

Representing the inconvenient: Greatest data collection hurdle

Presentation By:

Dr. Jack R. Gallagher

jgallagher@claritypharma.com



Key fieldwork challenge requires internal and external solutions

This brief presentation will focus on:

- The pervasive use of non-probability samples to collect pharmaceutical survey data.
- Why this is a vital problem needing transparency and attention.
- The need for increased cooperative action to encourage development and sharing of comprehensive, accurate master lists of physicians and other types of target survey participants (sampling frames) in EphMRA members' nations.

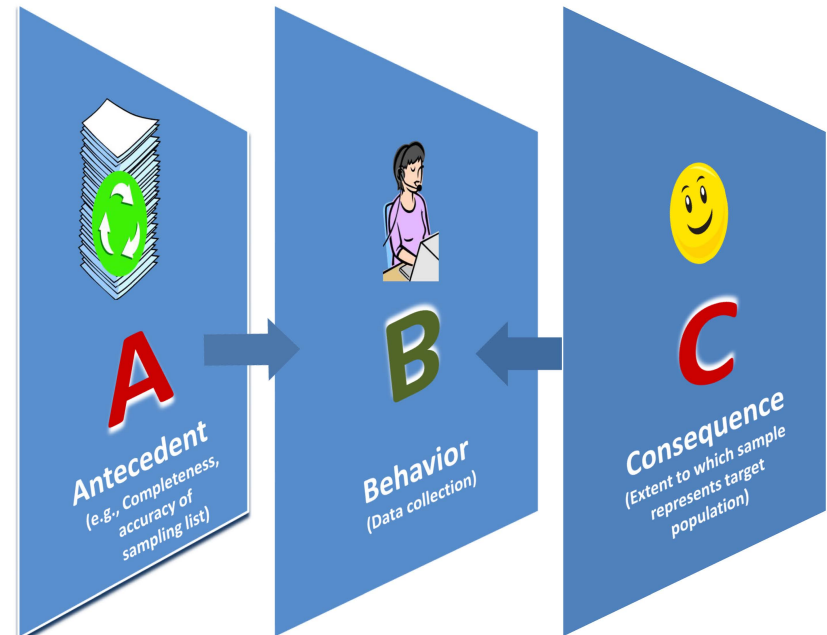


ABCs of influencing quality of data collection

As indicated in *Changing Behavior: How and Why*, a book by Dr. Jack R. Gallagher:

Behavior, including that involved in survey data collection, can be influenced at two points: “**A**” and “**C**.”

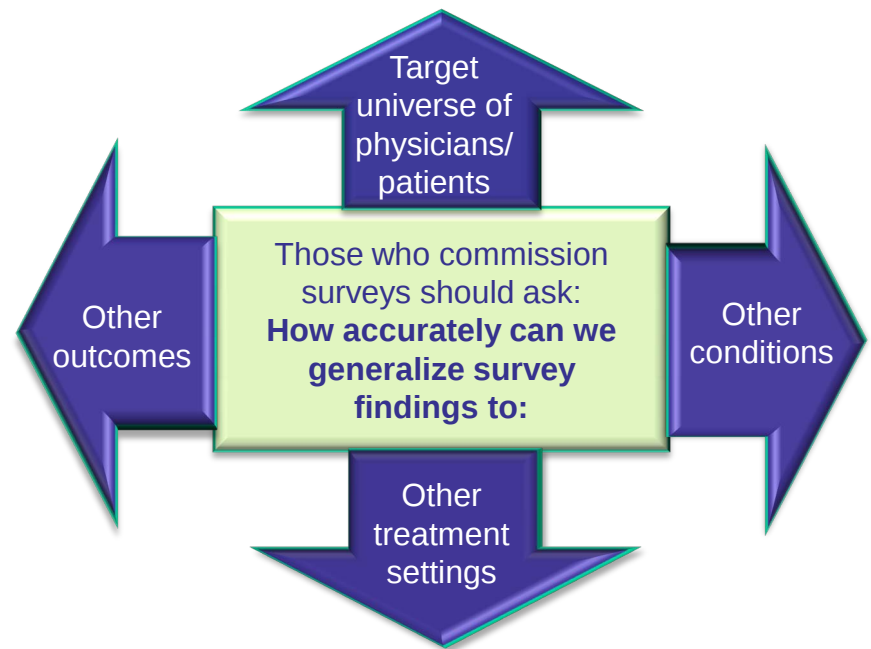
As an industry, we need to work together to better meet a **great challenge** at point “**A**,” e.g., in completeness and accuracy of the sampling list.



Clients expect to accurately generalize survey findings to a target universe

Unless a survey's objective is to obtain only **directional** findings, survey commissioners implicitly, if not explicitly, expect:

- the findings to yield accurate inferences beyond a survey's participants to a target population, including:
 - those who are “inconvenient” to select and/or recruit,
- in order to make correct strategic decisions.



Representativeness depends more on master list than sample size

A large unrepresentative sample can be **as misinforming** as a small unrepresentative sample.

A small representative sample can be **more informing** than a large unrepresentative sample.

Incomplete, inaccurate, outdated sampling master lists represent a vital obstacle to data collection in most countries!



Large Sample but Not Representative

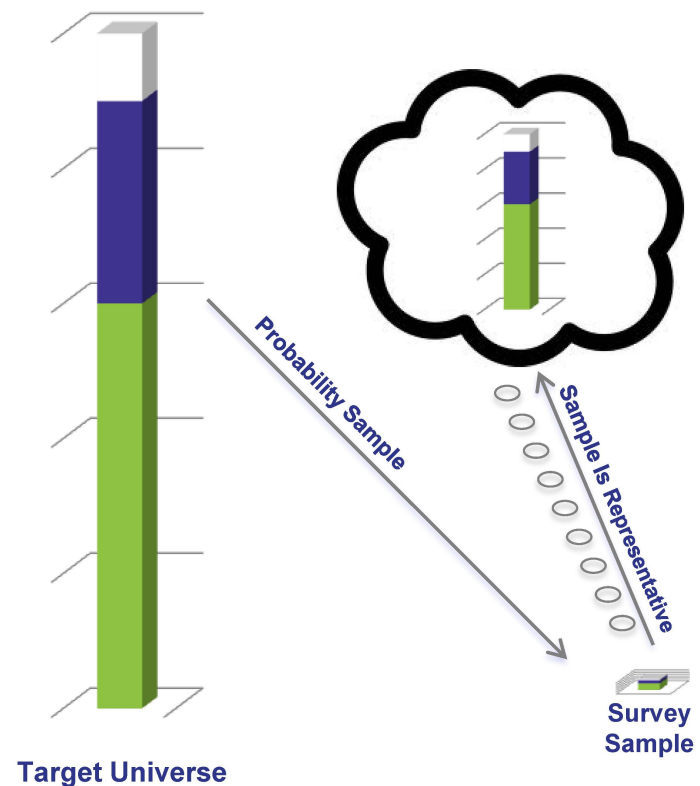


Small Sample and Representative

A probability sample is required for accurate representation

A **probability sample** is required for a sample to accurately represent the target universe.

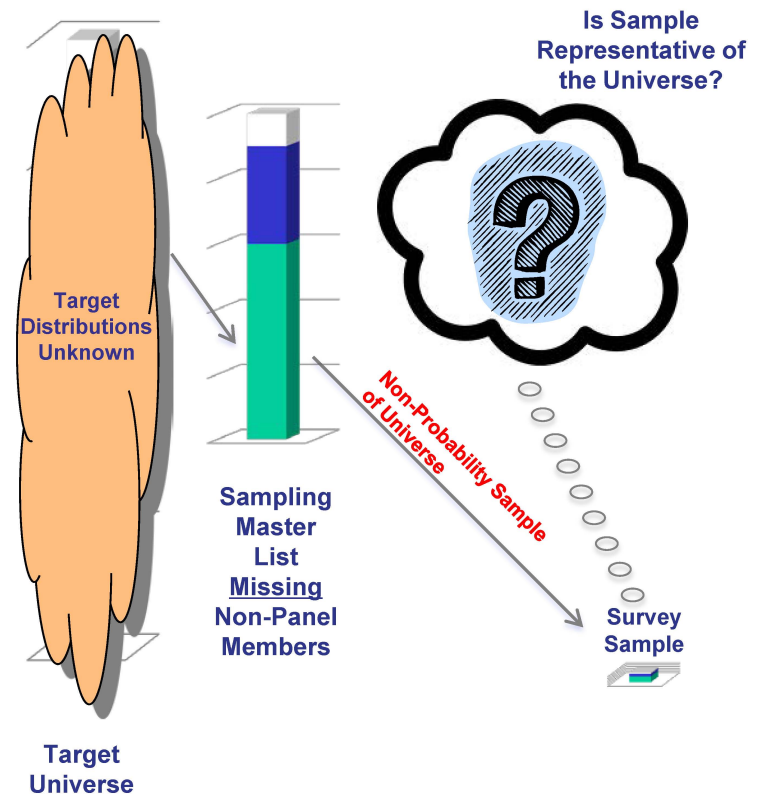
- Each physician/element of the entire target universe must have a **known** probability of selection.
- Otherwise, the use of **statistics** to determine the significance of findings **is not** permitted/appropriate.



A sample cannot accurately represent an unknown universe

The two greatest challenges to data collection in most of the world are illustrated here when:

- First, **no master sampling list** (complete, up-to-date view of the target universe) is available.
- Second, data usually are collected from a **non-probability**, convenience (panel) sample:
 - Not all members of the target universe have a known probability of being selected for the survey.



Prevalent non-representative sample bias noted at conference

The European Survey Research Association (ESRA) devoted part of its Warsaw conference to seeking solutions to the lack of a high-quality sampling frame and the bias inherent in the prevalently used opt-in convenience or panel survey sample. Session coordinator Lensvelt-Mulders noted:

*“It is not yet known how to deal with selection bias in non-random samples.”**

Convenience samples, e.g., panels made up of only some segments of a target population, are non-probability samples.

**“Selection bias in Internet panels: challenge or dead blow?” – G. Lensvelt-Mulders, ESRA conference, Warsaw, 2009, <http://www.europeansurveyresearch.org/sites/default/files/abstracts.pdf>

The predominantly used physician panel – a convenience sample that:



Panel members need to be represented – but not over-represented!

Making it more convenient to include the inconvenient

Meeting the need for comprehensive and accurate sampling frames is a very difficult task in all countries and may require the cooperation and assistance of national governments.

It is hoped that members, associate members, and EphMRA staff will continue to provide leadership and cooperative efforts to better meet this challenging obstacle to data collection.

