What is NCHS doing to get “fit”?

Amy M. Branum, PhD
Acting Associate Director for Science
National Center for Health Statistics
NCHS - who are we?

The US' federal statistical agency for health

Home of NHANES, NHIS, NSFG, health care surveys, and vital statistics
| ~100 scientific reports and analyses | Dozens of data files | Reams of technical documentation | Hours of technical expertise and guidance |
Production of high-quality, reliable, transparent statistics on the health of the US that educate, inform, and shape health policy.
NCHS’ COVID-19 index page has received 14.2 million hits since March

COVID-19 mortality pages have received 8 million hits
What is reliable? What is “fit”? Why did we need to revisit this?

Needed a more consistent approach across data divisions

Needed better guidance for staff and data users

Narrowly focused on standard errors as measures of variation

Statements from ASA and others regarding p-values
Internal workgroup formed

Recognized a need to have clear and transparent presentation criteria that can be broadly and efficiently implemented

Published this report as a culmination of that effort
estimate from a complex survey, the effective sample size, $n_e$, is defined as the sample size, $n$, divided by the design effect ($\gamma$). One approach used to calculate $n_e$ for sample survey is:

where, in this case, the design effect is:

Documentation for specific surveys should be consulted when calculating design effects, and for specific analytic purposes.

If the number of numerator events is 0 or equal to the denominator (the complement of 0 events), the estimated proportion is undefined. As a result, the estimated variance of the proportion will be 0, and the effective sample size is undefined. In these cases, the sample size should be used to determine whether the minimum size criterion is met, and it should also be used for CI calculations.

For complex sample surveys, due to sampling design and variability, there may be cases where the effective sample size is greater than the sample size. When sample size is greater than the sample size, the sample size should be used to determine the minimum sample size criterion is met, and it should be used for CI calculations.

**Standard**

- Estimated proportions should be based on a minimum denominator sample size and effective denominator sample size (when applicable) of 30. Estimates with either a denominator or an effective denominator sample size (when applicable) less than 30 should be suppressed.

- If the number of numerator events is 0 (or its complement), then the denominator sample size should be used to calculate any CI if applicable for CI calculations.

**Confidence Intervals**

The NCHS Data Presentation Standards for Proportions are based on the evaluation of absolute and relative 95% CI widths. Absolute CI widths are compared to a threshold of 0.307, and if the CI width is greater than 0.307, the proportion is considered unreliable and should not be calculated or presented. Relative CI widths are compared to a threshold of 120% of the proportion, and if the relative CI width is greater than 120% of the proportion, the proportion is considered unreliable and should not be calculated or presented.

For complex sample surveys, if the sample size and CI criteria are met for presentation of freedom, greater than 8, then the proportion should be flagged for statistical review. The review may result in either the presentation or the suppression of the proportion.

**Complementary Proportions**

The SE and width of the CI for the complement of a proportion (1 - $p$) are the same as those for the proportion. Consequently, there is a range of proportions where the CI criteria are met for presentation and suppression. However, the larger proportion may be the most salient measure, while for others, the smaller proportion may be the most important. Typically, both proportions are not shown (e.g., only the proportion with health insurance would be shown, not both the proportion with and the proportion without health insurance).

For complex sample surveys, if the sample size and CI criteria are met for presentation and suppression of freedom, greater than 8, then the proportion should be flagged for statistical review. The review may result in either the presentation or the suppression of the proportion.

**Statistical reviewer**

- Statistical reviewer

The NCHS Data Presentation Standards for Proportions are based on the evaluation of absolute and relative 95% CI widths. Absolute CI widths are compared to a threshold of 0.307, and if the CI width is greater than 0.307, the proportion is considered unreliable and should not be calculated or presented. Relative CI widths are compared to a threshold of 120% of the proportion, and if the relative CI width is greater than 120% of the proportion, the proportion is considered unreliable and should not be calculated or presented.

For complex sample surveys, if the sample size and CI criteria are met for presentation of freedom, greater than 8, then the proportion should be flagged for statistical review. The review may result in either the presentation or the suppression of the proportion.

**Confidence Intervals**

The NCHS Data Presentation Standards for Proportions are based on the evaluation of absolute and relative 95% CI widths. Absolute CI widths are compared to a threshold of 0.307, and if the CI width is greater than 0.307, the proportion is considered unreliable and should not be calculated or presented. Relative CI widths are compared to a threshold of 120% of the proportion, and if the relative CI width is greater than 120% of the proportion, the proportion is considered unreliable and should not be calculated or presented.

For complex sample surveys, if the sample size and CI criteria are met for presentation of freedom, greater than 8, then the proportion should be flagged for statistical review. The review may result in either the presentation or the suppression of the proportion.

**Complementary Proportions**

The SE and width of the CI for the complement of a proportion (1 - $p$) are the same as those for the proportion. Consequently, there is a range of proportions where the CI criteria are met for presentation and suppression. However, the larger proportion may be the most salient measure, while for others, the smaller proportion may be the most important. Typically, both proportions are not shown (e.g., only the proportion with health insurance would be shown, not both the proportion with and the proportion without health insurance).

For complex sample surveys, if the sample size and CI criteria are met for presentation and suppression of freedom, greater than 8, then the proportion should be flagged for statistical review. The review may result in either the presentation or the suppression of the proportion.
NCHS highly recommends that data users familiarize themselves with the data standards for proportions and implement them when using NCHS data.

Standards can be implemented with a number of software packages.
WHAT ABOUT THAT P-VALUE?

1. **ASSESSING SIGNIFICANCE**

   NCHS has a few challenges

2. **SIGNIFICANT BUT NOT MEANINGFUL CHANGES**

   In our “big data” systems, does a 0.1 percentage point change mean anything even if statistically significant at the p<0.05 level?

3. **BIG CHANGES, LIMITED STATISTICAL POWER**

   Some surveys have had decreasing sample sizes

   Harder to retain sufficient power to detect seemingly large differences

4. **NO CONCRETE GUIDANCE ON NAVIGATING SIGNIFICANCE**

   NCHS has not issued guidance to staff on how to assess statistical significance

   Work in progress