5. Result of PL survival estimation about number of sibling



Test of Equality over Strata			
Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	15.0353	1	0.0001
Wilcoxon	15.3315	1	<.0001
-2Log(LR)	16.2260	1	<.0001

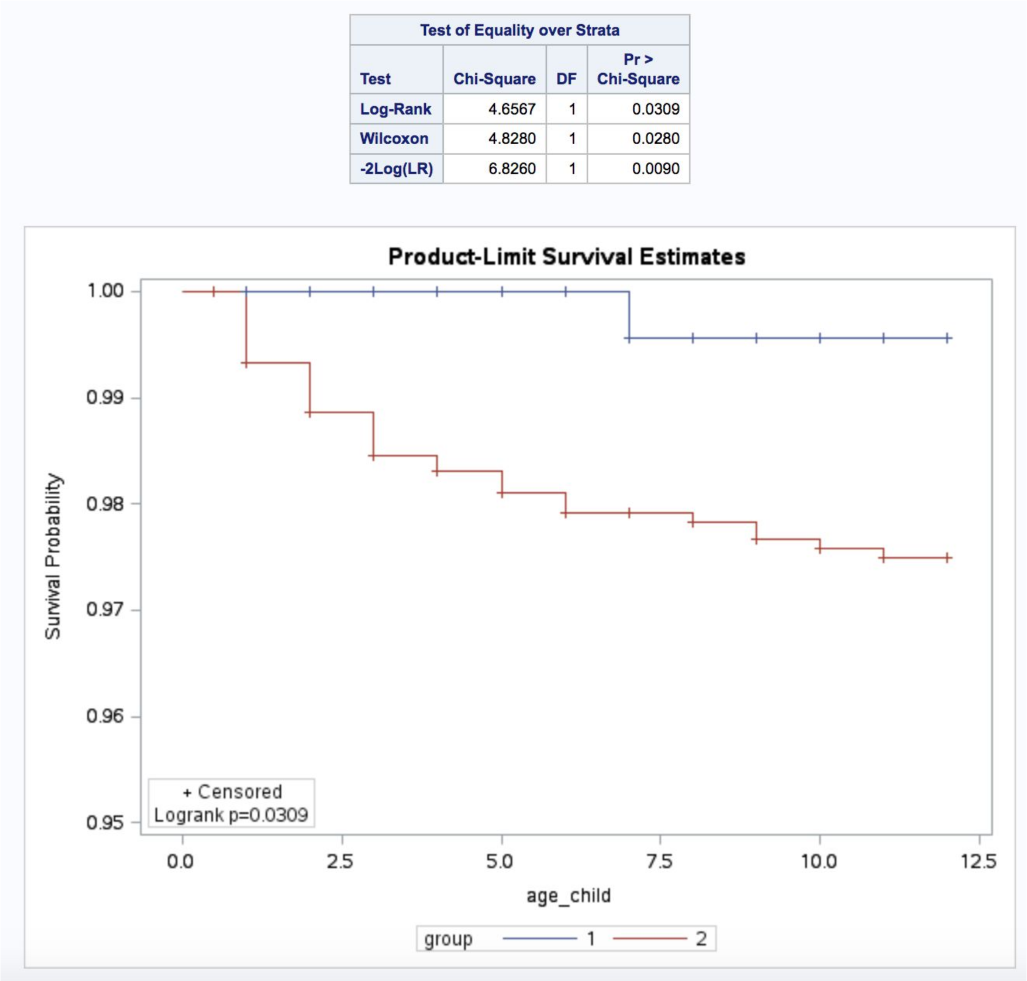
Figure 6. Result of PL survival estimation about Breastfeeding

4. Low risk group vs high risk group comparison

To proof the four important risk factors associated with the incidence of infant hospitalized pneumonia. We separated data into two groups: low risk group and high risk group to compare the difference of survival rate.

- Low risk group: mom’s age >= 23 years old, non cigarette, one or no sibling, breastfeeding.
- High risk group: mom’s age <23 years old, use cigarette, more the one siblings, non breastfeeding.

The high risk group decreased survival rate very fast, starting at one month of child’s age. Both log-rank test (p=0.031) and Wilcoxon test (p=0.028) shows significant difference between low risk group and high risk group, indicated in figure 7.



Test of Equality over Strata

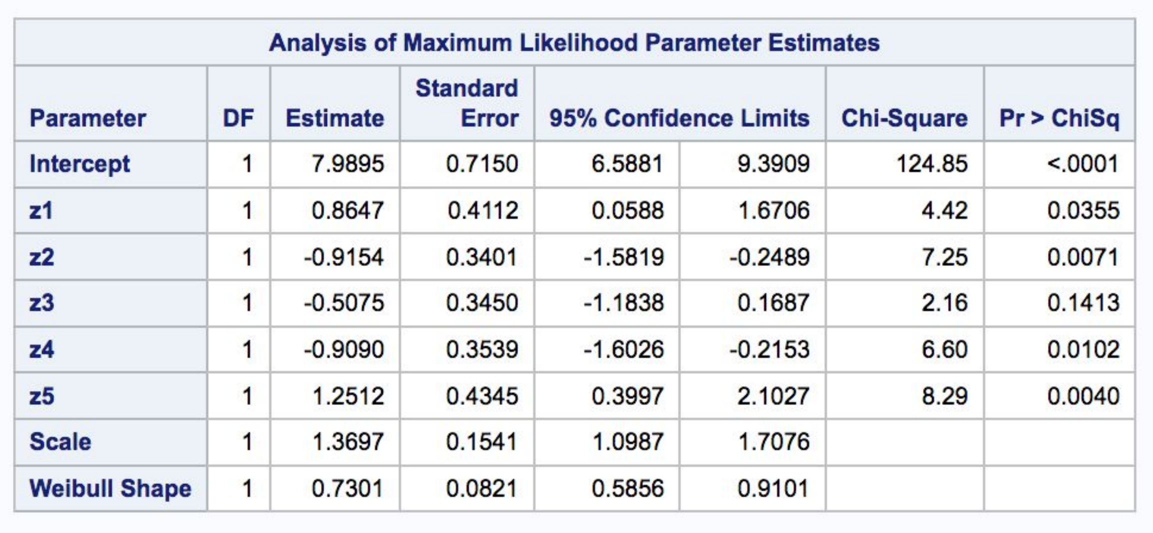
Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	4.6567	1	0.0309
Wilcoxon	4.8280	1	0.0280
-2Log(LR)	6.8260	1	0.0090

Figure 7. Result of PL survival estimation in low risk group vs high risk group

5. Parametric model development

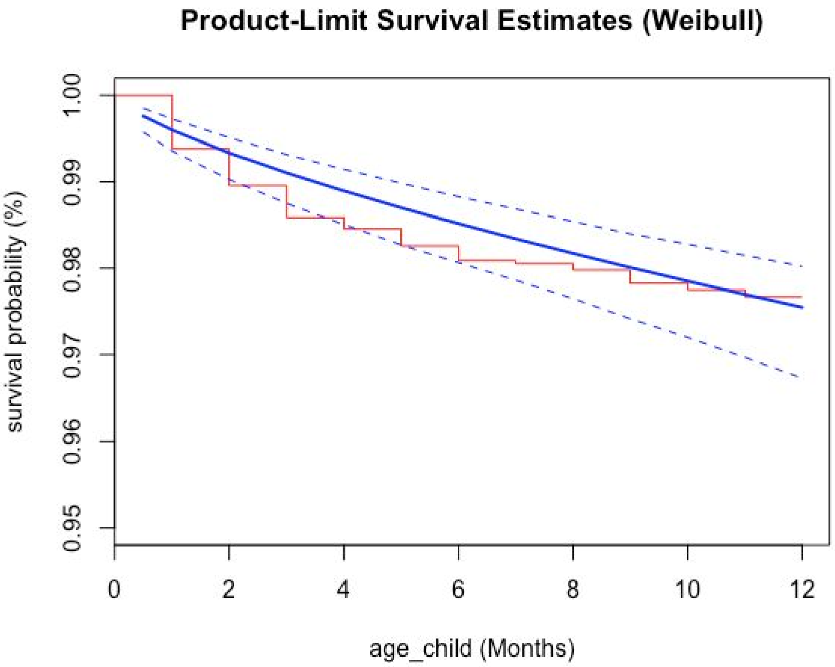
Using Weibull and llogistic modelling to fit with our multiple model with 5 risk factor (Z1: mom's age, Z2: cigarette, Z3: region 2, Z4: sibling and Z5: breastfeeding). Both of them verified that the mom' age, cigarette, number of sibling and breastfeeding are significant risk factors in our study (p-value < 0.05). Most of these results are consistent with the cox model fitting result.

Table 6. Result of the weibull model



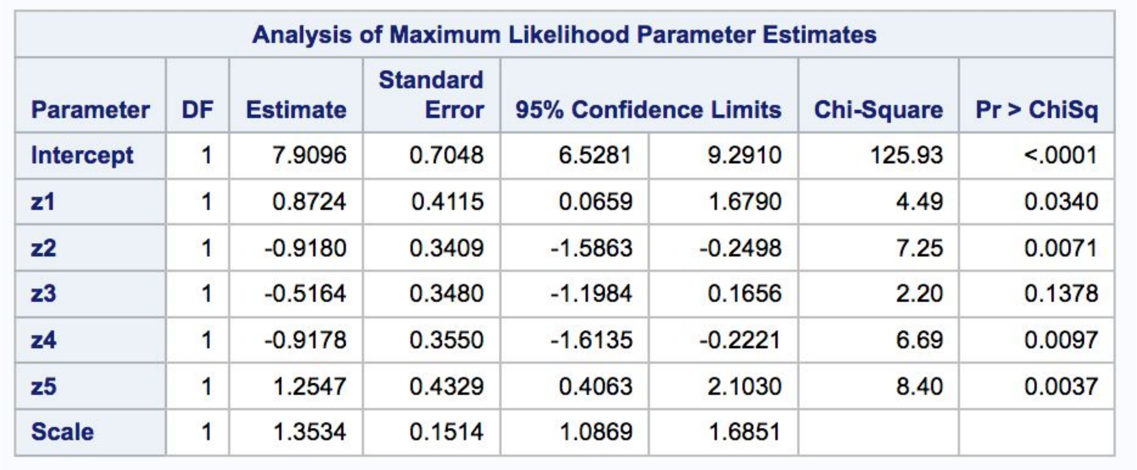
Analysis of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	7.9895	0.7150	6.5881	9.3909	124.85	<.0001
z1	1	0.8647	0.4112	0.0588	1.6706	4.42	0.0355
z2	1	-0.9154	0.3401	-1.5819	-0.2489	7.25	0.0071
z3	1	-0.5075	0.3450	-1.1838	0.1687	2.16	0.1413
z4	1	-0.9090	0.3539	-1.6026	-0.2153	6.60	0.0102
z5	1	1.2512	0.4345	0.3997	2.1027	8.29	0.0040
Scale	1	1.3697	0.1541	1.0987	1.7076		
Weibull Shape	1	0.7301	0.0821	0.5856	0.9101		

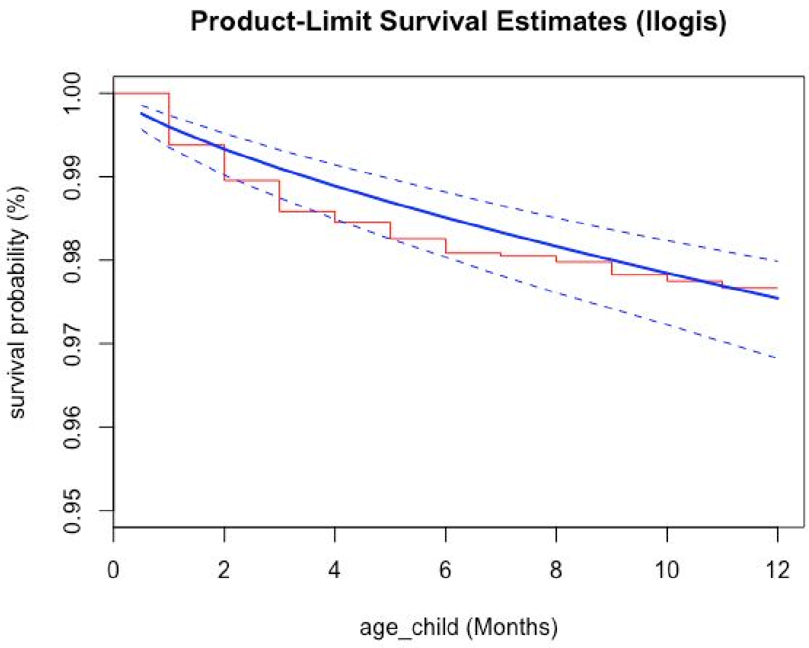


Product-Limit Survival Estimates (Weibull)

Table 7. Result of the llogistic model



Analysis of Maximum Likelihood Parameter Estimates							
Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	7.9096	0.7048	6.5281	9.2910	125.93	<.0001
z1	1	0.8724	0.4115	0.0659	1.6790	4.49	0.0340
z2	1	-0.9180	0.3409	-1.5863	-0.2498	7.25	0.0071
z3	1	-0.5164	0.3480	-1.1984	0.1656	2.20	0.1378
z4	1	-0.9178	0.3550	-1.6135	-0.2221	6.69	0.0097
z5	1	1.2547	0.4329	0.4063	2.1030	8.40	0.0037
Scale	1	1.3534	0.1514	1.0869	1.6851		



Conclusion

After study the distribution of 12 risk factors in the pneumonia data, we use semiparametric model to build multiple variance models. We use stepwise selection to pick significant factors, and choose the best model by the one with the smallest AIC value. Hence, we selected five predictors as the important risk factors in stepwise selection and the smallest AIC, including mom's age, cigarette, region, number of sibling and breastfeeding.

Moreover, we constructed each 5 factors to dummy variables and test the hypothesis H0: Beta=0 using the likelihood ratio, and Wald tests. Using Cox proportional hazards models, four factors were identified as the important risks associated with the hospitalized pneumonia incidence in infants. They are mom's age, cigarette, number of sibling and breastfeeding. We also use non-parametric model with log-rank and wilcoxon test to check and confirm the proportional hazards result.

Furthermore, we use two parametric model, Weibull and llogistic modelling to fit with our multiple model with 5 risk factor. Both of them verified that the mom' age, cigarette, number of sibling and breastfeeding are significant risk factors in our study.

Although region is not a significant factor if all four regions were considered, the significant difference between regions 2 (north central) compared with the other regions were observed. The rest factors were not identified as the significant ones in this study. Therefore, an infant with breastfeeding, one or no sibling, older and non-smoking mom has lower risk of hospitalized pneumonia. This study shows mother's behavior and children's environment are important factor associate to hospitalized for pneumonia.

References

1. Bulla2. A, Hitze KL. Acute respiratory infections: a review. Bull World Health Organ 1978;56:481-98. PMID:308414
2. Leowski3. J. Mortality from acute respiratory infections in children under 5 years of age: global estimates. World Health Stat Q 1986;39:138-44. PMID:3751104
3. Kramer MS, Kakuma R: The optimal duration of exclusive breastfeeding.Cochrane Database Syst Rev 2012, 8:CD003517.
4. Hanson LA, Korotkova M, Telemo E: Breast-feeding, infant formulas, and the immune system. Ann Allergy Asthma Immunol 2003, 90(6 Suppl 3):59-63.
5. Survival analysis : techniques for censored and truncated data, John P Klein, Melvin L Moeschberger, New York, 2003
6. Brogan, R. J. (Ed.). (2017, December). Pneumonia (for Parents). Retrieved from https://kidshealth.org/en/parents/pneumonia.html
7. Pneumonia in Children. (n.d.). Retrieved from https://www.cedars-sinai.org/health-library/diseases-and-conditions---pediatrics/p/pneumonia-in-children.html

Appendix

SAS code

```
data f113;
```

```
input age_child indic_hos Age_mom Urban Alcohol Cigarette Region Poverty Bweight
Race_mom Edu_mom N_sibling M_weaned M_solidfood Age_in_hos;
cards;
…
;
run;
data mydata;
set f113;
run;

/*observe the data*/
proc univariate data=mydata;
histogram;
run;

/*Multiple model selection*/
proc phreg data = mydata;
model age_child*indic_hos(0)= Age_mom Urban Alcohol Cigarette Region Poverty Bweight
Race_mom Edu_mom N_sibling M_weaned M_solidfood
/ ties=breslow;
run;

/*stepwise selection*/
proc phreg data = mydata;
model age_child*indic_hos(0)= Age_mom Urban Alcohol Cigarette Region Poverty Bweight
Race_mom Edu_mom N_sibling M_weaned M_solidfood
/selection=stepwise slentry=0.25 slstay=0.07 details ties=breslow;
run;

data mydata1;
set f113;
if (Age_mom ge 23)   then z1=1; else z1=0;
if (Cigarette ge 1)  then z2=1; else z2=0;
if (Region=2)              then z3=1; else z3=0;
if (N_sibling ge 1)  then z4=1; else z4=0;
if (M_weaned ge 1)   then z5=1; else z5=0;
run;

proc phreg data = mydata1;
model age_child*indic_hos(0)= z1 z2 z3 z4 z5 / ties=breslow;
run;

proc phreg data = mydata1;
```

```
model age_child*indic_hos(0)= z1/ ties=breslow;
run;

proc phreg data = mydata1;
model age_child*indic_hos(0)= z2/ ties=breslow;
run;

proc phreg data = mydata1;
model age_child*indic_hos(0)= z3/ ties=breslow;
run;

proc phreg data = mydata1;
model age_child*indic_hos(0)= z4/ ties=breslow;
run;

proc phreg data = mydata1;
model age_child*indic_hos(0)= z5/ ties=breslow;
run;

/*nonparametric test*/
proc lifetest data=mydata1 plots= LLS;
time age_child*indic_hos(0);
strata z1;
run;

proc lifetest data=mydata1 plots= LLS;
time age_child*indic_hos(0);
strata z2;
run;

proc lifetest data=mydata1 plots= LLS;
time age_child*indic_hos(0);
strata z3;
run;

proc lifetest data=mydata1 plots= LLS;
time age_child*indic_hos(0);
strata z4;
run;

proc lifetest data=mydata1 plots= LLS;
time age_child*indic_hos(0);
strata z5;
```

```
run;

/* K-M Plot for high risk group vs low risk group comparison*/
data mydata2;
set mydata;
if (Age_mom ge 23)  then group=1;
if (Cigarette=0)            then group=1;
if (N_sibling le 1) then group=1;
if (M_weaned ge 1)  then group=1;

if (Age_mom lt 23)   then group=2;
if (Cigarette ge 1)  then group=2;
if (N_sibling gt 1)  then group=2;
if (M_weaned le 1)   then group=2;

proc lifetest data=mydata2 plots= survival(test);
time age_child*indic_hos(0);
strata group;
run;

/*parametric model exploration*/
proc lifereg data = mydata1;
model age_child*indic_hos(0)=z1 z2 z3 z4 z5
/ covb distribution = weibull;
run;
proc lifereg data = mydata1;
model age_child*indic_hos(0)=z1 z2 z3 z4 z5
/ covb distribution = llogistic;
run;
```