Survey of Disability
Overview of Surveys and their design considerations

By

Rajendra Singh

**Presented at the Expert Group Meeting on
the Guidelines and Principles for the Development
of Disability Statistics, 12-14 July 2017**

UN Headquarters, New York

Survey of Disability

- Goals of the survey
- Population of interest (target population) and sampling frame
- Type of disability surveys
- Approaches of sampling
- Estimation of sample size
- Documentation
- Summary
- Issues for discussion

Survey of Disability

- Goals – important to design a survey
  - Prevalence rate – could be defined as:
    - At least one of the disabilities or by type of disability such as hearing, seeing walking, bathing, etc.
    - By demographic or socio-economic characteristics such as sex, age, employment status or income level, etc.
    - By level of geography – states, urban/non-urban, etc.
    - By any combination of above
    - Concepts of disability characteristics should be practical and could be implemented correctly

  - Precision – two options:
    - Coefficient of variation (CV)
    - Margin of error

Survey of Disability

- Population of interest (target population) and sampling Frame:
  - Objectives identify target population, for example,
    - 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 centers – population living in long term care centers
- Sampling Frame – data source(s) from which a sample is selected
- Sampling Frame should
  - Represent population of interest
  - Be complete
  - Be recent or current
  - Be accurate
  - If above conditions not satisfied, take steps to meet above conditions

Survey of Disability

- Approaches for sampling
  - Three main approaching for sampling:
  a) Standalone disability sample of HHs
    - A sample of enumeration areas (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

Survey of Disability

- Standalone sample of HHs (Continued ….)
- 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:
  - It is expensive
- Limitation:
  - Budget

Survey of Disability

- Main sampling approaches (continued …)
  - B) Incorporate a disability topical module in a survey
    - Before using the survey to attach a topical module one must:
      - Understand the sample design of the survey to be used for topical module (target population, oversample, etc.)
      - Understand the limitation for using the survey (sample size, number of disability questions for topical module, etc.)
      - Understand the effect on main survey
      - Understand the implication on disability data (precision, limitations on amount of data)

Survey of Disability

Incorporate a topical module (continued ...)

Advantages:
- 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
- Amount of data on disability

Survey of Disability

- Main sampling approaches (continued …)
  - C) Standalone sample using administrative list(s)

    Two types of lists – a list of persons and a list of institutions
    1) List of Persons
       - Use organizations (stakeholders) with knowledge of lists with disabled persons to obtain all lists to form complete target population
       - Combine multiple lists together into one list
       - Stratify the disabled persons on list(s)
         - By geographical location (province, urban, rural, etc.)
         - By type of disability even if rates are not needed by type of disability
       - Select random sample from each stratum

Survey of Disability

- Standalone sample using list (continued …)
  2) List of institutions
     - Use organizations with knowledge of lists of institutions
     - Create a combined list of persons for each type of institution
     - Stratify institutions
       – By their geographic location
       – By type of institutions (long term care center, home for assisted living or elderly, etc.)
     - Select
       – Simple random sample OR
       – Select a sample proportional to the number of disabled persons residing in each institution, and then select a random sample within the sampled institutions

Survey of Disability

- Standalone sample using list (continued…)
    - The following applies for both types of list(s)
        - Unduplicate persons that are in two or more lists or institutions to get correct selection probability to result in unbiased results
        - Bias results for disabled population if list incomplete or inaccurate
    - Correct for bias due to incomplete or inaccurate list
        - For an incomplete list
            » supplement list sample with general population sample
            » Select a larger sample from list frame and a smaller sample from general population
        - For inaccurate list, determine the source of inaccuracy and take steps to correct the list
        - Unduplicate persons that are in two or more lists – a difficult process unless persons have unique IDs

Survey of Disability

- Standalone sample using list(s) (continued …)
  - Advantages:
    - Good for target population such as persons with known disability, homes for elderly, long term care centers, home for assisted living, etc.
    - Easy to select simple random or systematic random sample
    - Possible to use stratified PPS sample to reduce cost
  - Disadvantages:
    - Requires preparatory work
      - » Check for list completeness
      - » Check list for accuracy
      - » Check list for being current
      - » Check if persons on the list can be located
      - » Creating frames by combing multiple lists (different formats, different order of field locations, etc.)
      - » Supplement list sample with general population HH sample if list is incomplete
  - Limitations:
    - Complexities may limit sharing and combining lists
    - Not always possible to get a complete and accurate list

Survey of Disability

- Sample size considerations – objectives and budget
  - Objectives
    - Disability prevalence rate
    - Precision of prevalence rate -- two types
      - Coefficient of variation (CV)
      - Margin of error
  - Budget
- Parameters needed to estimate sample size are
  - Rate of disability prevalence – if unknown, use the best guess based on the available information
  - Precision for prevalence rate
  - Estimate of design effect – if unknown, use the best guess based on available information

Survey of Disability

- Formula to calculate the sample size based on CV requirement;

$$n = \frac{q}{(CV)^2 p} Deff$$

Where

n	=	Sample size in terms of persons
p	=	Disability prevalence rate
q	=	1-p
CV	=	Coefficient of variation
Deff	=	Design effect

Finite population correction (FPC) factor is assumed to be 1 when n is very small compared to total population size. The sample size formula that include FPC will multiply sample size 'n' by FPC factor [(N-n)/N] where N population size.

Survey of Disability

- Design effect (Deff) is defined as
  - A factor by which the sampling variance for a survey is increased over that which would come about if a simple random sample was used with the same sample size.
  - Mathematically, it is defined as:
    - Deff = $1 + \rho$ (m - 1), where
    - ρ is the intraclass correlation and represents the clustering effect for the characteristic in question
    - m is the (average) size of the cluster
    - Deff is always ≥ 1.0 ;
    - Deff = 1 for only simple random sample

Survey of Disability

- Design effect role in sample size computation
  - Most of the countries use personal interviews to collect survey data
  - To save cost, generally simple random sample is not used instead a multi-stage stratified cluster sample is used
  - Cluster sample increases variance over simple random sample
  - Clustering effect high if characteristics under study is highly clustered
  - Disability is not expected to be highly clustered in general population survey -- multi-stage stratified cluster sample preferred
  - Disability is expected to be highly clustered in institutional population such as long term care centers, elderly housing, etc. – simple random sample preferred

Survey of Disability

- Estimation of sample size
  - Initially estimate the sample size for each target population that meets objectives
  - Budget may not support such a sample size, therefore,
  - Determine the largest sample that can be supported by the budget
  - Consider trade-offs to use sample size supported by budget
  - Consider changing objectives (prevalence rate or precision or both) to remain in budget
  - Consider getting additional budget to support larger sample
  - Most likely it would be an iterative process to reach at final sample size
  - Discuss with sponsor(s) of the survey about the implications of insufficient budget

Survey of Disability

- Survey documentation – an important aspect
  - Document for
    - Future references, and
    - To inform data users
  - Document should include
    - Sampling methodology
    - Estimation methodology
    - Quality of survey data including its strengths and weaknesses
    - Limitations of the data

Survey of Disability

- Summary
  - Main considerations in designing a reliable and affordable survey are:
    - Objectives
    - Precision
    - Budget
  - Additional information (distribution by type, sub-national, etc.) would require larger sample
  - Increase efficiency of design by
    - Designing a stratified cluster sample
    - Using list/administrative frames when possible and disabled persons on list can be easily located
    - Screening for disability on national survey or on census of population and housing

Survey of Disability

- Summary (continued …)
  - Ensure that the frame is complete, accurate and current, if not current, take steps to bring up to date
  - Use enumeration area (EA) as first stage of sampling from general population frame
  - Reduce bias when using list frames by
    - Supplementing list frame with general population frame
    - Unduplicate persons in different frames to get correct selection probability
  - Use multi-stage stratified sample when appropriate

Survey of Disability

- Summary (continued …)
  - Use Proportion to population size (PPS) sampling scheme to select first-stage sample units
  - Reduce design effect by reducing cluster size
  - Consider adjusting objectives if insufficient budget
  - Document sampling and estimation procedures
  - Document data quality including its strength and weaknesses
  - Document limitations of data

Survey of Disability

- Issues for discussion
  - What frame should be used? Census, list or list combined with general population frames.
  - Should disability survey use a standalone design?
  - Should topical module be used in a survey to collect disability data?
  - In case of using topical module, how to increase sample for disability module if primary survey does not have sufficient sample? (ex. collect data over several different time periods – months, quarters or years subject to primary survey design)

Survey of Disability

Thank you

Rajendra Singh