Diagnosing and Controlling Costly Sample Bias: A Gold-Standard Case History

Jack R. Gallagher, Ed.D., M.S., Clarity Pharma Research; Brent Schwartz, Allergan; Kylee Jean Heap, Clarity Pharma Research

One of the greatest unrecognized and undiagnosed problems in pharmaceutical survey research is faulty decision making that is caused by nonrepresentative, invalid survey samples.[1,2]

We believe correct survey-derived inferences require a probability sample, randomized and representative recruitment of target physicians, and high survey participation among recruited physicians.

Concern over growing use of nonprobability (biased) samples in survey research prompted the European Survey Research Association (ESRA) to devote a portion of its 2009 Warsaw conference to seeking solutions. In conference notes, ESRA observed that nonrandom survey samples are becoming more common in part because of the increasing use of panel and other convenience samples, and that as a consequence, "the validity of the results could be seriously corrupted and selection bias will become apparent."[3]

A probability sample requires each element to have a known nonzero probability of selection. This permits the use of statistics to make inferences about a population.[4] A probability sample and high participation rate reduce sample coverage bias and nonresponse bias.

High standards are required

Nothing else matters if the sample used for your research project does not have a high degree of representative precision, unless the objective is to obtain only directional insights. Even the most sophisticated analytics can lead you astray if based on a sample with hidden or unknown biases.[5]

Here we present 3 types of standards:

- on sample coverage,
- physician recruitment,
- and survey participation rates.

We believe these standards must be met to avoid or minimize nonrepresentative survey samples, and we provide a case history of a patient-records/treating physician study that followed these standards.

Case history met standards

Background

The study assessed treatment patterns for nonmuscle invasive bladder cancer (NMIBC) by a random sample of 259 U.S. urologists (stratified by region) who extracted information from the medical records of 1,010 patients. Urologists were allowed to provide up to their last 4 qualified cases, thus randomizing patients by physician. A major objective was to determine the extent to which patient care urologists used immediate postoperative instillation of intravesical chemotherapy (IPOIC) in patients with NMIBC. The target drug was in Phase 3 clinical trials.

Weighting of data

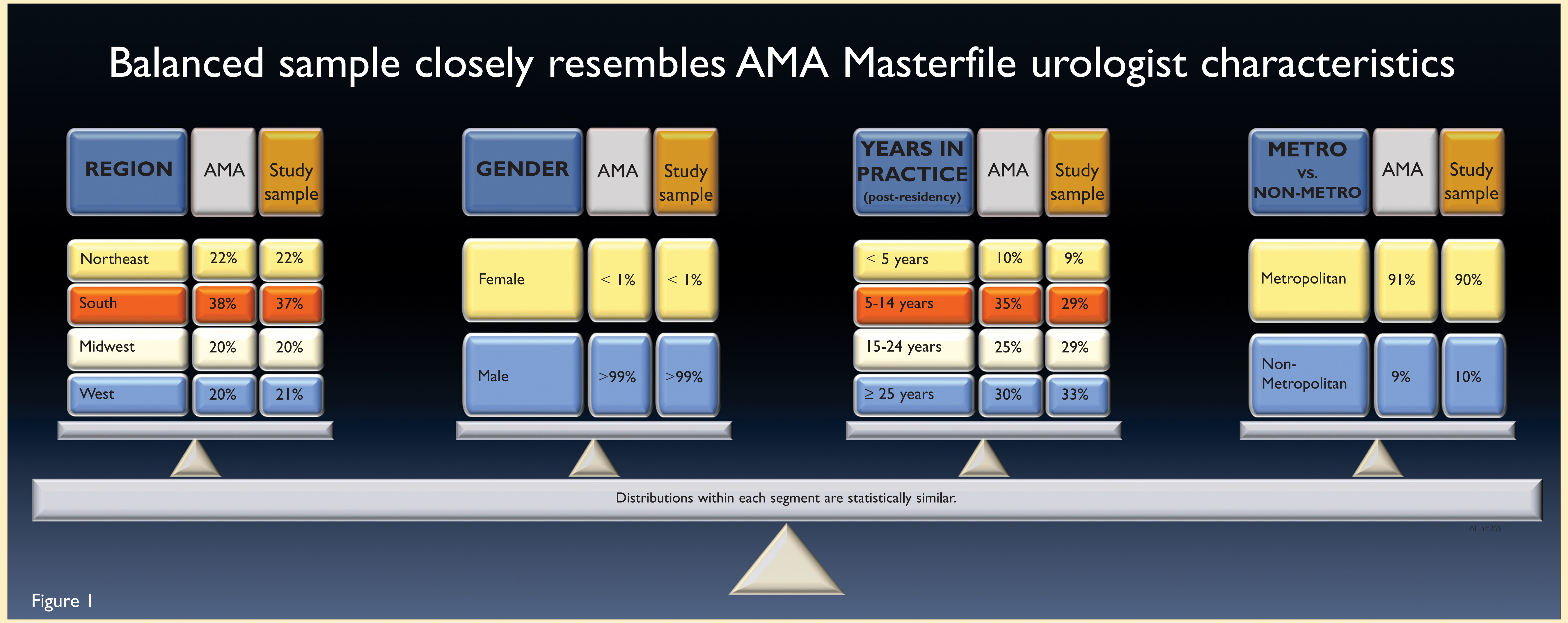
Weighting can be used to adjust for differences on key characteristics between a sample and sample frame. It also can be used to adjust for differences in patients' probability of being selected. In this study, statistical analyses were adjusted to account for individual differences among urologists in the volume of NMIBC patients they treated. Patients of urologists who treated few NMIBC cases had a greater chance of being among the last 4 treated (and selected) than those whose physicians treated more target patients. Thus, we applied an appropriate weight to ensure each patient's probability of being selected was accurately reflected in study results, i.e., each study patient represented the same number of patients with NMIBC across the U.S.

We also tried to minimize exaggeration that can occur at the margin (in self-reported patient volume estimates by "heavy treaters"), where even a small number of outliers can cause major distortion in results. To avoid such distortion, we converted data into deciles and assigned patients the median volume of their physicians' deciles.

Results

- ▶ All study objectives were achieved, including:
  - a greater understanding of NMIBC practice patterns:
    - IPOIC was found to be used infrequently, contrary to American Urological Association guidelines;
  - metrics to improve forecasting accuracy;
  - identifying strategic opportunity areas.
- ▶ Study insights helped in the redesign of clinical trials for the target drug.
- ▶ Findings merited acceptance and publication at or in 5 peer-review conferences or journals.

Following are our key sample validity standards[6] and evidence the standards were met:



Balanced sample closely resembles AMA Masterfile urologist characteristics

REGION	AMA	Study sample
Northeast	22%	22%
South	38%	37%
Midwest	20%	20%
West	20%	21%

GENDER	AMA	Study sample
Female	< 1%	< 1%
Male	>99%	>99%

YEARS IN PRACTICE (post-residency)	AMA	Study sample
< 5 years	10%	9%
5-14 years	35%	29%
15-24 years	25%	29%
≥ 25 years	30%	33%

METRO vs. NON-METRO	AMA	Study sample
Metropolitan	91%	90%
Non-Metropolitan	9%	10%

Distributions within each segment are statistically similar.

All n=259

Figure 1

Survey reports should contain written evidence regarding all these criteria.

1. Sample coverage standards

A. To correctly answer the strategic questions posed for a study, the correct population must be studied.

- ▶ The correct population, and the one studied in this project, is urologists who treat NMIBC. Once a patient is diagnosed with NMIBC, the patient is likely treated by a urologist in the U.S.[7] Nineteen of 20 NMIBC patients (95%) are treated by urologists in the U.S.[8]

B. To adequately study the target population, the master list from which the sample is selected (sample frame) must be complete, with accurate and up-to-date contact information and physician population metrics.

- ▶ The master list from which the case study sample was selected is the AMA Masterfile of target urologists nationally. Further:
  - The supplier had updated the Masterfile within the previous 3 months.
  - If a selected urologist had moved from the listed address, a researcher tracked the physician and updated the database.
  - This AMA Masterfile contained accurate metrics for patient care urologists.

2. Physician recruitment standards

A. Distributions of important identifiable characteristics of the sample are similar to corresponding ones in the total population (sample frame).

Without knowledge of the distribution of a target segment in the universe, a sample is likely to contain too few or too many participants in some target segments. A representative survey sample mirrors the target physician population on important practice, demographic, and attitudinal or other characteristics.

- ▶ The AMA Masterfile provides the nationwide distribution of patient care urologists by post-residency years of practice, gender, region, and whether a practice location is in a metropolitan area.
- ▶ As shown in Figure 1, the study sample closely paralleled the corresponding nationwide distribution for each of these characteristics.

B. Satisfactory efforts are made to maximize physician study participation and include specific procedures to obtain appropriate participation levels of:

1. marketing research-resisting physicians,
2. heavily burdened physicians, and
3. other hard-to-recruit physician segments.

A physician's decision about whether to participate in a study is determined by his or her assessment of both financial and psychic costs and benefits of such participation.

In our experience, physicians in general, and those who resist marketing research studies in particular, place high value on professional affiliations and on scientific work. For this study, prospective physicians learned:

- ▶ The study investigators are nationally known medical research scientists who contribute to peer-review medical journals.
- ▶ Study findings of interest would be shared through publications in medical journals.

Figure 2

To attract the hard-to-recruit physician, particularly the overly burdened or busy physician, we also employed the following strategies:

- multi-channels (Internet, phone, fax, mail);
- extensive efforts by researchers to accommodate participation obstacles physicians face;
- incentive ladder.

C. A substantial proportion of the sample is composed of hard-to-recruit physicians.

1. Report the percentage of physician participants who did not take part in market research in the previous 3 months.

To obtain a quantitative measure of whether the study was primarily composed of "easy to reach, glad to participate" physicians, we asked a random sample of participants if they took part in another market research study during the previous 3 months. Figure 2 shows hard-to-recruit physicians, who tend to be underrepresented in surveys, made up a substantial proportion of respondents. More than half the physicians asked (52%) if they had taken part in a study other than ours in the past 3 months said they had not.

3. Physician participation standards

A. High physician participation rate.

1. Report the physician participation rate (total physician participants / total qualified physicians contacted).

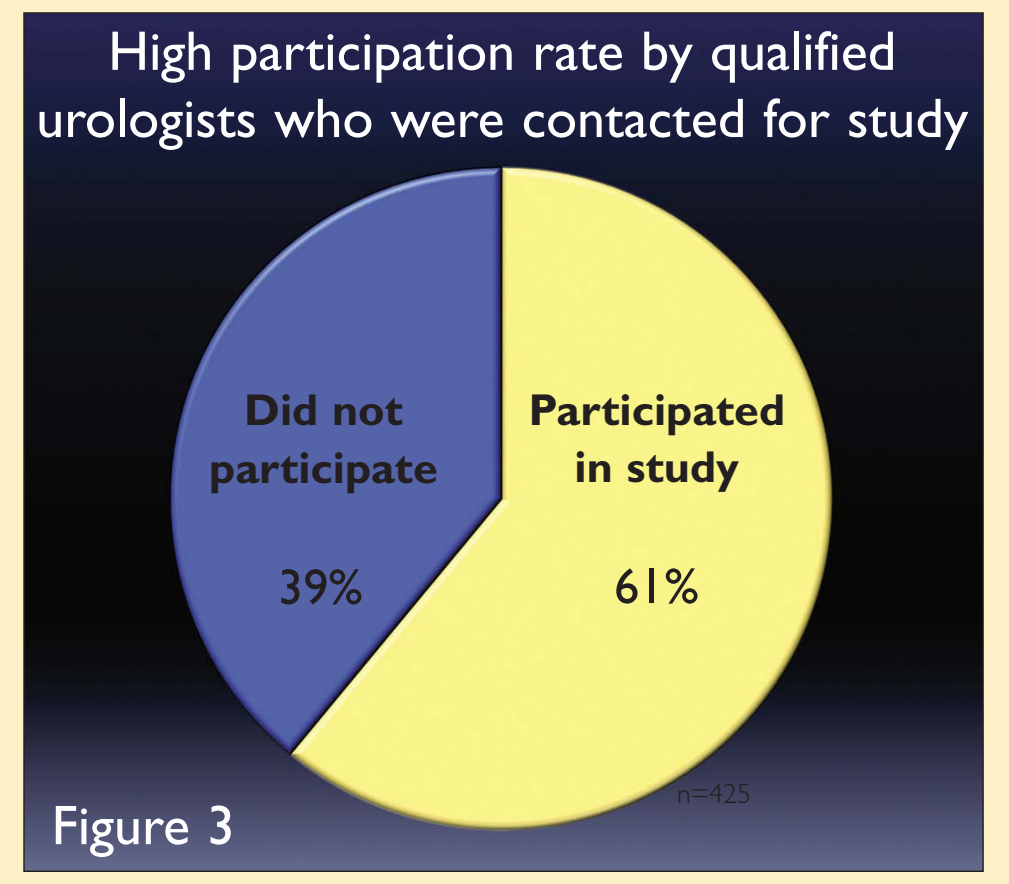


Figure 3

The "internal validity" of studies "can be threatened by self-selection bias resulting from differences between those who participate in a study and those who do not. .... [S]tudies with low levels of participation may be more vulnerable to self-selection bias than those with high participation," write Morton, Cahill, and Hartge.[9]

Draugalis and Plaza note, "Nonresponse error, or bias, occurs if data are not collected from each member of the sample. ... A response rate of 50%-60% or greater is optimal because nonresponse bias is thought to be minimal with that high of a response rate."[10]

We believe in transparency – reporting the rate and how it is calculated.

- ▶ Of 425 urologists invited to participate in this study, 259 *(61% response rate)* provided information on 1,010 NMIBC cases (see Figure 3).

B. Minimize item nonresponse.

To minimize the extent of missing information (item nonresponse), researchers had at least one follow-up conversation with each of the 259 physician participants to discuss and/or clarify information provided on the questionnaires.

- ▶ These extensive communications increased the likelihood that the intended and correct information was actually obtained.

C. Likely effects of nonresponse bias on study findings are appropriately reflected in recommendations.

Some bias is likely, as 39% of qualified physicians contacted did not participate in this study. We believe the likely effects of nonresponse bias on study recommendations were minimized through the extensive efforts described here.

This belief is further bolstered by an independent review of the study. Allergan retained an epidemiology consultancy to provide the independent assessment of the study's representativeness. The assessment found it to be so after looking at study physicians' demographics when compared to the AMA Masterfile, at the physicians' response rate, as well as at the patient sample size and that sample's representativeness of NMIBC patients in the U.S.

The consultancy review found the age distribution of patients in the study was "similar" to NMIBC incident + recurrent patients in the population reported by the Surveillance Epidemiology and End Results Program of the National Cancer Institute. The assessment also noted the study's proportion of incident vs. recurrent patients was "almost identical" to corresponding proportions calculated using population NMIBC epidemiology data from the National Oncology DataBase™, IMPAC Medical Systems. The independent review said the study was "well-designed to accomplish [its] primary objective" of understanding the NMIBC instillation market.[11]

Conclusion

Although pure representative sampling that precisely mirrors the target universe might be unattainable given time and money constraints, researchers should be committed to making a sample as representative as possible within the limits of the survey environment.▲

References

[1] Gallagher, JR, "Invalid patient studies: Not a bargain at any price," Journal of Health Care Marketing, March 1989; 9(1): 69-71.
[2] Gallagher, JR; Santos, J, "Preparing for a Successful Product Launch: Minimize Survey Sample Bias, the Silent Killer," poster, PBIRG, Chicago, February 2012.
[3] "New challenges in sampling," 3rd ESRA conference, Warsaw, July 2009, http://surveymethodology.eu/conferences/warsaw-2009/sessions/.
[4] Kish, L, "Survey Sampling," New York: John Wiley & Sons, Inc.; 1965.
[5] Lehman, EC, "Tests of significance and partial returns to mailed questionnaires," Rural Sociology, September 1963; 28: 284-9.
[6] Gallagher, JR, "Essential sample validity standards for physician surveys," white paper, Sept. 1, 2009.
[7] Kirkali, Z; et al., "Bladder cancer epidemiology, staging, grading and diagnosis," Urology, 2005; 66: 4.
[8] Internal Allergan report, 2009.
[9] Morton, LM; Cahill, J; Hartge, P, "Reporting Participation in Epidemiologic Studies: A Survey of Practice," American Journal of Epidemiology, 2006;163(3): 197-203.
[10] Draugalis, JR; Plaza, CM, "Best Practices for Survey Research Reports Revisited: Implications of Target Population, Probability Sampling, and Response Rate," American Journal of Pharmaceutical Education, 2009; 73(8): 142.
[11] Internal Allergan report, March 14, 2010.

Poster presented at the Sixth Annual PMRG Institute, Philadelphia, PA, Oct. 14-16, 2012; poster editing, design, graphics: Susan Carroll/Clarity Pharma Research.