Data Resource Center for Child and Adolescent Health
online, interactive data query estimate suppression and precision alert criteria
for the National Survey of Children's Health
(childhealthdata.org)

Background

Following the redesign of the survey in 2016, the annual sample size for the National Survey of is now smaller than in prior administrations of the survey. As a result, the availability of state-level estimates for some NSCH variables and survey items is more limited than in previous years. Smaller state-level sample sizes prevent development of precise estimates for some variables and survey items, particularly at the subgroup level within states.

The Data Resource Center for Child and Adolescent Health (DRC) website uses data display and suppression criteria set by the Maternal Child Health Bureau (MCHB) (sponsor of the NSCH). The DRC data query alerts data users when estimates are suppressed or flagged as imprecise using these criteria. This document describes how to interpret the precision of data query results when these criteria are applied.

What are confidence intervals?

Each DRC data query results table includes a confidence interval for each estimate included in the table. A "confidence interval" (CI) is a commonly used measure of precision when calculating statistical estimates. It alerts the data user that the estimate they are looking at is simply a sample-based **estimate**, rather than an unvarying or precise value.

Confidence intervals are typically expressed as a range with an upper and lower estimate (for example, CI 42.2-44.2) and the "true" population value is believed to lie within that range. Additionally, CIs are commonly expressed as "95% confidence intervals" – this indicates that if the population were sampled 100 times, the true value would be included in that range 95 times. Thus, there is a 95% probability that any given CI contains the true population prevalence.

The width of a CI indicates the precision of an estimate and is a function of the variability in the data as well as the sample size. As the sample size increases, a given estimate will have increasing precision and a smaller CI. Estimates with large CIs indicate less precision and may need to be interpreted with caution.

How confidence intervals are calculated for the DRC data query output tables?

The 95% confidence interval (CI) displayed on the DRC data query is calculated using a **logit transformation method**. This is the default methods used by both SPSS Complex Samples and SAS SUDAAN statistical software.

Let's take an example:

- In the 2016 NSCH, 43.2% of children had parents who reported they shared meals as a family every day of the week. The lower CI is 42.2% and the upper CI is 44.2%. Therefore, the estimate and the CI would commonly be written as **43.2% (95% CI 42.2-44.2).**

- The **absolute confidence interval** is calculated by subtracting the lower CI from the upper CI. In our example, this would be 44.2% - 42.2%, or 2.0%.
- The **relative confidence interval** is calculated by dividing the absolute CI by the estimate and multiplying by 100%. In our example, this would be 2.0 / 43.2% * 100%, or 4.6%.

What criteria are used to determine if data will be displayed or suppressed on the DRC?

The DRC uses data suppression and presentation criteria set by the Maternal Child Health Bureau (MCHB), which is a modified version of the [National Center for Health Statistics Data Presentation Standards for Proportions](#). A visual diagram of the application of these criteria for purposes of the DRC's data query output and display and suppression criteria is summarized in **Figure 1**.

**Which data are suppressed?** If an estimate's **unweighted denominator is less than 30**, the estimate is considered too unreliable for presentation. In these cases, estimates are suppressed and displayed as a dash (--).
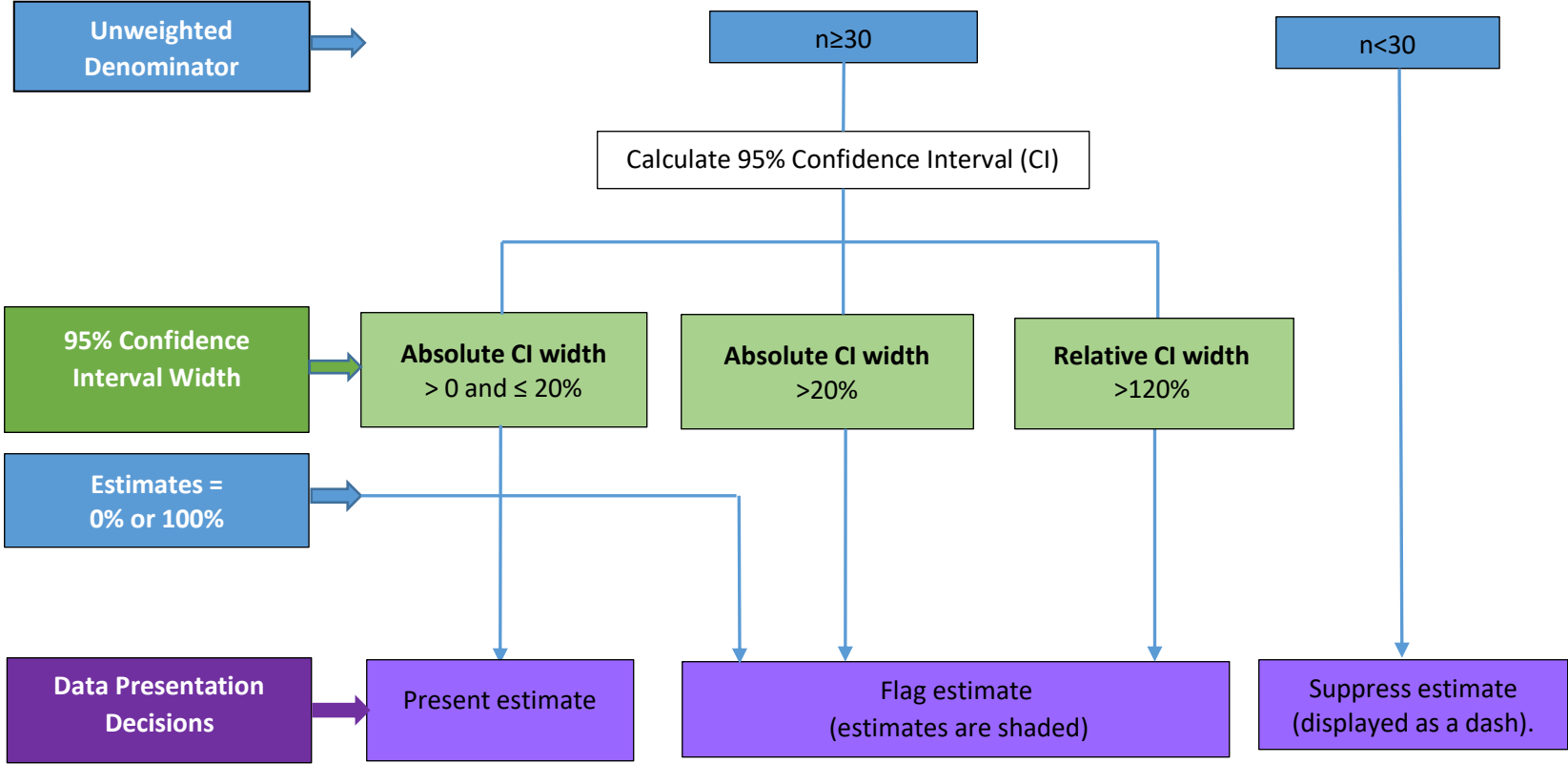
**Which data received an "interpret with caution" flag?** If 1) the **absolute CI width is greater than 20%**, 2) the **relative CI width is greater than 120% (1.2 times the estimate), or** 3) **the CI is inestimable due to no sample variance (estimate is exactly 0% or 100%)**, the estimate is displayed but **may not be reliable**. In these cases, **estimates are shaded** with a note to interpret with caution.

How will future NSCH data collection impact estimate precision?

Larger sample sizes typically improve estimate precision. Since the NSCH is now expected to be implemented annually, combining two or more sequential years together will enable increasingly precise estimates at the state and subgroup levels.

MCHB Data Suppression and Display Criteria

| Unweighted Denominator | → | | n≥30 | | | n<30 |

Calculate 95% Confidence Interval (CI)

| **95% Confidence Interval Width** | → | **Absolute CI width** > 0 and ≤ 20% | **Absolute CI width** >20% | **Relative CI width** >120% |

Estimates = 0% or 100% →

| **Data Presentation Decisions** | ⇒ | Present estimate | Flag estimate (estimates are shaded) | Suppress estimate (displayed as a dash). |