Methods

Indicator

H6. Percentage of children ages 5 to 17 years reported to have attention deficit/hyperactivity disorder, by sex, 1997-2017.
H7. Percentage of children ages 5 to 17 years reported to have a learning disability, by sex, 1997-2017.
H8. Percentage of children ages 5 to 17 years reported to have autism, 1997-2017.
H9. Percentage of children ages 5 to 17 years reported to have intellectual disability (mental retardation), 1997-2017.

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 neurodevelopmental disorders for children ages 5 to 17 from the NHIS 1997 to 2017 surveys. Indicator H6 gives the trends in the percentages of children reported to have attention deficit/hyperactivity disorder, stratified by sex. Indicator H7 gives the trends in the percentages of children reported to have a learning disability, stratified by sex. Indicator H8 gives the trends in the percentages of children reported to have autism. Indicator H9 gives the trends in the percentages of children reported to have intellectual disability (mental retardation), stratified by sex. For each indicator, the corresponding table H6a, H7a, H8a, and H9a gives the percentage of children reported to have the given neurodevelopmental disorder over the period 2014 to 2017, stratified both by age and sex. For each indicator, the corresponding table H6b, H7b, H8b, and H9b gives the percentage of children reported to have the given neurodevelopmental disorder over the period 2014 to 2017, 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). 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	H6. Percentage of children ages 5 to 17 years reported to have attention deficit/hyperactivity disorder (ADHD), by sex, 1997-2017. H7. Percentage of children ages 5 to 17 years reported to have a learning disability, 1997-2017. H8. Percentage of children ages 5 to 17 years reported to have autism, 1997-2017. H9. Percentage of children ages 5 to 17 years reported to have intellectual disability (mental retardation), 1997-2017.
Time Period	1997-2017
Data	Neurodevelopmental disorder prevalence in children ages 5 to 17

Indicator	H6. Percentage of children ages 5 to 17 years reported to have attention deficit/hyperactivity disorder (ADHD), by sex, 1997-2017. H7. Percentage of children ages 5 to 17 years reported to have a learning disability, 1997-2017. H8. Percentage of children ages 5 to 17 years reported to have autism, 1997-2017. H9. Percentage of children ages 5 to 17 years reported to have intellectual disability (mental retardation), 1997-2017.

Indicator	H6. Percentage of children ages 5 to 17 years reported to have attention deficit/hyperactivity disorder (ADHD), by sex, 1997-2017. H7. Percentage of children ages 5 to 17 years reported to have a learning disability, 1997-2017. H8. Percentage of children ages 5 to 17 years reported to have autism, 1997-2017. H9. Percentage of children ages 5 to 17 years reported to have intellectual disability (mental retardation), 1997-2017.

Years (1997-2003)	1997	1998	1999	2000	2001	2002	2003
Children	10,006	9,564	9,169	9,506	9,638	8,876	8,738
ADHD non-missing responses (%)	9,971 (99.6%)	9,536 (99.7%)	9,155 (99.8%)	9,481 (99.7%)	9,617 (99.8%)	8,845 (99.7%)	8,722 (99.8%)
ADHD missing responses (%)	35 (0.4%)	28 (0.3%)	14 (0.2%)	25 (0.3%)	21 (0.2%)	31 (0.3%)	16 (0.2%)
Learning disability non-missing responses (%)	9,974 (99.7%)	9,552 (99.9%)	9,155 (99.8%)	9,490 (99.8%)	9,624 (99.9%)	8,862 (99.8%)	8,724 (99.8%)
Learning disability missing responses (%)	32 (0.3%)	12 (0.1%)	14 (0.2%)	16 (0.2%)	14 (0.1%)	14 (0.2%)	14 (0.2%)
Autism non-missing responses (%)	9,996 (99.9%)	9,557 (99.9%)	9,165 (100.0%)	9,501 (99.9%)	9,633 (99.9%)	8,873 (100.0%)	8,730 (99.9%)
Autism missing	10 (0.1%)	7 (0.1%)	4 (0.0%)	5 (0.1%)	5 (0.1%)	3 (0.0%)	8 (0.1%)

Indicator	H6. Percentage of children ages 5 to 17 years reported to have attention deficit/hyperactivity disorder (ADHD), by sex, 1997-2017. H7. Percentage of children ages 5 to 17 years reported to have a learning disability, 1997-2017. H8. Percentage of children ages 5 to 17 years reported to have autism, 1997-2017. H9. Percentage of children ages 5 to 17 years reported to have intellectual disability (mental retardation), 1997-2017.						
responses (%)							
Intellectual disability non-missing responses (%)	9,991 (99.8%)	9,549 (99.8%)	9,165 (100.0%)	9,494 (99.9%)	9,628 (99.9%)	8,856 (99.8%)	8,728 (99.9%)
Intellectual disability missing responses (%)	15 (0.2%)	15 (0.2%)	4 (0.0%)	12 (0.1%)	10 (0.1%)	20 (0.2%)	10 (0.1%)
Years (2004-2010)	2004	2005	2006	2007	2008	2009	2010
Children	8,830	8,974	7,019	6,604	6,328	8,009	7,995
ADHD non-missing responses (%)	8,813 (99.8%)	8,952 (99.8%)	7,003 (99.8%)	6,595 (99.9%)	6,311 (99.7%)	7,994 (99.8%)	7,980 (99.8%)
ADHD missing responses (%)	17 (0.2%)	22 (0.2%)	16 (0.2%)	9 (0.1%)	17 (0.3%)	15 (0.2%)	15 (0.2%)
Learning disability non-missing responses (%)	8,823 (99.9%)	8,959 (99.8%)	7,004 (99.8%)	6,583 (99.7%)	6,319 (99.9%)	8,001 (99.9%)	7,986 (99.9%)
Learning disability missing responses (%)	7 (0.1%)	15 (0.2%)	15 (0.2%)	21 (0.3%)	9 (0.1%)	8 (0.1%)	9 (0.1%)
Autism non-missing responses (%)	8,825 (99.9%)	8,971 (100.0%)	7,012 (99.9%)	6,600 (99.9%)	6,328 (100.0%)	8,004 (99.9%)	7,987 (99.9%)

Indicator	H6. Percentage of children ages 5 to 17 years reported to have attention deficit/hyperactivity disorder (ADHD), by sex, 1997-2017. H7. Percentage of children ages 5 to 17 years reported to have a learning disability, 1997-2017. H8. Percentage of children ages 5 to 17 years reported to have autism, 1997-2017. H9. Percentage of children ages 5 to 17 years reported to have intellectual disability (mental retardation), 1997-2017.						
Autism missing responses (%)	5 (0.1%)	3 (0.0%)	7 (0.1%)	4 (0.1%)	0 (0.0%)	5 (0.1%)	8 (0.1%)
Intellectual disability non-missing responses (%)	8,828 (100.0%)	8,968 (99.9%)	7,015 (99.9%)	6,603 (100.0%)	6,322 (99.9%)	8,005 (99.9%)	7,990 (99.9%)
Intellectual disability missing responses (%)	2 (0.0%)	6 (0.1%)	4 (0.1%)	1 (0.0%)	6 (0.1%)	4 (0.1%)	5 (0.1%)
Years (2011-2017)	2011	2012	2013	2014	2015	2016	2017
Children	9,040	9,481	9,252	9,689	8,881	8,065	6,439
ADHD non-missing responses (%)	9,029 (99.9%)	9,464 (99.8%)	9,243 (99.9%)	9,668 (99.8%)	8,866 (99.8%)	8,049 (99.8%)	6,431 (99.8%)
ADHD missing responses (%)	11 (0.1%)	17 (0.2%)	9 (0.1%)	21 (0.2%)	15 (0.2%)	16 (0.2%)	8 (0.1%)
Learning disability non-missing responses (%)	9,031 (99.9%)	9,469 (99.9%)	9.244 (99.9%)	9,672 (99.8%)	8,872 (99.9%)	8,051 (99.8%)	6,428 (99.8%)
Learning disability missing responses (%)	9 (0.1%)	12 (0.1%)	8 (0.1%)	17 (0.2%)	9 (0.1%)	14 (0.2%)	11 (0.2%)

Indicator	H6. Percentage of children ages 5 to 17 years reported to have attention deficit/hyperactivity disorder (ADHD), by sex, 1997-2017. H7. Percentage of children ages 5 to 17 years reported to have a learning disability, 1997-2017. H8. Percentage of children ages 5 to 17 years reported to have autism, 1997-2017. H9. Percentage of children ages 5 to 17 years reported to have intellectual disability (mental retardation), 1997-2017.						
Autism non-missing responses (%)	9, 036 (100.0%)	9,474 (99.9%)	9,246 (99.9%)	9,682 (99.9%)	8,875 (99.9%)	8,055 (99.9%)	6,433 (99.9%)
Autism missing responses (%)	4 (0.0%)	7 (0.1%)	6 (0.1%)	7 (0.1%)	6 (0.1%)	10 (0.1%)	6 (0.1%)
Intellectual disability non-missing responses (%)	9, 037 (100.0%)	9,477 (100.0%)	9.246 (99.9%)	9,681 (99.9%)	8,872 (99.9%)	8,060 (99.9%)	6,438 (100.0%
Intellectual disability missing responses (%)	3 (0.0%)	4 (0.0%)	6 (0.1%)	8 (0.1%)	9 (0.1%)	5 (0.1%)	1 (0.0%)

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 1997-2017: Sample Child file samchild.dat, Person file personsx.dat, Family file familyxx.dat, Imputed Income files: 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 2006-2015, PSTRAT for 2016-2017), the pseudo-PSU (PSU for 1997-2005, PSU_P for 2006-2015, PPSU for 2016-2017), the race (RACE for 1997-1998, RACER_P for 1999, RACERP_I for 2000-2002, RACERPI2 for 2003-2017, using the 1997 OMB definitions), the Hispanic origin (ORIGIN for 1997-1999, ORIGIN_I for 2000-2017), and the detailed Hispanic origin (HISPAN_P for 1997-1998, HISPANCR

for 1999, HISPAN_I for 2000-2017). 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 the year 1997 to 2008 imputed income files we need the imputed poverty income ratio (RAT_CATI) which gives the ratio in categories. 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 2010-2017 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 (PX for 1997-2000), For the years 1997 to 2010, the questionnaire variables needed for these analyses are the responses to the following questions: “Has a doctor or health professional ever told you that <child’s name> had Attention Deficit Hyperactivity Disorder (ADHD) or Attention Deficit Disorder (ADD)?” “Has a doctor or health professional ever told you that <child’s name> had Autism?” “Has a doctor or health professional ever told you that <child’s name> had Mental Retardation?” and “Has a representative from a school or a health professional ever told you that <child’s name> had a learning disability?” For the years 2011 to 2013, the questionnaire variables needed for these analyses are the responses to the following questions: “Has a doctor or health professional ever told you that <child’s name> had Attention Deficit Hyperactivity Disorder (ADHD) or Attention Deficit Disorder (ADD)?” “Has a doctor or health professional ever told you that <child’s name> had Autism/Autism Spectrum Disorder?” “Has a doctor or health professional ever told you that <child’s name> had an intellectual disability, also known as mental retardation?” and “Has a representative from a school or a health professional ever told you that <child’s name> had a learning disability?” For the years 2014 to 2017, the questionnaire variables needed for these analyses are the responses to the following questions: “Has a doctor or health professional ever told you that <child’s name> had Attention Deficit Hyperactivity Disorder (ADHD) or Attention Deficit Disorder (ADD)?” “Has a doctor or health professional ever told you that <child’s name> had Autism, Asperger’s disorder, pervasive developmental disorder, or autism spectrum disorder?” “Has a doctor or health professional ever told you that <child’s name> had an intellectual disability, also known as mental retardation?” and “Has a representative from a school or a health professional ever told you that <child’s name> had a learning disability?”

-

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 neurodevelopmental disorder prevalence questions for children ages 5 to 17 years for the surveys from 1997 to 2017. The NHIS data were obtained from the NHIS website: http://www.cdc.gov/nchs/nhis.htm.

For these indicators we used the responses to the following questions. Attention Deficit/Hyperactivity Disorder: "Has a doctor or health professional ever told you that <child's name> had Attention Deficit/Hyperactivity Disorder (ADHD) or Attention Deficit Disorder (ADD)?" Learning disability: "Has a representative from a school or a health professional ever told you that <child's name> had a learning disability?" Autism for years 1997 to 2010: "Has a doctor or health professional ever told you that <child's name> had Autism?" Autism for years 2011 to 2013: "Has a doctor or health professional ever told you that <child's name> had Autism/Autism Spectrum Disorder?" Autism for years 2014 to 2017: "Has a doctor or health professional ever told you that <child's name> had Autism, Asperger's disorder, pervasive developmental disorder, or autism spectrum disorder?" (In previous years, the autism question was included in a set of questions about 10 health conditions, asked after the question about any other developmental delay. In 2014 to 2017, the autism question was a separate question asked before the question about any other developmental delay). Intellectual disability (mental retardation) for years 1997 to 2010: "Has a doctor or health professional ever told you that <child's name> had Mental Retardation?" Intellectual disability (mental retardation) for years 2011 to 2017: "Has a doctor or health professional ever told you that <child's name> had an intellectual disability, also known as mental retardation?"

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 to 2015, Blacks, Hispanics, and Asians were over-sampled. No over-sampling was implemented from 2016 to 2017. The publicly released data includes 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 and changed again in 2016. New strata were defined and PSUs were selected from these new strata. For example, pseudo-stratum 100 for 1997-2005 is unrelated to pseudo-stratum 100 for 2006-2015 and pseudo-stratum 100 for 2016. To properly treat the 2006-2015 data as independent from the 1997 to 2005 data, 1,000 was added to each of the year 2006 to 2015 pseudo-stratum numbers for these statistical analyses. To properly treat the 2016 to 2017 data as independent from the 1997 to 2015 data, 2,000 was added to each of the year 2016 to 2017 pseudo-stratum numbers for these statistical analyses.[i]

Race/Ethnicity and Family Income

For Tables H6b, H7b, H8b, and H9b, the prevalence percentages were calculated for demographic strata defined by the combination of race/ethnicity and family income.

[i] The addition of 1,000 for 2006 to 2015 was chosen to make the stratum numbers for 2005 and earlier distinct from the stratum numbers for 2006 to 2015. The addition of 2,000 for 2016 to 2017 was chosen to make the stratum numbers for 2016 to 2017 distinct from the stratum numbers for 2015 and earlier. This follows the recommendations in Appendix IV of the survey description document "2017 National Health Interview Survey (NHIS) Public Use Data Release Survey Description," CDC, June, 2018, http://www.cdc.gov/nchs/nhis/quest_data_related_1997_forward.htm.

The family income was characterized based on the poverty income ratio variable (POVRATI3 for 2010 and later), 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 2010 and later, the public release variable POVRATI3 gives the numerical value of PIR in thousandths. For prior years, the numerical values of PIR 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.
- Between 100% and 200% of Poverty Level: 1 ≤ PIR < 2.
- Above 200% of Poverty level: PIR ≥ 2.
- Above Poverty Level: PIR ≥ 1 (combines the previous two groups).
- Unknown Income: PIR is missing ("undefinable").[ii]

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. From 2007 to 2010, the income ranges were either 0-$34,999, $35,000-$49,999, $50,000-74,999, $75,000-$99,999, or $100,000 and above. For 2011 to 2017, the additional questions included questions about the size of the family and whether the income was above or below 100%, 138%, 200%, 250%, or 400% of the poverty threshold, and the income ranges were either 0-$34,999, $35,000-$49,999, $50,000-74,999, $75,000-$99,999, $100,000-$149,999, or $150,000 and above.. From 2007 to 2017, 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 were 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

[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.

[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/tecdoc17.pdf. August, 2018.

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 5 to 17 years for the years 2014 to 2017, the weighted percentage of children with imputed poverty income ratios was 17%.

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

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

[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.

- 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
- Other: RACERPI2 = 3, 5 or 6, ORIGIN_I = 2
    - American Indian, Alaska Native, Non-Hispanic: RACERPI2 = 3, ORIGIN_I = 2

The "Other" category includes non-Hispanic respondents whose race is not White, Black, or Asian, or who report multiple races.

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 5 to 17 years for the years 2014 to 2016, the weighted percentage of children with an imputed race or ethnicity was 9%. Among the sampled Hispanic (defined by ORIGIN_I) children ages 5 to 17 years for the years 2014 to 2016, the weighted percentage of children with an imputed specific Hispanic origin was 2%. To protect subject confidentiality, the public release version of NHIS for the year 2017 did not include information of which subjects had their race, ethnicity, or specific Hispanic origin imputed.

Calculation of Indicator

Indicator H6 is the percentage of children reported to have attention deficit/hyperactivity disorder. Indicator H7 is the percentage of children reported to have a learning disability. Indicator H8 is the percentage of children reported to have autism. Indicator H9 is the percentage of children reported to have intellectual disability (mental retardation). For each indicator, the corresponding table H6a, H7a, H8a, and H9a gives the percentage of children reported to have the given neurodevelopmental disorder during the period 2014 to 2017, stratified both by age and sex. For each indicator, the corresponding table H6b, H7b, H8b, and H9b gives the percentage of children reported to have the given neurodevelopmental disorder during the period 2014 to 2017, stratified both by race/ethnicity and family income.

To simply demonstrate the calculations, we will describe the calculations for the indicator H6 for 2010, using the NHIS 2010 responses to the question: "Has a doctor or health professional ever told you that <child's name> had Attention Deficit/Hyperactivity Disorder (ADHD) or Attention Deficit Disorder (ADD)?" We shall refer to this question as the ADHD question. The

calculations for the other indicators and supplementary tables use exactly the same method, except for the stratification by family income, which uses the five sets of imputed income values as demonstrated below. We have rounded all the numbers to make the calculations easier.

We begin with all the non-missing responses to the ADHD question in the NHIS 2010 survey for children ages 5 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. children in 2010 represented by that sampled child. For example, the first response for a child aged 5 to 17 years was No with a survey weight of 3,000, and so represents 3,000 children ages 5 to 17 years. A second child aged 5 to 17 years responded No with a survey weight of 9,000, and so represents 9,000 children ages 5 to 17 years. A third child aged 5 to 17 years responded Yes with a survey weight of 16,000, and so represents 16,000 children ages 5 to 17 years. The total of the survey weights for the sampled children equals 50 million, the total U.S. population of children ages 5 to 17 years for the year 2010.

To calculate the proportion of children ages 5 to 17 years with ADHD/ADD, we can use the survey weights to expand the data to the 2010 U.S. population of 50 million children ages 5 to 17 years. We have 3,000 No responses from the first child, 9,000 No responses from the second child, 16,000 Yes responses from the third child, and so on. Of these 50 million responses, a total of 5 million responses are Yes and the remaining 45 million are No. Thus 5 million of the 50 million children have ADHD/ADD, giving a proportion of about 10%.

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 7,980 of the 7,995 sampled children ages 5 to 17 years. (Don't know responses or refusals to answer are treated as missing). The survey weights for all 7,995 sampled children add up to 53.2 million, the total U.S. population of children ages 5 to 17 years. The survey weights for the 7,980 sampled children with non-missing responses add up to 53.1 million. Thus the available data represent 53.1 million children, which is more than 99 %, but not all, of the 2010 U.S. population of children ages 5 to 17 years. The survey weights for the Yes responses add up to 5.0 million, which is 9.5 % of the population with responses (5.0 million/53.1 million = 9.5 %). 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 in Tables H6b, H7b, H8b, and H9b, we use the five sets of imputed income values, which each give different results. For example, suppose we wish to estimate the proportion of White non-Hispanic children below the poverty level with ADHD/ADD in 2009-2012. Using the above calculation method applied for White non-Hispanic children below the poverty level for the combined set of years 2009 to 2012, the proportions for the five sets of imputed values are:18.8%, 19.1%, 19.0%, 19.0% and 19.1%. The estimated proportion of White non-Hispanic children below the poverty level with ADHD/ADD in 2009-2012 is given by the average of the five estimates, (18.8 + 19.1 + 19.0 + 19.0 + 19.1) /5 = 19.0 %.

Equations

The following equations give the mathematical calculations for the example of White non-Hispanic children below the poverty level using the ADHD question. Let w(i) denote the survey weight for the i'th surveyed White non-Hispanic child of ages 5 to 17 years. Exclude any surveyed children with a response other than Yes or No. For the ADHD 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 ADHD/ADD 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 ADHD/ADD:

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

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.

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 and changed again in 2016. New strata were defined and PSUs were selected from these new strata. For example, pseudo-stratum 100 for 1997-2005 is unrelated to pseudo-stratum 100 for 2006-2015 and pseudo-stratum 100 for 2016. To properly treat the 2006-2015 data as independent from the 1997 to 2005 data, 1,000 was added to each of the year 2006 to 2015 pseudo-stratum numbers for these statistical analyses. To properly treat the 2016 to 2017 data as independent from the 1997 to 2015 data, 2,000 was added to each of the year 2016 to 2017 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 is 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, Other. 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 two age groups: 5-10 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 neurodevelopmental disorder 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 each neurodevelopmental disorder in children ages 5 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 are used in Table 2 for the unadjusted comparisons between one 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 race/ethnicity/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 5 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 5 to 17 years.

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 5 to 17 years with neurodevelopmental disorders, between pairs of race/ethnicity groups, for 2014-2017.

			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)**
ADHD/ADD	White non-Hispanic	Black non-Hispanic	0.532	0.027	0.067	0.088	0.099	0.091
ADHD/ADD	White non-Hispanic	Asian non-Hispanic	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
ADHD/ADD	White non-Hispanic	Hispanic	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
ADHD/ADD	White non-Hispanic	Other	0.257	0.424	0.685	0.612	0.795	0.625
ADHD/ADD	Black non-Hispanic	Asian non-Hispanic	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
ADHD/ADD	Black non-Hispanic	Hispanic	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
ADHD/ADD	Black non-Hispanic	Other	0.156	0.033	0.095	0.092	0.207	0.132
ADHD/ADD	Asian non-Hispanic	Hispanic	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
ADHD/ADD	Asian non-Hispanic	Other	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
ADHD/ADD	Hispanic	Other	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
ADHD/ADD	Mexican	Puerto Rican	< 0.001	< 0.001	< 0.001	< 0.001	0.013	0.009
Learn disability	White non-Hispanic	Black non-Hispanic	0.024	0.993	0.172	0.200	0.762	0.780
Learn disability	White non-Hispanic	Asian non-Hispanic	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Learn disability	White non-Hispanic	Hispanic	0.207	0.001	< 0.001	< 0.001	0.351	0.415

[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	**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)**
Learn disability	White non-Hispanic	Other	0.487	0.914	0.472	0.508	0.930	0.815
Learn disability	Black non-Hispanic	Asian non-Hispanic	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Learn disability	Black non-Hispanic	Hispanic	0.002	0.007	0.002	0.001	0.356	0.415
Learn disability	Black non-Hispanic	Other	0.512	0.925	0.874	0.850	0.914	0.977
Learn disability	Asian non-Hispanic	Hispanic	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Learn disability	Asian non-Hispanic	Other	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Learn disability	Hispanic	Other	0.192	0.058	0.028	0.023	0.582	0.525
Learn disability	Mexican	Puerto Rican	< 0.001	< 0.001	< 0.001	< 0.001	0.025	0.022
Autism	White non-Hispanic	Black non-Hispanic	0.085	0.078	NA	NA	NA	NA
Autism	White non-Hispanic	Asian non-Hispanic	0.010	0.010	NA	NA	NA	NA
Autism	White non-Hispanic	Hispanic	< 0.001	< 0.001	NA	NA	NA	NA
Autism	White non-Hispanic	Other	0.867	0.955	NA	NA	NA	NA
Autism	Black non-Hispanic	Asian non-Hispanic	0.243	0.289	NA	NA	NA	NA
Autism	Black non-Hispanic	Hispanic	0.364	0.377	NA	NA	NA	NA
Autism	Black non-Hispanic	Other	0.179	0.181	NA	NA	NA	NA
Autism	Asian non-Hispanic	Hispanic	0.565	0.630	NA	NA	NA	NA
Autism	Asian non-Hispanic	Other	0.022	0.028	NA	NA	NA	NA
Autism	Hispanic	Other	0.015	0.015	NA	NA	NA	NA
Autism	Mexican	Puerto Rican	0.001	< 0.001	< 0.001	< 0.001	0.265	0.242
Int disability	White non-Hispanic	Black non-Hispanic	0.302	0.943	0.711	0.677	0.359	0.351
Int disability	White non-Hispanic	Asian non-Hispanic	0.011	0.011	0.043	0.050	0.064	0.069
Int disability	White non-Hispanic	Hispanic	0.267	0.983	0.867	0.915	0.985	0.958
Int disability	White non-Hispanic	Other	0.756	0.895	0.290	0.300	0.492	0.446
Int disability	Black non-Hispanic	Asian non-Hispanic	0.004	0.017	0.020	0.017	0.262	0.278
Int disability	Black non-Hispanic	Hispanic	0.895	0.960	0.569	0.575	0.435	0.396
Int disability	Black non-Hispanic	Other	0.915	0.876	0.222	0.222	0.284	0.250
Int disability	Asian non-Hispanic	Hispanic	0.004	0.015	0.043	0.041	0.093	0.091
Int disability	Asian non-Hispanic	Other	0.060	0.082	0.686	0.692	0.076	0.070
Int disability	Hispanic	Other	0.959	0.891	0.307	0.304	0.509	0.481
Int disability	Mexican	Puerto Rican	0.113	0.123	0.129	0.143	0.522	0.517

Table 2. Statistical significance tests comparing the percentages of children ages 5 to 17 years with neurodevelopmental disorders, between those below poverty level and those at or above poverty level, for 2014-2017.

		P-VALUES	
Variable	**Population**	**Unadjusted**	**Adjusted ***
ADHD/ADD	All	< 0.001	< 0.001
ADD/ADHD	White non-Hispanic	< 0.001	< 0.001
ADD/ADHD	Black non-Hispanic	0.001	< 0.001
ADD/ADHD	Asian non-Hispanic	0.067	0.068
ADD/ADHD	Hispanic	0.047	0.029
ADD/ADHD	Other	0.003	0.002
ADD/ADHD	Mexican	0.449	0.336
ADD/ADHD	Puerto Rican	0.004	0.004
Learn disability	All	< 0.001	< 0.001
Learn disability	White non-Hispanic	< 0.001	< 0.001
Learn disability	Black non-Hispanic	< 0.001	< 0.001
Learn disability	Asian non-Hispanic	0.154	0.159
Learn disability	Hispanic	0.003	0.002
Learn disability	Other	0.001	0.001
Learn disability	Mexican	0.193	0.142
Learn disability	Puerto Rican	0.012	0.012
Autism	All	0.958	0.371
Autism	White non-Hispanic	NA	NA
Autism	Black non-Hispanic	NA	NA
Autism	Asian non-Hispanic	NA	NA
Autism	Hispanic	NA	NA
Autism	Other	NA	NA
Autism	Mexican	0.088	0.093
Autism	Puerto Rican	0.071	0.061
Int disability	All	< 0.001	< 0.001
Int disability	White non-Hispanic	0.002	0.002
Int disability	Black non-Hispanic	0.001	0.001
Int disability	Asian non-Hispanic	0.006	0.005
Int disability	Hispanic	0.006	0.005
Int disability	Other	0.531	0.540
Int disability	Mexican	0.031	0.024
Int disability	Puerto Rican	0.063	0.062

* 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 5 to 17 years with neurodevelopmental disorders, for 2014-2017 (trends for 1997-2017 and autism trend for 1997-2013).

					P-VALUES	
Variable	**From**	**To**	**Against**	**Subset**	**Unadjusted**	**Adjusted***
ADHD/ADD	2014	2017	age		< 0.001	< 0.001
ADHD/ADD	2014	2017	income		< 0.001	< 0.001

					P-VALUES	
Variable	**From**	**To**	**Against**	**Subset**	**Unadjusted**	**Adjusted***
ADHD/ADD	2014	2017	sex		< 0.001	< 0.001
ADHD/ADD	1997	2017	year		< 0.001	< 0.001
ADHD/ADD	1997	2017	year	Boys	< 0.001	< 0.001
ADHD/ADD	1997	2017	year	Girls	< 0.001	< 0.001
Learn disability	2014	2017	age		< 0.001	< 0.001
Learn disability	2014	2017	income		< 0.001	< 0.001
Learn disability	2014	2017	sex		< 0.001	< 0.001
Learn disability	1997	2017	year		0.849	0.538
Learn disability	1997	2017	year	Boys	0.106	0.397
Learn disability	1997	2017	year	Girls	0.076	0.037
Autism	2014	2017	age		0.524	0.590
Autism	2014	2017	income		0.958	0.371
Autism	2014	2017	sex		< 0.001	< 0.001
Autism	1997	2013	year		< 0.001	< 0.001
Autism	1997	2017	year		< 0.001	< 0.001
Autism	1997	2017	year	Boys	< 0.001	< 0.001
Autism	1997	2017	year	Girls	< 0.001	< 0.001
Int disability	2014	2017	age		< 0.001	< 0.001
Int disability	2014	2017	income		< 0.001	< 0.001
Int disability	2014	2017	sex		< 0.001	< 0.001
Int disability	1997	2017	year		< 0.001	< 0.001
Int disability	1997	2017	year	Boys	< 0.001	< 0.001
Int disability	1997	2017	year	Girls	0.023	0.036

*For Against = ”age,” the comparison is between the age groups 5-10 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, 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 not missing, the comparison is the trend over different years, and the p-values are adjusted for age, race/ethnicity, and income.