MSQC 4.0 Case Sampling Methodology

Stacey Collins, MA, Ed Rothman, PhD
Michigan Surgical Quality Collaborative, University of Michigan, Ann Arbor, MI

Last year, MSQC 3.0 brought changes to the sampling methodology with the implementation of an equal allocation probability sample[1] in order to enhance statistical validity and support direct generalization to the overall hospital population. While the probability sampling methodology was implemented successfully, MSQC understands that not all procedure groups are equally as risky, or clinically as interesting, and thus should not necessarily be given equal consideration for clinical abstraction. Therefore, MSQC 4.0[2] brings two sampling methodology changes to address these concerns. First, hysterectomy and vascular procedures are now optional and require hospitals to declare participation on an individual basis. And across the board, MSQC 4.0 is moving to a Neyman's allocation sampling algorithm for selection of the required core procedure cases.

Neyman's allocation is a stratified sampling method in which the number of cases selected from each procedure group (stratum) is proportional to both the size of that procedure group and the previously calculated variance of the morbidity for that procedure group. Thus, instead of sampling the same number of cases from each core procedure, as was the case under equal allocation, more cases will be sampled from the more populated groups and the groups with more variability under Neyman's allocation.

Just like in MSQC 3.0, the MSQC 4.0 sampling methodology begins with a complete sampling frame of all MSQC eligible cases, based on the CPT Code inclusion criteria. As hysterectomy and vascular cases are now optional, they should only be included on the sampling frame if the individual hospital has opted in to collecting such cases. Once the sampling frame is uploaded to the workstation, Neyman's allocation is run on core procedure groups to determine the number of cases to include in the sample from each procedure group. (Please see appendix A for full mathematical formula of Neyman's allocation). Next, a rounding algorithm is used to ensure the number of cases to be sampled from each procedure group is a whole number, and that every procedure group represented on the sampling frame has at least 1 case selected for the sample. To complete the sample selection, cases from within each of the core procedure group are then selected randomly based on the number determined by the Neyman's allocation algorithm. Following that, optional cases from hysterectomy or vascular procedure groups are randomly selected based on each individual hospital's declared participation.

The number of cases selected from each procedure group will be different from cycle to cycle because the number of each procedure on the sampling frame will differ from cycle to cycle. For hospitals with low core volume that are also participating in one (or both) of the optional procedures, the MSQC 4.0 sampling algorithm includes a backfill component where additional optional cases will be

[1] Please see MSQC 3.0 Case Sampling Methodology for more information
[2] Beginning with operations performed January 1, 2017

sampled to account for any deficiency in the number of core cases available in a particular cycle, up to the workflow constraint of 25 cases per cycle[3].

In summary, changing the sampling methodology from equal allocation to Neyman's allocation will not affect any of the sampling gains that came with the implementation of a probability sample in MSQC 3.0. Instead, the new allocation algorithm is simply a move from one statistically sound method of stratified sampling to another, which allows MSQC to more directly target prevalent and variable procedures for clinical abstraction.

Appendix A

The formula for Neyman's allocation is as follows:

$$n_k = n_{core}\left(\frac{N_k S_k}{\sum_{i=1}^{K} N_i S_i}\right)$$

Where:

- n_k = number of cases to be selected from procedure group k
- n_{core} = total number of core cases to be selected for the sample
- N_k = total population of cases in procedure group k (on sampling frame)
- S_k = Square root of the variance (i.e., the Standard Deviation) for procedure group k and hospital size group
- K = total number of core procedure groups on sampling frame

References

Cochran, W. G. (1977) *Sampling Techniques* (3rd ed.). New York: John Wiley & Sons.

Kalton, G., (1983). *Introduction to survey sampling* (Sage University paper series on Quantitative Applications in the Social Sciences, No. 07-035). Newbury Park, CA: Sage.

Kish, L. (1965). *Survey Sampling.* New York: John Wiley & Sons.

Lohr, S. L. (2009). *Design and Analysis (*2nd ed.). Boston: Brooks/Cole.

[3] High volume hospitals do not have the backfill option, as they are required to abstract 25 core cases per cycle, and participation in hysterectomy or vascular procedures requires additional abstraction of 10 cases per cycle, up to a maximum of 35 cases per cycle.

Ott, R. L., & Longnecker, M. (2001). Using Surveys and Scientific Studies to Gather Data. In *An introduction to statistical methods and data analysis* (5th ed.) (pp. 18 – 38). Pacific Grove, CA: Wadsworth Group.