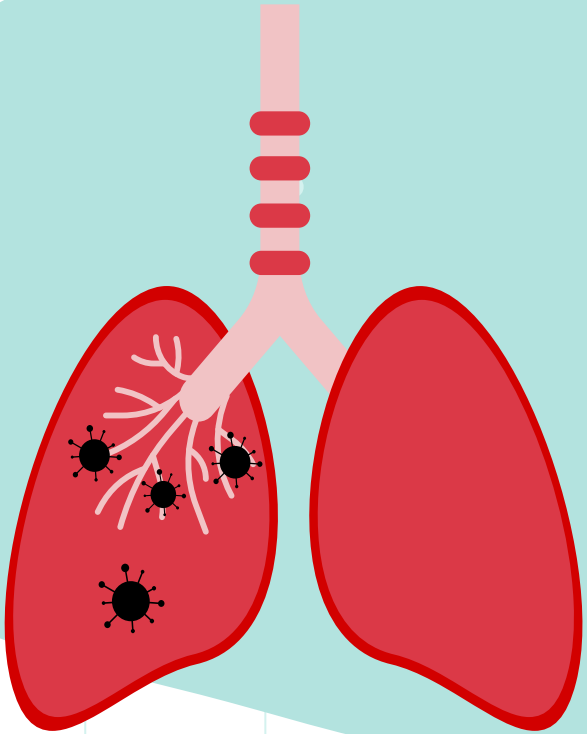
Impact of Sibling Count on Hospitalized Pneumonia Risk in Infants

Esther Ki | BSTA 662



PNEUMONIA

- The single largest infectious cause of death in children worldwide
- Mostly spreads by direct person-to-person contact via respiratory droplets
- According to WHO, the risk factors for pneumonia in children include
  - **Living in crowded homes**
  - **Indoor air pollution caused by cooking and heating with biomass fuels**
  - **Parental smoking**
  - **Undernourishment especially in infants not exclusively breastfed**

"Pneumonia in Children." *World Health Organization*, World Health Organization Fact Sheets, 11 Nov. 2022

DATA: TIME TO HOSPITALIZED PNEUMONIA IN YOUNG CHILDREN

- Obtained from our textbook (Klein 2nd)
  - Collected from annual personal interviews conducted as part of the National Longitudinal Survey of Youth (NLSY, 1995) from 1979 to 1986
- 3470 observations, 15 variables
- 3397 censored, 73 uncensored
- Information regarding the child, demographic characteristics and health behavior measures of the mother were recorded

chldage	Age child had pneumonia, months
hospital	Indicator for hospitalization for pneumonia (1=yes, 0=no
mthage	Age of the mother, years
urban	Urban environment for mother (1=yes, 0=no)
alcohol	Alcohol use by mother during pregnancy
smoke	Cigarette use by mother during pregnancy
region	Region of the country (1=northeast, 2=north central, 3=south, 4=west)

poverty	Mother at poverty level (1=yes, 0=no)
bweight	Normal birthweight (>5.5 lbs.) (1=yes, 0=no)
race	Race of the mother (1=white, 2=black, 3=other)
education	Education of the mother, years of school
nsibs	Number of siblings of the child
wmonth	Month the child was weaned
sfmonth	Month the child on solid food
agepn	Age child in the hospital for pneumonia, months

RESEARCH OBJECTIVE

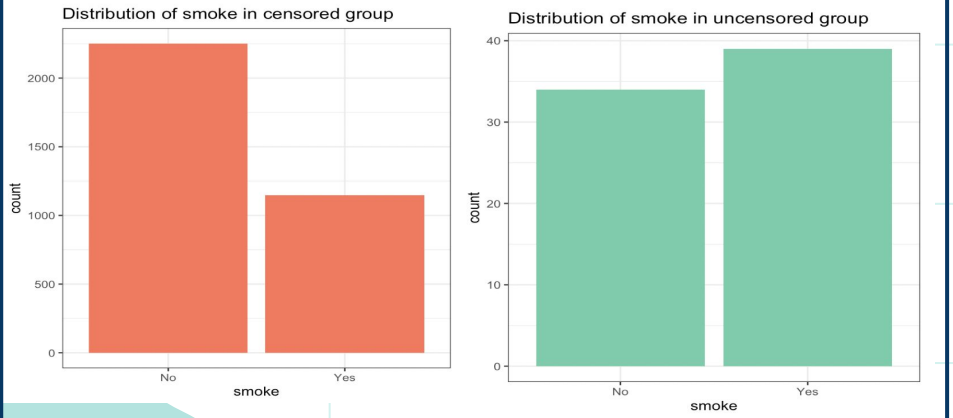
- Analyze time to hospitalized pneumonia in young children to see whether **the number of siblings** affects infants to have **hospitalized pneumonia in the first year of life**
- Observed Survival Time: From birth to age the child is in the hospital for pneumonia (agepn)
- Status variable: Indicator for hospitalization for pneumonia (hospital)

PRELIMINARY ANALYSIS

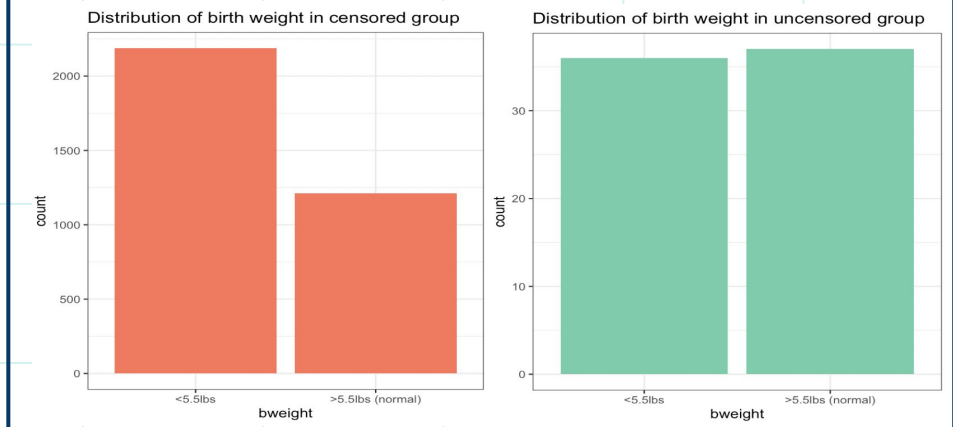
hospital	count	chldage_mean	mthage_mean	urban_mean	alcohol_mean	smoke_mean	poverty_mean	bweight_mean	nsibs_mean	wmonth_mean	sfmonth_mean	agepn_mean
0	3397	9.979394	21.65823	0.7624374	0.6655873	0.4362673	0.9225787	0.3564910	0.6717692	1.9552546	1.136885	7.847218
1	73	3.589041	20.82192	0.6712329	0.6164384	0.6849315	0.9041096	0.5068493	0.9452055	0.5479452	0.369863	8.671233

- Only **~2%** are uncensored → **Visualized censored and uncensored group separately**
  - Coral - censored, Green - uncensored

smoke

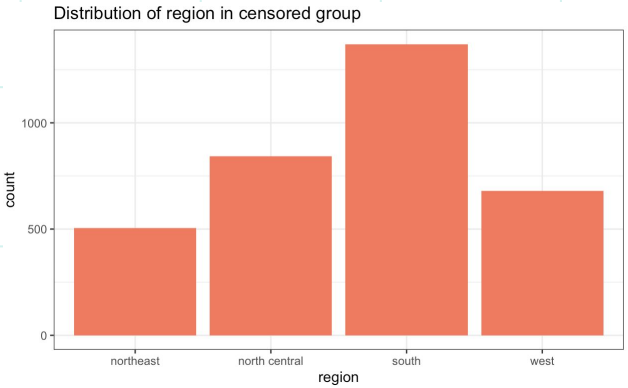


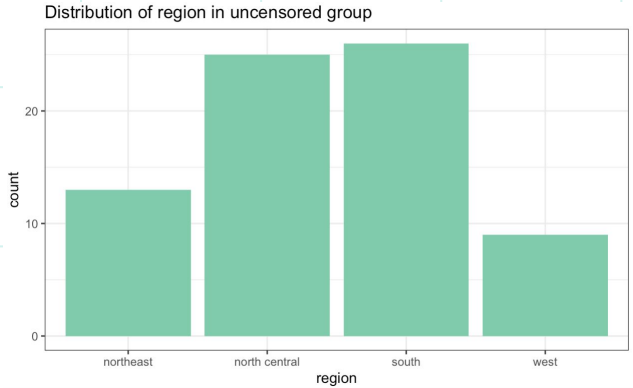
bweight



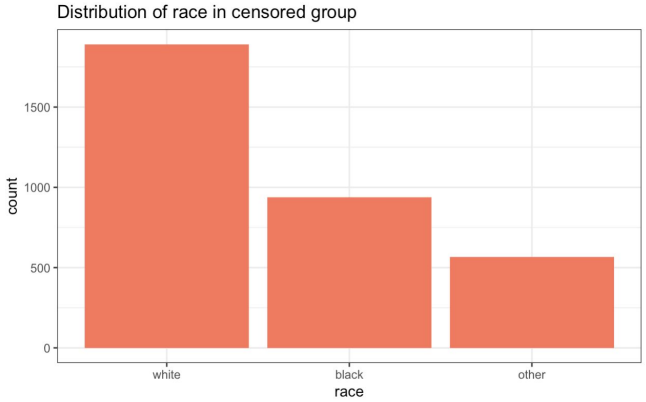
PRELIMINARY ANALYSIS: NOMINAL CATEGORICAL VARIABLES

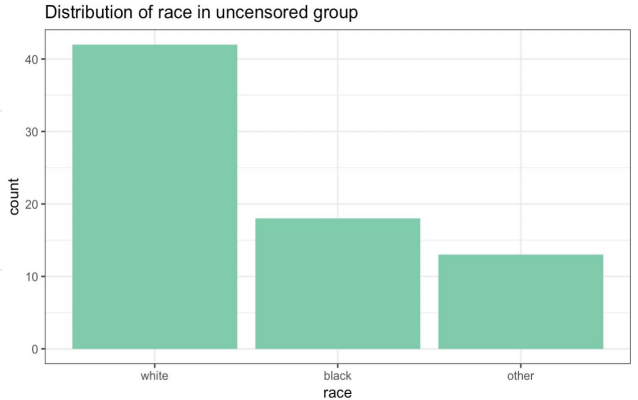
region





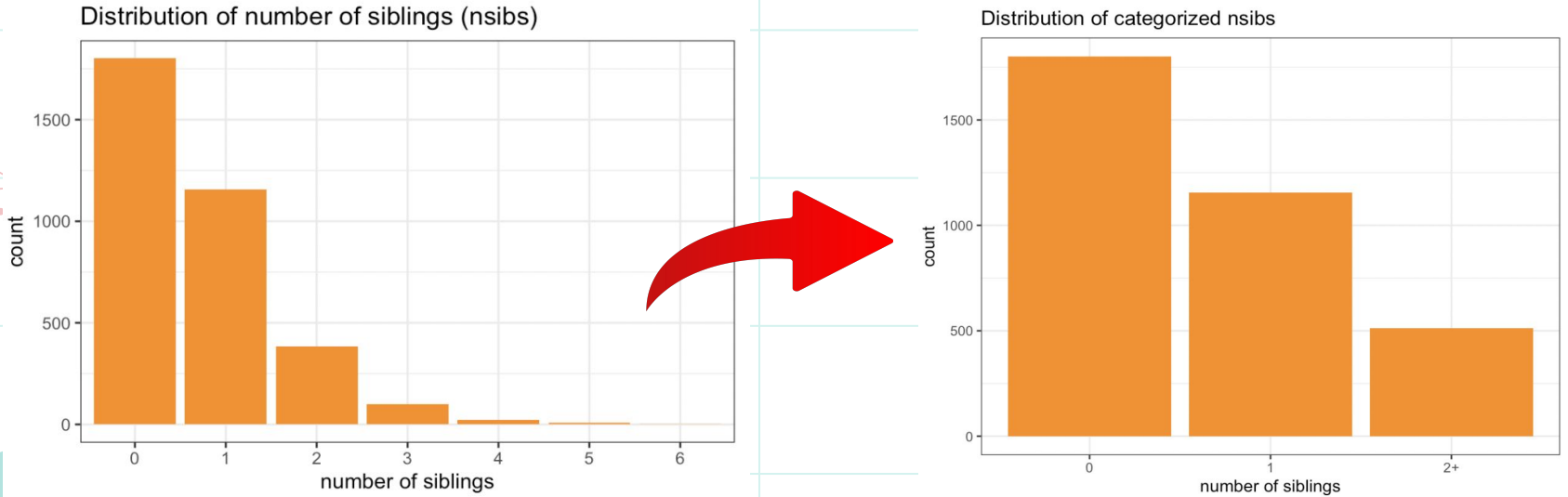
race





PRELIMINARY ANALYSIS: NUMBER OF SIBLINGS

- nsibs : integer variable ranging from 0 to 6
- Categorized into 3 levels for graphical checking → 0, 1, and 2+



NONPARAMETRIC TEST

Forward Stepwise Sequence of Chi-Squares for the Wilcoxon Test

Variable	DF	Chi-Square	Pr > Chi-Square	Chi-Square Increment	Pr > Increment
chldage	1	894.2	<.0001	894.2	<.0001
mthage	2	905.2	<.0001	10.9946	0.0009
nsibs_fct	3	910.9	<.0001	5.6819	0.0171
smoke	4	911.7	<.0001	0.8112	0.3678
urban	5	912.5	<.0001	0.7421	0.3890
sfmonth	6	912.7	<.0001	0.2090	0.6476
poverty	7	912.8	<.0001	0.1713	0.6790
education	8	913.0	<.0001	0.1401	0.7082
alcohol	9	913.0	<.0001	0.0290	0.8648
wmonth	10	913.0	<.0001	0.0160	0.8993
bweight	11	913.0	<.0001	0.00735	0.9317

Forward Stepwise Sequence of Chi-Squares for the Log-Rank Test

Variable	DF	Chi-Square	Pr > Chi-Square	Chi-Square Increment	Pr > Increment
chldage	1	1008.3	<.0001	1008.3	<.0001
mthage	2	1023.1	<.0001	14.7855	0.0001
nsibs_fct	3	1029.5	<.0001	6.3925	0.0115
smoke	4	1030.5	<.0001	1.0667	0.3017
urban	5	1031.4	<.0001	0.8689	0.3513
sfmonth	6	1031.7	<.0001	0.2786	0.5976
education	7	1031.8	<.0001	0.1169	0.7324
poverty	8	1031.9	<.0001	0.1144	0.7352
wmonth	9	1031.9	<.0001	0.0385	0.8445
alcohol	10	1032.0	<.0001	0.0323	0.8575
bweight	11	1032.0	<.0001	0.00542	0.9413

- smoke, bweight, wmonth, sfmonth.. not significant?
- From data description,
  - chldage: Age child had pneumonia, months
  - agepn (survival time): Age child in the hospital for pneumonia, months
  - **66% duplication rate**

NONPARAMETRIC TEST - WITHOUT CHLDAGE

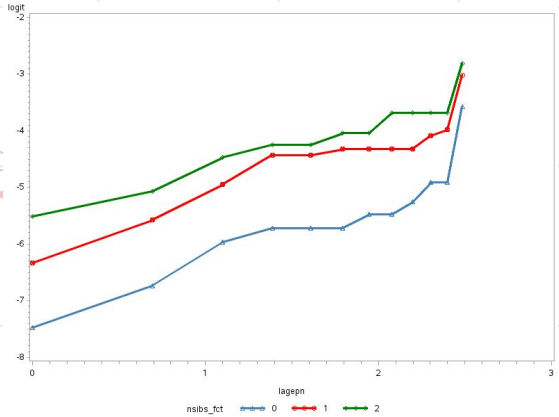
Forward Stepwise Sequence of Chi-Squares for the Wilcoxon Test					
Variable	DF	Chi-Square	Pr > Chi-Square	Chi-Square Increment	Pr > Increment
wmonth	1	10.7658	0.0010	10.7658	0.0010
nsibs_fct	2	19.4204	<.0001	8.6547	0.0033
smoke	3	26.4286	<.0001	7.0081	0.0081
education	4	29.0888	<.0001	2.6602	0.1029
urban	5	31.4074	<.0001	2.3186	0.1278
bweight	6	32.3733	<.0001	0.9659	0.3257
alcohol	7	33.1379	<.0001	0.7646	0.3819
sfmonth	8	33.4577	<.0001	0.3198	0.5717
mthage	9	33.6622	0.0001	0.2045	0.6511
poverty	10	33.6649	0.0002	0.00277	0.9580

Forward Stepwise Sequence of Chi-Squares for the Log-Rank Test					
Variable	DF	Chi-Square	Pr > Chi-Square	Chi-Square Increment	Pr > Increment
wmonth	1	10.5497	0.0012	10.5497	0.0012
nsibs_fct	2	19.0574	<.0001	8.5076	0.0035
smoke	3	26.0914	<.0001	7.0340	0.0080
education	4	28.7490	<.0001	2.6576	0.1031
urban	5	31.0416	<.0001	2.2926	0.1300
bweight	6	32.0222	<.0001	0.9806	0.3220
alcohol	7	32.7800	<.0001	0.7578	0.3840
sfmonth	8	33.1035	<.0001	0.3235	0.5695
mthage	9	33.2981	0.0001	0.1946	0.6591
poverty	10	33.3003	0.0002	0.00220	0.9626

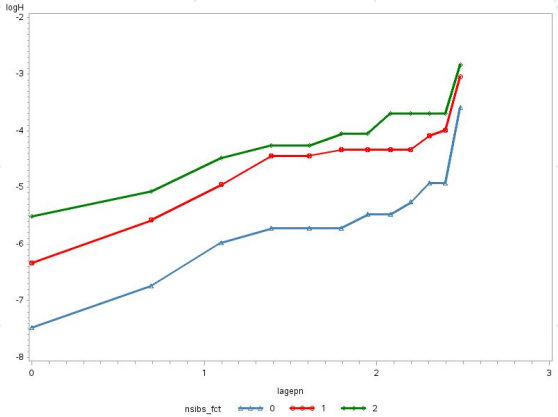
- smoke and wmonth became significant without chldage
- Matched better with the background knowledge and preliminary analysis
- Decided to move forward without chldage in further analysis

CHECKING THE DISTRIBUTION OF SURVIVAL TIME

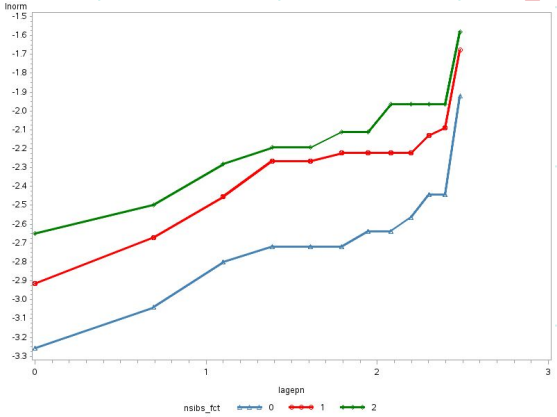
- Considering the effect of number of siblings



Log-logistic



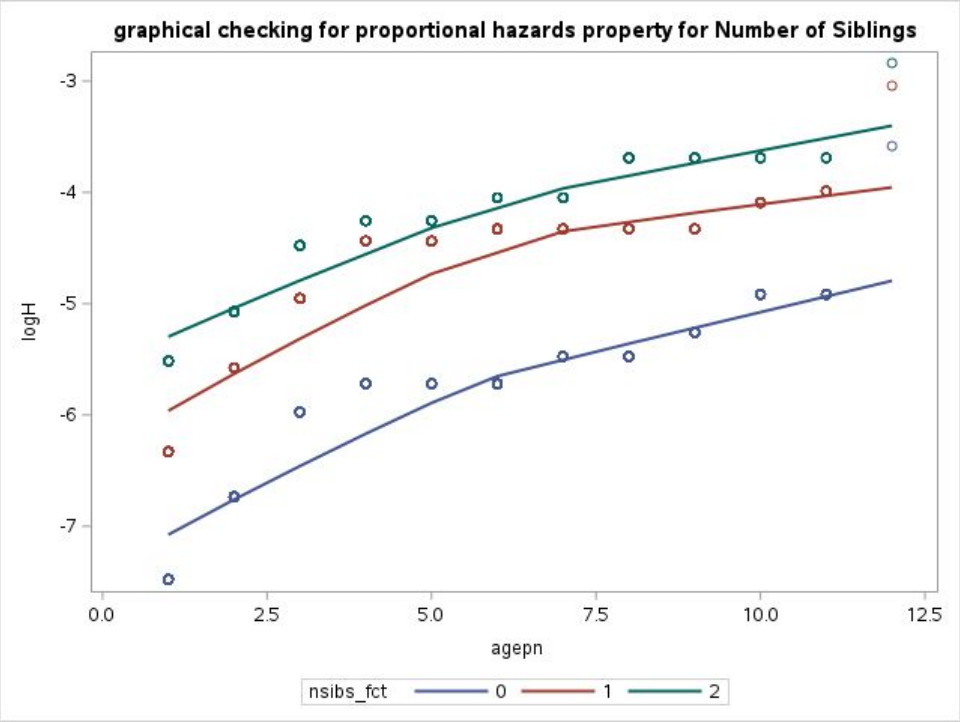
Weibull



Log-normal

None of them are appropriate due to the nonlinearity at the tails

CHECKING PROPORTIONAL HAZARDS PROPERTY



BUILDING COX REGRESSION MODEL

Analysis of Maximum Likelihood Estimates								
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label
nsibs		1	0.32047	0.11558	7.6886	0.0056	1.378	
smoke		1	0.38721	0.15148	6.5341	0.0106	1.473	
wmonth		1	-0.22563	0.08035	7.8842	0.0050	0.798	

Summary of Backward Elimination					
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq
1	poverty	1	11	0.0292	0.8643
2	sfmonth	1	10	0.0683	0.7938
3	education	1	9	0.1916	0.6615
4	alcohol	1	8	0.7653	0.3817
5	region	3	7	3.5858	0.3098
6	bweight	1	6	1.0699	0.3010
7	race	2	5	3.0634	0.2162
8	mthage	1	4	1.6073	0.2049
9	urban	1	3	2.1677	0.1409

- Backward elimination
  - **nsibs, smoke, wmonth** are significant variables

Final model with **nsibs, wmonth,** and **smoke**

CHECKING MODEL SIGNIFICANCE

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	31.8181	3	<.0001
Score	27.7234	3	<.0001
Wald	24.9247	3	<.0001

The final model is statistically significant in predicting survivorship.

CHECKING PROPORTIONAL HAZARDS PROPERTY

Supremum Test for Proportionals Hazards Assumption				
Variable	**Maximum Absolute Value**	**Replications**	**Seed**	**Pr > MaxAbsVal**
nsibs	0.6146	1000	976258287	0.3640
wmonth	1.0947	1000	976258287	0.1630
smoke	0.7773	1000	976258287	0.1790

Proportional hazards assumption holds for all three variables.

GOODNESS OF FIT TEST

Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	831.105	799.287
AIC	831.105	805.287
SBC	831.105	812.159

<Final model>

Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	831.105	786.376
AIC	831.105	818.376
SBC	831.105	855.023

<Full model>

H_0: final model with nsibs, wmonth, and smoke

H_1: full model with all covariates

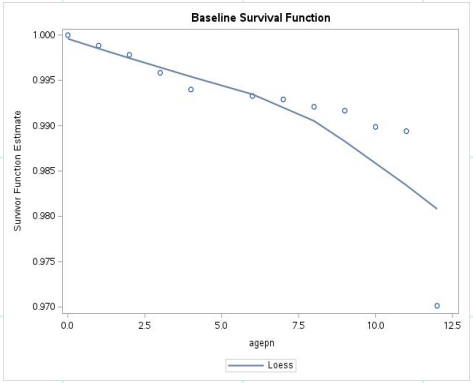
- $\Lambda = 2(LogL_1 - LogL_0) = [-2LogL_0]-[-2LogL_1] = 799.287 - 786.376 = 12.911$
- $df = df(H_1) - df(H_0) = 16 - 3 = 13$

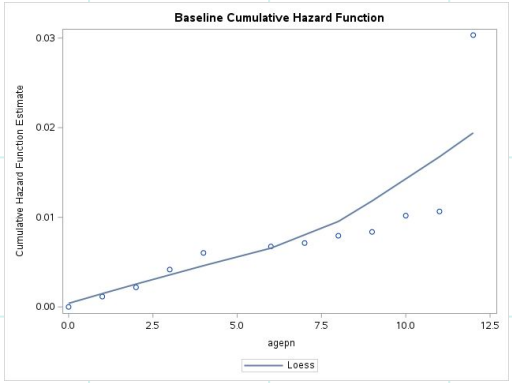
  $\rightarrow \Lambda = 12.911$ and df=13 gives the p-value = **0.455**

Therefore, the final model fits the data as well as the full model

FITTED FINAL COX REGRESSION MODEL

Analysis of Maximum Likelihood Estimates						
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
nsibs	1	0.32055	0.11560	7.6893	0.0056	1.378
wmonth	1	-0.22563	0.08036	7.8843	0.0050	0.798
smoke	1	0.38728	0.15149	6.5354	0.0106	1.473



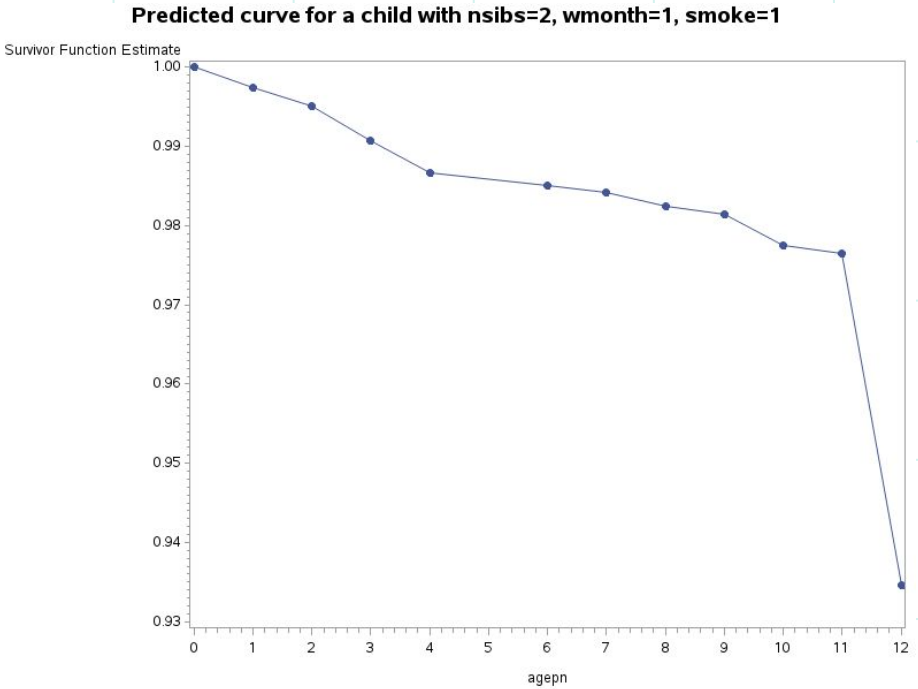


$$\log[h(y|x)] = \log[h_0(y)] + 0.321\text{nsibs} - 0.226\text{wmonth} + 0.387\text{smoke}$$

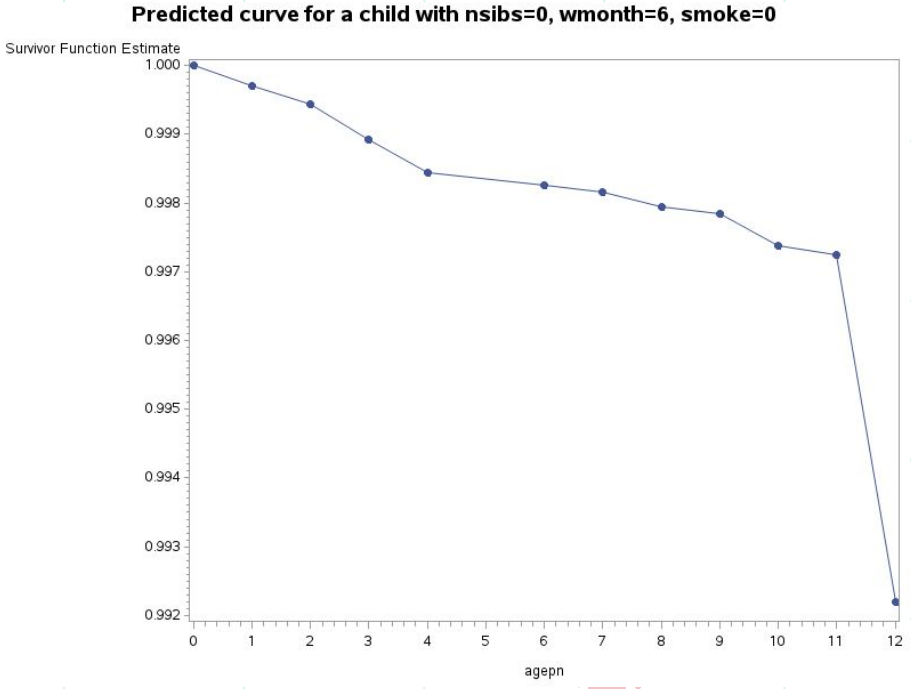
- RR of hospitalized pneumonia for a child is ~1.378 times the original risk for every additional sibling.
- RR of hospitalized pneumonia is ~0.798 times the original risk for every 1-month increase in breastfeeding period
- RR of hospitalized pneumonia is ~1.473 times the original risk if the mother used cigarette during pregnancy.

PREDICTED SURVIVAL FUNCTION

Child 1



Child 2



CONCLUSION

- The number of siblings affects the time to hospitalized pneumonia in children
  - Every additional sibling, the risk increases by ~37.8%
  - Increased person-to-person contact and air pollution level
- Duration of breastfeeding and maternal smoking during pregnancy are also highly significant factors
- Implications for public health interventions
  - Promote household hygiene practices and mitigate household environment factors
  - Promote breastfeeding initiation and continuation
  - Discourage maternal smoking during pregnancy

THANK YOU!