Sampling considerations for disability related surveys in the Pacific

There are two main types of surveys that are undertaken to study disability and each call for a different sampling strategy. To help to decide which type of survey should be implemented, it's important to consider the following questions: i) are you interested in measuring the prevalence of disability in the population; or ii) are you interested in examining the correlation between disability and various outcomes. The former requires significantly smaller samples than the latter, however it is noted that few data collections are only conducted to estimate prevalence and prevalence rates are often generated via disability questions (e.g., the Washington Group Short Set) in census.

There is general interest in providing information on the characteristics of people with disabilities, which justified the inclusion of disability questions in census/household surveys (or for the collection of socioeconomic data of interest in a disability survey). One type of disability survey is that which is intended to estimate both the prevalence of disability and the distribution and characteristics of all persons with disabilities. The other type of disability survey is that which is intended to study, the characteristics, attitudes, perceptions and/or needs only of known persons with disabilities (who can be identified through disability questions in a census, or through administrative records).

1. Scope and purpose objectives

When considering approaches to sampling for disability related surveys, it is first important to understand the objectives and to identify the population of interest. For example, target populations might include:

- Any disability in the population –entire population
- Disability rate for a province –population of province
- Disability rate for school children –population of school children
- Type of disability of persons living in long-term care centres –population living in long term care institutions

Depending on the objective of the survey, the following questions in relation to sampling should be considered:

i. Should disability survey use a standalone design and sampling methodology?
ii. Should disability modules be used in a household survey to collect disability data?

2. Sampling methodology

The sampling methodology to estimate disability prevalence and distribution of disabilities can be both complicated and challenging. The matter of sample design for a survey to measure disability rates and/or the distribution of disability by cause is a highly specialised topic. There are several reasons for this. First, a disability survey, unlike a general-purpose household survey such as one to study labour force activity or the general health conditions and characteristics of a population, is a limited-scope topic which requires dedicated, and perhaps even unusual, sample design procedures. In particular, the level of detail of information to be collected will affect the complexity of the sample design. Second, you need to choose how accurate you want your estimate to be. That is, what margin of error you are willing to have. Third, you need to determine how much confidence you want to have that the true prevalence rate in the population falls within your margin of error. Reducing the margin of error and increasing the level of confidence can significantly increase the size of the sample, so you should pick these parameters to align with your purpose for making the estimate.

This is chiefly because there are numerous conditions and circumstances, which are highly variable by country, that determine what an appropriate, practical sample design would be. These include the actual (or presumed) rate of disability in a country; the availability and quality of administrative records about persons with disabilities that might be of use in sampling; whether national or sub-national estimates (or both) are wanted; and whether an overall disability rate is the main objective or, instead, rates by type of disability are seen as taking precedence.

2.1. Sample size for a population of known persons with disabilities

The issue of sample size is perhaps the main consideration in designing a reliable and affordable disability survey when prevalence is more or less the principal measurement objective. A question often asked is what sample size is needed to generate a large enough sample of people with disabilities for quantitative analysis. The answer depends on a number of factors - the required sample size grows as:

- you are interested in more intersectionality, such as the influence of gender, age, ethnicity, region of residence, or other factors with disability status; and
- the more you want to look at the differences between people with different types of disability or degrees of disability.

To compute the sample size necessary to estimate disability prevalence you first need some basic information, or you need to make assumptions if that information is not known. You must have a rough idea of the standard deviation of the prevalence estimate for the population of interest. If your country has had reasonable disability questions in a census or household survey, you can use the standard deviation of those estimates. However, many surveys and censuses have questions that lead to low prevalence estimates because of poor questions.

The other parameters that are needed to calculate a sample size for a disability survey are the specified precision that the survey must attain, the confidence level required and an estimate of the sampling design effect.

Of course, national level, variables do not usually go far enough to satisfy most users or sponsors of a disability survey. Typically, a distribution is wanted (e.g., age groups, types of disability and the intersectionality of gender and disability), as well as sub-national geographical breakdowns. In those cases, the sample size would have to be multiplied by a factor to provide suitably reliable data to take account of the more detailed analytical requirements. And the more disaggregation you want to do by other characteristics, the bigger your sample size generally needs to be. If you want to compare, for example, rural women with disabilities to urban women with disabilities in terms of their levels of education or poverty status, you must have a reasonable number of each type of woman.

If there is interest in a small sub-population, like people with disabilities, then it may require very large sample sizes to generate a sub-population large enough to make efficient estimates for – especially if you are interested in comparing sub-populations of the sub-population – comparing people with physical vs. mental vs. sensory disabilities.

Sample sizes for analysing the relation of disability to various outcomes requires larger samples than simply estimating disability prevalence.

2.2. Advantages of a standalone disability survey (as opposed to one that is integrated in another household survey)

Advantages

- ✓ More complete in terms of target population
- ✓ Would collect detailed data on disability as its primary goal
- ✓ Would collect demographic and/or economic characteristics of disabled persons as needed

- ✓ Provide more insight about the disabled persons' conditions
- ✓ Greater flexibility

Disadvantages

- ❖ Expensive

2.3. Steps in selecting a standalone disability sample of HHs

- ✓ A sample of PSUs (EAs)
- ✓ Stratify EAs to form strata
- ✓ Select EAs within stratum proportional to their population size
- ✓ Select a sample of households (HHs) within selected EAs
- ✓ Identify HHs with at least one person with disability
- ✓ Partitions sampled HHs into two strata
    - One with HHs with at least one identified disabled person
    - Second with HHs with no identified disabled person
- ✓ Select sample of HHs from both strata
    - Select a large sample of HHs from strata with disabled person
    - Select a small sample of HHs from strata with no disabled person

As mentioned above in relation to sample size., the requisite sample size can quickly become implausibly large on several accounts. Additional information about the distribution, types and causes of disability, create even greater requirements for large sample sizes.

Consequently, it is important to seek ways of increasing the efficiency of sample design.

2.4. Increase efficiency

- ✓ Designing a stratified cluster sample
- ✓ Using list/admin frames where possible ad disabled persons or list
- ✓ Screen for disability on national survey or on census of population ad housing

2.5. Sampling issues to consider when including a disability module in an existing survey

Before attaching a disability module to a survey, one must:

i. Understand the sample design of the survey to be used for the disability module (target population, oversample, etc.)
ii. Understand the limitations for using the survey (sample size, number of disability questions in disability module)
iii. Understand the effect on main survey
iv. Understand the implication on disability data (precision, limitations on amount of data)

Advantages of including a disability module in an existing survey:

- ✓ Allows comparison of disabled persons with general population
- ✓ It is economical

Disadvantages

- ❖ Respondent burden may adversely affect primary survey response rate
- ❖ May provides fewer details on disability questions since it's not a primary disability survey
- ❖ Sample size constraint due to main survey sample size
- ❖ Less flexibility

Limitations

- ❖ Sample size
- ❖ Data on disability is restricted as it cannot dominate the core survey

2.6. Use of sampling frames

A sampling frame for disability related surveys require data source(s) from which a sample is selected.

The quality of the sampling frame or frames—the materials from which the sample is selected—is vital for any household survey. A sampling frame should always adhere as closely as practicable to certain basic principles. The frame must have a strong correspondence with the target population of the survey in order for probability sampling to be achieved. This usually means that the frame must be as accurate and as complete as possible. It also means that the frame should be current, or else provision should be made to bring it up to date prior to sample selection.

A sampling frame should:

- ✓ Represent population of interest
- ✓ Be complete
- ✓ Be recent or current
- ✓ Be accurate
- ✓ If above conditions are not satisfied, take steps to meet above conditions

As a result of the low disability rate in most countries, a general-purpose sample design (e.g., SRS or two-stage sampling) is likely to yield very few cases of persons with disabilities for analysis, unless the sample is very large. For this reason, it is argued that sampling plans should encompass certain techniques that are specifically dedicated to the problem of disability measurement, in order to improve the efficiency of the design. In the Pacific usually the most recently population and housing census is used as the sample frame. In some cases, it is a good idea to sort the frame according to the regional population structure. For example a population census may enable you to identify households with members with disabilities. This way, a systematic sample will retain the appropriate population distribution across regions. To achieve this, it's important to include in the census a well-defined set of questions to determine who has a disability.