1. Introduction

The U.S. Consumer Product Safety Commission is a government regulatory agency that is charged with protecting the public from unreasonable risk of serious injury or death from thousands of types of consumer products under the agency's jurisdiction. Because some products pose a fire risk, agency staff makes annual estimates of fires, fire injuries, fire deaths and property damage from unintentional, consumer product-related fires that are reported to fire departments. One of the questions that occurred to agency analysts years ago was the extent to which fire department-reported (and for the most part, fire department-attended) incidents represented the total number of unwanted fires.

To address this question, agency staff commissioned an in-person survey in 1974, conducted by the U.S. Bureau of the Census (U.S. Consumer Product Safety Commission, 1978). The survey, in April of that year, asked respondents if they had a fire in or around their home at any time in the previous 12 months. Estimates from the survey data were 5.57 million residential fires in the 12-month period, of which 733,000 or 13 percent were reported to fire departments. A later reanalysis of the data corrected the estimates to 13 million fires, of which 1.2 million or 9 percent were reported to fire departments (University of Wisconsin, 1977). A second survey, this time a telephone survey, was conducted in 1984 (Audits and Surveys, 1985). The 1984 survey estimated 23.7 million fires, with 800,000 or 3 percent of the total reported to fire departments.

The most important substantive finding of the two surveys was the large number of unreported fires.

An important methodological finding from both the University of Wisconsin reanalysis and the 1984 survey was that respondents were unable to recall some fire incidents. Differences in methodology, especially the length of recall periods in the surveys, were likely to have accounted for the difference in the estimates of the number of fires.

In 2004-2005, U.S. Consumer Product Safety Commission (CPSC) staff conducted the third survey of fire department-attended and unattended residential fires. In the last 20 years, statistics reported from fire departments (official data) have shown that fire incidence and fire loss in general have decreased. For 2005, the most recent year for which official data are available, CPSC staff estimated that there were 375,100 unintentionally caused, consumer product-related, fire department-attended residential structure fires, resulting in 2,630 fire deaths, 12,820 fire injuries, and $6.2 billion in property loss (Chowdhury, Greene and Miller, 2008). One critical question is if fires that were not reported to or attended by U.S. fire departments had also decreased as much as reported fires.

Retrospective surveys ask respondents to describe incidents occurring during some defined period. In this paper, that period is called the "qualifying period." Some or all of the data from the qualifying period can be used for making annual estimates. The length of the period used to make annual estimates depends on how long researchers believe people to have reliably recalled incidents. That period is called the "recall period," and it can vary from a few days from the interview up to the qualifying period.

Like many other surveys, memory and recall problems are among the most common non-sampling errors encountered (Tourangeau, Rips and Rasinski, 2000, Chapter 4). The common theme in the literature is that recall periods are often short, sometimes no more than a few weeks. Moreover, the literature indicates that the length of the recall period depends on the severity of the incident.

For example, a 1994 study compared injury rates based on a two-week recall period from the National Health Interview Study with estimates from the Child Health Supplement that had an annual recall period, showing decreases in recall over time and increasing recall with severity (Harel, Overpeck, Jones, Scheidt, Bujur, Trumble and Anderson, 1994). Cummings, Rivara, Thompson and Reid (2005) demonstrated decreasing recall of children's injuries over time by comparing a telephone survey with written records from their Health Maintenance Organization. Landen and Hendricks (1995) found annual estimates of occupational injuries based on a four-week recall period were 32 percent higher than estimates based on a one-year recall period. They also found that injuries resulting in lost work days were recalled longer than those without lost work days.

Failure to recall also occurred in the two household fire surveys. In the 1974 survey, respondents were asked to provide information on all fire events occurring up to 12 months before the interview. A reanalysis of the data showed that respondents were likely to have failed to recall fire incidents and that recall decreased with increasing time from the interview (University of Wisconsin, 1977). The subsequent reanalysis used only a three-month recall period and resulted in an annual estimate for the number of fires that was more than twice the original estimate. The 1984 survey asked respondents to describe incidents occurring in the three-month (qualifying) period but made fire estimates using a recall period of the month before the telephone interview.

While a shorter recall period results in more accurate recall, according to the literature cited above, there is a tradeoff. As the length of the recall period decreases, the sample size decreases and the sampling variance increases. Moshiro, Heuch, Astrom, Setel and Kvale (2005) in recommending a short recall period, called for large sample sizes to offset the increased sampling variance. This tradeoff between sample size and bias from incomplete recall was formalized by Warner, Schenker, Heinen and Fingerhut (2005). Defining the loss due to recall as the "bias," the mean square error (MSE) was calculated as the sum of the square of the bias and the sampling variance. They recommended that the recall period be selected to minimize the MSE. Their paper contains other suggestions for remedying concerns raised in this literature including procedures for imputation of missing dates and separating the analyses into strata that are defined by incident severity.

2. Methods

2.1 Survey Description

The sampling frame for the 2004-2005 Residential Fire Survey consisted of all U.S. residential telephone numbers, i.e., all U.S. households with at least one land-line telephone in the home. The frame was developed using the GENESYS sampling system (GENESYS is a product of the Marketing Systems Group, Fort Washington, PA). GENESYS is a computer program and data system that is used to create random digit dialing (RDD) single-stage probability samples of telephone numbers. GENESYS contains telephone exchange level estimates for demographic variables such as age, income, home ownership, education, race, whether the block belongs to a metropolitan (urban) or non-metropolitan (non-urban) region, etc.; a feature then allows designing a sample that can be stratified to over- or under-sample households along certain demographic variables. This was useful in meeting a survey objective of providing sufficient representation of key demographic subgroups such as Native Americans, African Americans, households in rural areas, households of Hispanic origin, and the elderly. The sample was designed as a single stage, stratified probability sample.

Telephone interviews began on June 4, 2004, and were completed on September 5, 2005. To be eligible to participate in the study as a fire household, the respondent had to be 18 years of age or older and to have had an eligible fire within the past 90 days. Eligible fires were defined in the beginning of the survey where the respondent was asked the following:

> *We are interested in learning about any fires – large or small—that you have had in or around your home. By "fire" I mean any incident – large or small—that resulted in unwanted flame or smoke and could have caused damage to life or property if left unchecked.*

Telephone interviewers were aware that respondents might forget that they had a fire during the past 90 days. When respondents indicated that they did not have a fire, then interviewers asked them to recall if they had at least one or more of common fire type incidents such as unwanted flaming or smoking on the stove or another cooking appliance, a smoking electrical appliance, burning or smoldering clothing, etc. Respondents who recalled that fire incidents occurred were then asked about the date(s) of the fire incidents and a series of questions describing the incident, fire defenses (e.g., smoke alarms and extinguishers), housing and household characteristics. Of the households screened that did not report having a fire in the past 90 days, a subset of 2.5 percent (1 in 40) were selected randomly as a comparison sample of non-fire households. Information was collected from these non-fire households on housing and household characteristics and fire defenses.

There were 916 respondents reporting 961 fire incidents and a comparison sample of 2,161 non-fire households in the survey. To obtain this sample, approximately 580,000 numbers were dialed. The AAPOR response rates (RR) were as follows: RR2 22.5 percent, RR4 31.6 percent (American Association for Public Opinion Research, 2000). Sampling weights were created as the inverse of the sampling probability.

2.2 Assessing the Need for Bias Corrections and Imputation

In keeping with much of the literature on the relationship between incident severity and recall, fire incidents were first stratified by a variety of indicators that were thought to have influenced recall. The final set of indicators appeared to represent the severity of the incident. Indicators distinguishing high severity from low severity were that at least one of the following events occurred at a fire: a smoke alarm sounded, somebody attempted to put out the fire using a fire extinguisher, people left or tried to leave the residence during the fire, the fire department attended, or there was some reported flame damage.

Survey respondents were able to provide the date of the incident for 649 of the 961 incidents. Using these completely specified fire dates, the weighted estimated weekly fires were calculated for each week from the day of the interview. These values, stratified by severity, are shown in Figure 1.

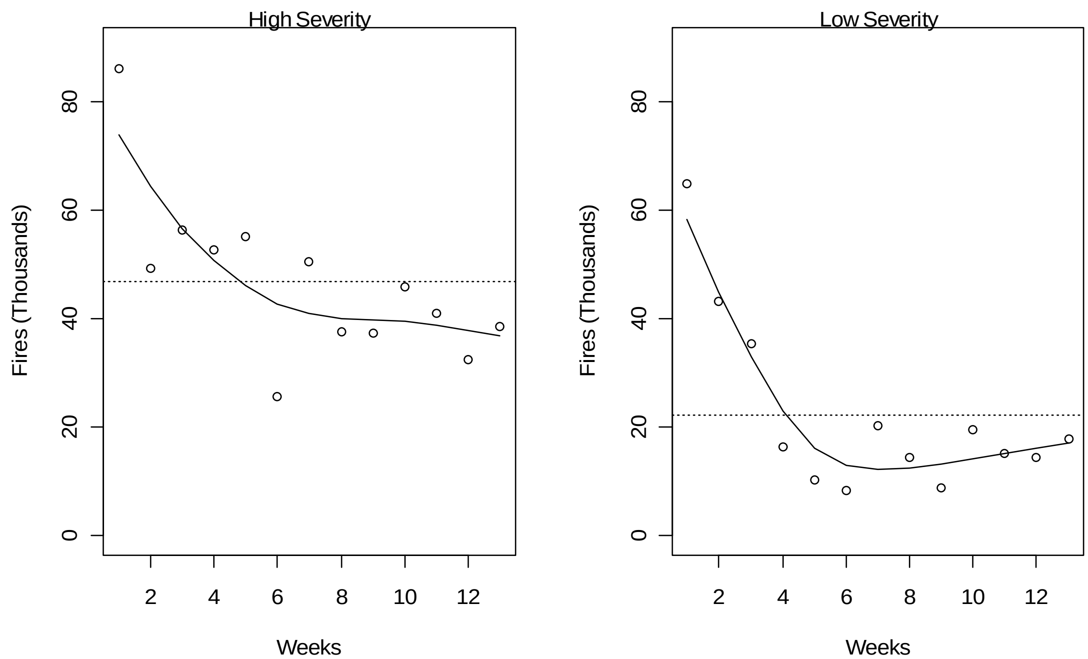

Figure 1: Estimated high (left panel) and low severity (right panel) fires by weeks from the interview. The solid lines are nonlinear regression estimates using b-splines with 4 degrees of freedom. The dashed horizontal lines are the average weekly fire estimates over the period, 46,769 (high severity) and 22,150 (low severity). (All nonlinear regressions in this paper used routines in R. R is a free software environment for statistical computing and graphics that runs on a wide variety of platforms. For information on R see http://www.r-project.org/.)

Both panels in Figure 1 show decreasing numbers of fires by week as the number of weeks from the interview increased. The decrease in incidents with increasing weeks is much steeper in the low severity plot than the high severity plot indicating that, relatively speaking, low severity incidents were not recalled as long on average as high severity incidents.

2.3 Imputation Procedure

Missing fire dates were imputed by selecting an elapsed time between interview and fire date and then computing the fire date from the possible elapsed times. Similar to Warner et al (2005), a two-stage strategy was used:

Stage 1. When respondents reported a single fire where the month but not the day of fire was known, the elapsed time between interview and fire date was selected randomly (i.e., following a uniform distribution) out of the possible elapsed times between the beginning of the month (or the day of the interview, whichever was closer) and the end of the month (or the end of the 90-day recall period, again whichever was closer). The imputed fire date was then calculated by subtracting the imputed elapsed time from the interview date. These imputed dates were classified as belonging to high or low severity incidents based on the definition above, but severity did not play a role in this stage of imputation.

Stage 2. For respondents who reported a single fire where the month and day were unknown, imputed elapsed times were selected at random with replacement by severity level from the imputed elapsed times in stage 1. The imputed fire date was then also calculated by subtracting the elapsed time from the interview date.

Six survey respondents reported two fires each without specifying either month or day. Missing fire dates were imputed by sampling from the uniform distribution for possible elapsed times, then assigning the shortest to the first reported fire and the second shortest to the other fire.

The imputation process described above was repeated 15 times, producing 15 datasets with imputed dates. The literature suggests a minimum of five imputation datasets, but more datasets are useful when the imputation variance might be large (Schaeffer, 1997).

2.4 Mean Square Error Analysis

Analysis of the multiple imputation data sets then proceeded by computing the mean square error (MSE) for various recall periods and then selecting the recall periods with the lowest value of the MSE. Separate computations were made for the two different severity levels, to allow the possibility of different recall periods for high and low severity incidents. Annual estimates were made for each hypothetical recall period by adding the weighted estimates where the elapsed time between interview and fire date fell into the recall period, then scaling by the proportion of the year in the recall period. This is a form of poststratification. Symbolically, the revised weight w'_{ij} is defined as

$$w'_{ij} = \left(\frac{52}{r}\right) w_{ij} I_{ij}(r) \tag{1}$$

where

w_{ij} is the original survey weight of incident *j* in stratum *i*
r is the length of the recall period in weeks
$I_{ij}(r)$ is 1 if incident *j* in stratum *i* has occurred within recall period *r;* 0, otherwise

The estimate for the annual number of fires using recall period *r* involves summing equation (1) over incidents and strata.

A cubic smoothing spline with four degrees of freedom was fit to the plot of estimated annual fires against recall period length (Hastie and Tibshirani, 1990; R software was used). The fitted value of the smoothing spline for the 14-day recall period was used as the reference value for making the bias estimate, that is it was assumed to be the “true” value. The choice of the 14-day reference period was in keeping with Warner et al (2005) and much of the literature. The use of the smoothing spline instead of a linear regression was a departure from Warner et al.

The MSE was then calculated for each hypothetical recall period, from the bias estimate and the variance. Calculations for annual estimates and the sampling variance were made in SAS® using Proc Surveymeans (SAS Institute, 2004). The total variance, including the imputation variance was calculated using Proc MIAnalyze.

3. Results

Table 1 below shows the analysis for the high severity incidents. The first column is the hypothetical recall period, first row including the incidents reported only during the first week from the interview, second row containing both the first and second weeks, etc. The second and third columns show the estimated incidents and the spline fit to those estimates. The standard error in column 4 includes both the contribution of sampling and imputation. The bias is the difference between the spline fit for the two-week recall reference period and the spline fit for any given recall period. The Root Mean Square Error (RMSE) is the square root of the sum of the squared standard error and the squared bias.

Table 1: Estimated Annual High Severity Fire Incidents, Bias, and Root Mean Square Error by Cumulative Weeks from the Interview and Two-Week Reference Period (Thousands of Fires)

Hypothetical Recall Period	*Estimated Annual Fires*	*Spline Fit to Annual Fires*	*Standard Error*	*Bias*	*RMSE*
1	5418	5094	851	365	926
1-2	4507	4728	552	0	552
1-3	4268	4418	434	-310	534
1-4	4112	4184	377	-544	662
1-5	4098	4021	324	-708	779
1-6	3861	3909	280	-819	866
1-7	3884	3838	260	-891	928
1-8	3809	3792	241	-936	967
1-9	3753	3763	221	-965	990
1-10	3770	3745	209	-984	1006
1-11	3754	3730	197	-998	1017
1-12	3690	3718	186	-1010	1027
1-13	3725	3708	176	-1020	1036

As shown in Table 1, the RMSE decreases with increasing cumulative weeks from the date of the interview using recall periods of one, one to two, and one to three weeks, and then increases with increasing recall periods again. The minimum RMSE is achieved using a hypothetical recall period of the first three weeks.

Table 2 shows the same analysis for the low severity incidents. Consistent with Figure 1, the best recall period for the low severity incidents is shorter than the high severity incidents. In this case, the best recall period is same as the reference period, namely two weeks.

Table 2: Estimated Annual Low Severity Fire Incidents, Bias, and Root Mean Square Error by Cumulative Weeks from the Interview and Two-Week Reference Period (Thousands of Fires)

Hypothetical Recall Period	*Estimated Annual Fires*	*Spline Fit to Annual Fires*	*Standard Error*	*Bias*	*RMSE*
1	3701	3574	704	367	794
1-2	3162	3207	462	0	462
1-3	2855	2863	358	-344	496
1-4	2508	2558	294	-648	712
1-5	2250	2307	241	-899	931
1-6	2066	2113	208	-1093	1113
1-7	1998	1971	189	-1235	1250
1-8	1891	1868	170	-1339	1350
1-9	1778	1792	156	-1414	1423
1-10	1751	1739	147	-1468	1475
1-11	1706	1699	138	-1508	1514
1-12	1676	1667	131	-1540	1545
1-13	1654	1639	124	-1568	1573

Table 3 below shows the estimates for the annual number of fires based on the analysis above, using the three- and two-week recall periods.

Table 3: 2004-2005 Fire Estimates by Fire Department Attendance (Number of fires per household based on 113,343,000 households in 2004-2005)

Fire Department Attendance	Estimated Fires per Year (95% Confidence Interval)	Fires per 100 Households (95% Confidence Interval)
Both	7,430,069 (6,195,938 - 8,664,199)	6.56 (5.46 - 7.64)
Attended Only	254,441 (65,165 - 443,716)	0.22 (0.06 - 0.39)
Unattended Only	7,175,628 (5,933,397 - 8,417,859)	6.33 (5.23 - 7.42)

4. Discussion

One of the objectives for the 2004-2005 survey was to compare the decrease in unreported fire incidence from the 1984 survey with the decrease in reported fire incidence. Some have suggested that newer technology, such as more widely used and better quality smoke alarms, would make it possible for residents to detect and extinguish fires when the fire was smaller, thus reducing or eliminating the need for fire department assistance (see Audits and Surveys, 1985, page 20). This would then result in a greater decrease in fire department-attended fires than unattended fires. The results from the survey suggest that this conjecture may not be true. CPSC staff estimated there were 655,500 unintentional fire department-attended residential structure fires in 1980 (Mah, 2001), while in 2005, there were an estimated 375,100 fires. This is a decrease of 43 percent. On the other hand, the number of unreported/unattended fires has dropped from 22.9 million in the 1984 survey to 7.2 million, a decrease of 68.7 percent. Note that estimates for 1984 are not available, but as the number of fires has been decreasing over the years, there would have been fewer fires in 1984 than in 1980, and the associated percentage decrease would have been less than 43 percent. Thus it is safe to conclude that unattended fires have dropped faster than attended fires.

The decrease in the number of unreported fires is even more striking because the number of households has increased from 84 million (Audits and Surveys, 1985) to 113.3 million (U.S. Bureau of the Census, 2009 at http://www.census.gov/population/socdemo/hh-fam/cps2005/tabH2-all.csv). This is an almost one-third increase in the number of households in 20 years. Taking this increase into account, the 1984 survey estimated an annual household incidence rate of 28.3 (reported and unreported) fires per 100 households per year. The 2004-2005 survey showed that the household fire incidence estimate dropped by 76.8 percent to 6.6 fires per 100 households per year.

Our advice to those planning a retrospective survey follows along the lines suggested by Warner et al (2005). This includes stratifying by factors associated with recall, imputing missing dates, defining a hypothetical recall period as the reference period, computing sampling and imputation variance and bias, and finally making decisions about the recall period using the mean square error (MSE). Survey respondents often tend to forget dates and events, thus short recall periods tend to be the most accurate. The unfortunate consequence of this finding is that short recall periods result in small sample sizes, and there may be a need to contact a large number of respondents to develop an adequate size sample. As a result, retrospective surveys of adverse events can be expensive.

Acknowledgements

The CPSC staff study team consisted of Linda E. Smith (now enjoying retirement), the two co-authors, and William W. Zamula of the Directorate for Economic Analysis. Drafts of the report were read and commented on by Kathleen A. Stralka, Director, Division of Hazard Analysis, and Russell H. Roegner, Associate Executive Director, Directorate of Epidemiology. Assistance with interpreting fire data was provided by Rohit Khanna, Fire Protection Engineer, Directorate for Engineering Sciences. Erlinda Edwards of the Office of Hazard Identification and Reduction provided extremely helpful editorial comments.

In addition to funding from the Consumer Product Safety Commission, survey funding was also provided by the Division of Unintentional Injury Prevention of the National Center for Injury Prevention and Control in the Centers for Disease Control and Prevention (CDC), and the United States Fire Administration, a component agency of the Department of Homeland Security.

The telephone survey was conducted by Synovate, Inc. Alan Roshwalb designed the sampling plan, the sample weighting, and prepared the SAS® dataset used for the final analysis. Tim Amsbury and John Lavin were instrumental along with CPSC staff in the design of the questionnaire and supervised the data collection. The project was supervised by Corporate Vice President, W. Burleigh "Leigh" Seaver.

References

American Association for Public Opinion Research (2000), "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys," AAPOR, Ann Arbor, MI.

Audits and Surveys, Inc. (1985), "1984 National Sample Survey of Unreported, Residential Fires." Final Technical Report Prepared for the U.S. Consumer Product Safety Commission. Princeton, NJ.

Chowdhury R, Greene M and Miller D (2008), "2003-2005 Residential Fire Loss Estimates," U.S. Consumer Product Safety Commission, Washington, DC.

Cummings P, Rivara FP, Thompson RS and Reid RJ (2005), "Ability of Parents to Recall the Injuries of Their Young Children," *Injury Prevention*, 11, pp. 43-47.

Harel Y, Overpeck MD, Jones DH, Scheidt PC, Bijur PE, Trumble AC and Anderson J (1994), "The Effects of Recall on Estimating Annual Nonfatal Injury Rates for Children and Adolescents," *American Journal of Public Health*, 84, 4, 599-605.

Landen DD and Hendricks S (1995), "Effect of Recall on Reporting of at-Work Injuries," *Public Health Reports*, 110:3, pp. 350-354.

Mah J (2001), "1998 Residential Fire Loss Estimates: U. S. National Estimates of Fires, Deaths and Property Losses from Non-Incendiary, Non-Suspicious Fires." U.S. Consumer Product Safety Commission, Washington, DC.

Mock C, Acheampong F, Adei S and Koepsell T (1999), "The Effect of Recall on Estimation of Incidence Rates for Injury in Ghana," *International Journal of Epidemiology*, 28, 4, pp. 750-755.

Moshiro C, Heuch I, Astrom AN, Setel P and Kvale G (2005), "Effect of Recall on Estimation of Non-Fatal Injury Rates: A Community Based Study in Tanzania," *Injury Prevention*, 11, pp 48-52.

SAS Institute Inc. (2004), *SAS/STAT® 9.1 Users Guide*. SAS Institute, Cary, NC.

Schaefer JL (1997), *Analysis of Incomplete Multivariate Data*, Chapman and Hall, New York.

Tourangeau R, Rips U and Rasinski K (2000), *The Psychology of Survey Response*, New York, Cambridge University Press.

Warner M, Schenker N, Heinen MA and Fingerhut LA (2005), "The Effects of Recall on Reporting Injury and Poisoning Episodes in the National Health Interview Survey," *Injury Prevention*, 11, pp. 282-287.

U.S. Consumer Product Safety Commission (1978), "Special Report: Results of National Household Fire Survey." HIA Special Report, U.S. Consumer Product Safety Commission, Washington, DC.

University of Wisconsin, Department of Statistics (1977), "Statistical Analysis of the National Household Fire Survey," Madison, WI.