Courtesy of Photoshare

Sampling and Weighting Guide for Non-household Data

Introduction

The Health Accounts methodology tabulates health spending data—from government, donors, NGOs, insurance providers, employers, and households—in a series of two-dimensional tables to show the flow of funds through the system. Collecting quality health spending data to populate the Health Accounts tables can be time and resource intensive. For this reason, good quality secondary data[1] should be used wherever available. However, in some countries, secondary data may not be available for all data sources. When this is the case, health accountants often must turn to primary data collection—survey data.

Survey data are valuable only as far as they reliably reflect the reality of a country's health system. During the planning stages, Health Accounts teams must decide how to go about collecting the necessary data because these decisions can impact the quality and validity of the data. For example, the more comprehensive the data collection, the closer the team gets to calculating total health spending—but the more it costs to undertake and the longer it takes to complete. On the other hand, small-scale surveys can cost very little but often produce results that cannot be generalized to the population as a whole. The objective is to strike an appropriate balance between cost of data collection and accuracy of health spending. Making reliable expenditure estimates with survey data requires understanding the potential pitfalls in the use of these data and familiarity with methods to overcome their inherent weaknesses.

---

[1] Common secondary data sources include: government estimates of revenue and expenditure, chambers of commerce, umbrella organizations or supervisory authorities for insurers, existing national household expenditure surveys, donor data available online, and annual reports from NGOs and/or providers.

[2] This guide provides sampling recommendations for a country with basic data availability. For countries with more sophisticated data systems, cluster sampling, multistage sampling, and other sampling methods are also possible. For countries where their statistics office already does regular surveys of employers, NGOs, etc., the surveys would already take these sampling considerations into account.

Objective and target audience of this guide

This guide provides a high-level overview of common sampling and weighting methods used by Health Accounts teams in the estimation of a country's total health spending.[2] It is designed for Health Accounts technical team members who use survey data as part of their estimation of total health spending. Teams with prior experience conducting Health Accounts can use this guide to assess whether there may be more efficient surveying and weighting options than what may have been used previously. For teams starting Health Accounts for the first time, it is intended to guide the Health Accounts technical team members during the early stages of study design. This guide addresses sampling and weighting methods commonly applied to *non-household* survey data, including: NGO, employer, insurer, and donor data.

Please note: this guide is not intended to be a comprehensive statistical manual. An entire body of literature is devoted to the process of designing and drawing a sample and to the inferential statistics that come from the resulting survey. This guide attempts to provide a Health Accounts team with a practical starting point in determining how to begin thinking about data collection and to make informed choices around handling sampling and weighting. This guide provides a foundation from which team members can solicit additional support as needed from technical and statistical experts.

Getting started: Learning from previous experience

The improvement of surveys (and ultimately your country's Health Accounts data) requires continuous learning and effort. For countries that have conducted previous Health Accounts, it is important that the team review the successes and areas for improvement as a basis for making productive changes and use that information to build off of their previous work. For example, in your previous Health Accounts, was the sample size you used large enough? If not, how can you make adjustments this time around? Additionally, did you get good response rates? If not, what changes will you make to strengthen responses (Hint: Consider the length of time needed to complete the survey; whether proper sensitization was done to inform respondents of the value of Health Accounts; more efficient ways to distribute the questionnaires, etc.)? Teams that are

just starting their Health Accounts for the first time should learn from limitations and identify areas where they could improve their data collection and analysis in future exercises.

Sampling methods

In the scenario where primary data collection was not limited by time and resources, Health Accounts teams would survey all entities in each of the four target populations[3] (donors, NGOs, employers, insurers), receive completed surveys from all entities surveyed, and use the collected data to calculate total health expenditure. In reality, time and resource constraints may prevent the team from collecting survey data from all entities. When this is the case, the Health Accounts team must decide how to collect a sample of data in order to estimate the total spending. In the section below, we describe four common sampling techniques used in resource tracking exercises and identify when each technique is most appropriate.

Census

"In a census, each and every unit in the population is sought and surveyed for the information being collected. In theory, everything is known, and nothing must be guessed. Population censuses and economic censuses are well known examples of this type of data collection, but there are others. Government budget documents that reflect every government unit and every expenditure made by those units also are censuses" (World Bank, WHO, USAID 2003; p. 98).

Census	
Pros	**Cons**
• Everyone in the population is invited to respond. • A census often results in enough respondents to have a high degree of statistical confidence in the survey results.	• A full list of the entire population to be studied must be available. • Requires considerable time and financial resources • Complex to organize and execute (except when the universe is very small) • Can result in survey fatigue—the more times you survey the entire list, the lower the response rate generally is.

[3]For the purpose of this guide, population refers to the statistical definition i.e., the entire set of organizations that could be surveyed.

Non-probability sampling: purposive sampling

A purposive sample is one in which a researcher deliberately selects a sample based on their knowledge about the population and according to the needs of the study (Statistics How To 2018). Once the overall population is identified, the researcher picks the number of entities that can be afforded to sample and that represent the largest portion of health spending in the population. For example, the purposive sample may consist of a handful of companies that the surveyor believes to be the major contributors to health spending—for example, the main parastatal and private enterprises that are already known to be the major spenders on health.

Purposive sampling	
Pros	**Cons**
• Requires a low commitment of resources • Ensures you collect information about the specific entity of interest (e.g., the largest NGOs operating in the health sector or the largest donors) • Can be used when it is impossible to list the full population	• Because purposive sampling is not random, the reliability and validity of the information obtained can be questionable because of selection bias. • The results of purposive samples may not be representative or generalizable to the larger population.

Simple random sampling

In cases where the list of entities with health expenditure is long, one can consider simple random sampling. In a simple random sample, everyone in the population has an equal chance of being selected for the sample. This method is likely to underestimate health spending because small organizations in the population (who also have lower health spending) would be just as likely to be sampled than large ones. Therefore, teams should always conduct stratified random sampling and use simple random sampling only where the data do not permit the division into strata.

Simple random sampling	
Pros	**Cons**
• Relatively easy to assemble the sample • Cheaper and often faster than a full census • The sample is generalizable and represents an accurate statistical measure of the larger population.	• A full list of the entire population to be studied must be available. • Researchers must acquire responses from an adequate number of respondents to ensure sampling bias does not occur—time or budget constraints may limit this. • Large-scale random-sample surveys can require considerable time and financial resources.

Stratified random sampling

In cases where the list of entities with health expenditure is long and characteristics of the entities correlated to health spending are known, one can consider stratified random sampling. Stratified random sampling involves first dividing the population into smaller groups called strata. The strata represent characteristics of the population (e.g., entity size, industry, or ownership) that relate to the probability and level of health spending. Random samples are then selected from each stratum.

Stratified random sampling	
Pros	**Cons**
• Stratification ensures that each subgroup of a given population is adequately represented within the whole sample population. • Provides greater precision than a simple random sample of the same size • Cheaper and often faster than a full census	• A full list of the entire population to be studied must be available. • Enough must be known about the population to allocate every entity in the population to one, and only one, strata. • A stratified random sample typically requires a larger sample size than a simple random sample.

Determining your sample size

The size of a sample has implications for survey precision. Generally, teams should survey the largest sample that is feasible because larger sample sizes result in smaller sampling error and more precise data. There are several factors that go into determining your sample size, including survey power, the level of sampling error, and the confidence interval. Refer to the additional resources on page 5 for assistance with sampling.

Sampling Steps: Sampling employers using stratified random sampling

Figure 1 demonstrates how a Health Accounts team may go about using stratified random sampling to select a stratified random sample of employers.

Figure 1. Steps for sampling employers using stratified random sampling

Define the population

In this example, our population is all licensed employers operating in the country during the year of study

Identify the census population

The census population allows us to identify all employers in the country from which to sample. A complete list of employers in the country can usually be obtained through a regulatory body like the Chamber of Commerce or a Board of Trade. This full list of employers serves as your sampling frame. For this example, let's assume there are 100 licensed employers operating in the country.

Choose the relevant strata

The appropriate number of groups or "strata" depends on the total number of institutions (e.g., total number of employers) in the country and the expected number of institutions in each group (e.g., total number of employers with health spending in each size). In this example, the Health Accounts researchers believe that there may be differences in the types of spending on health between large (101+ employees), medium (50-100 employees), and small (<50 employees) employers. Strata should be chosen such that all actors in the strata behave in a similar way, in this case provide health benefits to their employees in a similar way or at a similar value. Thus, we will use these three strata on the size of the employer to assist in our sampling. Another common stratum for employers is industry type in order to distinguish between construction or manufacturing industries that tend to have higher health benefits for employees, and other sectors.

List the population according to the chosen strata

In order to sort the census list of employers by the three size strata, the Health Accounts team will need to ensure that the census list contains information on the number of employees at each company when requested from the Chamber of Commerce. In this example, we find that 40 employers are large, 40 employers are medium, and 20 employers are small.

Choose your sample size

The Health Accounts team should work with a statistician to conduct a sample size calculation to determine a proper sample size that adequately represents the population. In reality, budget and time limitations may also help determine the size of your sample. For this example, our sample size is 30 employers.

Calculate a proportional stratification

We need to ensure that the number of employers selected for the sample from each stratum is proportionate to the number of large, medium, and small employers in the population. To achieve this, we first multiply the desired sample size by the proportion of units in each stratum. Therefore, to calculate the number of large employers required in our sample, we multiply 30 by 0.40 (i.e., 40% of the population of employers), which gives us a total of 12 large employers. If we do the same for medium employers, we get 12 employers (i.e., 40% of employers are medium, where 30 x 0.40 = 12). If we do the same for small employers, we get 6 employers (i.e., 20% of employers are small, where 30 x 0.20 = 6). This This means that we need to select 12 large employers, 12 medium employers, and 6 small employers for our sample of 30 employers.

Use a simple random sample to select your sample

Now that we have chosen to sample 12 large, 12 medium, and 6 small employers, we will use simple random sampling to select the employers to survey from the list.

Weighting methods

In a typical survey covering a range of large and small organizations, the team may decide to guarantee the inclusion of the five or six largest organizations and sample from the remaining smaller organizations. The respondents who are selected with certainty represent only themselves, while each of the sampled smaller organizations represents a number of others that we hope are like it. We estimate the national total by the number of organizations it represents. Organizations chosen with certainty have a weight of one, because each represents only itself. If we sampled one-fifth of the medium-sized organizations, each would have a weight of five. If we sampled one in fifty of the smallest organizations, each would have a weight of fifty. Because some organizations do not respond to survey invitations, we base these weights on the actual number of respondents.

As the example in Figure 2 shows, deciding to weight the data and the method by which the weights are applied can result in very different estimates of health spending, ranging from $800 thousand to $2 million. Therefore, to estimate the level of spending on health in a country, it is crucial that Health Accounts teams apply weights correctly.
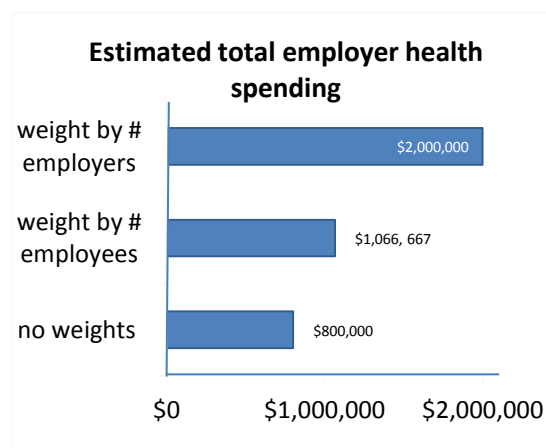
Additional resources

- Contact your country's central statistical office or ministry dealing with population or labor and employment matters as these ministries have statistical experts among their staff who can assist Health Accounts teams with sampling design and weighting.
- Researchers within the School of Public Health or School of Economics at local universities often have statistical expertise that can support Health Accounts teams with sampling design.
- There are several free websites that provide sample size and confidence interval calculators:
    - https://www.surveysystem.com/sscalc.htm
    - http://www.raosoft.com/samplesize.html
    - https://www.qualtrics.com/blog/calculating-sample-size/

The following are additional resources around surveys and sampling:

- World Bank, WHO, and USAID. 2003. Guide to Producing National Health Accounts with Special Applications for Low-income and Middle-income Countries. Accessed August 20, 2018: http://www.who.int/health-accounts/documentation/English_PG.pdf (Chapter 8 focuses on developing and using surveys in health accounts).
- Henry, G. T. 1990. "Practical Sampling." Applied Social Research Methods Series, Volume 21. Sage Publications.
- Kish, L. 1965. Survey Sampling. New York: John Wiley and Sons (contains a fairly detailed model of survey error).
- Salant P., and D. A. Dilman. 1994. How to Conduct Your Own Survey. John Wiley & Sons, Inc.
- Schonlau, M., R. D. Fricker, and M. N. Elliott. Conducting Research Surveys via E-mail and the Web. ISBN: 0-8330-3110-4, MR-1480-RC, Rand, 2001. Available at http://www.rand.org/publications/MR/MR1480

Figure 2. Different weighting methods can result in very different estimates of total health spending

Employer	Employer size (# employees)	Health expenditure reported in survey
A	50	No response
B	20	No response
C	30	No response
D	100	$300,000
E	200	$500,000
Total	400	$800,000

Average expenditure per employer = 800,000 / 2 = 400,000

Average expenditure per employee = 800,000/300 = 2,667

Estimated total employer health spending

- weight by # employers — $2,000,000
- weight by # employees — $1,066, 667
- no weights — $800,000

Deciding which weighting logic to use depends on the process followed to sample the institution and on the amount of information available on the census list. Because sampling and weighting approaches are linked to each other, they must be established before data collection begins. The next section discusses weighting options for Health Accounts.

Weighting options

Table 1 summarizes options for sampling and weighting for each type of institution surveyed in the Health Accounts estimation. For teams using the Health Accounts Production Tool (HAPT) software to assist with estimation of your country's health spending, application of weights occurs in Module 4 "Data Import" of the HAPT (World Health Organization n.d.).

Table 1. Sampling and Weighting Recommendations by Institution Type

Entity Type	Sample type (in order of preference)	Data needed for census or sampling	Weight type (in order of preference)	Data needed for weighting
Donors, NGOs, and Insurers	**Census**: If the number of entities operating in the health sector is low (as can be the case with donors, NGOs, and insurers), it is recommended that Health Accounts teams survey all entities in the country rather than take a sample.	• A complete list of donors/NGOs/insurers in the country • Contact information for all entities so that surveys can be delivered to them	**Donors—no weighting**: When expenditures are collected from all entities in the population, there is no need for weighting. In the event that there is a low response rate, it is still not recommended to weight the donor data, as health spending from one donor to the next can vary greatly. It is better to underestimate donor spending on health than apply weights and make the unsubstantiated assumption that spending across donors is consistent.	Not applicable
			NGOs and insurers: When expenditures are collected from all entities in the population, there is no need for weighting. In the event that there is a low response rate and information on the size and other relevant characteristics of the sample is available, refer to the weighting instructions below for "stratified random sample."	Refer to weighting options for "stratified random sampling" below.
	Stratified random sample: In cases where the list of entities with health expenditure is long, one can consider random sampling stratified by size or other characteristics that relate to the probability and level of health spending.	• A complete list of donors/NGOs/insurers operating in the health sector • Information on stratification criteria (e.g., size, industry, ownership) available for the complete list of entities from which you sample. For donors, size could be by the approximate size of their programs, and for insurers this could be by the approximate number of their insurance policy holders • Contact information for all sampled donors/NGOs/insurers so that surveys can be delivered to them	**Donors and NGOs—weight by number of donors[3]/NGOs stratified by size, type, or other characteristic:** The assumption is that similar entities (by type, size, etc.) will likely allocate resources for health in similar ways and amounts.	• A complete list of all entities operating in the health sector • Entity size, type, or other stratification characteristic included in the complete list
			Insurance—weight by number of individuals covered stratified by type or other characteristic: When data on the number of individuals covered by a health insurer are available and data on the insurer type are available, Health Accounts teams can also choose to weight based on the size of the insurer within each type.	• A complete list of all insurers in the country • Insurer size , type, or other stratification characteristic included in the complete list
	Simple random sample: In cases where the list of entities with health expenditure is long and no other data are available on the characteristics of the entities, one can conduct a simple random sample.	• A complete list of NGOs operating in the health sector • Contact information for all sampled NGOs so that surveys can be delivered to them	**Donors—no weighting:** Because donors are not homogenous in the nature of their health spending, it is not advised to weight solely by the number of donors. In these cases, it is better to acknowledge that donor spending on health is underestimated.	Not applicable
			Weight by number of NGOs/insurers only: This should be used with caution and only if stratification methods are not feasible.	• List of NGOs/insurers in the health sector
	Purposive sample: As a last resort, Health Accounts teams should identify and purposively sample those major donors/NGOs/insurers known to support the largest health programs and to have the most health spending.	• A complete list of donors/NGOs/insurers in the country • Contact information for all entities so that surveys can be delivered • Information on the general magnitude of health programs by entity to inform sampling; expert opinion can be used to identify the biggest donors and NGOs; this information may be available from an umbrella organization or supervisory authority for insurers.	**No weighting**: By nature, purposive samples tend to be small, which makes it likely that the team can follow up with and get responses from each of the surveyed entities. Because purposive sampling is not random, weighting should not be used to estimate the population spending because the health spending from the purposive sample is unlikely to represent spending of donors/NGOs/insurers that weren't sampled. It is better to underestimate spending on health than apply weights and make the unsubstantiated assumption that spending across entities is consistent. In the event of a low response rate within the purposive sample and the entities in the purposive sample are of a similar magnitude, the team could consider weighting for the non-responses in the sample but not up to the population.	Not applicable

[3] HAPT software does not allow for weighting of donor data within the tool itself. Therefore, any donor weighting would need to be calculated outside of the HAPT software and uploaded.

Entity Type	Sample type (in order of preference)	Data needed for census or sampling	Weight type (in order of preference)	Data needed for weighting
Employers	**Stratified random sample**: Health Accounts teams typically conduct random sampling of employers. The health benefits package employers provide to their employees is usually competitive and similar in nature, which makes them suitable to consider sampling. Samples should be stratified by characteristics believed to relate to probability and level of expenditure. For example, spending by enterprises for health care typically varies by industry, enterprise size, ownership (public and private sectors, foreign or domestic), etc.	• A complete list of employers in the country • Information on stratification criteria (e.g., size, industry, ownership) available for the complete list of entities from which you sample • Contact information for all employers so that surveys can be delivered to them	**Weight by number of full-time employees stratified by size, type, or other characteristic:** When data on the number of employees are available and data on the industry sector are available, Health Accounts teams can also choose to weight based on the size of the company within each industry.	• A complete list of employers in the country • The number of full-time employees listed for each employer in the full list • Each employer's industry or other stratification criteria identified in the complete list of employers
	Simple random sample: In cases where no other data are available on the characteristics of the employers, one can conduct a simple random sample.	• A complete list of employers • Contact information for all sampled employers so that surveys can be delivered to them	**Weight by number of employers only:** This should be used with caution and only if stratification methods are not feasible.	• A complete list of employers
	Purposive sample: If time or resource constraints prohibit the use of a random sample, Health Accounts teams should identify and survey a purposive sample of the largest employers in the country known or suspected to have the most health spending. In particular, the team should target employers that may have their workplace programs, health facilities, and contracts with providers (e.g., anything other than health insurance benefits, because in a worst case scenario, the health insurance contributions from employers can be triangulated from the insurance data).	• A complete list of employers in the country • Contact information for all employers so that surveys can be delivered to them • Information on the largest employers in the country to inform sampling (this information may be available through the Chamber of Commerce)	**No weighting**: By nature, purposive samples tend to be small, which makes it likely that the team can follow up with and get responses from each of the surveyed entities. Because purposive sampling is not random, weighting should not be used to estimate the population spending because the health spending from the purposive sample is unlikely to represent spending of employers that were not sampled. In addition, because you have purposively sampled the largest employers, in theory, you have captured the majority of non-insurance health benefits provided by private companies. While you may risk underestimating employer spending on health, the amount of underestimation should be low. In the event of a low response rate within the purposive sample and the entities in the purposive sample are of a similar magnitude, the team could consider weighting for the non-responses in the sample but not up to the population.	Not applicable
	Census: Because the number of employers in a country is usually large, the time and financial requirements to conduct a census make it infeasible and not recommended.	Not applicable	Not applicable	Not applicable

Weighting steps: Example of how to apply weights to your data

The following example builds off of the employer sampling example from page 4, assuming that the country is using weighting by employer size (e.g., large, medium, and small). In our example country's health system, there are 100 employers of varying sizes. Forty of the employers are large, 40 are medium size, and 20 are small. Table 2 summarizes the survey data of these 100 employers.

When taking the categorical approach to weighting (most common), weights are calculated according to the following basic formula:

$$Weight = \frac{total\ \#\ (number\ of\ institutions\ in\ the\ population)}{total\ \#\ (of\ institutions)\ that\ returned\ a\ survey}$$

Weights are calculated for each employer group and then multiplied by the amount of spending collected through surveys in that group. These products are added together to estimate the total spending by the population of employers. For our example country, the employer weighting will be:

Table 2: Example of unweighted and weighted employer survey data

Employer size	Employers in health system (P)	Sample during the survey	Employers responding to the survey (K)	Total amount spent (m)	Weight = P/K	Total weighted health expenditure by employer size = (P/K)*m
Large	40	12	10	$6,500	40/10 = 4.0	$26,000
Medium	40	12	5	$4,000	40/5 = 8.0	$32,000
Small	20	6	5	$1,700	20/5 = 4.0	$6,800
			Total (unweighted) =	$12,200	Total (weighted) =	$64,800

As you can see in this example, the total unweighted amount collected through surveys was $12,200. Once the weights were applied, the weighted value of health spending in the country is $64,800. This weighted value accounts for estimated health spending across the entire population of employers. Without weighting, the team in this example would underestimate employer contributions to health spending.

Conclusion

This is an introductory guide to help Health Accounts teams with their data planning process. Health Accounts teams are not limited to their own expertise in generating data—most central statistical offices and ministries dealing with population or labor and employment matters have statistical experts among their staff that the team should engage where needed (see resources on page 5). These colleagues can help the Health Accounts team identify sampling frames and sample sizes. The time invested upfront in collecting the necessary information to conduct proper sampling and weighting will pay off longer term, leading to a more accurate picture of your country's health spending.

References

Ashagari, T., S. Nakhimovsky, A. Zida, and H. Cogswell. 2014. Health Accounts Production Tool SHA 2011: User Guide, Version 2.0. Accessed August 20, 2018, https://www.hfgproject.org/health-accounts-production-tool-sha-2011/

Statistics How To. 2018. Purposive Sampling (Deliberate Sampling), http://www.statisticshowto.com/purposive-sampling/

World Bank, WHO, USAID. 2003. Guide to producing national health accounts with special applications for low-income and middle-income countries, Accessed August 20, 2018 http://www.who.int/health-accounts/documentation/English_PG.pdf

World Health Organization. n.d. Health Accounts Production Tool (HAPT), Accessed September 17, 2018 http://www.who.int/health-accounts/tools/HAPT/en/