

NCHS Data Presentation Standards for Rates and Counts

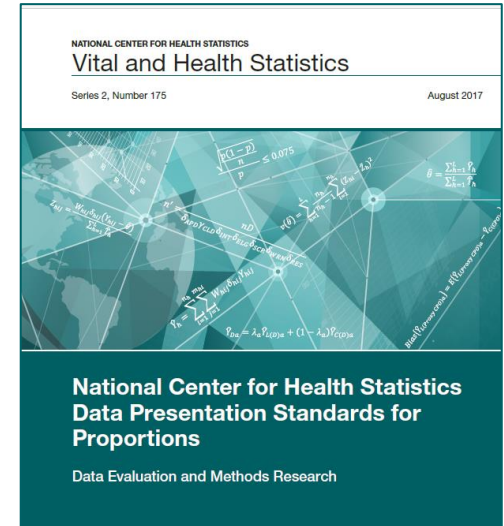
Jennifer D. Parker
Director, Division of Research and Methodology
National Center for Health Statistics

Meeting of the NCHS Board of Scientific Counselors
February 10, 2022



NCHS Data Presentation Standards for Proportions released August 2017

- Standard
 - Criteria based on minimum sample size and absolute and relative width of exact confidence interval
 - Estimated proportions with few degrees of freedom or with suppressed complementary proportions are to be evaluated individually
- Implementation generally, although not always, led to more estimates presented compared to prior criteria



NCHS Data Presentation Standards for Rates (and Counts) workgroup formed 2018

- Workgroup was formed in 2018 and includes representatives from DRM, DAE, DVS, DHCS, OCD and DRM contractors (mathematical statisticians)
- Objective
 - Identify current practices for data suppression of rates and counts
 - Review and examine statistical principles, particularly properties of intervals calculated for rates and counts
 - Create and evaluate presentation standards for rates and counts

Provisional guidance for vital rates and counts presented to BSC January 2020

- Rates with (assumed) constant numerators and non-constant denominators from vital statistics, including:
 - national and state death rates, birth and fertility rates, age-adjusted death rates; rates for subpopulations
- Strategy
 - Identify and compare current presentation guidelines, methods of interval estimation, relative standard errors

Structure of Standards

- National Vital Statistics System
 - Counts
 - Crude rates with denominator assumed constant
 - Age-adjusted death rates with denominator assumed constant
 - Rates with non-constant denominator (e.g., from ACS)
- Population Health Surveys
 - Counts
 - Rates with denominator assumed constant
 - Rates with non-constant denominator

Presentation Criteria

- Minimum sample size or effective sample size
 - Minimum sample size and effective sample size (when applicable) of 10 in the numerator (rates and counts) and in the denominator (rates only)
- Maximum relative width of confidence interval
 - If the sample size criteria are met, calculate a 95% two-sided confidence interval using the appropriate method and obtain its relative width. Estimated rates should have a relative confidence interval width of 160% or lower
- Minimum degrees of freedom
 - When applicable, if the degrees of freedom are fewer than 8 then flag for statistical review

Presentation Criteria – Confidence Intervals for Rates from NVSS (1 of 2)

- Denominators assumed constant
 - Calculate gamma interval where the lower limit is the 0.025 quantile of the standard gamma with $\alpha=x$ and $\beta=1$ and the upper limit is the 0.975 quantile of the standard gamma with $\alpha=x+1$ and $\beta=1$.
 - The adjustment used by [Fay-Feuer](#) should be applied for age-adjusted vital rates.
- Denominators assumed non-constant
 - Calculate a Student's t interval for the logarithm of the rates with variance estimated using method supplied by source. Parameters for age-adjusted intervals can be formed using weighted combinations of age-specific estimates.

Presentation Criteria – Confidence Intervals for Rates from NVSS (2 of 2)

- Denominators from other population surveys and assumed nonconstant
 - Calculate a Student's t interval for the logarithm of the rate. Parameters for intervals for age-adjusted intervals can be formed using weighted combination of age-specific intervals.
- Denominator from births file, which is subject to random variation and assumed non-constant*
 - Calculate a Student's t confidence interval for the logarithm of the rate

* e.g., [period-linked, cohort-linked infant mortality files](#)

Presentation Criteria – Confidence Intervals for Rates from Population Health Surveys

- Denominator assumed constant
 - Calculate a Student's t confidence interval for the logarithm of the rate
- Denominator assumed non-constant
 - Calculate a Student's t confidence interval for the logarithm of the rate
 - The interval accounts for the sampling variability of the denominator

Evaluations for Vital Statistics

- Distributional assumptions: comparison of 'exact' gamma interval width thresholds, relative standard errors (RSE), and counts
- Age-adjusted death rates: simulation-based comparison of Anderson-Rosenberg method, used at [NCHS](#) and in CDC WONDER, with approaches by Fay/Feuer and Tiwari et al, used in U.S. Cancer registries.
- Simulation to examine proposed and previous standards for county-level data by race/Hispanic origin and cause of death
- Simulations evaluating log Student's t interval for non-constant denominator, including correlation assumptions for numerator and denominator

Evaluations for National Health Care Surveys

- Simulations based on NAMCS data structure
 - Evaluation of confidence intervals when the denominator is constant
 - Evaluation of confidence intervals by nominal and effective sample sizes and survey years when the denominator is constant
 - Evaluation of the design effect using row proportions and using totals or counts when the denominator is constant
 - Evaluation of rates and age-adjusted rates when both the numerator and the denominator are non-constant
- Simulations based on sampling small samples from a large combined file of multiple years of NAMCS
 - Evaluation estimates presented using proposed and prior standards, interval coverage, effective sample size, design effect, RSE
 - Focus on detailed age groups, race/ethnicity groups, reason for visit

NCHS Data Presentation Standard for Rates and Counts

- ✓ Review by the workgroup
- ✓ Review by the NCHS ADS from each division
- ✓ Review by additional mathematical statisticians

- Next Steps
 - Consolidate comments and prepare for official NCHS clearance
 - Three presentations to NCHS: Overview (February), Technical Development and Evaluations (April), and Implementation (late spring/summer)

Data Presentation Standards Workgroup

- **Division of Research and Methodology**
 - Jennifer Parker
 - Katherine Irimata
 - Don Malec (retired)
 - Guangyu Zhang
 - Makram Talih (SIS)
 - Frances McCarty (SIS, Virginia Tech)
- **Office of Center Director**
 - Amy Branum
- **Division of Vital Statistics**
 - Brady Hamilton
 - Ken Kochanek
- **Division of Health Care Statistics**
 - Alex Strashny
 - Danielle Davis
- **Division of Analysis and Epidemiology**
 - Barnali Das

Questions?

Jennifer Parker
jdparker@cdc.gov

