Estimation Methodology for Weekly Surveys of Influenza Vaccination Rates

Kennon R. Copeland[1], Nicholas Davis[1], Lin Liu[1], Nadarajasundaram Ganesh[1], James A. Singleton[2], and Tammy A. Santibanez[2]
[1]NORC at the University of Chicago, 4350 East-West Highway, Bethesda, MD 20814
[2]National Center for Infectious and Respiratory Diseases, Centers for Disease Control and Prevention, 1600 Clifton Road, N.E., Atlanta, GA 30333

Abstract
Surveys of influenza vaccination coverage that enable production of estimates within the influenza season often involve collection of weekly survey data, typically based on relatively small sample sizes yielding relatively high variability. Such variability also adversely affects the stability of estimates across time, the result being estimated trends that may show occasional declines, even though the true population trends are by definition non-decreasing. Composite estimation, utilizing data and combining estimates across time periods, offers the opportunity for more stable estimates of coverage levels and trends as well as estimated trends less subject to period-to-period decreases. Use of survival analysis techniques is another alternative that ensures non-decreasing estimated trends. This paper profiles variability associated with direct estimates of levels and trends associated with the influenza module of the National Immunization Survey, proposes a composite estimation and a survival analysis approach for combining data across time, assesses the variability associated with composite and survival estimates of weekly influenza vaccination rates, and discusses potential error associated with use of data collected in different survey periods.

Key Words: composite estimation, National Immunization Survey, trend estimates, survival analysis.

1. Introduction

The Advisory Committee on Immunization Practices (ACIP) recommends that all persons aged ≥6 months receive annual influenza vaccination (CDC [(1)]). The ability to monitor influenza vaccination coverage enables the Centers for Disease Control and Prevention (CDC) to evaluate the effectiveness of the influenza vaccination program. The availability of estimates within the influenza season allows for the direction of efforts toward priority and under-vaccinated groups while the availability of estimates soon after the end of the influenza season allows for the guiding of improvements for subsequent seasons. The availability of weekly estimates within the influenza season has also been used to gain a better understanding of vaccine demand fluctuations within a season, and is particularly important during emergency situations such as an influenza pandemic or influenza vaccination shortage.

CDC uses data from the National Immunization Survey (NIS) to measure influenza vaccination coverage of children in the United States, with the target populations being children 19-35 months (NIS-Child) and 13-17 years (NIS-Teen). The objectives of the NIS influenza surveys are to obtain timely, precise, and representative influenza vaccination coverage estimates and influenza-related measures for the United States for each influenza season. The NIS influenza surveys consist of three components to enable the production of estimates for all children 6 months to 17 years old. While the NIS-Child collects parent report of childhood influenza vaccination for children 19-35 months at the time of interview and the NIS-Teen for children 13-17 years, there was a gap in parent reported influenza vaccination coverage data for children 6-18 months and children 3-12 years old; these data were needed to provide influenza vaccination coverage estimates for all children aged 6 months through 17 years in the United States. Thus, the NIS-Child Influenza Module (NIS-CIM) was conducted beginning in the 2009-10 influenza season for households with children 6-18 months and children 3-12 years old that are identified during the screening for the NIS and NIS-Teen. Combining the data from these three sources allows for the production of influenza estimates for all children 6 months-17 years.

To provide timely estimates from the NIS influenza surveys, CDC contracted with NORC at the University of Chicago to design and implement the NIS influenza surveys for the 2011-2012 flu season. Weekly national estimates were generated within the four to six days following the end of each survey week. Monthly estimates at both national and state level were generated within the 11-13 days following the end of each survey month.

Final season and, at times, interim monthly estimates are posted on the FluVaxView website (CDC [(2)]). For calculation of these official influenza vaccination coverage estimates posted to FluVaxView, Kaplan-Meier survival analysis (discussed later) is used to determine the cumulative influenza vaccination coverage (≥1 dose) during July or August through May using monthly interview data collected September or October through June. However, CDC also relies upon the composite measures calculated by NORC for internal examination of coverage. The composite measure (described later) was also used for the child estimates in the November coverage online report posted in conjunction with National Influenza Vaccination Week (http://www.cdc.gov/flu/fluvaxview/nifs-estimates-nov2012.htm). Direct estimates are used when modeling studies are conducted to examine factors associated with influenza vaccination coverage of children.

2. Survey Design

The NIS consists of a random-digit-dialed telephone survey based on quarterly samples of landline and cellular telephones contacted to identify residential households (referred to as the NIS sample). The sampling strata for the NIS sample are comprised of 56 core NIS awardee areas, in addition to optional sub-state areas that change from year to year. The 56 core NIS awardee areas are: New York City; Philadelphia County, PA; the District of Columbia; the city of Chicago, IL; Bexar County, TX; the city of Houston, TX; and the remaining states and sub-state areas. The United States Virgin Islands was added in 2009. Additional sampling strata in 2011 were: Prince George's County, MD; Dallas County, TX; and El Paso County, TX. Additional sampling strata in 2012 were: Dallas County, TX; and El Paso County, TX. Sampling probabilities among these strata are calculated to achieve a target coefficient of variation of 7.5 percentage points for NIS and 6.5 percentage points for NIS-Teen estimates of a 50% vaccination coverage rate.

The NIS-Flu consists of three nested questionnaires administered by computer-aided telephone interview (CATI) to households identified from the NIS sample. Both the landline and cell-phone samples are screened for eligibility for: the NIS, which targets children age 19 to 35 months; the NIS-Teen, which targets teens age 13 to 17 years; and the Child Influenza Module (CIM), which targets the remaining age groups 0 to 18 months[1] and 3 to 12 years. Separate samples are not drawn for these three questionnaires; all three operate on the same quarterly dual-frame samples. The CIM was not implemented in the United States Virgin Islands.

The three surveys are combined to cover all children aged <18 years at the time of screening. First, identified sample households are screened for eligibility for the NIS. Influenza vaccination coverage is obtained by this survey for all children in these households that are age 19-35 months at the time of screening. Next, a subsample of households is screened for eligibility for the NIS-Teen survey, and influenza vaccination coverage information is collected for one randomly selected teen among all teens that are age 13-17 years. Finally, all households are screened for the post-NIS Flu module, and influenza vaccination status is collected for one randomly selected child from among all children who are either age 0-18 months or 3-12 years at the time of screening. By combining completed interviews from the three surveys, we can produce influenza vaccination coverage estimates for the population of children age <18 years.

All NIS-Child and NIS-CIM interviews and a large majority of NIS-Teen interviews are conducted based on recall. In these interviews, the respondent was asked whether the child had received an influenza vaccination “since July 1, 2011.” If so, the respondent was asked whether the child received either one dose or two or more doses in that time period. The month and year of the first influenza vaccination since July 1, 2011 was collected and, if the respondent reported two doses, the date of the second dose was collected as well. The survey instruments were designed such that vaccination dates outside of the 2011-12 flu season could not be entered. When a parental respondent to the NIS-Teen survey reported vaccination status using a shot record, all influenza vaccinations received in the 12 months prior to the interview were reported along with the date of each dose. Using the reported dates, we identified which vaccines, if any, were received during the 2011-12 flu season (i.e., July 1, 2011 through June 30, 2012) to determine the teen’s vaccination status.

Estimates for a survey month were based upon all completed interviews from survey weeks with end dates within the survey month (e.g., November 2011 survey month consisted of survey weeks ending Nov 5, Nov 12, Nov 19, and Nov 26; sample completed Nov 27-30 were part of survey week ending Dec 3 and thus included in December 2011 survey month).

3. Direct Estimator

Completed interviews for children age 6 months - 17 years from a given survey week (defined as Sunday through Saturday) can be used to produce an estimate of the seasonal influenza vaccination coverage rate among children for each survey week. The NIS-Flu survey for the 2011-12 influenza season consisted of 44 such weeks, with an average number of completed child interviews per week of 2,324.

Prior to computing estimates, completed interviews from the survey week for children aged 6 months and older are weighted to reflect:

- probability of sample selection;
- number of telephone lines by which the household can be reached;
- random selection of child within household (NIS-Teen and CIM only);

[1] While estimates encompassed only children 6 months and older, sampling included all ages under 18 years. CDC analysis further restricted eligibility by excluding all children under 6 months as of November 1, 2011.

- combination of landline telephone and cell-phone interviews with a compositing factor for the overlapping group of dual landline and cell-phone users;
- post-stratification and ratio adjustments to population control totals including HHS regions, age by sex, and race/ethnicity;

Direct estimates of influenza vaccination coverage can be derived on a weekly basis using the sample weights and reported vaccination status of sample children with completed interviews during the survey week, w. To summarize the statistical methodology by which weekly direct vaccination coverage rates and their standard errors are estimated from the sample, let Y_{hij} be an indicator of vaccination status for the j^{th} child in the i^{th} sampled household in the h^{th} stratum of the NIS sampling design, which is equal to 1 if the child is reported as having received an influenza vaccinations and 0 otherwise. Also, let W_{hij} denote the final sampling weight for this sampled child. Letting

$$\hat{Y}_{wh} = \sum_{i,j\in w} W_{hij} Y_{hij}$$

and

$$\hat{T}_{wh} = \sum_{i,j\in w} W_{hij}$$

Then the national estimator of the influenza vaccination coverage rate for week w may be expressed as

$$\hat{\theta}_w = \frac{\sum_h \hat{Y}_{wh}}{\sum_h \hat{T}_{wh}}.$$

Letting

$$Z_{whij} = \frac{W_{hij}\left(Y_{hij} - \hat{\theta}_w\right)}{\hat{T}_{wh}}, i \in h, w, j \in i$$

$$Z_{whi} = \sum_{j\in i} Z_{whij}, i \in w,$$

and

$$\bar{Z}_{wh} = \frac{\sum_{i\in h,w} Z_{whi}}{m_{wh}},$$

an estimator of the variance of the estimated influenza vaccination coverage rate for week w can be expressed as

$$v\left(\hat{\theta}_w\right) = \sum_h \frac{m_{wh}}{m_{wh} - 1} \sum_{i\in h,w} \left(Z_{whi} - \bar{Z}_{wh}\right)^2.$$

In these equations, m_{wh} denotes the number of sampled households containing persons with completed interviews for week w in the h^{th} stratum. Note that these formulae extend to estimates for any subgroup, s, of the population.

Monthly direct estimates, $\hat{\theta}_m$, are derived in a similar manner through aggregation of the data across all survey weeks, w, within the survey month, m. Monthly weights are derived by combining the weekly weights for all of the survey weeks within the survey month, with the sum of the weights for each week being equal to the total population of children 6 months to 17 years. When the weights are combined across survey weeks, they are adjusted by dividing the weights by the number of weeks in the survey month, then re-raked to population estimates (with state replacing HHS region). In this way, the monthly weights sum to the total population of children 6 months to 17 years in the U.S.

Given the direct weekly estimates are based upon completed interviews obtained throughout the survey week, and respondents are asked to report influenza vaccinations received up to the interview date, the weekly/monthly vaccination coverage estimates do not represent the vaccination coverage as of the end of the survey week/month, but rather roughly represent the vaccination coverage as of the mid-point of the survey week/month.

4. Data Collection and Performance of the NIS-Flu Survey

Table 1 contains summary data collection performance statistics for the 2011-12 NIS-Flu. Resolution rates (the percentage of selected telephone numbers identifiable as residential, non-residential, or non-working numbers) and working residential number (WRN – for landline numbers)/assigned personal cell number (APCN – for cell numbers) were similar across NIS-Child, NIS-Teen, and CIM, as would be expected. Resolution rates were on the order of 82% for landline sample and 49-51% for cell sample. WRN rates were on the order of 16%-17% and APCN rates were on the order of 39%-43%.

When looking at the NIS-Child sample, which targets children 19 to 35 months, among identified landline residential telephones, the percentage completing the NIS screener to determine the presence of an eligible child 19-35 months was 90.1%, with 81.9% of screened and eligible households having at least one completed NIS-Child interview. The product of the resolution rate, the screener completion rate, and the interview completion rate, known as the CASRO response rate, was 60.6% for the landline NIS-Child sample. Among the cell-phone sample, 76.4% of active personal cell-phone numbers completed the screener, and 74.1% of eligible households had at least one complete NIS-Child interview, leading to a CASRO rate of 28.8%.

For the NIS-Teen subsample, among identified landline residential telephones in the NIS-Teen subsample, the percentage completing the NIS-Teen screener to determine the presence of an eligible teen 13-17 years was 84.4%, with 80.4% of screened and eligible households having a completed NIS-Teen interview. The CASRO response rate was 55.7% for the landline NIS-Teen subsample. Among the cell-phone sample, 69.7% of active personal cell-phone numbers completed the screener, and 66.5% of eligible households had a complete NIS-Teen interview, for a CASRO rate of 23.1%.

In 2011, 50% of all members of the NIS sample were flagged to receive the CIM. Because NIS-Flu surveillance did not begin until September, 2011, not all Q3/2011 sample cases were eligible to receive the module; only sample cases that had not yet been finalized by September 1 were eligible for the CIM. The flagging rate was increased in Q1/2012 to 75% for the landline sample, then in Q2/2012 increased to 100% for both the landline and cell-phone samples in Q2/2012. Among identified landline residential telephones in the CIM sample, 71.7% completed the screener of the module for the presence of a child either 0-18 months or 3-12 years old, and 88.2% of screened and eligible households had a complete CIM interview. The module's CASRO response rate was 51.8% for the landline sample. Among the cell-phone sample, the screener completion rate for the CIM was 60.6%, and the interview completion rate was 87.6%. For the cell-phone sample, the CASRO rate was 26.3%.

As shown in Table 2, a total of 102,254 completed interviews for children 6 months to 17 years old were obtained from the NIS sample from the first week of September 2011 through the last week of June 2012; of which 13,406 (13.1%) were NIS-Child completes, 31,097 (30.4%) were for NIS-Teen completes, and 57,751 (56.5%) were CIM completes. The cell-phone sample accounted for 29.3% of the total NIS-Flu completes.

The number of completed interviews by week obtained for the NIS-Flu survey is shown in Figure 1. The number generally ranged between 2,000 and 2,500. The two low points within the season occurred during weeks in which a separate National Flu Survey was conducted, as interviewing resources were focused on that survey during those time periods.

Direct weekly seasonal influenza vaccination coverage estimates and associated 95% CIs for the total population of children 6 months to 17 years old for the 2011-12 flu season are shown in Figure 2. While the underlying population vaccination coverage trend for an influenza season must be non-decreasing (outside the small chance of an effect due to child mortality), the vaccination coverage estimate trend is not, due to independent samples from week-to-week and sampling variability.

Direct monthly seasonal influenza vaccination coverage estimates have narrower confidence intervals and the trends are more stable than the weekly estimates, as shown in Figure 3. However, the monthly trends, like the weekly trends, are not restricted to being non-decreasing.

5. Approaches to Derive More Stable Estimates

As seen above, direct estimates of vaccination coverage are subject to sampling variability which affects stability of the weekly and, to a lesser extent, monthly trends, and are not constrained to being non-decreasing. In an attempt to provide more stable estimates of the vaccination coverage rates, both for individual time periods and across time, two approaches were undertaken to combine data across weeks. The challenge in combining data across weeks is that the end date of the reference period varies, as respondents were asked their vaccination status as of the date of the interview. However, respondents reporting having received a vaccination were also asked month of vaccination, which provides a set of fixed reference points which is used in the two approaches.

5.1 Composite Estimator

The first approach entailed generation of composite estimates (Schaible [3], Wolter [4]) of monthly vaccination coverage for each month prior to the current survey week using data from the current survey week along with all other survey weeks following the end of each individual month, and incorporating those monthly coverage estimates with weekly estimates of the interim vaccination coverage to a specific survey week to generate what was referred to as an "enhanced" weekly vaccination coverage estimate. The scenario for data availability and utilization for estimating vaccinations received by month is illustrated in Figures 4 and 5.

As seen in Figure 4, data on reported influenza vaccinations obtained in survey week ending January 21 are classified into one of five months, based on reports of month vaccination was received.

As seen in Figure 5, all survey weeks beginning with week ending October 8 contribute information about vaccinations received in September. September survey weeks cannot contribute information for the month of September as the reports provide incomplete information on vaccination status as of end of September. Week ending October 1 does not contribute information for September as the majority of the days in the survey week fell in September and thus would provide incomplete information as to vaccination status as of end of September.

Survey weeks contributing information for other months are identified in a similar manner. All survey weeks from December contribute information for September and October. However, survey week ending December 3 does not contribute information for November as the majority of the days in the survey week fell in November. Finally, the current survey week, week ending January 21, is used to determine the estimate for vaccinations received to date in the current month, January.

Extending the notation provided for the direct weekly estimates, let $\hat{\theta}_{mw}$ be the estimated proportion of children receiving an influenza vaccination in a prior month, m, based upon sample data from survey week w, and n_w be the number of completed interviews from survey week w. Then the composite vaccination coverage estimate for month m can be expressed as

$$\hat{\theta}_m = \sum_{w \in W(m)} \frac{n_w}{\sum_{w' \in W(m)} n_{w'}} \hat{\theta}_{mw},$$

where $W(m)$ represents the set of survey weeks following month m for which data are available, and the estimated variance of the composite estimate for month m can be expressed as

$$v(\hat{\theta}_m) = \sum_{w \in W(m)} \left(\frac{n_w}{\sum_{w' \in W(m)} n_{w'}} \right)^2 v(\hat{\theta}_{mw}).$$

The composite influenza vaccination coverage rate estimate for survey week w can then be expressed as

$$\hat{\theta}_w = \sum_{m \in M(w)} \hat{\theta}_m + \hat{\theta}_{vw},$$

where $M(w)$ represents the set of months prior to survey week w, and $\hat{\theta}_{vw}$, is the direct estimate of the proportion of children receiving a vaccination in the current month, based upon the sample interviewed in survey week w. The estimated variance of the composite vaccination coverage rate for survey week w can be expressed as (under the assumption of zero correlation between monthly estimates)

$$v(\hat{\theta}_w) = \sum_{m \in M(w)} v(\hat{\theta}_m) + v(\hat{\theta}_w).$$

5.2 Survival Estimator

This section discusses a Kaplan-Meier survival approach to generate monthly and weekly vaccination coverage for children age6 months to 17 years . Unlike the official monthly (or weekly) estimates of vaccination coverage, which are for the mid-point of the survey month (or week), the Kaplan-Meier approach yields vaccination coverage estimates as of the last day of the calendar month (or week), and furthermore, the Kaplan-Meier estimate of vaccination coverage is guaranteed to be non-decreasing.

Although the flu survey is conducted on a weekly basis, for children reported as having received a vaccination, only the month of vaccination is requested, not the week or actual date. Moreover, if a child is not vaccinated at the time of the weekly survey, the child could still potentially be vaccinated in a subsequent week, and thus, unvaccinated children are considered to be right censored observations. Since the Kaplan-Meier approach estimates the "survival rate", the estimator for vaccination coverage, $Z(t)$, is one minus the Kaplan-Meier estimator, $S(t)$ (Klein and Moeschberger [5]). That is,

$$Z(t) = 1 - S(t) \tag{1}$$

$$S(t) = \begin{cases} 1 & if\ t < t_1 \\ \prod_{t_i \le t} \left[1 - \frac{d_i}{Y_i}\right] & if\ t \ge t_1 \end{cases} \tag{2}$$

where t is the month (or week) for which vaccination coverage is being estimated, t_1 denotes the month (or week) after which flu vaccines become available (i.e., vaccination coverage prior to this month (or week) is 0), $t_i \le t$ denotes all months (or weeks) prior to month (or week) t, d_i is the weighted (using survey weights) estimate for the number of children who were vaccinated at time t_i, and Y_i is the weighted estimate for number of children who are "at risk" at time t_i. "At risk" refers to the weighted estimate for the total number of children excluding the weighted estimate for children who were vaccinated or censored prior to time t_i.

As associated standard error for $\hat{Z}(t)$ (Klein and Moeschberger [5]) is given by

$$se[Z(t)] = \sqrt{[S(t)]^2 \sum_{t_i \le t} \frac{d_i}{Y_i(Y_i - d_i)}} \tag{3}$$

where $S(t), Y_i,\ d_i$ are as defined above.

For the Kaplan-Meier approach that generates monthly vaccination coverage, the "censor" and "time-to-event" (i.e., time at which vaccination or censoring occurs) variables are defined as follows:

i. If a child was vaccinated in a month prior to the month of interview, then the "censor" variable was set to "not censored" and the "time-to-event" variable was set to the month of vaccination.
ii. If a child was vaccinated in the same month as the month of interview, then the "censor" variable was set to "censored" and the "time-to-event" variable was set to the month preceding the month of interview.
iii. If a child was not vaccinated as of the day of the interview, then the "censor" variable was set to "censored" and the "time-to-event" variable was set to the month preceding the month of interview.

For children who were not vaccinated as of the day of the interview, the "time-to-event" variable was set to the month preceding the month of interview as it is possible for these children to be vaccinated after the day of the interview but prior to the end of the month of interview. To be consistent with this definition, children who were vaccinated in the month of interview were also defined to be not vaccinated (i.e., "censored") as of the month preceding the month of interview. However, for the most recent month (for example, June) in which interviews were conducted, the above approach does not yield an estimate for vaccination coverage. Thus, in order to obtain an estimate for the most recent interview month, for children who were interviewed in that month, the censor and time-to-event variables were defined as follows:

iv. If a child was vaccinated in the most recent month in which interviews were conducted, then the "censor" variable was set to "not censored" and the "time-to-event" variable was set to the most recent month (June).

v. If a child was not vaccinated as of the day of the interview, then the "censor" variable was set to "censored" and the "time-to-event" variable was set to the most recent month (June).

As mentioned above, survey weights were taken into account when using the Kaplan-Meier approach. The monthly weights from each survey month were appropriately normalized using the number of interviewed children in each state and month. These normalized weights and the previously defined "censor" and "time-to-event" variables were used in SUDAAN to produce the monthly estimates of vaccination coverage based on the Kaplan-Meier estimator given by (1)-(2), and an associated standard error given by (3).

In order to generate weekly Kaplan-Meier estimates, week of vaccination is needed. However, as mentioned previously, if a child is vaccinated, only the month of vaccination is known. Based on the month of vaccination and other reported information, an imputation method for week of vaccination could be applied.

6. Estimator Comparisons

As seen in Figure 6, composite estimates of weekly influenza vaccination coverage rates showed more stable trends and narrower confidence intervals than did the direct estimates. However, there were several occurrences of a week-to-week decrease in the composite estimates as the composite estimator is not constrained to be non-decreasing.

As seen in Figure 7, Kaplan-Meier estimates of monthly influenza vaccination coverage rates showed more stable trends and narrower confidence intervals than did the direct estimates. As the Kaplan-Meier estimator is constrained to be non-decreasing, there were no occurrences of month-to-month decrease in the Kaplan-Meier estimates.

Finally, as seen in Figure 8, Kaplan-Meier estimates of monthly influenza vaccination coverage rates showed slightly more stable trends and narrower confidence intervals than did the composite estimates.

Table 3 presents comparative information on the stability and precision associated with direct, composite, and Kaplan-Meier estimates of influenza vaccination coverage rates. As can be seen, both the direct and composite estimates are subject to period-to-period changes. Standard errors for direct estimates are two to four times those of composite and Kaplan-Meier estimates, while standard errors for composite estimates are slightly greater than those for Kaplan-Meier estimates. Finally, direct estimates differ noticeably from the corresponding composite and Kaplan-Meier estimates; this difference does not appear directional for weekly estimates, but is negative for monthly estimates. There is little difference between composite and Kaplan-Meier estimates.

7. Considerations and Limitations

The NIS-Flu Survey measures vaccination coverage during the vaccination period. As a result, respondents have a non-zero likelihood of receiving a vaccination after their interview. This in turn leads to direct estimates not representing vaccination coverage rates as of the end of the survey period and makes the Kaplan-Meier estimates desirable.

The direct and composite estimators are not constrained to yield non-decreasing trends, while trends for the population must be non-decreasing. The Kaplan-Meier estimator by definition yields non-decreasing estimates. Constraints could be developed for the composite estimator and, with more assumptions, the direct estimator. Examples of such constraints include:

Simple: $\tilde{\theta}_w = max(\hat{\theta}_w, \hat{\theta}_{w-1})$
Model-based: $\tilde{\theta}_w = max(\hat{\theta}_w, pred(\hat{\theta}_w), \tilde{\theta}_w)$,
$pred(\hat{\theta}_w)$= time series forecast for $\hat{\theta}_w$

Composite estimates for prior periods are revised across time during the influenza season, which results from the incorporation of additional data about prior periods obtained from survey weeks moving forward through time. Kaplan-Meier monthly estimates are likewise subject to revision. For the 2011-2012 influenza season, the average absolute revision for Sep-Mar weekly composite estimates was 1.4 percentage points, with most revisions (25 of 31 weeks) revised downward from their original estimates. This downward revision (as illustrated in Figure 9) may reflect recall bias as to when vaccination was received in later survey weeks (Bilgen, et al [(6)]).

Given the benefits of the Kaplan-Meier over the composite estimator, ideally, weekly vaccination coverage rate estimates would be generated using the Kaplan-Meier approach. Consideration should be given to implementing the Kaplan-Meier approach for weekly estimates, using a weekly time unit along with multiple imputation of vaccination week. However, obtaining information about exact date of vaccination would be problematic in terms of accuracy and potential impact on item nonresponse and potential for interview break-offs. Thus, consideration could be given to developing and implementing a multiple imputation procedure for assigning week of vaccination within the reported month of vaccination, using the empirical data from the survey weeks within a given month. In developing such an approach, good predictors for week of vaccination will be required and it will be necessary to balance complexity and timeliness against improved precision.

While offering improved stability and precision, the complexity of the composite and Kaplan-Meier estimators make it difficult to explain the NIS-Flu Survey methodology to non-statistical audiences. Composite or Kaplan-Meier estimates can be used to generate stratified estimates, but not directly for modeling the association of influenza vaccination with sociodemographic and other factors. As an extension of the Kaplan-Meier approach, Cox proportional hazards models can be used to assess factors associated with vaccination, if the proportional hazards assumption holds.

The definition of a target population for the NIS-Flu Survey is somewhat dynamic across time, as the age at interview does not always correspond to the age at some specific date within the influenza season. Thus, analysts must be cautious in defining the population to which estimates apply. While eligibility for the NIS-Flu Survey is determined by age at interview, many estimates of interest are produced for domains defined by child age as of November 1, 2011. Because children age <6 months are not used for analyses, and because CDC is interested in the population of children 6 months or older as of a particular point in the flu season, the screening process for the 2013-14 influenza season has been to screen for children 6 months or older as of December 1, 2013.

As with most surveys, the NIS-Flu Survey is subject to item nonresponse. For purposes of deriving composite and Kaplan-Meier estimates, month of vaccination must be listed for all records for which the child was reported as having received an influenza vaccination. Hot deck imputation (Kalton and Kasprzyk [(7)]) is used to derive month of vaccination when missing.

Potential reporting error based upon parents' erroneous recall of their child's vaccination status and month of vaccination may lead to increased bias as the period of recall increases across the NIS-Flu Survey data collection period. These issues, reported on earlier by Bilgen, et al [6], warrant further investigation. Work is being carried out to address misclassification of vaccination status and other biases via a total survey error simulation model.

Consideration could also be given to creating a panel component to the survey, i.e., following-up on some portion of the weekly sample who report being "not vaccinated". This would result in the weekly estimates being correlated and reduce the likelihood (though not eliminate it) of the estimates decreasing with time. One downside of this approach is the potential for the original interview influencing behavior and resulting in a vaccination being obtained that otherwise would not have occurred. Another possible approach is to limit the recall period by restricting the weeks of interview used to estimate for a given past month; this would be easier to implement for the composite approach but would reduce precision due to reduced sample sizes.

8. Discussion

The NIS Flu Survey provides estimates of influenza vaccination coverage rates on a weekly and monthly basis through the influenza season. Direct estimates are subject to high variability, along with unstable and non-conforming trends. Alternative estimation approaches using composite estimation and Kaplan-Meier analysis yield estimates with higher precision and greater stability.

The CDC uses the Kaplan-Meier method for both within season and end of season monthly estimates. When needed for rapid reporting of within season estimates, the composite approach has been used. CDC's estimates of season-specific influenza vaccination coverage are posted online at http://www.cdc.gov/flu/fluvaxview/.

For weekly within-season estimates, the composite estimator approach is preferred over direct estimates, as it maximizes the use of all available data and yields greater precision and smoother vaccine uptake curves. However, more effort is needed to educate end users on the credibility of this approach.

References

[1] Centers for Disease Control and Prevention. (2013). "Prevention and control of influenza with vaccines: recommendations of the Advisory Committee on Immunization Practices (ACIP)--United States, 2012-13 influenza season." Morbidity and Mortality Weekly Report, Aug 17;61(32):613-618.

[2] Centers for Disease Control and Prevention. (2012). FluVaxView report. Available at http://www.cdc.gov/flu/professionals/vaccination/coverage_1112estimates.htm

[3] Schaible, W.L. (1978). "Choosing Weights for Composite Estimators for Small Area Statistics." Proceedings of the Section on Survey Research Methods, Joint Statistical Meeting, pp. 741-746.

[4] Wolter, K.M. (1979). "Composite Estimation in Finite Populations," *Journal of the American Statistical Association*, 4: 604-613.

[5] Klein, J.P. and Moeschberger, M.L. (1997). Survival Analysis: Techniques for Censored and Truncated Data. New York: Springer.

[6] Bilgen, I., Copeland, K.R., Santibanez, T.A., and Davis, N. (2012). "Examination of Recall Error in Reports of H1N1 and Seasonal Flu Vaccination." Presented at the American Association for Public Opinion Research Annual Conference.

[7] Kalton, G. and Kasprzyk, D. (1986). "The Treatment of Missing Survey Data," *Survey Methodology*, 12: 1-16.

Table 1: Data Collection and Performance Statistics, 2011-12 NIS-Flu

	NIS-Child (19-35 months)		NIS-Teen (13-17 years)		Child Influenza Module (0-18 months, 3-12 years)	
	Landline	*Cell*	*Landline*	*Cell*	*Landline*	*Cell*
Total Telephone Numbers Released	7,010,547	1,861,505	3,965,018	1,034,497	3,648,510	1,211,310
Resolution Rate	82.1%	50.7%	82.1%	49.7%	81.9%	49.5%
Working Residential Numbers (Landline)	17.0%	---	16.7%	---	16.4%	---
Assigned Personal Cell Number (Cell)	---	41.1%	---	42.7%	---	39.2%
Screener Completion Rate	90.1%	76.4%	84.4%	69.7%	71.7%	60.6%
Age Eligibility Rate	2.3%	4.3%	7.6%	8.6%	12.2%	16.3%
Interview Completion Rate	81.9%	74.1%	80.1%	66.5%	88.2%	87.6%
CASRO Response Rate	60.6%	28.8%	55.5%	23.1%	51.8%	26.3%

Table 2: Distribution of Completed Interviews by Questionnaire and Sample Frame, 2011-12 NIS-Flu

	Number	Percentage
Total Number of NIS-Flu Interviews	102,254	-
Number of NIS-Child Interviews	13,406	13.1
Number of NIS-Child Landline Telephone Interviews	8,167	60.9
Number of NIS-Child Cell Phone Interviews	5,239	39.1
Number of NIS-Teen Interviews	31,097	30.4
Number of NIS-Teen Landline Telephone Interviews	24,068	77.4
Number of NIS-Teen Cell Phone Interviews	7,029	22.6
Number of Child Influenza Module Interviews	57,751	56.5
Number of Child Influenza Module Landline Telephone Interviews	40,023	69.3
Number of Child Influenza Module Cell Phone Interviews	17,728	30.7
Number of Landline Telephone Interviews	72,258	70.7
Number of Cell Phone Interviews	29,996	29.3

Table 3: Performance of Vaccination Coverage Rate Estimators, 2011-12 NIS-Flu

	Number of Trend Decreases	Standard Error		Average Deviation*	
		Median	Average	Actual	Absolute
Weekly Estimates					
Direct	16	2.3	2.4	0.1	1.1
Composite	7	0.6	0.8	-	-
Monthly Estimates					
Direct	1	1.2	1.2	-1.8	1.8
Composite	0	0.5	0.5	-0.2	0.3
Kaplan-Meier	0	0.5	0.4	-	-

*Average Deviation is calculated as the mean of the differences from the composite estimate for weekly estimates and from the Kaplan-Meier estimate for monthly estimates

Figure 1. Number of Completed Interviews by Week, 2011-12 Influenza Season, Children 6 mo - 17 yr: NIS-Flu

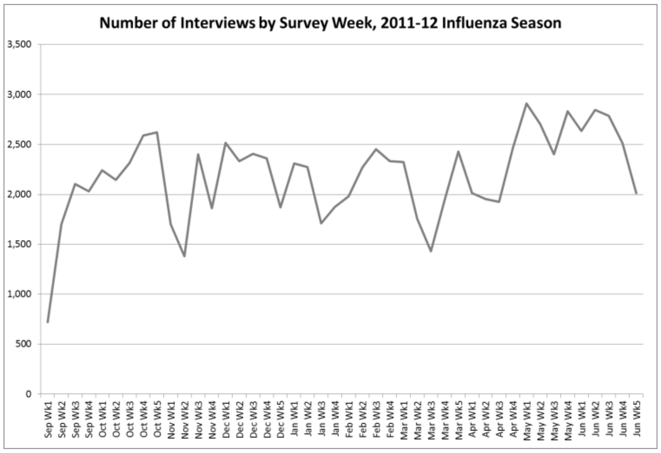

Figure 2. Weekly Direct Estimates of Influenza Vaccination Coverage Rates, 2011-12 Influenza Season, Children 6 mo - 17 yr: NIS-Flu

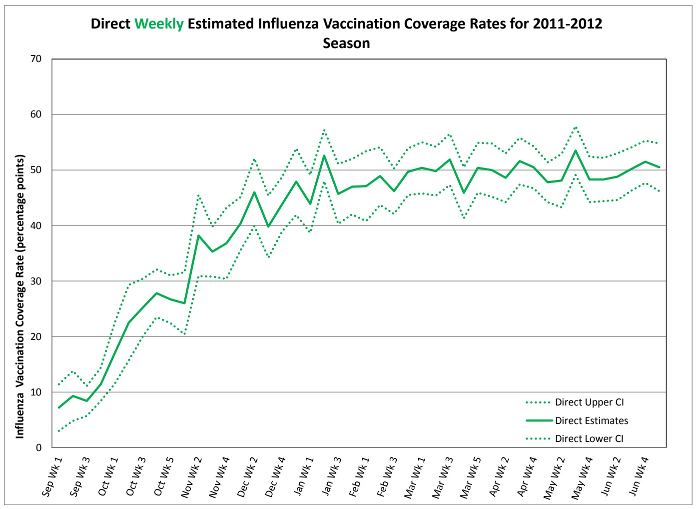

Figure 3. Weekly and Monthly Direct Estimates of Influenza Vaccination Coverage Rates, 2011-12 Influenza Season, Children 6 mo - 17 yr: NIS-Flu

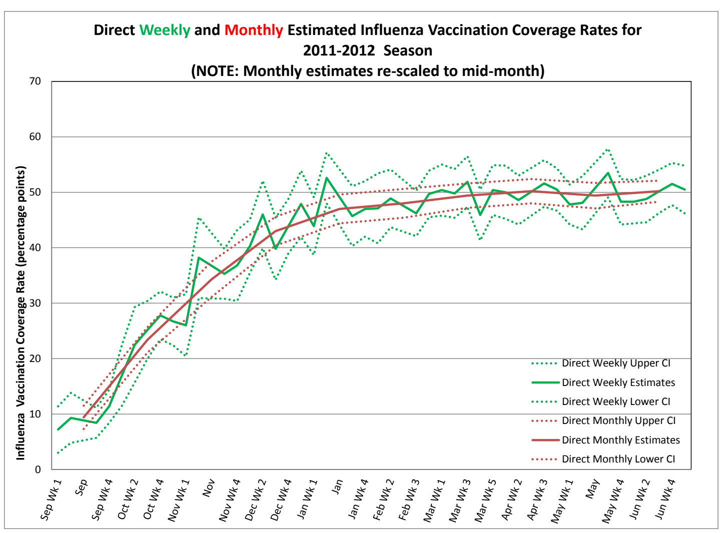

Figure 4. Illustration of Segmentation of Reported Vaccinations Into Month of Vaccination

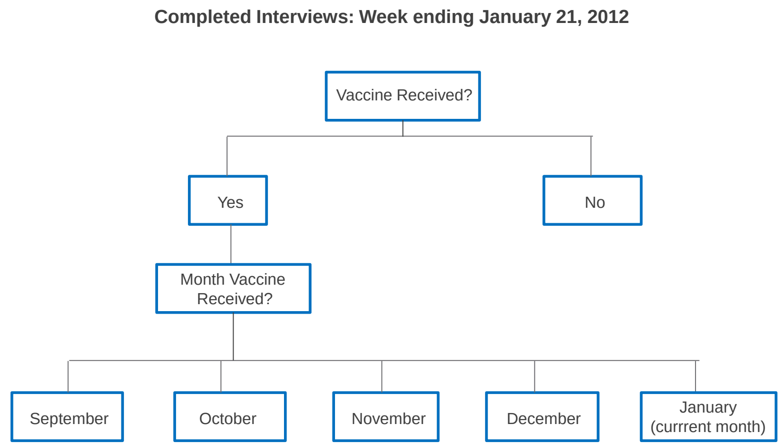

Figure 5. Illustration of Combination of Weekly Data on Reported Month of Vaccination

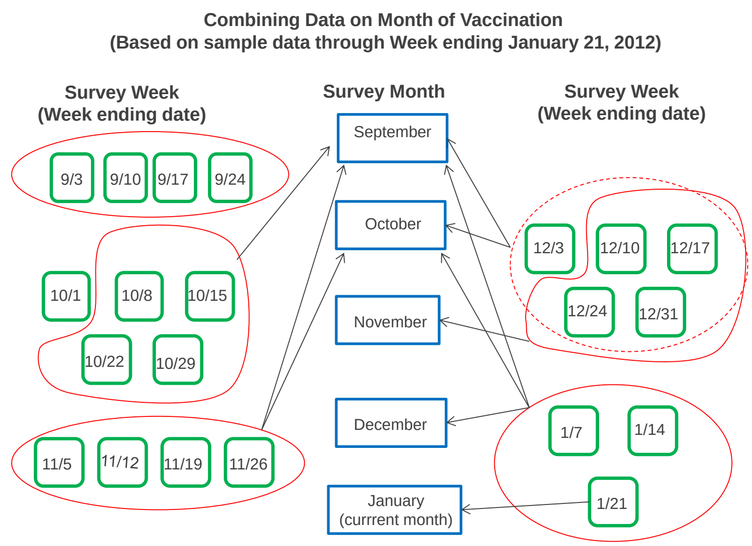

Figure 6. Weekly Direct and Composite Estimates of Influenza Vaccination Coverage Rates, 2011-12 Influenza Season, Children 6 mo - 17 yr: NIS-Flu

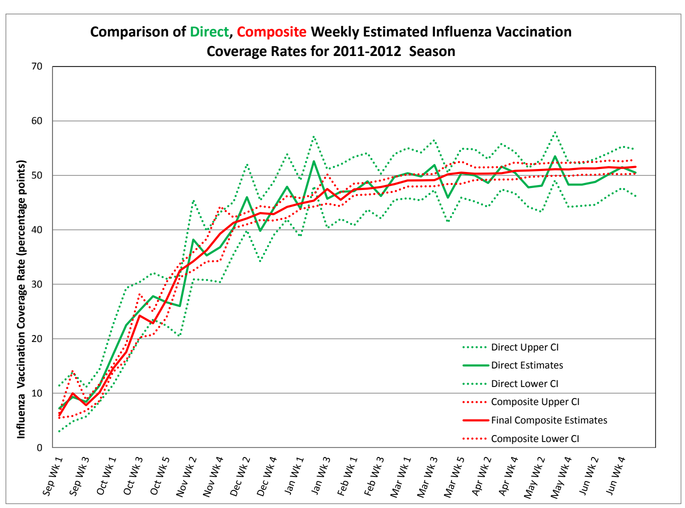

Figure 7. Monthly Direct and Kaplan-Meier Estimates of Influenza Vaccination Coverage Rates, 2011-12 Influenza Season, Children 6 mo - 17 yr: NIS-Flu

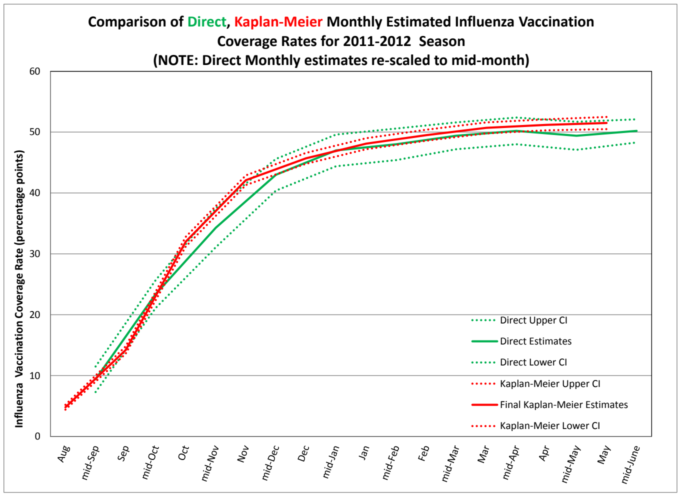

Figure 8. Monthly Composite and Kaplan-Meier Estimates of Influenza Vaccination Coverage Rates, 2011-12 Influenza Season, Children 6 mo - 17 yr: NIS-Flu

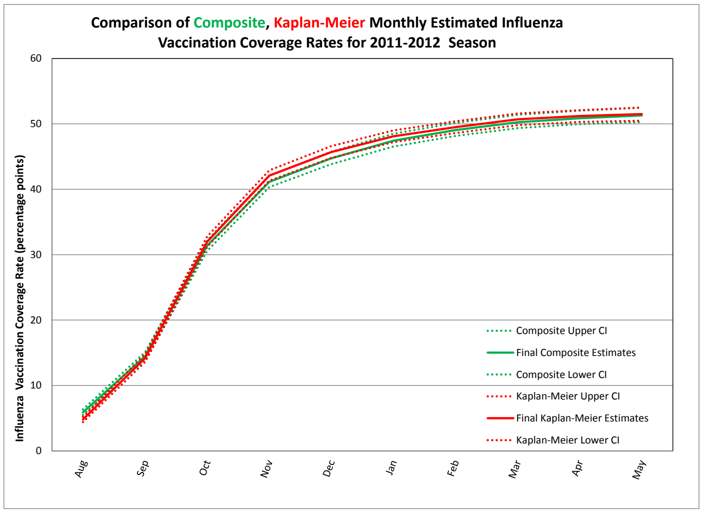

Figure 9. Composite Estimates of Influenza Vaccination Coverage Rates for October Week 2, 2011-12 Influenza Season, Children 6 mo - 17 yr: NIS-Flu