Draft Suppression/Presentation Guidelines for Proportions

Jennifer Parker
for the
Data Suppression/Presentation Workgroup

NCHS Board of Scientific Counselors Meeting
January 22, 2015

Background (1 of 2)

- Purpose: to propose updated guidelines for data suppression/presentation for routinely published estimates
  - Intended for publications with numerous estimates from possibly many data sources and little space for standard errors or other measures of precision, e.g. Health, United States, Healthy People 2020
  - Current guidelines/practice differ across data divisions and programs

SAFER • HEALTHIER • PEOPLE™

CDC
Monitoring the Nation's Health

Vital and Health Statistics

HEALTHY PEOPLE 2010

Number 24

Statistical Notes

CDC
SAFER • HEALTHIER • PEOPLE™

July 2002

Healthy People 2010 Criteria for Data Suppression

Richard J. Klein, M.P.H.; Suzanne E. Proctor, M.S.P.H.; Manon A. Boudreault, M.P.H.; and Kathleen M. Turczyn, M.P.H.

Table 1. Summary of criteria for data suppression for the *Healthy People 2010* major data systems

Data system	Number of HP2010 objectives tracked	Criteria for DSU[1]	Population groups always displayed as DSU in HP2010[2]
		Sample surveys	
Behavioral Risk Factor Surveillance System (BRFSS)	9	Denominator < 50	(None)
Continuing Survey of Food Intake by Individuals (CSFII)	6	Means: Sample size < 30•VIF or CV > 30% Estimates 25–75%: Sample size < 30•VIF or CV > 30% Estimates < 25% or > 75%: Smaller of np or n(1 – p) < 8•VIF	American Indian or Alaska Native Asian or Pacific Islander Asian Native Hawaiian or Other Pacific Islander
Medical Expenditure Panel Survey (MEPS)	6	Denominator < 70; or RSE > 30%	American Indian or Alaska Native Asian or Pacific Islander Asian Native Hawaiian or Other Pacific Islander
Monitoring the Future Study (MTF)	5	No formal criteria	American Indian or Alaska Native Asian
National Ambulatory Medical Care Survey (NAMCS)	7	Numerator < 30; or RSE > 30%	American Indian or Alaska Native Asian or Pacific Islander Asian Native Hawaiian or Other Pacific Islander Hispanic or Latino
National Crime Victimization Survey (NCVS)	6	Numerator < 10	(None)
National Health Interview Survey (NHIS)	67	Denominator < 50; or RSE > 30%	(None)
National Health and Nutrition Examination Survey (NHANES)	35	Denominator < 30; or RSE > 30%	American Indian or Alaska Native Asian or Pacific Islander Asian Native Hawaiian or Other Pacific Islander Hispanic or Latino
National Hospital Ambulatory Medical Care Survey (NHAMCS)	6	Numerator < 30; or RSE > 30%	American Indian or Alaska Native Asian or Pacific Islander Asian Native Hawaiian or Other Pacific Islander Hispanic or Latino

Background (2 of 2)

- Workgroup formed in Spring 2013
- Workgroup includes representatives for major data programs
  - OAE: Jennifer Parker, Makram Talih, Dedun Ingram
  - ORM: Don Malec, Vlad Beresovsky, Joe Fred Gonzalez, Iris Shimizu
  - DHIS: Chris Moriarty
  - DHNES: Margaret Carroll
  - DVS: Brady Hamilton, Ken Kochanek
- What follows represents the majority view of the Workgroup, but not a consensus of all Workgroup participants.

Scope (1 of 3)

- The workgroup focused on developing suppression/presentation criteria to be applied to proportions from survey data that will appear in standard data products with multiple tables and stand-alone estimates, such as *Health United States* or *Healthy People 2020*, or in other data products where estimates require readily applied and transparent suppression/presentation standards.

  – No specific recommendations for means, percentiles and rates or recommendations for vital statistics were made

Scope (2 of 3)

- The workgroup expects that data analysts and Division ADSs producing topic-specific publications understand the methodology underlying suppression/presentation criteria for the standard publications and, in combination with subject matter expertise, will choose appropriate suppression/presentation criteria for their specific product.
  - Generally, the workgroup recommends that confidence intervals be presented alongside all types of estimates (proportions, means, percentiles, rates) in these other types of data products whenever possible.

Scope (3 of 3)

- Each of the Center's data systems has unique features and constraints. As a result, the workgroup recognized that Division ADSs may need to apply and recommend additional standards or calculation methods for their data system.

Proposed guidelines (1 of 2)

- For proportions derived from survey data, discontinue use of RSE as the suppression/presentation criterion.
- Effective sample size (nominal sample size divided by design effect) should be greater than or equal to 30. We make no recommendation on numerator size for proportions.
- When calculating **age-adjusted** proportions, the same criteria used for crude estimates should be used

Proposed guidelines (2 of 2)

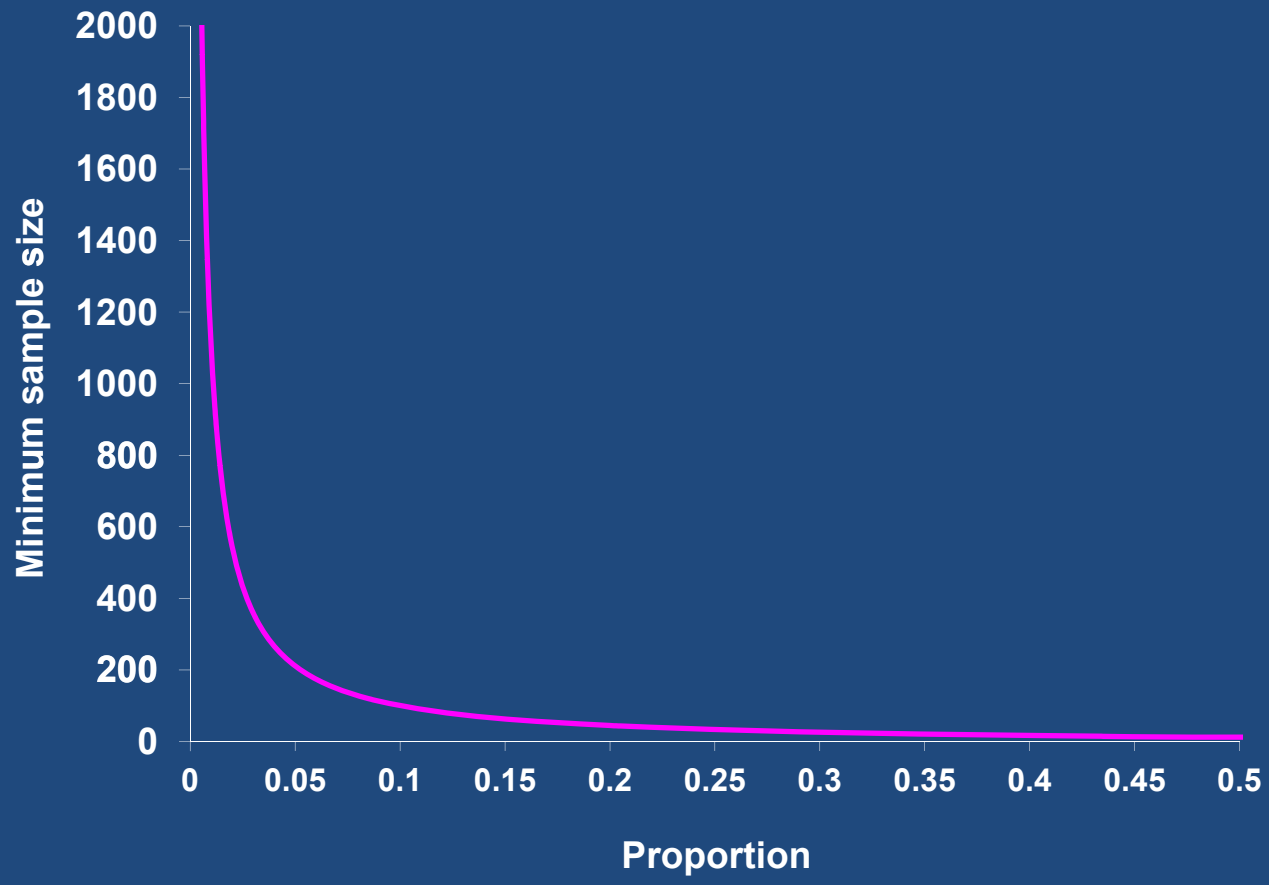
- Clopper Pearson confidence intervals should be estimated using the approach described by Korn and Graubard for complex surveys (see above). The confidence intervals will generally be asymmetric.
  - Calculate the absolute width of the CI as the difference between the upper and lower bound. Calculate the relative width by dividing the absolute width by the estimated proportion and multiplying by 100%.
  - Estimated proportions (percents) with absolute confidence interval widths **less than 0.06 (6%)** should not be suppressed or identified as unreliable.
  - Estimated proportions (percents) with absolute confidence interval widths **greater than 0.20** (20%) should always be suppressed or identified as unreliable.
  - Otherwise, estimated proportions with relative confidence interval widths **greater than 120%** should be suppressed or identified as unreliable.

How did we get there? (1 of 3)

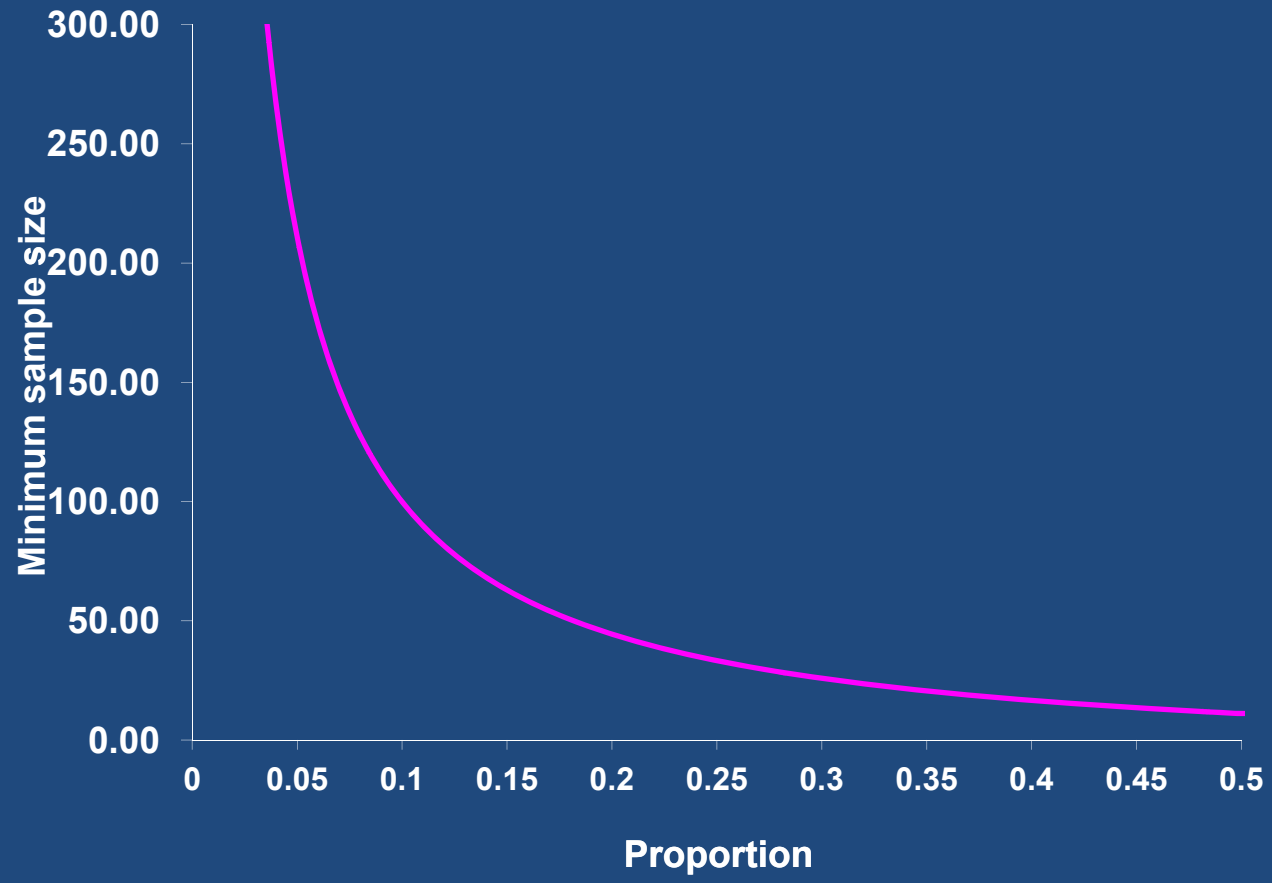
- The relative standard error (RSE*) for proportions performs poorly for very small proportions, large proportions and those in the middle.
  - RSE differs for p and 1-p so analysts need to decide to suppress the larger estimate if the smaller estimate (which can be obtained by subtraction) would be suppressed.

RSE=100%(SE/estimate)

Minimum sample size for relative standard error 30% (no design effects)



Minimum sample size for relative standard error 30% (no design effects)



How did we get there? (2 of 3)

- Good way to assess precision of an estimated proportion is with a confidence interval.

- However, commonly used Wald confidences intervals are known to perform poorly for proportions.
  - Under-coverage
  - Range can be outside 0-1

How did we get there? (3 of 3)

We will skip the formulas and the numerous assessments of different confidence interval methods

Korn-Graubard confidence interval

- Calculation of the interval more fully incorporates information from the survey design, including the design effects and effective sample size

- Based on our simulations, coverage of a 95% interval was close to 95%.
  - Coverage, on average, was conservative meaning that the interval included the true estimate more than 95% of the time.

- The width of a CI (the difference between the upper and lower bound) can be used to set standards for proportions by establishing a threshold using the **absolute width** and/or by establishing a threshold based on the **width relative to the estimate**d proportion.

Comparison to currently used RSE

- Relative confidence interval width of approximately 117% corresponds to RSE < 30%
- The relative confidence interval width has similar shortcomings as the RSE, too conservative for small p and too liberal for large p.
- Guidelines based on both the relative and absolute intervals were developed.

Application to NHANES: High blood pressure among children, NHANES 2009-2012

	Num	Den	%	RSE	Lower CI limit	upper CI limit	Abs CI width	Rel CI width	CI width <6	CI width >20	Rel CI width >120
Boys											
8-17 years	32	1720	1.7	24	1.0	2.7	1.7	100.0	TRUE	FALSE	FALSE
8-12 years	15	898	1.4	35	0.6	2.9	2.3	164.3	TRUE	FALSE	TRUE
13-17 years	17	822	1.9	28	1.0	3.3	2.3	121.0	TRUE	FALSE	TRUE
Girls											
8-17 years	23	1641	1.6	13	1.2	2.1	0.9	56.3	TRUE	FALSE	FALSE
8-12 years	16	910	2.1	29	1.0	3.7	2.7	128.6	TRUE	FALSE	TRUE
13-17 years	7	731	1.2	47	0.3	3.0	2.7	225.0	TRUE	FALSE	TRUE

How did we get there? Minimum sample size

- Based on properties of the Central Limit Theorem, it was decided to set a minimum denominator size of 30, adjusted for the design effect.
  - The NHANES Analytic Guidelines currently recommends one minimum sample size for mid-range proportions and provides other recommended sample sizes for smaller and larger proportions.
  - Although some supported a recommendation based on the size of the proportion, most workgroup members supported the single guideline.

Concerns

- Because the evaluations showed that relative CI criteria performs similarly to a RSE criterion for many situations, some thought that the extra effort required to apply the CI criteria is unnecessary and too complicated for some users to implement and understand.
  - The ease of use is an important advantage to a RSE based criteria.
- Length of a CI, either in absolute or relative terms, is less commonly used than the RSE, which may be confusing or off-putting to some users
- The concept of "relative confidence interval", as defined herein, may not be a standard statistical concept.
- The absolute criterion, designed to facilitate the presentation of small proportions, may be too liberal.

Degrees of Freedom

- No recommendations for a required number of DF were made. However, users are urged to assess estimates based on fewer than 8 DF.

- Because the variance of the SE estimate is related to the DF, estimated standard errors for estimates based on a small number of DF may be unreliable
  - This consideration is greater for subgroup estimates from NHANES and for some NHIS state estimates than for other national estimates.
  - Specifically, the RSE of the SE can be approximated as 100*sqrt(2/DF).
  - Estimated proportions with fewer than 8 DF have standard errors with RSE of 50% or more.
  - Although the confidence interval approach described above incorporates the DF, there will be instances where estimates based on very low DF meet the suppression/presentation standard and are reasonable and other instances where they are not.

Steps to implementation

- 2016?
  - 2015 JSM panel on data suppression
  - Seminars to staff
  - Computer code for SUDAAN, SAS, Stata
  - Online documentation (short)
  - Series report or other expanded report
  - Fuller write-ups of some of the simulations and evaluations

Clopper-Pearson confidence intervals adapted by Korn and Graubard for complex surveys

$$p_L(x,n) = \frac{v_1 F_{v_1 v_2}(\alpha/2)}{v_2 + v_1 F_{v_1 v_2}(\alpha/2)}$$

$$p_u(x,n) = \frac{v_3 F_{v_3 v_4}(1-\alpha/2)}{v_4 + v_1 F_{v_3 v_4}(1-\alpha/2)}$$

Where

$v_1 = 2x, v_2 = 2(n-x+1), v_3 = 2(x+1), v_4 = 2(n-x)$, and $F_{d_1 d_2}(\beta)$ is the β quantile of the F distribution with d_1 and d_2 degrees of freedom.

The degrees of freedom adjusted effective sample size used in the above is calculated as:

$$n^*_{df} = \frac{\hat{p}(1-\hat{p})}{\widehat{var}(\hat{p})}\left(\frac{t_{n-1}(1-\alpha/2)}{t_d(1-\alpha/2)}\right)^2$$

where *d* is often taken to be the number of sampled clusters (psu) minus the number of strata