Making Sense of Biostatistics: Types of Nonprobability Sampling

By Kathleen Mathieson

Probability sampling methods, such as simple random sampling, as discussed in the last column, are the best way to reduce sampling bias and achieve a representative sample of patients for a clinical study. These methods, however, require a comprehensive list of eligible patients, which is rarely available in practice. Therefore, clinical studies must often use methods of nonprobability sampling, in which subjects are chosen on a basis other than random selection. This column describes several common types of nonprobability sampling methods and factors that can reduce bias and increase generalizability. Regardless of the sampling method used, assignment to treatment arms should still be random.

Convenience sampling entails selecting patients on the basis of their availability. For example, patients may be selected from an easily accessible geographic region or specific clinical site. Patients with certain characteristics are more likely to be chosen via convenience sampling. For example, in community-based research, healthier patients are often more accessible than those with more serious illness. This factor and others can lead to over- or under-representation of certain population attributes that are being studied, and therefore decrease the generalizability of study results.

The type of convenience sampling used most commonly in clinical studies is consecutive sampling. In consecutive sampling, each consecutive eligible patient who presents for care is approached for enrollment. Consecutive sampling provides some structure and thus additional rigor in that it includes all patients who are accessible within the defined study time period. The resulting sample is thus more likely to represent the target population than one resulting from simple convenience sampling. When consecutive sampling is used, researchers must ensure that the study period spans a long enough time period to avoid potential bias. For example, in studies where seasonal variations may be important (e.g., allergy, depression, or arthritis trials), it is important for the study period to be long enough to achieve a broad, generalizable sample of patients.

Another way to enhance the rigor of convenience sampling is to use quota sampling to obtain a certain number of participants with specific characteristics. For example, many more women than men are diagnosed with lupus. If researchers wish to test the efficacy of a treatment in both men and women, they may need to set quotas by sex to ensure appropriate statistical power. Using consecutive sampling, researchers may enroll patients of both sexes until the quota for women is filled, and then continue recruiting men until their quota is filled.

Purposive sampling is a precise form of convenience sampling in which patients are hand-picked based on certain criteria when a researcher is interested in a very specific group of patients for reasons of feasibility or efficiency. Thus, studies with very specific eligibility criteria are engaging in purposive sampling; it's a matter of degree.

For studies of rare conditions or for difficult-to-reach populations, snowball sampling may be an efficient and cost-effective method. This method involves identifying an initial individual or set of individuals who meet study criteria, then using those individuals to find and contact others for the study. In the case of a rare illness, a researcher might identify the initial “seed” of patients via a support group and then use those patients to reach other eligible patients.
Convenience methods that can be expected to bias the sample should be avoided. For example, recruiting patients that live near the research site might not introduce substantial bias, but selecting patients based on their Medicare/Medicaid status might.

Like all study design decisions, choosing a sampling technique for clinical research involves trade-offs. When choosing a technique, researchers must balance the objective of generalizability with real-world factors, such as cost, efficiency, feasibility and practicality. Although probability sampling is the gold standard for ensuring that the study sample accurately represents the target population, in practice it is usually more of a yardstick against which more feasible methods can be measured. Nonprobability sampling methods, such as those described in this column, are frequently used in clinical research, and while not as rigorous as probability methods, can provide efficient and cost-effective ways to recruit patients.

References

Author
Kathleen Mathieson, PhD, CIP, is an Associate Professor in the Doctor of Health Sciences Program at A.T. Still University (ATSU) in Mesa, Arizona. She is also a Vice-Chair of the ATSU-Arizona Institutional Review Board and a Still Research Institute Scientist. Contact her at 1.602.573.6547 or kmathieson@atsu.edu.