

Comparing Results from Cox Proportional Hazards Models using SUDAAN® and SAS® Survey Procedures to a Logistic Regression Model for Analysis of Influenza Vaccination Coverage

Zhai, Y., Kahn, K.E., O'Halloran, A., Leidos, Inc., Santibanez, T.A., CDC

ABSTRACT

The National Immunization Survey-Flu (NIS-Flu) is an ongoing, national telephone survey of households with children in the United States used to measure influenza vaccination coverage. The data collected by NIS-Flu have similarities to data typically analyzed using survival analytic procedures. Estimates of vaccination coverage from the NIS-Flu survey are calculated using Kaplan-Meier survival analysis procedures to account for censoring of the data. However, multivariable models to examine socio-demographic characteristics associated with receipt of influenza vaccination using NIS-Flu data have typically been done using logistic regression rather than using survival analytic methods. The logistic regression approach ignores the time-to-event and censoring characteristics of the influenza data and assumes that censoring throughout the survey period occurs equally among the comparison groups of interest. If this assumption is untrue, measures of association for receipt of influenza vaccine could be biased. Another approach used to address the censoring issues in NIS-Flu data is to restrict the logistic regression analysis to interviews conducted at the end of the vaccination period (i.e., March-June) when it is unlikely that many respondents would be vaccinated after the time of interview. However, this approach ignores a large amount of data, results in a reduced precision of estimates, and potentially exacerbates recall bias.

This project assessed the feasibility, methods, and advantages of using a Cox proportional hazards model as opposed to a logistic regression model using full NIS-Flu 2013-14 season data and a logistic regression model using end of vaccination period data. This project also compared the results of Cox proportional hazards model from SUDAAN SURVIVAL and from SAS SURVEYPHREG procedures.

Despite the slight underestimate of the associations between vaccination status and demographic characteristics, the logistic model remains a reasonable alternative to the Cox proportional hazards model in analyzing the NIS-Flu data. The SAS SURVEYPHREG and the SUDAAN SURVIVAL produced nearly identical Cox proportional hazards model results. Conclusions drawn based on the results from logistic regression and either of the Cox proportional hazards models using full or post-vaccination period NIS-Flu data are comparable.

INTRODUCTION

The influenza vaccination data collected by the National Immunization Survey-Flu (NIS-Flu) has similarities to data typically analyzed using survival analytic procedures. Children for whom parents answer "No" to the question regarding whether their child had received an influenza vaccination are censored at the date of the telephone interview, indicating the child was not vaccinated by the time of interview. Children can, and some do, receive an influenza vaccination later in the influenza season after the NIS-Flu interview, especially for interviews occurring during the active vaccination months. However, the NIS-Flu does not follow-up with children to determine if vaccination occurred after interview. To account for this censoring in the data, estimates of vaccination coverage from the NIS-Flu are calculated using Kaplan-Meier survival analysis procedures (CDC, 2010a, 2010b, 2013b; Lu, Rodriguez-Lainz, O'Halloran, Greby, & Williams, 2014). However, multivariable models to examine socio-demographic characteristics associated with receipt of influenza vaccination using data from vaccination surveys have typically been done using logistic regression in SUDAAN with the RLOGIST procedure rather than using survival analytic methods (Lu et al., 2014). The logistic regression model, although familiar to most analysts and researchers, ignores the time-to-event and censoring characteristics of the influenza data. The approach also assumes that censoring throughout the survey period occurs equally among the comparison groups of interest. If this assumption is untrue, measures of association for receipt of influenza vaccine could be biased. Another approach that has been used for dealing with the censoring issues with influenza survey data is to restrict the logistic regression analysis to interviews conducted at the end of the vaccination period, when it is unlikely that many respondents would be vaccinated after the time of interview (Santibanez, 2012). However, this approach ignores a large amount of collected data, resulting in a reduced sample size and reduced precision of estimates, and potentially exacerbates recall bias since the interviews included in the analyses take place months after vaccination for many respondents.

The SURVIVAL procedure in SUDAAN utilizes the Cox proportional hazards model with complex survey data (RTI, 2009). SAS has extended its software capabilities for handling complex survey data; after releasing SURVEYMEANS, SURVEYREG, SURVEYFREQ, and SURVEYLOGISTIC in version 8.0 and 9.0, in 2010, an experimental SURVEYPHREG procedure was released with the SAS/STAT 9.22 giving SAS the capability of performing Cox proportional hazards regressions for complex survey data (SAS, 2010). Cheng compared logistic

regression procedures in SAS (SURVEYLOGISTIC) and SUDAAN (RLOGIST), and found that SAS and SUDAAN produced identical parameters and variance estimates for complex survey data with large sample size (Chen, 2006). However, to our knowledge, the similarities between the Cox proportional hazards model in SAS (SURVEYPHREG) and SUDAAN (SURVIVAL) have not been explored and reported.

This paper assessed the feasibility, methods, and advantages of using a Cox proportional hazards model as opposed to a logistic regression model to examine socio-demographic variables associated with receipt of influenza vaccination. Using NIS-Flu influenza and socio-demographic data from the 2013-14 influenza season, results of a Cox proportional hazards model were compared to those from a logistic regression model using all of the months of NIS-Flu interviews (October-June interviews) and a logistic regression model using only interviews from the post-vaccination period (March-June interviews). This paper also describes how to perform the Cox proportional hazards model using both the SUDAAN SURVIVAL and the SAS SURVEYPHREG procedures and compares the results from these two procedures. The implications of these comparisons may extend to other complex survey data sources used for assessing associations with influenza vaccination, such as the National Health Interview Survey (NHIS) and Behavioral the Risk Factor Surveillance System (BRFSS).); however, the comparisons should be tested before use.

DESCRIPTION OF THE SURVEY DATA ANALYZED FOR THIS STUDY

The National Immunization Survey-Flu (NIS-Flu) is an ongoing, national, list-assisted random-digit-dialed dual frame (land line and cellular) telephone survey of households with children in the United States used to measure influenza vaccination coverage. It includes three components: the NIS for children 19–35 months, the NIS-Teen for children 13–17 years, and the NIS child influenza module for children 6–18 months and 3–12 years identified during the screening of households for the NIS and NIS-Teen. The Council of American Survey and Research Organizations (CASRO) response rates for the NIS-Flu for the 2013-2014 influenza season ranged from 58.6%–63.4% for landline, and 32.1%–33.5% for cellular telephones across the three survey components (CDC, 2013a, 2014; Frankel, 1983). The sample included a total of 107,550 children 6 months to 17 years. To make the results comparable to an ongoing full vaccination coverage project, only 65,848 children 6 months to 8 years were included in the analyses.

Respondents ≥ 18 years were asked if their child had received an influenza vaccination since July 1, 2013 and, if so, in which month and year. For vaccinated children with missing month and year of vaccination (7.7%), this information was imputed from donor pools matched for week of interview, age group, state of residence, and race/ethnicity. Children reported to have received influenza vaccine since July 1, 2013 were considered vaccinated. For logistic regression analysis, the outcome was simply whether children were vaccinated (1="Yes", 0="No"). For the Kaplan-Meier survival analysis, the vaccination coverage estimates were calculated with the event of being vaccinated and event time defined as the month when a subject received the first dose of influenza vaccine. Censoring was handled in the following way: 1) if reported unvaccinated at the time of interview, subjects were censored and the time was reset as the month prior to the interview month; 2) if reported vaccinated during the interview month, subjects were censored and the censored time was set to the month prior to interview month (treated as unvaccinated prior to interview month). These same definitions of event, event time, and censoring were likewise used for the Cox proportional hazard analyses. The NIS-flu data for the 2013-14 influenza season included interviews conducted from October 2013 to June 2014. We considered the post-vaccination period data to include only interviews conducted from February 2014 to June 2014. The covariates of interest included child's age, child's race/ethnicity, mother's education, income/poverty level, metropolitan statistical area (MSA), and number of children in the household. Wald chi-square tests were used to test the significance of the associations between influenza vaccination status and demographic characteristics. All estimates, along with 95% confidence intervals (CIs), were performed using SAS 9.3 (SAS institute, Inc., Cary, NC) and SAS-callable SUDAAN 11.0.0 (Research Triangle Institute, Research Triangle Park, NC). All tests were two-sided with a significance level of 0.05.

LOGISTIC REGRESSION MODELS SUBPOPULATION SYNTAX

Performing analyses only on a subpopulation of a survey sample is very common. In this study, for all analyses we included the subset of children 6 months to 8 years, and for one of the logistic regression models performed additional subsetting to include only children with interviews in the post-vaccination period. Subsetting the survey sample in a data step prior to the analytic procedure or using a 'where' or 'by' statement inside the analytic procedure to restrict analysis to certain subpopulation ignores the randomness and variability of the subpopulation across the strata of the sample design. This generally results in underestimating the variances (Cochran, 1997). The correct approach involves using the entire sample in the analysis and also taking into account the sample size of the subpopulation of interest. All SAS survey procedures offer a DOMAIN statement for subpopulation analysis. With a DOMAIN statement in the survey procedures, SAS first runs a regular analysis with the entire dataset, then performs a subpopulation analysis in each of the domains while including the entire dataset in variance estimates. If only one subpopulation analysis is desired, a SAS technique that assigns a minuscule weight to all non-domain cases in a dataset step prior to the analysis limits the analysis to a selected domain and eliminates unnecessary processing time for analyzing the entire sample and remaining subpopulations (Lewis, 2012). Correspondingly, SUDAAN offers a

SUBPOPN statement for the subpopulation analysis. Unlike SAS, the SUBPOPN statement in SUDAAN restricts the analysis for specific subpopulations only while obtaining the correct variance estimates.

LOGISTIC REGRESSION SYNTAX AND RESULTS

The syntax for performing logistic regression analyses with SAS SURVEYLOGISTIC and SUDAAN RLOGIST is shown in Appendix 1. Chen demonstrated how to specify options to obtain identical beta coefficients, p-values, odds ratios and nearly identical confidence intervals from SAS SURVEYLOGISTIC and SUDAAN RLOGIST (Chen, 2006) and we modelled our code after Chen's. One of the output options of the SUDAAN RLOGIST procedure is the calculation of Adjusted Prevalence (AP) and Adjusted Prevalence Ratios (APR) which is often preferable in studies when the outcome of interest is not rare such as the outcome of influenza vaccination status; the option to output APRs is currently not available with SAS SURVEYLOGISTIC but might be incorporated into the procedure in the future. The results of two logistic regression models from SUDAAN RLOGIST, one using the full data set (interviews from October through June) and another using the subset of data from the post-vaccination period (interviews February through June), are shown in table 1. Most of the confidence intervals around the APRs from the full model appear smaller than those from the post-vaccination subset model. Two variables that were significant with the full model (Mother's education level [<12 years] and number of children in the household [2 children]) were no longer significant with the subset model, likely due to the larger variance due to the smaller sample size.

COX PROPORTIONAL HAZARDS MODELS INTERVAL-CENSORING AND TIES HANDLING

Influenza vaccinations are available to the public in the United States beginning in July each influenza season with vaccinations continuing until June of the following calendar year; however, most vaccinations are received between August and December each influenza season. Interviews for the NIS-Flu are conducted from October through June to estimate flu vaccination coverage from July through May. With the NIS-Flu, only the month (not the exact date) of vaccination was collected for children reported as vaccinated because parents are more likely to remember the month of vaccination as oppose to the exact date or week of vaccination. Because of this, both event and censor time were measured as months which requires the data to be treated as discrete, interval-censoring data with mass ties in each of the intervals. Both SAS and SUDAAN offer EFRON and BRESLOW likelihood methods to handle ties. BRESLOW is the default method in SAS while the EFRON is the default in SUDAAN. The EFRON method is preferred over BRESLOW for instances when there are many ties and censoring, hence it was chosen in our analyses (Hertz-Picciotto & Rockhill, 1997). In addition, SUDAAN offers INTERVALS and LAMBDA statements to define the discrete proportional hazards model.

EVALUATION OF THE PROPORTIONALITY ASSUMPTION

The semi-parametric Cox proportional hazards model requires no assumption about the form of the baseline hazard. However, the proportionality of the hazards had to be assessed before the model could be employed. One commonly used method is the plot of $\log[-\log(\text{survival})]$ versus $\log$ of survival time (log-negative log plot). If the categorical dependent variable satisfies the proportional hazards assumption, the plot should result in roughly parallel lines. Another easily applied statistical method is to include interactions of the dependent variables and the survival time or log-transferred survival time as time-dependent variables in the model. If the hazard is proportional, the Wald F-test of these time-dependent variables will be insignificant in the model. Time-dependent variables should be tested separately and together to ensure the adjusted Wald F-test is also insignificant. Both plot and time-dependent variables methods can be easily preformed in SAS using LIFETEST AND SURVEYPHREG procedures, yet not so easily in SUDAAN SURVIVAL. The KAPMEIER/SURVIVAL procedure in SUDAAN does not offer output of survival plots or log-negative log plot, plus the SURVIVAL procedure does not support creating time-dependent variables within the procedure like SAS does. Instead, the SURVIVAL procedure in SUDAAN offers Computation of Schoenfeld residuals and Score residuals to allow users to evaluate "goodness of fit" and the validity of the proportional hazards assumption which is somewhat complicated to apply. All independent variables in our models satisfied the proportional hazards assumption with p-values from both adjusted and adjusted Wald F-test being greater than 0.05. The log-negative log plot showed roughly parallel lines for all independent variables. The syntax of testing the proportional hazards assumption using time-dependent covariates and log-negative log plot is shown in appendix 3. The test results are shown in table 2 and all log-negative log plots are shown in figure 1.

COX MODEL SYNTAX AND RESULTS

The syntax of performing Cox proportional hazards regression in SAS SURVEYPHREG and in SUDAAN SURVIVAL is shown in appendix 2. The results of Cox proportional hazards regression in SAS SURVEYPHREG and in SUDAAN SURVIVAL are shown in table 1. The results from the SAS SURVEYPHREG and the SUDAAN SURVIVAL are nearly identical in terms of Adjusted Hazard Ratio (AHR), confidence interval, and Chi-Square test p-values; the AHR differed only when carried to the second decimal place. Only one factor had different conclusions based on the different procedures. Based on the results from SAS, children of mothers with less than a high school education were significantly more likely to be vaccinated than children of mothers with >12 years of education, but no college degree ($p=0.03$), but the difference between these groups was not significant based on the results from

SUDAAN, ($p=0.07$). The APR's from the logistic regression models using full NIS-Flu data are generally slightly smaller than those AHR's from the Cox proportional hazards models with SAS SURVEYPHREG. The only difference in terms of significance was that, in the SAS SURVEYPHREG results, children from a family with unknown poverty status were significantly more likely to be vaccinated than children from a family above poverty but earned $\leq \$75,000$ annually ($p=0.05$), yet this difference was insignificant in the logistic regression model ($p=0.09$).

CONCLUSION

Cox proportional hazards models and logistic regression models are closely related and sometimes interchangeable. In recent years, the Cox proportional hazards model has gained in popularity over commonly used methods based on logistic regression due to the ability of the model to account for the distribution of time-to-event and incorporate censoring. In our study, the AHR's from the Cox proportional hazards model were generally larger than the APR's from the logistic model. If an AHR and an APR are both interpreted as the odds of receiving influenza vaccination, which is very common in practice, the results from the logistic model seem to underestimate the associations between vaccination status and demographic characteristics. Power to detect differences may be gained by accounting for the time- to-event and censoring in the NIS-Flu data. The SAS SURVEYPHREG and the SUDAAN SURVIVAL produced nearly identical results. Previous findings have indicated that when the follow-up period is short and the event occurrence rate is low, the regression coefficients of the logistic regression model approximate those of the Cox proportional hazards model (Abbott, 1985). From our observation, the conclusions drawn based on the results from logistic regression and Cox proportional hazards models using full or post-vaccination period NIS-Flu data are comparable. The logistic regression model seems to be a reasonable alternative to the Cox proportional hazards model in analyzing the NIS-Flu data, especially where the proportional hazards assumption is violated or difficult to verify. However, in situations when the data is in survival format and the proportional hazards assumption is met, Cox proportional hazards model is preferred over Logistic regression.

Restricting the logistic regression analysis to interviews conducted at the end of the vaccination period (i.e., March-June) could theoretically address the censoring issues. We observed slightly smaller APR's with wider confidence intervals from the post-vaccination period results which implied a reduced precision of estimates compared to full NIS-Flu data, but there was not enough evidence to indicate whether the model was more accurate with regards to parameter estimates than the model using full NIS-Flu data. In addition, since the full NIS-Flu data model did find significant differences that the post-vaccination model did not, any benefits gained from using post-vaccination data might be offset by the fact that the model ultimately loses power to detect differences. Future research is needed to examine the benefits of using post-vaccination period data rather than the full data.

APPENDIX 1: SURVEY LOGISTIC REGRESSION

SAS syntax:

```
PROC SURVEYLOGISTIC DATA=finaldata;
  STRATA hhs_region ;
  CLUSTER caseid;
  WEIGHT KM_WT_norm_I;
  CLASS sex_i(REF='Female') age_grp_4(REF='13-17y') raceeth4_i(REF='Non-Hispanic, White Only')
        education_mom(REF='Some College') inc_pov(REF='< $75,000, Above Poverty') msa3_i(REF=last)
        n_under18(REF='3+ children')/ORDER=INTERNAL PARAM=REF;
  MODEL new_flu_flag(EVENT='1') = sex_i age_grp_4 raceeth4_i education_mom inc_pov msa3_i n_under18/VADJUST=NONE;
  ODS OUTPUT ODDSRATIOS=ORS;

RUN;
```

SUDAAN syntax:

1. Full NIS-Flu data

```
PROC RLOGIST DATA=finaldata DESIGN=WR FILETYPE=SAS;
  NEST hhs_region caseid;
  WEIGHT KM_WT_norm_I;
  SUBPOPN age_dom=1;
  CLASS sex_i age_g3 raceeth4_i education_mom inc_pov msa3_i n_under18;
  REFLEVEL sex_i =2 age_g3 =3 raceeth4_i=3 education_mom=3 inc_pov=2 msa3_i=3 n_under18=3 ;
  MODEL new_flu_flag = sex_i age_g3 raceeth4_i education_mom inc_pov msa3_i n_under18;
  PREDMARG sex_i(2) age_g3(3) raceeth4_i(3) education_mom(3) inc_pov(2) msa3_i(3) n_under18(3)/ADJRR;
  TEST ADJWALDF;
  OUTPUT / BETAS=all FILENAME="BETA" FILETYPE=sas REPLACE;
  OUTPUT / RISK=all FILENAME="ORL" FILETYPE=sas REPLACE;
  OUTPUT / PREDRISK=all FILENAME="APR" FILETYPE=sas REPLACE;
  OUTPUT / PRED_MRG=all FILENAME="ADP" FILETYPE=sas REPLACE;

RUN;
```

2. Post-vaccination period NIS-Flu data (interview Feb-Jun)

```
PROC RLOGIST DATA=finaldata DESIGN=WR FILETYPE=SAS;
  NEST hhs_region caseid;
  WEIGHT KM_WT_norm_I;
  SUBPOPN age_dom=1 and month>=2 and month<=6;
  CLASS sex_i age_g3 raceeth4_i education_mom inc_pov msa3_i n_under18;
  REFLEVEL sex_i =2 age_g3 =3 raceeth4_i=3 education_mom=3 inc_pov=2 msa3_i=3 n_under18=3 ;
  MODEL new_flu_flag = sex_i age_g3 raceeth4_i education_mom inc_pov msa3_i n_under18;
  PREDMARG sex_i(2) age_g3(3) raceeth4_i(3) education_mom(3) inc_pov(2) msa3_i(3) n_under18(3)/ADJRR;
  TEST ADJWALDF;
  OUTPUT / BETAS=all FILENAME="BETA2" FILETYPE=sas REPLACE;
  OUTPUT / RISK=all FILENAME="ORL2" FILETYPE=sas REPLACE;
  OUTPUT / PREDRISK=all FILENAME="APR2" FILETYPE=sas REPLACE;
```

```

OUTPUT / PRED_MRG=all FILENAME ="ADP2" FILETYPE=sas REPLACE;
RUN;

```

APPENDIX 2: COX PROPORTIONAL HAZARDS REGRESSION

SAS syntax:

```

PROC SURVEYPHREG DATA=finaldata VARMETHOD=TAYLOR;
  DOMAIN age_dom;
  CLASS sex_i(REF='Female') age_g3(REF='5-8y') raceeth4_i(REF='Non-Hispanic, White Only')
    education_mom(REF='Some College') inc_pov(REF='< $75,000, Above Poverty') msa3_i(REF=last)
    n_under18(REF='3+ children')/ORDER=INTERNAL;
  STRATA hhs_region;
  CLUSTER caseid;
  WEIGHT KM_WT_norm_I;
  MODEL TIME1_I*censor1_I(0) = sex_i age_g3 raceeth4_i education_mom inc_pov msa3_i n_under18 /RISKLIMIT TIES=EFRON;
  ODS OUTPUT ParameterEstimates=paraest;
RUN;

```

SUDAAN syntax:

```

PROC SURVIVAL DATA=finaldata DESIGN=WR;
  SUBPOPN age_dom=1;
  NEST hhs_region caseid;
  WEIGHT KM_WT_norm_I;
  CLASS sex_i age_g3 raceeth4_i education_mom inc_pov msa3_i n_under18;
  REFLEVEL sex_i=2 age_g3=3 raceeth4_i=3 education_mom=3 inc_pov=2 msa3_i=3 n_under18=3;
  EVENT censor1_I;
  MODEL TIME1_I = sex_i age_g3 raceeth4_i education_mom inc_pov msa3_i n_under18/intervals=11;
  TEST WALDCHI SATADJCHI;
  OUTPUT / RISK=default FILENAME=HR REPLACE;
  OUTPUT / TESTS=default FILENAME=WTEST REPLACE;
  OUTPUT / BETAS=default FILENAME=BETAS REPLACE;
RUN;

```

APPENDIX 3: EVALUATE THE PROPORTIONALITY ASSUMPTION

1. Time-dependent variable method (e.g. sex)

```

PROC SURVEYPHREG DATA = finaldata1 VARMETHOD=TAYLOR;
  CLASS sex_i(REF='Female') age_g3(REF='5-8y') raceeth4_i(REF='Non-Hispanic, White Only') education_mom(REF='Some College')
    inc_pov(REF='< $75,000, Above Poverty') msa3_i(REF=last) n_under18(REF='3+ children')/ORDER=INTERNAL;
  STRATA hhs_region;
  CLUSTER caseid;
  WEIGHT KM_WT_norm_I_d;
  MODEL TIME1_I*censor1_I(0)=sex_i age_g3 raceeth4_i education_mom inc_pov msa3_i n_under18 sext /RISKLIMIT TIES=EFRON;
    sext = sex_i*TIME1_I;
    test_proportionality: TEST sext;

```

```
ODS OUTPUT ParameterEstimates=paraest2;  
ODS OUTPUT TESTS3= sext;  
RUN;
```

2. log-negative log plot method (e.g. sex)

```
PROC LIFETEST DATA=finaldata METHOD=KM PLOTS=LLS NOTABLE;  
  TIME TIME1_I*censor1_I(0);  
  STRATA sex_i;  
  FREQ KM_WT_norm_I;  
RUN;
```

Table 1. Comparison of logistic regression and Cox proportional hazards regression in assessing association of demographic characteristics with influenza vaccination coverage^a among children 6 months to 8 years, National Immunization Survey-Flu (NIS-Flu), United States, 2013–2014 influenza season

Demographic characteristics	Logistic model (Oct-Jun interviews)		Logistic model (Feb-Jun interviews)		Cox PH in SUDAAN (Oct-Jun interviews)		Cox PH in SAS (Oct-Jun interviews)	
	APR ^b (95% CI) ^c	p-value	APR (95% CI)	p-value	AHR ^d (95% CI)	p-value	AHR (95% CI)	p-value
Gender								
Male	1.00 (0.98, 1.03)	0.76	1.01 (0.97, 1.04)	0.7	1.02 (0.97, 1.07)	0.4	1.01 (0.97, 1.06)	0.58
Female	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
Age ^e								
6-23 months	1.21 (1.17, 1.25)	<0.001	1.15 (1.10, 1.19)	<0.001	1.24 (1.17, 1.32)	<0.01	1.26 (1.20, 1.34)	<0.01
2-4 years	1.07 (1.04, 1.11)	<0.001	1.06 (1.02, 1.10)	<0.01	1.16 (1.10, 1.23)	<0.01	1.15 (1.08, 1.21)	<0.01
5-8 years	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
Child's race/ethnicity								
Hispanic	1.16 (1.11, 1.21)	<0.001	1.15 (1.10, 1.20)	<0.001	1.22 (1.14, 1.30)	<0.01	1.23 (1.15, 1.31)	<0.01
Black, non-Hispanic	1.01 (0.96, 1.07)	0.63	1.01 (0.96, 1.07)	0.63	0.97 (0.90, 1.05)	0.49	0.98 (0.91, 1.06)	0.62
White, non-Hispanic	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
Other/multiple races, non-Hispanic	1.11 (1.06, 1.16)	<0.001	1.09 (1.04, 1.15)	<0.001	1.15 (1.07, 1.23)	<0.01	1.15 (1.07, 1.24)	<0.01
Mother's education level								
<12 years	1.07 (1.01, 1.14)	0.04	1.07 (1.00, 1.15)	0.06	1.10 (0.99, 1.21)	0.07	1.11 (1.01, 1.23)	0.03
12 years	1.04 (0.99, 1.09)	0.12	1.02 (0.97, 1.08)	0.39	1.05 (0.97, 1.13)	0.23	1.06 (0.98, 1.14)	0.15
>12 years, not college graduate	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
College graduate	1.16 (1.12, 1.21)	<0.001	1.19 (1.14, 1.24)	<0.001	1.29 (1.22, 1.38)	<0.01	1.29 (1.21, 1.38)	<0.01
Poverty status ^f								
Above poverty, >\$75,000/year	1.14 (1.10, 1.18)	<0.001	1.11 (1.07, 1.16)	<0.001	1.22 (1.15, 1.29)	<0.01	1.22 (1.16, 1.30)	<0.01
Above poverty, ≤\$75,000/year	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
At or below poverty level	1.06 (1.01, 1.11)	0.02	1.05 (1.00, 1.11)	0.04	1.11 (1.03, 1.19)	<0.01	1.11 (1.03, 1.20)	<0.01
Unknown	1.05 (0.99, 1.11)	0.09	1.07 (1.00, 1.13)	0.06	1.10 (1.01, 1.20)	0.03	1.09 (1.00, 1.19)	0.05

Household in MSA ^g								
MSA, Principle City	1.08 (1.03, 1.13)	<0.001	1.05 (1.00, 1.10)	0.06	1.14 (1.06, 1.23)	<0.01	1.15 (1.07, 1.24)	<0.01
MSA, Not Principle City	1.07 (1.03, 1.11)	<0.01	1.06 (1.01, 1.10)	0.01	1.12 (1.05, 1.19)	<0.01	1.13 (1.06, 1.20)	<0.01
Non-MSA	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
No. of children < 18 in household								
1 child	1.01 (0.98, 1.05)	0.49	0.99 (0.95, 1.03)	0.66	1.06 (1.00, 1.12)	0.07	1.05 (0.99, 1.11)	0.11
2 children	1.04 (1.00, 1.08)	0.04	1.02 (0.98, 1.06)	0.42	1.07 (1.01, 1.13)	0.02	1.06 (1.00, 1.13)	0.03
≥ 3 children	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent

^a Children were considered vaccinated if they received one or more doses of influenza vaccine during July 2013-May 2014.

^b APR = Adjusted Prevalence Ratio. Estimates in bold are statistically significantly different from the referent (P < 0.05).

^c CI=confidence interval.

^d AHR=Adjusted Hazard Ratio.

^e Age as of November 1st 2013.

^f Defined using household reported income, number of people living in the household, and US Census Bureau poverty Thresholds (<http://www.census.gov/hhes/www/poverty/data/threshld/>).

^g MSA=Metropolitan Statistical Area.

Table 2. Test of Proportionality for Cox proportional hazards model in assessing associations of demographic characteristics with influenza vaccination coverage^a among children 6 months to 8 years, National Immunization Survey-Flu (NIS-Flu), United States, 2013–2014 influenza season.

Demographic characteristics	Univariate test		Adjusted test	
	Chi-square	p-value	Chi-square	p-value
Gender	3.60	0.06	3.46	0.06
Age ^b	1.41	0.24	2.14	0.14
Child's race/ethnicity	0.12	0.73	1.13	0.29
Mother's education level	1.06	0.30	1.48	0.22
Poverty status ^c	1.03	0.31	0.54	0.46
Household in MSA ^d	0.17	0.68	0.33	0.57
No. of children < 18 in household	0.93	0.34	1.71	0.19

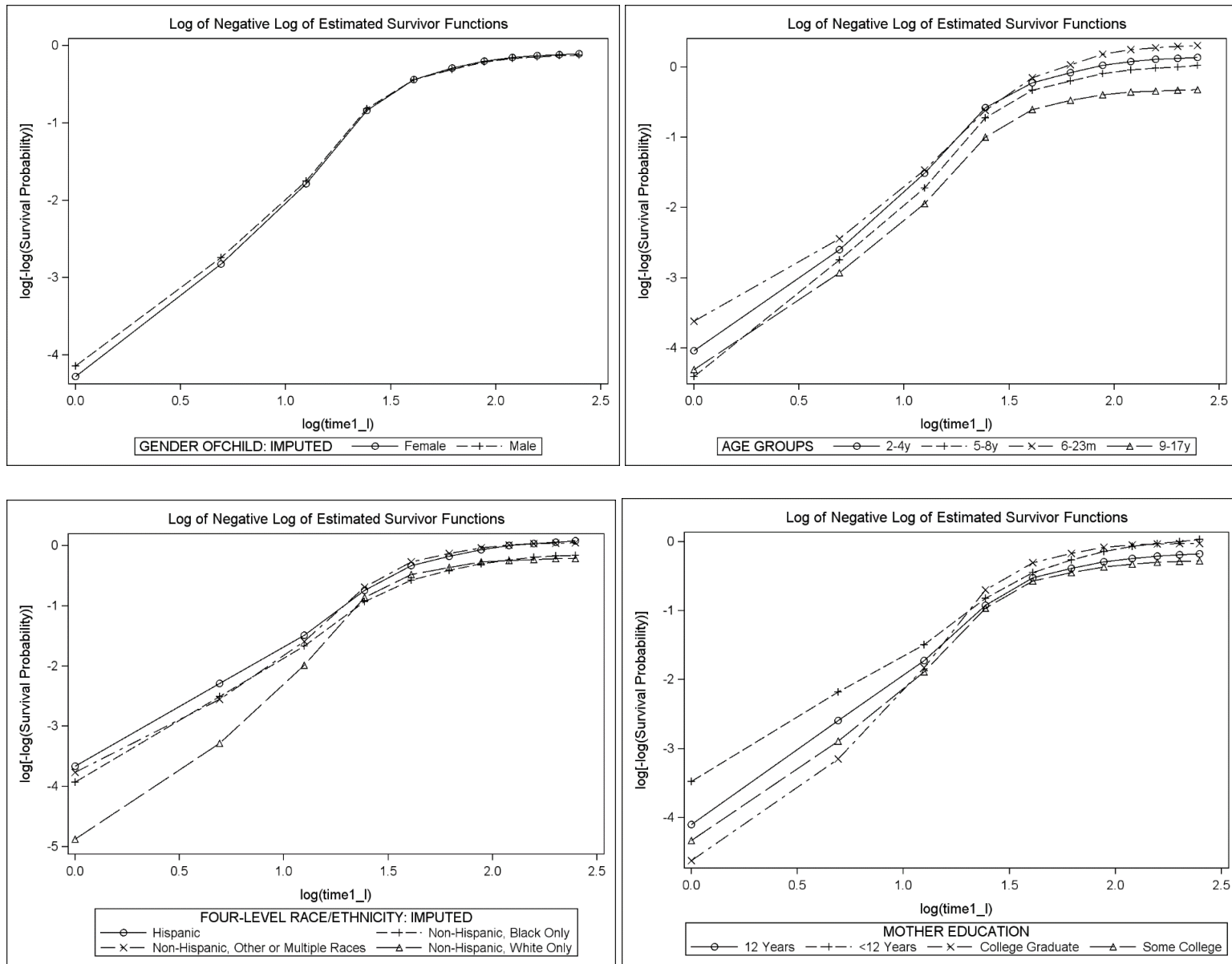
^a Children were considered vaccinated if they received one or more doses of influenza vaccine during July 2013–May 2014.

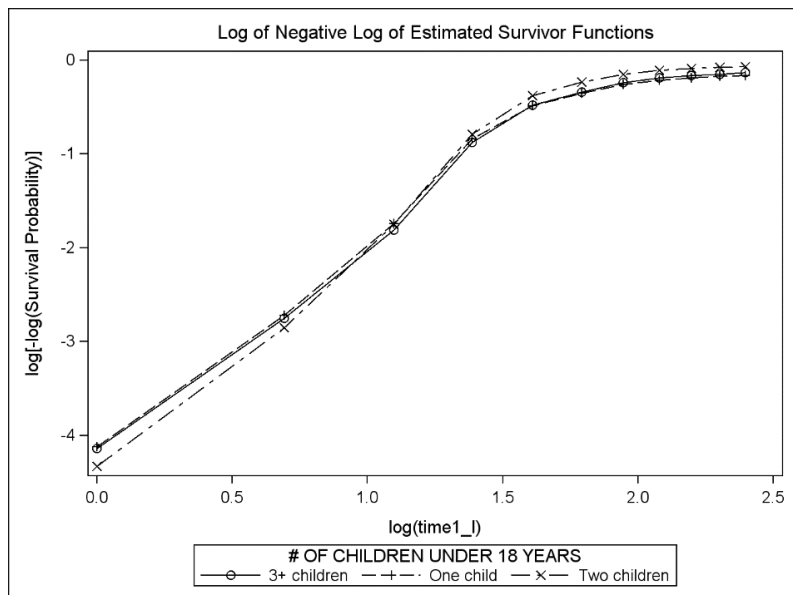
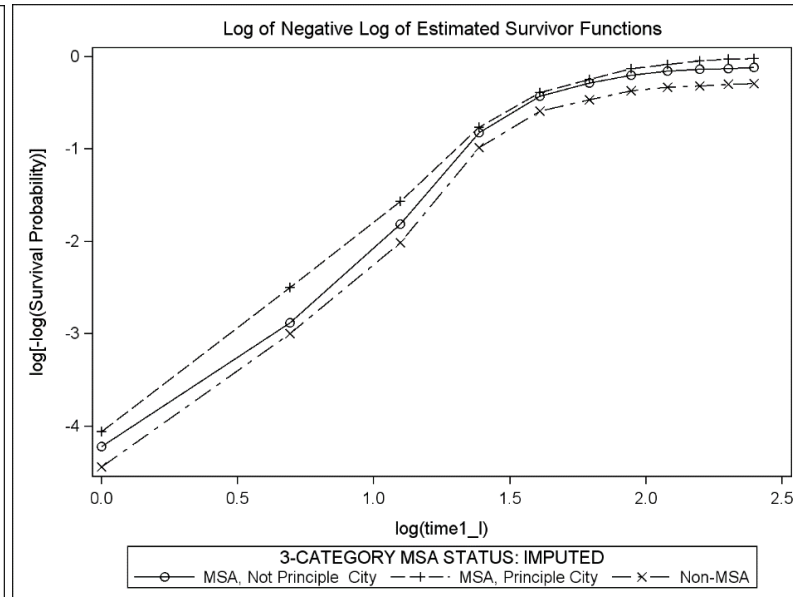
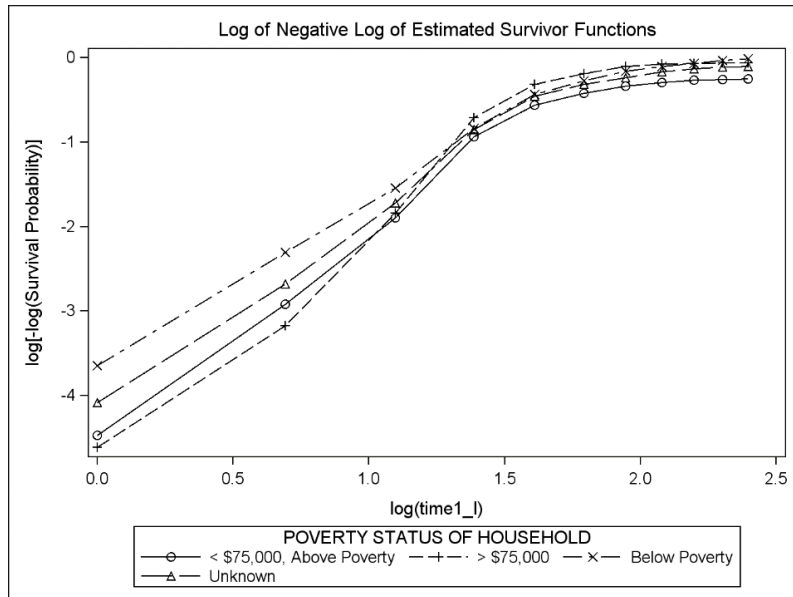
^b Age as of November 1st 2013.

^c Defined using household reported income, number of people living in the household, and US Census Bureau poverty thresholds (<http://www.census.gov/hhes/www/poverty/data/threshld/>).

^d MSA=Metropolitan Statistical Area.

Figure 1. Log-negative log plot for each of the demographic characteristic variables in the Cox proportional hazards model using data for children 6 months to 8 years, National Immunization Survey-Flu (NIS-Flu), United States, 2013–2014 influenza season.





REFERENCES

- Abbott, R. D. (1985). Logistic regression in survival analysis. *Am J Epidemiol*, 121(3), 465-471.
- CDC. (2010a). State-specific influenza A (H1N1) 2009 monovalent vaccination coverage—United States, October 2009–January 2010. *MMWR*, 59, 363-368.
- CDC. (2010b). state-specific seasonal influenza vaccination coverage—United States, August 2009–January 2010. *MMWR* 2010;59:477–84. *MMWR*, 59(477-84).
- CDC. (2013a). Flu Vaccination Coverage, United States, 2012-13 Influenza Season. 2013. Available at <http://www.cdc.gov/flu/fluview/cov-1213estimates.htm> [accessed Jun 17, 2014].
- CDC. (2013b). Surveillance of Influenza Vaccination Coverage — United States, 2007–08 Through 2011–12 Influenza Seasons. *MMWR*, 62(4), 1-28.
- CDC. (2014). CDC. Flu Vaccination Coverage, United States, 2013-14 Influenza Season. 2014. Available at <http://www.cdc.gov/flu/fluview/cov-1314estimates.htm> [accessed Feb 27, 2015].
- Chen, X. (2006). Survey Logistic Regression: Some SAS® and SUDAAN® Comparisons, *Preceedings of NESUG 19, Paper pos12*.
- Cochran, W. C. (1997). *Sampleing Techniques* (Third Edition ed.): John Wiley and Sons, NY.
- Frankel, L. (1983). The report of the CASRO task force on response rates. *Wiseman F, editor. Improving data quality in sample surveys. Cambridge, MA, Marketing Science Institute.*
- Hertz-Picciotto, I., & Rockhill, B. (1997). Validity and efficiency of approximation methods for tied survival times in Cox regression. *Biometrics*, 53(3), 1151-1156.
- Lewis, T. (2012). Modeling Complex Survey Data, *Preceedings of MWSUG 2012, Paper SA-07-2012*.
- Lu, P. J., Rodriguez-Lainz, A., O'Halloran, A., Greby, S., & Williams, W. W. (2014). Adult vaccination disparities among foreign-born populations in the U.S., 2012. *Am J Prev Med*, 47(6), 722-733.
- RTI (Ed.). (2009). *SUDAAN User's Manual, Release 10.0*. . Research Triangle Park, NC: Research Triangle Institute; 2009.
- Santibanez, T. A., Singleton, J.A., Santibanez, S.S., et al. . (2012). Socio-demographic differences in opinions about 2009 pandemic influenza A (H1N1) and seasonal influenza vaccination and disease among adults during the 2009-10 influenza season. *Influenza and Other Respiratory Diseases, published online DOI: 10.1111/j.1750-2659.2012.00374.x*.
- SAS (Ed.). (2010). *SAS/STAT® 9.22 User's Guide*. Cary, NC: SAS Institute Inc. Copyright © 2010.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the Author at:

Yusheng Zhai
 Affiliated with Leidos, Inc. Atlanta, Georgia, USA
 Centers for Disease Control and Prevention (CDC)
 National Center for Immunization and Respiratory Diseases (NCIRD)
 1600 Clifton Road, NE; Mail Stop A -19
 Atlanta, GA 30329
 E-mail: xds3@cdc.gov
 Phone: +1 404 639-8506

Trademark Citation: SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are registered trademarks or trademarks of their respective companies.
 SUDAAN is a registered trademark of the Research Triangle Institute, North Carolina.