

DETERMINING SAMPLE SIZE

Calculating a statistically valid sample size for a chart review follows steps adapted from statistical techniques used for descriptive studies. The process uses a nomogram, or table, to identify the desired number:

1. Estimate the expected proportion within the population that will have the measure of interest.

If you have a benchmark from literature or prior studies, use it. Otherwise, consult with colleagues or experts in the field to determine an estimate. The tables generally require this proportion to be 50 percent or less. If more than 50 percent of the population is expected to have the characteristic, then base your sample size calculation on the proportion *without* the characteristic.

2. Specify the width of the confidence interval you wish to use.

All empirical estimates based on a sample have a certain degree of uncertainty associated with them. It is necessary, therefore, to specify the desired width of the confidence interval (W). This gives a range of values that you can be confident contains the true value. In most cases, an appropriate width is 0.20 (that is, plus or minus 10 percent).

3. Set the confidence level.

This is a measure of the precision or level of uncertainty. Typically 95 percent is used, meaning that we are 95 percent certain that the interval includes the true value. This is arbitrary, however, and other levels of confidence can be used. The table shown below is for a 95-percent confidence level. The narrower the width of the confidence interval and the higher the confidence level, the larger the sample size.

4. Use the nomogram (below) to estimate sample size.

Sample size for a descriptive study of a dichotomous variable 95-percent confidence interval					
Width of the confidence interval (W)	0.10	0.15	0.20	0.25	0.30
Expected proportion (P)					
0.10	138	61			
0.15	196	87	49	31	
0.20	246	109	61	39	27
0.25	288	128	72	46	32
0.30	323	143	81	52	36
0.40	369	164	92	59	41
0.50	384	171	96	61	43

Adapted with permission from Hulley SB, et al. *Designing Clinical Research*, 3rd ed. Philadelphia: Wolters Kluwer Health; 2006:91.

An example

According to *HEDIS 2007 Audit Means, Percentiles and Ratios*, the NCQA's annual report of health plan performance data, 68.9 percent of women age 40 to 69 had a mammogram during 2006. This makes the expected proportion of those without screening 31.1 percent. We choose a width of the confidence interval of 0.20 (plus or minus 10 percent) and a confidence level of 95 percent. This means that we want to be 95 percent confident that the result falls between 58.9 percent and 78.9 percent. Using the nomogram to determine the sample size, we read down the left column of figures for the expected proportion without the characteristic (0.30 is the closest value to 31.1 percent) and then across to the chosen width of the confidence interval (0.20). When we follow the column down, we find the required sample size (81). If the number required is too large to be completed, we can recalculate with a lower confidence level or wider interval; this will produce a smaller sample size.



FPM Toolbox To find more practice resources, visit <https://www.aafp.org/fpm/toolbox>.

Developed by Barbara H. Gregory, MPH, MA, Cheryl Van Horn, RN, and Victoria S. Kaprielian, MD. Copyright © 2004 American Academy of Family Physicians. Physicians may duplicate or adapt for use in their own practices; all other rights reserved. Related article: <https://www.aafp.org/fpm/2008/0700/pa3.html>.