# Methods (H1 and H2)

### Indicator

H1. Percentage of children ages 0 to 17 years with asthma, 1997-2010.
H2. Percentage of children ages 0 to 17 years with current asthma, by race/ethnicity and family income, 2007-2010.

### Summary

Since 1957, the National Center for Health Statistics, a division of the Centers for Disease Control and Prevention, has conducted the National Health Interview Survey (NHIS), a series of annual U.S. national surveys of the health status of the noninstitutionalized civilian population. These indicators use responses to questions on asthma for children ages 0 to 17 years from the NHIS 1997 to 2010 surveys; these questions have changed over time. Indicator H1 gives the percentages of children ever diagnosed with asthma that also had an asthma attack in the previous 12 months (1997-2010), and of children that currently have asthma (2001-2010). Indicator H2 uses responses to questions on asthma for children ages 0 to 17 years from the NHIS 2007 to 2010 surveys. Indicator H2 gives the percentages of children that currently have asthma, stratified both by race/ethnicity (using NHIS information on race and Hispanic origin) and family income (using reported or imputed NHIS poverty-income ratio data for each respondent). Table H1a gives the percentages of children with asthma in the previous 12 months for 1997-2010, by sex. Table H1b gives the percentages of children with asthma in the previous 12 months for 1980-1996. Table H1c gives the percentages of children with current asthma that had an asthma attack in the past 12 months for 2001 to 2010. Table H2a gives the percentages of children that currently have asthma for 2007-2010, stratified both by age group and sex. Percentages are calculated by combining positive responses to the relevant questions with the survey weights for each respondent. The survey weights are the annual numbers of children in the noninstitutionalized civilian population represented by each respondent.

### Data Summary

| Indicator | H1. Percentage of children ages 0 to 17 years with asthma, 1997-2010. H2. Percentage of children ages 0 to 17 years with current asthma, by race/ethnicity and family income, 2007-2010. | | | | | | |
|---|---|---|---|---|---|---|---|
| Time Period | 1997-2010 | | | | | | |
| Data | Asthma prevalence in children ages 0 to 17 years. | | | | | | |
| Years (1997-2003) | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 |
| Children | 14,290 | 13,645 | 12,910 | 13,376 | 13,579 | 12,524 | 12,249 |
| Asthma attack non-missing responses (%) | 14,242 (99.7 %) | 13,608 (99.8 %) | 12,685 (99.8 %) | 13,350 (99.8%) | 13,556 (99.9%) | 12,492 (99.8%) | 12,224 (99.8%) |
| Asthma attack | 48 | 37 | 25 | 26 | 23 | 32 | 25 |

| Indicator | H1. Percentage of children ages 0 to 17 years with asthma, 1997-2010. H2. Percentage of children ages 0 to 17 years with current asthma, by race/ethnicity and family income, 2007-2010. | | | | | | |
|---|---|---|---|---|---|---|---|
| missing responses (%) | (0.3%) | (0.2%) | 0.2%) | (0.2%) | (0.1%) | (0.2%) | (0.2%) |
| Current asthma non-missing responses (%)* | | | | | 13,534 (99.7%) | 12,475 (99.6%) | 12,207 (99.7%) |
| Current asthma missing responses (%) | | | | | 45 (0.3%) | 49 (0.4%) | 42 (0.3%) |
| Years (2004-2010) | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 |
| Children | 12,424 | 12,523 | 9.837 | 9,417 | 8,815 | 11,156 | 11,277 |
| Asthma attack non-missing responses (%) | 12,395 (99.8 %) | 12,500 (99.8 %) | 9,810 (99.8 %) | 9,401 (99.9%) | 8,798 (99.8%) | 11,141 (99.9%) | 11,256 (99.8%) |
| Asthma attack missing responses (%) | 29 (0.2%) | 23 (0.2%) | 27 (0.2%) | 16 (0.1%) | 17 (0.2%) | 15 (0.1%) | 21 (0.2%) |
| Current asthma non-missing responses (%) | 12,386 (99.7 %) | 12,496 (99.8 %) | 9,797 (99.6 %) | 9,394 (99.8%) | 8,793 (99.7%) | 11,129 99.8%) | 11,253 (99.8%) |
| Current asthma missing responses (%) | 38 (0.3%) | 27 (0.2%) | 40 (0.4%) | 23 (0.2%) | 22 (0.3%) | 27 (0.2%) | 24 (0.2%) |

* This survey question was first asked in 2001.

**Overview of Data Files**

The following files are needed to calculate these indicators. All these files together with the survey documentation and SAS programs for reading in the data are available at the NHIS website: http://www.cdc.gov/nchs/nhis.htm.

- NHIS 1997-2010: Sample Child file samchild.dat. Person file personsx.dat, Family file familyxx.dat, Imputed Income files 2007-2010: incmimp1.dat, incmimp2.dat, incmimp3.dat, incmimp4.dat, and incmimp5.dat. The Sample child file is an ASCII file containing interview data for children ages 0 to 17 years. For children ages 0 to 17 years, the responses were obtained from a knowledgeable adult family member residing in the household. Demographic data is obtained from the Person and Family files. The demographic variables needed for this indicator are the sample child survey weight (WTFA_SC), age (AGE_P), sex (SEX), the pseudo-stratum (STRATUM for 1997-2005, STRAT_P for 2006-2010), the pseudo-PSU (PSU for 1997-2005, PSU_P for 2006-2010), the race (RACERPI2, using the 1997 OMB definitions), the Hispanic origin (ORIGIN_I), and the detailed Hispanic origin HISPAN_I. The pseudo-stratum and pseudo-PSU variables provide an approximation to the exact sample design variables, and were

created by CDC by combining stratum information in a manner to protect the confidentiality of the publicly released data. From each of the year 2007 and 2008 imputed income files we need the imputed poverty income ratio (RAT_CATI), which gives the poverty income ratio category calculated from the reported exact family income, if available, or else gives the imputed category randomly generated by multiple imputation using regression models. From the year 2009 imputed income files we need the imputed poverty income ratio (POVRATI2) which gives the numerical value of the poverty income ratio in hundredths. From the year 2010 imputed income files we need the imputed poverty income ratio (POVRATI3) which gives the numerical value of the poverty income ratio in thousandths. The files are sorted and merged using the identifiers HHX, FMX, and FPX. The questionnaire variables needed for these analyses are the responses to the following questions: "Has a doctor or other health professional ever told you that <child's name> had asthma?" (CASHMEV) and if yes, "During the past 12 months, has <child's name> had an episode of asthma or an asthma attack?" (CASHYR). For 2001-2010 another needed variable is the response to the question: "Does <child's name> still have asthma?" (CASSTILL).

- NHIS 1980-1996. Condition file conditon.dat. This file is an ASCII file that contains the age (AGE), condition number (CNUM), survey weight (WTFA), and the parent's response to "Did <child's name> have this condition in the past 12 months?" (CPAST12). Data for children ages 17 and under and for the asthma condition were extracted. Used only for Table H1b.

**National Health Interview Survey (NHIS)**

Since 1957, the National Center for Health Statistics, a division of the Centers for Disease Control and Prevention, has conducted the National Health Interview Survey (NHIS), a series of annual U.S. national surveys of the health status of the noninstitutionalized civilian population. These indicators use responses from a knowledgeable adult family member residing in the household to asthma prevalence questions for children ages 0 to 17 years for the surveys from 1997 to 2010. The NHIS data were obtained from the NHIS website: http://www.cdc.gov/nchs/nhis.htm.

For 1997-2010, the first asthma question was: "Has a doctor or other health professional ever told you that <child's name> had asthma?" (CASHMEV). If the response was Yes, then the second question "During the past 12 months, has <child's name> had an episode of asthma or an asthma attack?" was asked (CASHYR). For 2001-2010, Yes responders to the CASHMEV question were also asked "Does <child's name> still have asthma?" (CASSTILL). For all three questions, responses other than Yes or No were treated as missing data. For the CASHYR and CASSTILL questions, responders who said No to the CASHMEV question were, for these analyses, treated as also responding No to the CASHYR and CASSTILL questions, even though they were not asked those questions. For 1980 to 1996, the asthma survey question was "Did <child's name> have asthma in the past 12 months?"

The NHIS uses a complex multi-stage, stratified, clustered sampling design. Certain demographic groups have been deliberately over-sampled. Oversampling is performed to

increase the reliability and precision of estimates of health status indicators for these population subgroups. From 1997 to 2005, Blacks and Hispanics were over-sampled. From 2006, Blacks, Hispanics, and Asians were over-sampled. The publicly released data include survey weights to adjust for the over-sampling, non-response, and non-coverage. The statistical analyses used the sample child survey weights (WTFA_SC, 1997 and later) to re-adjust the responses to represent the national population.

The sample design was changed in 2006. New strata were defined and PSUs were selected from these new strata. For example, pseudo-stratum 1 for 1997-2005 is unrelated to pseudo-stratum 1 for 2006-2010. To properly treat the 2006-2010 data as independent from the 2005 data, 1,000 was added to each of the year 2006, 2007, 2008, 2009, and 2010 pseudo-stratum numbers for these statistical analyses.[i]

**Race/Ethnicity and Family Income**

For Indicator H2, the prevalence percentages were calculated for demographic strata defined by the combination of race/ethnicity and family income.

The family income was characterized based on the RAT_CATI variable (POVRATI2 for 2009, POVRATI3 for 2010), which gives the level of the ratio of the family income to the poverty level. The National Center for Health Statistics obtained the family income for the respondent's family during the family interview. The U.S. Census Bureau defines annual poverty level money thresholds varying by family size and composition. The poverty income ratio (PIR) is the family income divided by the poverty level for that family. For 2006 to 2008, the public release variable RAT_CATI gives the value of the PIR for various ranges, Under 0.5, 0.5-0.74, 0.75 to 0.99, …, 4.50-4.99, 5.00 and Over. For 2009, the RAT_CATI category was computed from the public release variable POVRATI2 which gives the numerical value of PIR in hundredths. For 2010, the RAT_CATI category was computed from the public release variable POVRATI3 which gives the numerical value of PIR in thousandths. For prior years, the numerical values of POVRATI2 can be obtained from the Supplemental Imputed Income files available from the NHIS website: http://www.cdc.gov/nchs/nhis.htm.

Family income was stratified into the following groups:

- Below Poverty Level: PIR < 1, i.e., RAT_CATI = 1, 2, or 3.
- Between 100% and 200% of Poverty Level: 1 ≤ PIR < 2, i.e., RAT_CATI = 4, 5, 6, or 7.
- Above 200% of Poverty level: PIR ≥ 2, i.e., RAT_CATI = 8, 9, 10, 11, 12, 13 or 14.
- Above Poverty Level: PIR ≥ 1 (combines the previous two groups).
- Unknown Income: PIR is missing ("undefinable"), i.e., RAT_CATI = 96.[ii]

[i] The addition of 1,000 was chosen to make the stratum numbers for 2005 and earlier distinct from the stratum numbers for 2006 and later. This follows the recommendations in Appendix III of the survey description document "2010 National Health Interview Survey (NHIS) Public Use Data Release NHIS Survey Description," CDC, June, 2011, http://www.cdc.gov/nchs/nhis/quest_data_related_1997_forward.htm.

[ii] Although missing values of family income were statistically imputed for the vast majority of respondents, there were a few respondents that still had an unknown income after the income imputation.

Approximately 30% of families did not report their exact family income. From 1997 to 2006, the majority of these families either reported their income by selecting from two categories (above or below $20,000) or from 44 categories. For 2007 and later, the income questions were revised, so that families not reporting an exact income were first asked to report their income as the two categories above or below $50,000, and were then asked appropriate additional questions to refine the income range as either 0-$34,999, $35,000-$49,999, $50,000-74,999, $75,000-$99,999, or $100,000 and above. From 2007 to 2010, between 91% and 95% of families either gave the exact income or a categorical response.

NCHS reports[iii] evidence that the non-response to the income question is related to person-level or family-level characteristics, including items pertaining to health. Therefore treating the missing responses as being randomly missing would lead to biased estimates. To address this problem, NCHS applied a statistical method called "multiple imputation" to estimate or "impute" the family income based on the available family income and personal earnings information and on responses to other survey equations. A series of regression models was used to predict the exact family income from the available responses. Five sets of simulated family income values were generated for each family that did not report their exact family income. In this manner, NCHS generated five data sets, each containing a complete set of family income values (either the reported or the imputed values). The poverty income ratio categories or values were calculated from the income values and the family size and composition variables. An estimated prevalence percentage was computed for each of the five data sets. The overall estimated prevalence percentage is the arithmetic mean of the five estimates.

The poverty income ratios were calculated by NCHS using the exact family income, if available, or otherwise were calculated from the imputed family income. Among the sampled children ages 17 years and under for the years 2007 to 2010, the weighted percentage of children with imputed poverty income ratios was 21%.

Race was characterized using the race variable for the 1997 OMB standards,[iv] RACERPI2. The possible values of this variable are:

- 1. White only
- 2. Black / African American only
- 3. American Indian Alaska Native (AIAN) only
- 4. Asian only
- 5. Race group not releasable
- 6. Multiple race

[iii] "Multiple imputation of family income and personal earnings in the National Health Interview Survey: Methods and Examples," http://www.cdc.gov/nchs/data/nhis/tecdoc_2010.pdf. August, 2011.

[iv] Revised race standards were issued by the Office of Management and Budget in 1997 and were to be fully implemented across the federal statistical system by January 2003. Under the new standards, the minimum available race categories include: White, Black, AIAN, Asian, and Native Hawaiian or Other Pacific Islander (NHOPI). A very important change was that under the new standards, respondents may select more than one race category.

The Native Hawaiian or Other Pacific Islander (NHOPI) race group is not specified in the public release version due to confidentiality concerns. Respondents with the single race NHOPI have RACERPI2 = 5 and respondents of multiple races including NHOPI have RACERPI2 = 6.

The ORIGIN_I variable indicates whether or not the ethnicity is Hispanic or Latino. ORIGIN_I = 1 if the respondent is Hispanic or Latino. ORIGIN_I = 2 if the respondent is not Hispanic or Latino.

The HISPAN_I variable indicates the specific Hispanic origin or ancestry.

- 00 Multiple Hispanic
- 01 Puerto Rico
- 02 Mexican
- 03 Mexican-American
- 04 Cuban/Cuban American
- 05 Dominican (Republic)
- 06 Central or South American
- 07 Other Latin American, type not specified
- 08 Other Spanish
- 09 Hispanic/Latino/Spanish, non-specific type
- 10 Hispanic/Latino/Spanish, type refused
- 11 Hispanic/Latino/Spanish, type not ascertained
- 12 Not Hispanic/Spanish origin

The race/ethnicity was defined based on RACERPI2, ORIGIN_I, and HISPAN_I:

Race/ethnicity:

- White non-Hispanic: RACERPI2 =1, ORIGIN_I = 2
- Black or African-American, Non Hispanic: RACERPI2 = 2, ORIGIN_I = 2
- Asian non-Hispanic: RACERPI2 = 4, ORIGIN_I = 2
- Hispanic: ORIGIN_I = 1
    - Mexican: ORIGIN_I = 1 and HISPAN_I = 02, 03
    - Puerto Rican: ORIGIN_I = 1 and HISPAN_I = 01
- All Other Races: RACERPI2 = 3, 5 or 6, ORIGIN_I = 2
    - American Indian, Alaska Native, Non-Hispanic: RACERPI2 = 3, ORIGIN_I = 2

The "All Other Races" category includes all other races not specified, together with those individuals who report more than one race.

Some respondents gave missing or incomplete answers to the race/ethnicity questions. In those cases NCHS applied a statistical method called "hot-deck imputation" to estimate or "impute" the race or ethnicity based on the race/ethnicity responses for other household members, if available, or otherwise based on information from other households. The NHIS variables ORIGIN_I , HISPAN_I, and RACERPI2 use imputed responses if the original answer was

missing or incomplete. Among the sampled children ages 17 years and under for the years 2007 to 2010, the weighted percentage of children with an imputed race or ethnicity was 9%. Among the sampled Hispanic (defined by ORIGIN_I) children ages 17 years and under for the years 2007 to 2010, the weighted percentage of children with an imputed specific Hispanic origin was 2%.

**Calculation of Indicator**

Indicator H1 is the percentage of children ages 17 years or under for whom the response was Yes to the asthma attack in the last 12 months or current asthma questions, as detailed in the section "National Health Interview Survey (NHIS)." Indicator H2 is the percentage of children ages 17 years or under for whom the response was Yes to the current asthma question, stratified by race/ethnicity and family income. Table H2a is the percentage of children ages 17 years or under for whom the response was Yes to the current asthma question, stratified by age and sex.

To simply demonstrate the calculations, we will describe the calculations for the indicator H2, and will use the NHIS 2007-2010 responses to the CASSTILL question asking if the child still had asthma for White non-Hispanic children of all incomes. This question was only asked if the response was Yes to the CASHMEV question about whether the child was ever diagnosed with asthma. As described above, the question of interest is whether the child was ever diagnosed with asthma and still had asthma. We shall call this combined question the current asthma question. This question is answered Yes if CASHMEV = 1 (Yes) and CASSTILL = 1 (Yes). This question is answered No if either CASHMEV = 1 (Yes) and CASSTILL = 2 (No), or if CASHMEV = 2 (No). Otherwise the response is missing. We have rounded all the numbers to make the calculations easier:

We begin with all the non-missing responses to the current asthma question in the NHIS 2007-2010 surveys for White non-Hispanic children ages 0 to 17 years. Assume for the sake of simplicity that Yes or No responses were available for every sampled child. Each sampled child has an associated survey weight that estimates the total number of U.S. White non-Hispanic children in 2007-2010 represented by that sampled child. For example, the first response for a White non-Hispanic child aged 17 years or under was No with a survey weight of 10,000, and so represents 10,000 White non-Hispanic children ages 17 years or under. A second White non-Hispanic child aged 17 years or under responded No with a survey weight of 7,000, and so represents 7,000 White non-Hispanic children ages 17 years or under. A third White non-Hispanic child aged 17 years or under responded Yes with a survey weight of 21,000, and so represents 21,000 White non-Hispanic children ages 17 years or under. The total of the survey weights for the sampled White non-Hispanic children equals 160 million, the total U.S. population of White non-Hispanic children ages 17 years or under summed over all four years; thus the annual population is about 40 million.

To calculate the proportion of White non-Hispanic children with current asthma, we can use the survey weights to expand the data to the total four-year U.S. White non-Hispanic population of 160 million White non-Hispanic children ages 0 to 17 years. We have 10,000 No responses from the first child, 7,000 No responses from the second child, 21,000 Yes responses from the third child, and so on. Of these 160 million responses, a total of 13 million responses are Yes and the

remaining 147 million are No. Thus 13 million of the 160 million White non-Hispanic children have current asthma, giving a proportion of about 8%.

In reality, the calculations need to take into account that Yes or No responses were not reported for every respondent, and to use exact rather than rounded numbers. There were non-missing responses for 17,692 of the 17,738 sampled White non-Hispanic children ages 0 to 17 years over the four-year period. ("Don't know" responses or refusals to answer are treated as missing). The survey weights for all 17,738 sampled children add up to 164.8 million, the total four-year U.S. population of White non-Hispanic children ages 0 to 17 years. The survey weights for the 17,692 sampled White non-Hispanic children with non-missing responses add up to 164.4 million. Thus the available data represent 164.4 million children, which is more than 99%, but not all, of the four-year U.S. population of White non-Hispanic children ages 0 to 17 years. The survey weights for the Yes responses add up to 13.5 million, which is 8.2% of the population with responses (13.5 million/164.4 million = 8.2%). Thus we divide the sum of the weights for participants with Yes responses by the sum of the weights for participants with non-missing responses. These calculations assume that the sampled children with non-missing responses are representative of the children with missing responses.

For calculation of prevalence by income group, we use the five sets of imputed income values, which each give different results. Suppose we wish to estimate the proportion of White non-Hispanic children below the poverty level with current asthma. Using the above calculation method applied for White non-Hispanic children below the poverty level, the proportions for the five sets of imputed values are: 10.6%, 10.7%, 10.5%, 10.6%, and 10.7%. The estimated proportion of White non-Hispanic children below the poverty level with current asthma is given by the average of the five estimates, (10.6 + 10.7 + 10.5 + 10.6 + 10.7) /5 = 10.6%.

Equations

The following equations give the mathematical calculations for the example of White non-Hispanic children below the poverty level. Let w(i) denote the survey weight for the i'th surveyed White non-Hispanic child of ages 0 to 17 years. Exclude any surveyed children with a response other than Yes or No. For the current asthma question, let the response indicator c(i) = 1 if the i'th surveyed White non-Hispanic child had a Yes response and let c(i) = 0 if the i'th surveyed White non-Hispanic child had a No response. Let the income indicator d(i, j) = 1 if the i'th surveyed White non-Hispanic child was below the poverty level according to the j'th set of imputed values and let d(i, j) = 0 if the i'th surveyed White non-Hispanic child was not below the poverty level according to the j'th set of imputed values.

1. Fix j = 1, 2, 3, 4 or 5. Sum (over i) all the survey weights multiplied by the income indicators to get the total weight W(j) for set j:

$$W(j) = \Sigma\, w(i) \times d(i, j)$$

2. Fix j = 1, 2, 3, 4 or 5. Sum (over i) all the survey weights multiplied by the response indicators and multiplied by the income indicators to get the total weight D(j) for set j for White non-Hispanic children below the poverty level with a Yes response:

$$D(j) = \Sigma\ w(i) \times c(i) \times d(i, j)$$

3. Divide D(j) by W(j) to get the percentage of children with asthma in set j:

$$\text{Percentage (j)} = (D(j) / W(j)) \times 100\%$$

4. Average the percentages across the 5 sets to get the estimated percentage of children with current asthma:

$$\text{Percentage} = [\text{Percentage (1)} + \text{Percentage (2)} + \text{Percentage (3)} + \text{Percentage (4)} + \text{Percentage (5)}] / 5$$

If the demographic group of interest includes all incomes, then the percentages will be equal for all five sets of imputed values, so the calculation in steps 1 to 3 need only be done for j =1, and step 4 is not required.

Table H1c gives the percentages of children with current asthma that had an asthma attack in the past 12 months for 2001 to 2010. For each year, these percentages were calculated by dividing the percentage of children that both have current asthma and had an asthma attack in the past 12 months by the percentage of children with current asthma:

Percentage of children with current asthma who had an asthma attack in the past 12 months =
Percentage of children with Yes responses to both the current asthma and asthma attack questions
/ Percentage of children with Yes responses to the current asthma question
× 100%

Relative Standard Error

The uncertainties of the percentages were calculated using SUDAAN® (Research Triangle Institute, Research Triangle Park, NC 27709) statistical survey software. SUDAAN was used to calculate the estimated percentages and the standard errors of the estimated percentages. The standard error is the estimated standard deviation of the percentage, and this depends upon the survey design. The standard error calculation also incorporates the extra uncertainty due to the multiple imputations of the income variables (based on the variation between the estimated percentages from each of the five sets of imputations). For this purpose, the public release version of NHIS includes the variables STRATUM and PSU, which are the Masked Variance Unit pseudo-stratum and pseudo-primary sampling unit (pseudo-PSU). For approximate variance estimation, the survey design can be approximated as being a stratified random sample with replacement of the pseudo-PSUs from each pseudo-stratum; the true stratum and PSU variables are not provided in the public release version to protect confidentiality.

The sample design was changed in 2006. New strata were defined and PSUs were selected from these new strata. For example, pseudo-stratum 1 for 2005 is unrelated to pseudo-stratum 1 for 2006-2010. To properly treat the 2006-2010 data as independent from the 2005 data, 1,000 was added to each of the year 2006, 2007, 2008, 2009, and 2010 pseudo-stratum numbers for these statistical analyses.

The relative standard error is the standard error divided by the estimated percentage:

$$\text{Relative Standard Error (\%)} = [\text{Standard Error (Percentage)} / \text{Percentage}] \times 100\%$$

Percentages with a relative standard error less than 30% were treated as being reliable and were tabulated. Percentages with a relative standard error greater than or equal to 30% but less than 40% were treated as being unstable; these values were tabulated but were flagged to be interpreted with caution. Percentages with a relative standard error greater than or equal to 40%, or without an estimated relative standard error, were treated as being unreliable; these values were not tabulated and were flagged as having a large uncertainty.

**Statistical Comparisons**

Statistical analyses of the percentages of children with a positive response to the question of interest were used to determine whether the differences between percentages for different demographic groups were statistically significant. Using a logistic regression model, the logarithm of the odds that a given child has a positive response is assumed to be the sum of explanatory terms for the child's age group, sex, income group and/or race/ethnicity. The odds of a positive response are the probability of a positive response divided by the probability of a negative response. Thus if two demographic groups have similar (or equal) probabilities of a positive response, then they will also have similar (or equal) values for the logarithm of the odds. Using this model, the difference in the percentage between different demographic groups is statistically significant if the difference between the corresponding sums of explanatory terms is statistically significantly different from zero. The uncertainties of the regression coefficients were calculated using SUDAAN® (Research Triangle Institute, Research Triangle Park, NC 27709) statistical survey software to account for the survey weighting and design. A p-value at or below 0.05 implies that the difference is statistically significant at the 5% significance level. No adjustment is made for multiple comparisons.

For these statistical analyses we used two income groups, below poverty level, and at or above poverty level. The small number of children with unknown (and unimputed) incomes were included in the at or above poverty level group. For the main analyses we also used five race/ethnicity groups: White non-Hispanic; Black non-Hispanic; Asian non-Hispanic; Hispanic; All Other Races. In addition, for specific comparisons between the Mexican and Puerto Rican subgroups, we applied a similar statistical analysis using three ethnicity groups: Mexican; Puerto Rican; Other Hispanic or Non-Hispanic. We also used three age groups: 0-5, 6-10, and 11-17.

For each type of comparison, we present unadjusted and adjusted analyses. The unadjusted analyses directly compare a percentage between different demographic groups. The adjusted

analyses add other demographic explanatory variables to the statistical model and use the statistical model to account for the possible confounding effects of these other demographic variables. For example, the unadjusted race/ethnicity comparisons use and compare the percentages between different race/ethnicity pairs. The adjusted analyses add age, sex, and income terms to the statistical model and compare the percentages between different race/ethnicity pairs after accounting for the effects of the other demographic variables. For example, if White non-Hispanics tend to have higher family incomes than Black non-Hispanics, and if the prevalence of a disease strongly depends on family income only, then the unadjusted differences between these two race/ethnicity groups would be significant but the adjusted difference (taking into account income) would not be significant.

Comparisons of the prevalence of current asthma in children ages 0 to 17 years between pairs of race/ethnicity groups and between the two income groups are shown in Tables 1 and 2, respectively. For the unadjusted "All incomes" comparisons, the only explanatory variables are terms for each race/ethnicity group. For these unadjusted comparisons, the statistical tests compare the percentage for each pair of race/ethnicity groups. For the adjusted "All incomes (adjusted for age, sex, income)" comparisons, the explanatory variables are terms for each race/ethnicity group together with terms for each age, sex, and income group. For these adjusted comparisons, the statistical test compares the pair of race/ethnicity groups after accounting for any differences in the age, sex, and income distributions between the race/ethnicity groups.

In Table 1, for the unadjusted "Below Poverty Level" and "At or Above Poverty Level" comparisons, the only explanatory variables are terms for each of the 10 race/ethnicity/income combinations (combinations of five race/ethnicity groups and two income groups). For example, in row 1, the p-value for "Below Poverty Level" compares White non-Hispanics below the poverty level with Black non-Hispanics below the poverty level. The same set of explanatory variables is used in Table 2 for the unadjusted comparisons between a race/ethnicity group below the poverty level and the same race/ethnicity group at or above the poverty level. The corresponding adjusted analyses include extra explanatory variables for age and sex, so that for each race/ethnicity group, the two income groups are compared after accounting for any differences due to age or sex. Also in Table 2, the unadjusted p-value for the population "All" compares the percentages for children ages 0 to 17 years below poverty level with those at or above poverty level, using the explanatory variables for the two income groups. The adjusted p-value includes adjustment terms for age, sex, and race/ethnicity in the model.

Additional comparisons are shown in Table 3. The Against = "age" unadjusted p-value compares the percentages for different age groups. The adjusted p-value includes adjustment terms for income, sex, and race/ethnicity in the model. The Against = "sex" unadjusted p-value compares the percentages for boys and girls. The adjusted p-value includes adjustment terms for age, income, and race/ethnicity in the model. The Against = "income" unadjusted p-value compares the percentages for those below poverty level with those at or above poverty level. The adjusted p-value includes adjustment terms for age, sex, and race/ethnicity in the model. The Against = "year" p-value examines whether the linear trend in the percentages is statistically significant; the adjusted model for trend adjusts for demographic changes in the populations from year to year by including terms for age, sex, income, and race/ethnicity. The Subset column specifies the demographic group of interest. For the Against = "age," "sex," and "income" comparisons, the

comparisons are for all children and so no Subset is defined. For the Against= "year" trend analyses, results are given for the overall trend (Subset = missing) and for the trends in each sex group, so that, for example, the Subset = "Boys" examines whether there is a statistically significant trend for boys ages 0 to 17 years. For the Asthma attack variable, the "Against" = "year" trend analysis for the Subset "Current asthma" evaluates where there is a statistically significant trend in the percentage with an asthma attack in the past 12 months among children with current asthma.

For more details on these statistical analyses, see the memorandum by Cohen (2010).[v]

Table 1. Statistical significance tests comparing the percentages of children ages 0 to 17 with current asthma, between pairs of race/ethnicity groups, for 2007-2010.

| | | | P-VALUES | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Variable** | **First race/ethnicity group** | **Second race/ethnicity group*** | **All incomes** | **All incomes (adjusted for age, sex, income)** | **Below Poverty Level** | **Below Poverty Level (adjusted for age, sex)** | **At or Above Poverty Level** | **At or Above Poverty Level (adjusted for age, sex)** |
| Current asthma | White non-Hispanic | Black non-Hispanic | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| Current asthma | White non-Hispanic | Asian non-Hispanic | 0.098 | 0.115 | 0.016 | 0.013 | 0.444 | 0.554 |
| Current asthma | White non-Hispanic | Hispanic | 0.531 | 0.183 | 0.092 | 0.114 | 0.432 | 0.656 |
| Current asthma | White non-Hispanic | Other | < 0.001 | < 0.001 | 0.025 | 0.026 | 0.003 | 0.001 |
| Current asthma | Black non-Hispanic | Asian non-Hispanic | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| Current asthma | Black non-Hispanic | Hispanic | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| Current asthma | Black non-Hispanic | Other | 0.019 | 0.057 | 0.254 | 0.237 | 0.073 | 0.116 |
| Current asthma | Asian non-Hispanic | Hispanic | 0.197 | 0.378 | 0.049 | 0.039 | 0.750 | 0.746 |
| Current asthma | Asian non-Hispanic | Other | < 0.001 | < 0.001 | 0.002 | 0.001 | 0.010 | 0.009 |
| Current asthma | Hispanic | Other | < 0.001 | < 0.001 | 0.001 | 0.002 | 0.001 | 0.001 |
| Current asthma | Mexican | Puerto Rican | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |

* "Other" represents the "All Other Races" category, which includes all other races not specified, together with those individuals who report more than one race.

Table 2. Statistical significance tests comparing the percentages of children ages 0 to 17 years with current asthma, between those below poverty level and those at or above poverty level, 2007-2010.

| | | P-VALUES | |
|---|---|---|---|
| **Variable** | **Population*** | **Unadjusted** | **Adjusted**** |
| Current asthma | All | < 0.001 | < 0.001 |
| Current asthma | White non-Hispanic | 0.002 | < 0.001 |

[v] Cohen, J. 2010. *Selected statistical methods for testing for trends and comparing years or demographic groups in ACE NHIS and NHANES indicators.* Memorandum submitted to Dan Axelrad, EPA, 21 March, 2010.

| | | P-VALUES | |
|---|---|---|---|
| **Variable** | **Population*** | **Unadjusted** | **Adjusted**** |
| Current asthma | Black non-Hispanic | < 0.001 | < 0.001 |
| Current asthma | Asian non-Hispanic | 0.143 | 0.133 |
| Current asthma | Hispanic | 0.109 | 0.058 |
| Current asthma | Other | 0.095 | 0.102 |
| Current asthma | Mexican | 0.416 | 0.612 |
| Current asthma | Puerto Rican | < 0.001 | < 0.001 |

* "Other" represents the "All Other Races" category, which includes all other races not specified, together with those individuals who report more than one race.
** Comparison for "All" is adjusted for age, sex, and race/ethnicity; comparisons for race/ethnicity categories are adjusted for age and sex.

Table 3. Other statistical significance tests comparing the percentages of children ages 0 to 17 years with asthma, for 2007-2010 (trends for 1997-2010).

| | | | | | P-VALUES | |
|---|---|---|---|---|---|---|
| **Variable** | **From** | **To** | **Against** | **Subset** | **Unadjusted** | **Adjusted*** |
| Current asthma | 2007 | 2010 | age | | < 0.001 | < 0.001 |
| Current asthma | 2007 | 2010 | sex | | < 0.001 | < 0.001 |
| Current asthma | 2007 | 2010 | income | | < 0.001 | < 0.001 |
| Current asthma | 1997 | 2010 | year | | 0.001 | 0.001 |
| Current asthma | 1997 | 2010 | year | Boys | 0.004 | 0.005 |
| Current asthma | 1997 | 2010 | year | Girls | 0.051 | 0.035 |
| Asthma attack | 1997 | 2010 | year | | 0.683 | 0.331 |
| Asthma attack | 2001 | 2010 | year | Current asthma | <0.001 | <0.001 |

*For Against = "age," the comparison is between the age groups 0-5, 6-10, and 11-17, and the p-values are adjusted for sex, race/ethnicity, and income.
For Against = "sex," the comparison is between boys and girls, and the p-values are adjusted for age, race/ethnicity, and income.
For Against = "income," the comparison is between those below the poverty level and those at or above the poverty level and the p-values are adjusted for age, sex, and race/ethnicity.
For Against = "year," where Subset is missing or "Current asthma," the comparison is the trend over different years, and the p-values are adjusted for age, sex, race/ethnicity, and income.
For Against = "year," where Subset is "Boys" or "Girls," the comparison is the trend over different years, and the p-values are adjusted for age, race/ethnicity, and income.

# Methods (H3)

**Indicator H3**

Children's emergency room visits and hospitalizations for asthma and other respiratory causes, ages 0 to 17 years, 1996-2008.

**Summary**

*Emergency Room Visits*
Since 1992, the National Center for Health Statistics, a division of the Centers for Disease Control and Prevention, has conducted the National Hospital Ambulatory Medical Care Survey (NHAMCS), a series of annual U.S. national surveys of visits to the emergency departments and outpatient departments of noninstitutional general and short-stay hospitals, exclusive of federal, military, and Veteran's Administration hospitals. For emergency room visits, this indicator uses the first diagnosis ICD-9 code to count emergency room visits for asthma and all other respiratory causes, asthma, and all respiratory causes other than asthma (composed of the following subcategories: upper respiratory conditions, pneumonia or influenza, and other lower respiratory conditions besides asthma) . The national numbers of emergency room visits by children ages 17 years and under are calculated by combining visits for each respiratory disease diagnosis with the survey weights for each child patient. The survey weights are the numbers of hospital emergency room visits by children ages 17 years and under in the noninstitutionalized civilian population represented by each patient visit in the survey database. This indicator shows the rate of emergency room visits per 10,000 children, calculated by dividing the national number of emergency room visits by the total U.S. population of noninstitutionalized civilian children ages 17 years and under. Table H3a provides the rate of emergency room visits by children 17 years and under, stratified by race/ethnicity, for the years 2005-2008. Table H3b provides the rate of emergency room visits by children 17 years and under, stratified by age group, for the years 2005-2008.

*Hospitalizations*
Since 1965, the National Center for Health Statistics, a division of the Centers for Disease Control and Prevention, has conducted the National Hospital Discharge Survey (NHDS), a series of annual U.S. national surveys of hospital discharges from non-federal short-stay hospitals. This indicator uses the first diagnosis ICD-9 code to count hospital discharges for asthma and all other respiratory causes, asthma, and all respiratory causes other than asthma (composed of the following subcategories: upper respiratory conditions, pneumonia or influenza, and other lower respiratory conditions besides asthma). The national numbers of hospital discharges by children ages 17 years and under are calculated by combining hospital discharges for each respiratory disease diagnosis with the survey weights for each child patient. The survey weights are the numbers of hospital discharges by children ages 17 years and under in the noninstitutionalized civilian population represented by each hospital discharge in the survey database. This indicator shows the rate of hospitalizations per 10,000 children, calculated by dividing the national number of hospital discharges by the total U.S. population of noninstitutionalized civilian children ages 17 years and under. Table H3c provides the rate of hospitalizations by children ages 17 years and under, stratified by race, for the years 2005-2008. Table H3d provides the rate

of hospitalizations by children ages 17 years and under, stratified by age group, for the years 2005-2008.

**Data Summary**

| Indicator | H3. Children's emergency room visits and hospitalizations for asthma and other respiratory causes, ages 0 to 17 years, 1996-2008. | | | | | |
|---|---|---|---|---|---|---|
| Time Period | 1996-2008 | | | | | |
| Data | Emergency room visits and hospitalizations by children ages 0 to 17 years. | | | | | |
| Years (1996-2001) | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 |
| Emergency room visits sampled | 5,777 | 5,690 | 6,153 | 5,072 | 6,264 | 8,386 |
| Hospital discharges sampled | 60,708 | 64,681 | 65,546 | 62,561 | 65,043 | 68,370 |
| Years (2002-2007) | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
| Emergency room visits sampled | 8,849 | 9,725 | 8,642 | 8,159 | 9,231 | 7,929 |
| Hospital discharges sampled | 65,868 | 65,536 | 72,585 | 71,402 | 69,847 | 67,757 |
| Years (2008) | 2008 | | | | | |
| Emergency room visits sampled | 7,438 | | | | | |
| Hospital discharges sampled | 25,506 | | | | | |

**Overview of Data Files**

The following files are needed to calculate this indicator.

*Emergency Room Visits*

- NHAMCS 1996-2008: EDXXXX.exe, where XXXX denotes the four-digit year. Each file is a compressed executable file that when decompressed gives an ASCII file containing emergency room visit data for a survey year. These files were obtained from the ftp site:

ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHAMCS/

The variables needed for this indicator are the survey year, age, physician's diagnosis #1 (DIAG1), and the following sampling design information: the patient visit weight (PATWT), masked stratum (STRATM for years 1996 to 2001, CSTRATM for years 2002 to 2008), and masked primary sampling unit (PSUM for years 1996 to 2001, CPSUM for years 2002 to 2008) The masked variables provide an approximation to the exact sample design variables, and were created by CDC by combining stratum information in a manner to protect the confidentiality of the publicly released data. For the supplemental table, the patient race and ethnicity variables RACE (RACEIM for 2007 and 2008) and ETHNIC (ETHIM for 2007 and 2008) are also needed.

- Census data.[vi] For the years 1996 to 1999, the national noninstitutionalized civilian populations were obtained from the url:

  http://www.census.gov/popest/data/national/asrh/1990s/nat_monthly_noninstitutional.html

  For the years 2000 to 2008, the national noninstitutionalized civilian populations were obtained from the url:

  http://www.census.gov/popest/data/national/asrh/2009/2009-nat-ni.html

  In each case, the file for each year includes the required variables: month, year, age, total U.S. population. The "month" gives the date for the population estimate. For these analyses, data for month = 7 were selected, corresponding to the populations as of July 1.

  For Table H3a, populations stratified by race and ethnicity were obtained using the detailed population data in the same census files for the years 2005 to 2008, as detailed below.

*Hospitalizations*

- NHDS 1996-2008: NHDSXX.PU.TXT, where XX denotes the two-digit year (NHDS96.ASC and NHDS97,ASC for 1996 and 1997). Each file is downloadable as a compressed file that decompresses into an ASCII file containing hospital discharge data for a survey year. These files were obtained from the ftp site:

  ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHDS/

[vi] Calculation of the indicator values for both emergency room visit and hospitalizations use the "civilian noninstitutionalized population" as the population denominator. An alternative approach that has been used by the CDC is to use the "civilian population." The choice of population denominator has a small impact on the calculated rates; thus the results reported in ACE are slightly different from those in CDC publications such as The State of Childhood Asthma, United States, 1980–2005, http://www.cdc.gov/nchs/data/ad/ad381.pdf.

This site only contains the files from 1996 onwards. The variables needed for this indicator are the survey year, age, physician's diagnosis #1 (DIAG1), and the analysis weight. For Table H3c, the patient race variable RACE is also needed.

- Census data. For the years 1996 to 1999, the national noninstitutionalized civilian populations were obtained from the url:

  http://www.census.gov/popest/data/national/asrh/1990s/nat_monthly_noninstitutional.html

  For the years 2000 to 2008, the national noninstitutionalized civilian populations were obtained from the url:

  http://www.census.gov/popest/data/national/asrh/2009/2009-nat-ni.html
  In each case, the file for each year includes the required variables: month, year, age, total U.S. population. The "month" gives the date for the population estimate. For these analyses, data for month = 7 were selected, corresponding to the populations as of July 1.

  For Table H3c, populations stratified by race were obtained using the detailed population data in the same census files for the years 2005 to 2008, as detailed below.

**National Hospital Ambulatory Medical Care Survey (Emergency Room Visits)**

The National Hospital Ambulatory Medical Care Survey (NHAMCS) is conducted by the National Center for Health Statistics, a division of the Centers for Disease Control and Prevention. The complex multi-stage survey is designed to collect data on ambulatory care services in hospital emergency and outpatient departments; these analyses only used the emergency department visits. Sampled hospitals are noninstitutional general and short-stay hospitals located in all states and Washington DC, but exclude federal, military, and Veteran's Administration hospitals. Data from sampled visits are obtained on the demographic characteristics, expected source(s) of payments, patients' complaints, physician's diagnoses, diagnostic and screening services, procedures, types of health care professionals seen, and causes of injury.

These analyses focused on visits to emergency rooms by children ages 17 years and under for respiratory diseases. Emergency room data was selected by using the ED files only. The age variable was used to select visits by children ages 17 years and under. The respiratory disease categories were selected based on the first physician's diagnosis code (DIAG1) using the International Classification of Diseases (ICD-9), first three characters:

- Asthma and all other respiratory causes: codes 460-466, 480-488, 490-496
- All respiratory causes other than asthma: codes 460-466, 480-488, 490-492, 494-496
- Upper respiratory: codes 460-466
- Pneumonia or influenza: codes 480-488
- Other lower respiratory: codes 490-492, 494-496

- Asthma: code 493

The NHAMCS uses a complex multi-stage, stratified, clustered sampling design. The statistical analyses used the patient visit survey weights (PATWT) to re-adjust the sample of visits to represent the total national population of emergency room visits in each calendar year.

**National Hospital Discharge Survey (Hospitalizations)**

The National Hospital Discharge Survey (NHDS) is conducted by the National Center for Health Statistics, a division of the Centers for Disease Control and Prevention. The complex multi-stage survey is designed to collect data on inpatients discharged from non-federal short-stay hospitals. Sampled hospitals are short-stay general or children's general hospitals located in all states and Washington DC, with an average length of stay of fewer than 30 days and six or more beds staffed for patients use. Federal, military, and Veteran's Administration hospitals are excluded, as are hospital units of institutions. Data from sampled visits are obtained on the demographic characteristics and physician's diagnoses.

These analyses focused on hospital discharges by children ages 17 years and under for respiratory diseases. The age variable was used to select visits by children ages 17 years and under. The respiratory disease categories were selected based on the first physician's diagnosis code (DIAG1) using the International Classification of Diseases (ICD-9), first three characters:

- Asthma and all other respiratory causes: codes 460-466, 480-488, 490-496
- All respiratory causes other than asthma: codes 460-466, 480-488, 490-492, 494-496
- Upper respiratory: codes 460-466
- Pneumonia or influenza: codes 480-488
- Other lower respiratory: codes 490-492, 494-496
- Asthma: code 493

The NHDS uses a complex multi-stage, stratified, clustered sampling design. The statistical analyses used the survey analysis weights to re-adjust the sample of discharges to represent the total national population of hospital discharges in each calendar year.

Although the available data were collected for hospital discharges, we assume for these analyses that admission and discharge rates are equal.

**Calculation of Indicator**

*Emergency Room Visits*
Indicator H3 shows the rate of emergency room visits by noninstitutionalized civilian children ages 17 years or under that were for a given respiratory disease.

For each year and respiratory disease, we carried out the following calculations:

1. We extracted the NHAMCS survey data for all the emergency room visits by children ages 17 years or under for the given respiratory disease. We selected all visits where the age was between 0 and 17 and the first three characters of the diagnosis code were:

- Asthma and all other respiratory causes: codes 460-466, 480-488, 490-496
- All respiratory causes other than asthma: codes 460-466, 480-488, 490-492, 494-496
- Upper respiratory: codes 460-466
- Pneumonia or influenza: codes 480-488
- Other lower respiratory: codes 490-492, 494-496
- Asthma: code 493

For each visit, the patient weight (PATWT) denotes the national number of patient visits represented by that visit.

2. We summed the NHAMCS patient weights for all the selected visits to estimate the total number of emergency room visits by children ages 17 or under for the given respiratory disease:

$$\text{Total Number of Visits} = \Sigma \text{ PATWT, summed over all selected visits}$$

3. Using the census data, we calculated the total population of children ages 17 years or under by summing the populations for the ages 0, 1, 2, … 17:

$$\text{Population} = \Sigma \text{ Population (age A), summed over ages 0, 1, 2, … 17}$$

4. We divided the total number of visits (NHAMCS data) by the total population (census data) to get the rate per 10,000 of children's visits for the respiratory disease:

$$\text{Rate per 10,000} = [\text{Total Number of Visits} / \text{Population}] \times 10000$$

For Table H3b, rates stratified by age group were tabulated for the years 2005, 2006, 2007, 2008, and for the four-year period 2005-2008. These rates were calculated using the same procedure as above, except that the visits and populations were summed across the children in each age group.

Race/Ethnicity

For Table H3a, rates stratified by race/ethnicity[vii] group were tabulated for the years 2005, 2006, 2007, 2008, and for the four-year period 2005-2008. These rates were calculated using the same procedure as above, except that the visits and populations were summed across the children in each race/ethnicity group and year.

The race/ethnicity groups were defined using the variables RACEIM (RACE for 1996 to 2006) and ETHIM (ETHNIC for 1996 to 2006) in the NHAMCS files. These are the patient's race and ethnicity, and are given statistically imputed values in the database if they are not reported. For

[vii] These data are not stratified by income because the NHAMCS data do not give the patient's income. Since 2006, NHAMCS reports the median family income for the patient's zip code, which would poorly match the available census income data (the patient's zip code is not available in the publicly released NHAMCS data).

children ages 17 years and under in 2005 to 2008, a weighted percentage of 16% of the values for race, 21% of the values for ethnicity, and 26% of the values for race and ethnicity combined were imputed.

These variables were coded as follows:

ETHIM: Patient ethnicity (Hispanic/Non-Hispanic)
1 = Hispanic or Latino
2 = Not Hispanic or Latino

RACEIM: Patient race
1 = White only
2 = Black/African American only
3 = Asian only
4 = Native Hawaiian/Other Pacific Islander only
5 = American Indian/Alaska Native only
6 = More than one race reported

Using these variables, the following race/ethnicity groups were defined for the NHAMCS emergency room visits data:

- All: RACEIM = any, ETHIM = any
- White non-Hispanic: RACEIM = 1, ETHIM = 2
- Black non-Hispanic: RACEIM = 2, ETHIM = 2
- American Indian/Alaska Native, Non-Hispanic: RACEIM = 5, ETHIM = 2
- Asian and Pacific Islander, Non-Hispanic: RACEIM = 3 or 4, ETHIM = 2
- Hispanic: ETHIM = 1

The associated populations were computed from the post-censal 2000 noninstitutionalized civilian population files for the year 2009 at:

http://www.census.gov/popest/data/national/asrh/2009/2009-nat-ni.html
For each month, year, and age, the file provides the total population as well as the populations by age, sex, race and ethnicity. Populations are provided for male Hispanics, female Hispanics, male Non-Hispanics, and female Non-Hispanics of the following race combinations:

RACENUM (Census data)

1. White alone
2. Black alone
3. American Indian/Alaska Native alone
4. Asian alone
5. Hawaiian or Pacific Islander alone
6. Two or more races

(Other specific multiple race combinations are also provided in the dataset, such as "White alone, or in combination with another race").

Thus the total census populations corresponding to the selected NHAMCS race/ethnicity groups are obtained by summing the populations as follows:

- All: Total population
- White non-Hispanic: RACENUM = 1, Non-Hispanic, Age <= 17, Gender = male or female
- Black non-Hispanic: RACENUM = 2, Non-Hispanic, Age <= 17, Gender = male or female
- American Indian/Alaska Native non-Hispanic: RACENUM = 3, Non-Hispanic, Age <= 17, Gender = male or female
- Asian and Pacific Islander non-Hispanic: RACENUM = 4 or 5, Non-Hispanic, Age <= 17, Gender = male or female
- Hispanic: RACENUM = 1 to 6, Hispanic, Age <= 17, Gender = male or female

Using the same four steps described above under "Calculation of Indicator," the total number of visits and total population for each race/ethnicity group are used to get the rate per 10,000 of children's visits for the respiratory disease:

$$\text{Rate per 10,000} = [\text{Total Number of Visits} / \text{Population}] \times 10000$$

Relative Standard Error

The uncertainties of the rates were calculated using SUDAAN® (Research Triangle Institute, Research Triangle Park, NC 27709) statistical survey software. SUDAAN was used to calculate the estimated percentages and the standard errors of the estimated percentages. The standard error is the estimated standard deviation of the percentage, and this depends upon the survey design. For this purpose, the public release version of NHAMCS includes the following variables:

- Masked Stratum (CSTRATM)
- Masked Primary Sampling Unit (CPSUM)

These variables are "Masked" so that the sample design represented by these variables is an approximation to the true sample design, which was not made publicly available in order to protect confidentiality. Note that starting in 2003, the public release version does not include masked sampling design variables beyond the first stage of sampling. For approximate variance estimation, the survey design can be approximated as being a multi-stage random sample where the first stage samples with replacement the masked primary sampling units from the masked strata.

The survey software was used to estimate the standard deviation of the total number of visits by children ages 17 or under for the given respiratory disease, SD (Total Visits).

The rate of visits is calculated as:

$$\text{Rate per 10000} = [\text{Total Number of Visits} / \text{Population}] \times 10000$$

Treating the census population estimates as having negligible uncertainty, we get the standard error of the rate by dividing the standard deviation of the total by the population:

$$\text{Standard Error (Rate)} = [\text{SD (Total Visits)} / \text{Population}] \times 10000$$

The relative standard error is the standard error divided by the estimated rate:

$$\text{Relative Error (\%)} = [\text{Standard Error (Rate)} / \text{Rate}] \times 100\%$$

Rates with a relative error less than 30% and with at least 30 sampled visits (for the given disease) sampled were treated as being reliable and were tabulated. Rates with a relative error greater than or equal to 30% but less than 40% and with at least 30 sampled visits were treated as being unstable; these values were tabulated but were flagged to be interpreted with caution. Rates with a relative error greater than or equal to 40% or missing or with at most 29 sampled visits were treated as being unreliable; these values were not tabulated and were flagged as having a large uncertainty.

*Hospitalizations*

Indicator H3 also shows the rate of hospitalizations by civilian children ages 17 or under that were for a given respiratory disease.

For each year and respiratory disease, we carried out the following calculations:

1. We extracted the NHDS survey data for all the hospital discharges by children ages 17 years or under for the given respiratory disease. We selected all hospital discharges where the age was between 0 and 17 years and the first three characters of the diagnosis code were:

- Asthma and all other respiratory causes: codes 460-466, 480-488, 490-496
- All respiratory causes other than asthma: codes 460-466, 480-488, 490-492, 494-496
- Upper respiratory: codes 460-466
- Pneumonia or influenza: codes 480-488
- Other lower respiratory: codes 490-492, 494-496
- Asthma: code 493

For each hospital discharge, the survey analysis weight (WEIGHT) denotes the national number of hospital discharges represented by that discharge.

2. We summed the NHDS analysis weights for all the selected hospital discharges to estimate the total number of hospital discharges by children ages 17 years or under for the given respiratory disease:

Total Number of Hospital Discharges = Σ WEIGHT,
summed over all selected discharges

3. Using the census data, we calculated the total population of children ages 17 years or under by summing the populations for the ages 0, 1, 2, … 17:

Population = Σ Population (age A), summed over ages 0, 1, 2, … 17

4. We divided the total number of hospital discharges (NHDS data) by the total population (census data) to get the estimated rate of children's hospitalizations for the respiratory disease:

Rate per 10000 = [Total Number of Hospital Discharges / Population] × 10000.

For Table H3d, rates stratified by age group were tabulated for the years 2005, 2006, 2007, 2008, and for the four-year period 2005-2008. These rates were calculated using the same procedure as above, except that the hospital discharges and populations were summed across the children in each age group.

Race

For Table H3c, rates stratified by race group were tabulated for the years 2005, 2006, 2007, 2008, and for the four-year period 2005-2008. These rates were calculated using the same procedure as above, except that the hospital discharges and populations were summed across the children in each race group and year.

The race groups were defined using the variable RACE in the NHDS files. There is no variable for Hispanic ethnicity in NHDS.

For the years 2005-2008, this variable was coded as follows:

RACE: Patient race
1 = White
2 = Black/African American
3 = American Indian/Alaskan Native
4 = Asian
5 = Native Hawaiian or other Pacific Isldr
6 = Other
8 = Multiple race indicated
9 = Not stated

Using this variable, the following race groups were defined for the NHDS hospital discharge data:

- All: RACE = any
- White: RACE = 1
- Black: RACE = 2
- American Indian/Alaska Native: RACE = 3
- Asian and Pacific Islander: RACE = 4 or 5
- All Other Races: RACE = 6, 8 or 9

Note that following NCHS recommendations, due to concerns about high uncertainty, detailed results are not presented for the American Indian/Alaskan Native, Asian and Pacific Islander, and All Other Races categories. However those three categories were used to define the race groups for the statistical comparisons presented in the "Statistical Comparisons" section below.

The associated populations were computed from the post-censal 2000 noninstitutionalized civilian population files for the year 2009 at:

http://www.census.gov/popest/data/national/asrh/2009/2009-nat-ni.html

For each month, year, and age, the file provides the total population as well as the populations by age, sex, race, and ethnicity. Populations are provided for various combinations including males and females of the following race combinations:

RACENUM (Census data)

1. White alone
2. Black alone
3. American Indian/Alaska Native alone
4. Asian alone
5. Hawaiian or Pacific Islander alone
6. Two or more races

(Other specific multiple race combinations are also provided in the dataset, such as "White alone, or in combination with another race").

Thus the total census populations corresponding to the selected NHDS race groups are obtained by summing the populations as follows:

- All: Total population
- White: RACENUM = 1, Age <= 17, Gender = male or female
- Black: RACENUM = 2, Age <= 17, Gender = male or female
- American Indian/Alaska Native: RACENUM = 3, Age <= 17, Gender = male or female
- Asian and Pacific Islander: RACENUM = 4 or 5, Age <= 17, Gender = male or female
- All Other Races: RACENUM = 6, Age <= 17, Gender = male or female

Note that following NCHS recommendations, due to concerns about high uncertainty, detailed results are not presented for the American Indian/Alaskan Native, Asian and Pacific Islander, and All Other Races categories. However those three categories were used to define the race groups for the statistical comparisons presented in the "Statistical Comparisons" section below.

Using the same four steps described above under "Calculation of Indicator," the total number of discharges and total population for each race group are used to get the estimated rate per 10,000 of children's hospitalizations for the respiratory disease:

Rate per 10,000 = [Total Number of Hospital Discharges / Population] × 10000

Relative Standard Error

The uncertainties of the rates were computed for the years 1996 to 2008 using approximate relative standard error equations provided in the file documentation for each year. The documentation is provided at the ftp site:

ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHDS/

The equation provided in the documentation is of the form:

Relative Standard Error (Total Discharges) =
$\sqrt{(a + b / \text{Total Discharges})} \times 100\%$

The relative standard error is defined as the standard deviation divided by the estimated value:

Relative Standard Error (Total Discharges) =
[Standard Deviation (Total Discharges) / Total Discharges] × 100%

To derive error estimates for public release that would be applicable to a wide variety of statistics, NCHS produced numerous estimates and their variances. NCHS then used a regression model to produce best-fit curves, based on the empirically determined relationship between the size of an estimate X and its relative variance. The regression intercepts a and slopes b were tabulated by NCHS for various population subgroups and selected statistics.

The NCHS tabulated parameters a and b for the first-listed diagnosis for the Under 15 age group are listed in the following table.

| Year | a | b |
|---|---|---|
| 1996 | 0.017 | 229.443 |
| 1997 | 0.0147 | 181.262 |
| 1998 | 0.013772 | 221.956 |
| 1999 | 0.016494 | 223.072 |
| 2000 | 0.021332 | 284.1142 |
| 2001 | 0.019559 | 255.6805 |

| Year | a | b |
|---|---|---|
| 2002 | 0.0211 | 241.964 |
| 2003 | 0.02189 | 278.306 |
| 2004 | 0.02165 | 252.708 |
| 2005 | 0.02222 | 211.185 |
| 2006 | 0.02734 | 220.637 |
| 2007 | 0.036972 | 167.01187 |
| 2008 | 0.05044 | 516.705 |

The rate equals the total discharges divided by the total population:

$$\text{Rate per 10000} = [\text{Total Number of Hospital Discharges} / \text{Population}] \times 10000$$

The relative standard error of each rate is the estimated standard deviation of the rate divided by the estimated rate. Assuming that the uncertainty of the census populations is negligible, the relative standard error of the rate is equal to the relative standard error of the total discharges:

$$\text{Relative Standard Error (Rate)} = \sqrt{(a + b / \text{Total Discharges})} \times 100\%$$

For the rates for the years 2005-2008 combined, the calculation is more complicated.

1. Use the above equations for each year, 2005, 2006, 2007 and 2008 to obtain the standard deviation for the total discharges in that year:

$$\text{SD (Total Discharges, Year Y)} = [\text{Relative Standard Error (Total Discharges)} \times \text{Total Discharges}] / 100 = \sqrt{(a + b / \text{Total Discharges})} \times \text{Total Discharges}$$

2. Calculate the variance, Var, for each year:

$$\text{Var (Total Discharges, Year Y)} = [\text{SD (Total Discharges, Year Y)}]^2$$

3. Estimate the total discharges for years 2005 to 2008 by summing the four annual estimates:

$$\text{Total Discharges (2005-2008)} = \text{Total Discharges (2005)} + \text{Total Discharges (2006)} + \text{Total Discharges (2007)} + \text{Total Discharges (2008)}$$

4. Estimate the total population for years 2005 to 2008 by summing the four annual populations:

$$\text{Total Population (2005-2008)} = \text{Population (2005)} + \text{Population (2006)} + \text{Population (2007)} + \text{Population (2008)}$$

5. Estimate the rate for years 2005-2008 by dividing the total discharges by the total population:

$$\text{Rate per 10000 (2005-2008)} =$$

[Total Discharges (2005-2008) / Total Population (2005-2008)] × 10000

6. Estimate the variance of the total discharges for 2005-2008. Assuming that the annual estimates are (approximately) independent, the variance of the sum equals the sum of the variances, which gives:

Var (Total Discharges (2005-2008)) = Var (Total Discharges, 2005) +
Var (Total Discharges, 2006) + Var (Total Discharges, 2007)
+ Var (Total Discharges, 2008)

(This uses the results of the second step).

7. Calculate the standard deviation of the total discharges for 2005-2008:

SD (Total Discharges, 2005-2008) = √[Var (Total Discharges, 2005-2008)]

8. Calculate the relative standard error of the total discharges using the results of the third and seventh steps:

Relative Standard Error (Total Discharges, 2005-2008) =
[SD (Total Discharges, 2005-2008) / (Total Discharges, 2005-2008)] × 100%

9. Calculate the relative standard error of the rate of discharges for 2005-2008. Assuming the populations have negligible uncertainty, it again follows that the relative standard error of the rate equals the relative standard error of the total discharges, which is given in the eighth step:

Relative Standard Error (Rate per 10000, 2005-2008) =
Relative Standard Error (Total Discharges, 2005-2008)

Rates with a relative error less than 30% and at least 30 sampled hospital discharges (for the given disease) were treated as being reliable and were tabulated. Rates with a relative error greater than or equal to 30% but less than 40% and with at least 30 sampled hospital discharges were treated as being unstable; these values were tabulated but were flagged to be interpreted with caution. Rates with a relative error greater than or equal to 40% or missing or with at most 29 sampled hospital discharges were treated as being unreliable; these values were not tabulated and were flagged as having a large uncertainty.

**Questions and Comments**

Questions regarding these methods, and suggestions to improve the description of the methods, are welcome. Please use the "Contact Us" link at the bottom of any page in the America's Children and the Environment website.

**Statistical Comparisons**

Statistical analyses of the emergency room visit rates or hospitalization rates were used to determine whether the differences between rates for different demographic groups were statistically significant. For these analyses, the rates and their standard errors were calculated for each combination of age group, sex, and race/ethnicity or race group using the method described in the corresponding "Relative Standard Error" section. For emergency room visits, rates and their standard errors are calculated for each combination of age group, sex, and race/ethnicity. For hospitalizations, rates and the relative standard errors of the rates are calculated for each combination of age group, sex, and race. The standard error of the rate is given by the product of the rate and its relative standard error. These calculated standard errors account for the survey weighting and design.

Using a weighted linear regression model, the rate was assumed to be the sum of explanatory terms for age, sex, and/or race/ethnicity or race and a random error term; the error terms were assumed to be approximately independent and normally distributed with a mean of zero and a variance equal to the square of the standard error. Using this model, the difference in the value of a rate between different demographic groups is statistically significant if the difference between the corresponding sums of explanatory terms is statistically significantly different from zero. A p-value at or below 0.05 implies that the difference is statistically significant at the 5% significance level. No adjustment is made for multiple comparisons.

For each type of comparison, we present unadjusted and adjusted analyses. The unadjusted analyses directly compare a rate between different demographic groups. The adjusted analyses add other demographic explanatory variables to the statistical model and use the statistical model to account for the possible confounding effects of these other demographic variables. For example, the unadjusted race/ethnicity comparisons for emergency room visits use and compare the visit rates between different race/ethnicity pairs. The adjusted race/ethnicity comparisons use the rates for each age/sex/race/ethnicity combination. The adjusted analyses add age and sex terms to the statistical model and compare the rates between different race/ethnicity pairs after accounting for the effects of the other demographic variables. For example, if Hispanic children tend to be younger than White non-Hispanics, and if the visit rate strongly depends on age only, then the unadjusted differences between these two race/ethnicity groups would be significant but the adjusted difference (taking into account age) would not be significant.

Comparisons of emergency room visit rates for asthma and other respiratory causes between pairs of race/ethnicity groups are shown in Table 1. Comparisons of hospitalization rates for asthma and other respiratory causes between Whites and Blacks are also shown in Table 1. In Table 1, for the "Unadjusted" comparisons, the only explanatory variables are terms for each race/ethnicity or race group. For these unadjusted comparisons, the statistical tests compare the percentiles for each pair of race/ethnicity or race groups. For the "Adjusted for age, sex" comparisons, the explanatory variables are terms for each race/ethnicity or race group together with terms for each age group and sex. For these adjusted comparisons, the statistical test compares the pair of race/ethnicity or race groups after accounting for any differences in the age and sex distributions between the race/ethnicity or race groups.

Additional comparisons are shown in Table 2 for emergency room visits and in Table 3 for hospitalizations. The Against = "age" unadjusted p-value compares the rates between all the age groups. The adjusted p-value includes adjustment terms for sex and race/ethnicity or race in the model. The Against = "year" unadjusted p-value compares the trends in the rates by regressing against the calendar year. The adjusted p-value includes adjustment terms for age, sex and race/ethnicity or race in the model.

For the analyses of emergency room visits, the race/ethnicity groups used were: White non-Hispanic; Black non-Hispanic; API non-Hispanic; AIAN non-Hispanic; Hispanic; Other. API denotes either Asian or Native Hawaiian or Pacific Islander. AIAN denotes American Indian or Alaska Native. For these data the "Other" race/ethnicity category denotes children reporting multiple races and was not an available category for the years 1996 to 1998. For the analyses of hospitalizations, the race groups used were: White; Black; API; AIAN; All Other Races. API denotes either Asian or Native Hawaiian or Pacific Islander. AIAN denotes American Indian or Alaska Native. For these data the "All Other Races" category includes children of Other races,[viii] children of multiple races (for 2000 or later), and children with a race that was not stated. For the analyses of emergency room visits and hospitalizations, the age groups used were: < 12 months, 1 to < 2 years, 2 to < 3 years, 3 to < 6 years, 6 to < 11 years, 11 to < 16 years, and 16 to < 18 years.

For more details on these statistical analyses, see the memorandum by Cohen (2011). [ix]

Table 1. Statistical significance tests comparing the rates of emergency room visits or hospitalizations for asthma and other respiratory causes by children ages 0 to 17 years, between pairs of race/ethnicity groups, for 2005-2008.

| | | | P-VALUES | | |
|---|---|---|---|---|---|
| **Indicator** | **Variable** | **First race/ethnicity group** | **Second race/ethnicity group** | **Unadjusted** | **Adjusted for age, sex** |
| Emergency room visits | Asthma and all other respiratory causes | White non-Hispanic | Black non-Hispanic | < 0.001 | < 0.001 |
| Emergency room visits | Asthma and all other respiratory causes | White non-Hispanic | API non-Hispanic | 0.745 | 0.032 |
| Emergency room visits | Asthma and all other respiratory causes | White non-Hispanic | AIAN non-Hispanic | 0.142 | < 0.001 |
| Emergency room visits | Asthma and all other respiratory causes | White non-Hispanic | Hispanic | 0.030 | 0.084 |
| Emergency room visits | Asthma and all other respiratory causes | Black non-Hispanic | API non-Hispanic | < 0.001 | < 0.001 |
| Emergency room visits | Asthma and all other respiratory causes | Black non-Hispanic | AIAN non-Hispanic | < 0.001 | < 0.001 |
| Emergency room visits | Asthma and all other respiratory causes | Black non-Hispanic | Hispanic | < 0.001 | < 0.001 |
| Emergency room visits | Asthma and all other respiratory causes | API non-Hispanic | AIAN non-Hispanic | 0.318 | 0.881 |
| Emergency room visits | Asthma and all other respiratory causes | API non-Hispanic | Hispanic | 0.421 | 0.007 |
| Emergency room | Asthma and all other | AIAN non-Hispanic | Hispanic | 0.005 | < 0.001 |

[viii] Although the NHDS hospital discharge data includes Other races as a possible category, the corresponding census population data only provides estimates for the single race groups: White, Black,, Asian, AIAN, Hawaiian and Pacific Islander; and for multiple races.

[ix] Cohen, J. 2011. *Selected statistical methods for testing for trends and comparing years or demographic groups in other ACE health-based indicators.* Memorandum from J. Cohen, ICF to Dan Axelrad, EPA, 16 June, 2011.

| | | | P-VALUES | | |
|---|---|---|---|---|---|
| **Indicator** | **Variable** | **First race/ethnicity group** | **Second race/ethnicity group** | **Unadjusted** | **Adjusted for age, sex** |
| visits | respiratory causes | | | | |
| Hospitalizations | Asthma and all other respiratory causes | White | Black | < 0.001 | < 0.001 |

Table 2. Other statistical significance tests comparing the rates of emergency room visits for asthma and other respiratory causes by children ages 0 to 17 years for 2005 to 2008 (trends for 1996-2008).

| | | | | P-VALUES | |
|---|---|---|---|---|---|
| **Variable** | **From** | **To** | **Against** | **Unadjusted** | **Adjusted*** |
| Asthma and all other respiratory causes | 2005 | 2008 | age | < 0.001 | < 0.001 |
| Asthma and all other respiratory causes | 1996 | 2008 | year | 0.381 | < 0.001 |
| Other respiratory causes | 1996 | 2008 | year | 0.679 | 0.056 |
| Asthma | 1996 | 2008 | year | 0.023 | 0.004 |

*For Against = ”age,” the p-values are adjusted for sex and race/ethnicity.
For Against = “year,” the p-values are adjusted for age, sex, and race/ethnicity.

Table 3. Other statistical significance tests comparing the rates of hospitalizations for asthma and other respiratory causes by children ages 0 to 17 years for 2005 to 2008 (trends for 1996-2008).

| | | | | P-VALUES | |
|---|---|---|---|---|---|
| **Variable** | **From** | **To** | **Against** | **Unadjusted** | **Adjusted*** |
| Asthma and all other respiratory causes | 2005 | 2008 | age | < 0.001 | < 0.001 |
| Asthma and all other respiratory causes | 1996 | 2008 | year | < 0.001 | < 0.001 |
| Other respiratory causes | 1996 | 2008 | year | 0.001 | < 0.001 |
| Asthma | 1996 | 2008 | year | < 0.001 | < 0.001 |

*For Against = ”age,” the p-values are adjusted for sex and race.
For Against = “year,” the p-values are adjusted for age, sex, and race.