Institute for Health Metrics and Evaluation

Data Release Information Sheet

Data Summary

Dataset name: Panama Salud Mesoamérica Initiative Baseline Census and Household Survey 2013

Project name: Salud Mesoamérica Initiative Evaluation

Date of release: February 27, 2019

Summary:

The Salud Mesoamérica Initiative (SMI) focuses on reducing inequalities in maternal and child health in Mesoamerica. This dataset is the product of an SMI impact evaluation. It includes results of a baseline household census and baseline household survey conducted in two provincial-level indigenous regions, or comarcas, in Panama: Emberá-Wounaan and Guna Yala. The census captured basic demographic characteristics of all usual household occupants and was used to produce a sample of households containing eligible women (ages 15-49) and children (ages 0-59 months) for the household survey. In the household survey, any household heads and all eligible women were interviewed. Information was collected on additional demographic characteristics; healthcare use, access, and expenditures; and perceived quality of key interventions for women of reproductive age and children. Height, weight, and anemia measurements were taken for children under age 5.

Relevant publications and visualizations:

Mokdad AH, Colson KE, Zúñiga-Brenes P, Ríos-Zertuche D, Palmisano EB, Alfaro-Porras E, et al. Salud Mesoamérica 2015 Initiative: design, implementation, and baseline findings. *Popul Health Metr*. 2015 Feb 7; 13:3. doi: 10.1186/s12963-015-0034-4.

Acknowledgments

Contributing organizations:

- Institute for Health Metrics and Evaluation (IHME)
- Foundation for Education and Social Development (FES)

Funders:

- Bill and Melinda Gates Foundation (BMGF)
- Carlos Slim Health Institute
- Spanish Agency for International Development Cooperation (AECID)
- Inter-American Development Bank (IDB)

File Information

Data Files

File Name	Description	Data structure
IHME_SMI_PAN_HHS_2013_CENSO_EMBERA_Y2019M02D27 [CSV, DTA]	PAN Emberá-Wounaan census	Each row represents one household, unique identifier “seg” + “nhogar”
IHME_SMI_PAN_HHS_2013_CENSO_GUNA_Y2019M02D27 [CSV, DTA]	PAN Guna Yala census	Each row represents one household, unique identifier “seg” + “nhogar”
IHME_SMI_PAN_HHS_2013_FECHA_EMBERA_Y2019M02D27 [CSV, DTA]	PAN Emberá-Wounaan dates: 1 survey per household	Each row represents one household, unique identifier “seg” + “nhogar”
IHME_SMI_PAN_HHS_2013_FECHA_GUNA_Y2019M02D27 [CSV, DTA]	PAN Guna Yala dates: 1 survey per household	Each row represents one household, unique identifier “seg” + “nhogar”
IHME_SMI_PAN_HHS_2013_MOD1_EMBERA_Y2019M02D27 [CSV, DTA]	PAN Emberá-Wounaan household module 1: 1 survey per household	Each row represents one household, unique identifier “seg” + “nhogar”
IHME_SMI_PAN_HHS_2013_MOD1_GUNA_Y2019M02D27 [CSV, DTA]	PAN Guna Yala household module 1: 1 survey per household	Each row represents one household, unique identifier “seg” + “nhogar”
IHME_SMI_PAN_HHS_2013_MOD2A_EMBERA_Y2019M02D27 [CSV, DTA]	PAN Emberá-Wounaan household module 2a: 1 interview with every eligible woman aged 15-49 years in the household	Each row represents one woman, unique identifier “seg” + “nhogar” + “id_woman”

File Name	Description	Data structure
IHME_SMI_PAN_HHS_2013_MOD2A_GUNA_Y2019M02D27 [CSV, DTA]	PAN Guna Yala household module 2a: 1 interview with every eligible woman aged 15-49 years in the household	Each row represents one woman, unique identifier “seg” + “nhogar” + “id_woman”
IHME_SMI_PAN_HHS_2013_MOD2B_EMBERA_Y2019M02D27 [CSV, DTA]	PAN Emberá-Wounaan household module 2b: 1 interview per live birth in the last 5 years for each eligible woman aged 15-49 years in the household	Each row represents one delivery, unique identifier “seg” + “nhogar” + “id_woman” + “lb_num”
IHME_SMI_PAN_HHS_2013_MOD2B_GUNA_Y2019M02D27 [CSV, DTA]	PAN Guna Yala household module 2b: 1 interview per live birth in the last 5 years for each eligible woman aged 15-49 years in the household	Each row represents one delivery, unique identifier “seg” + “nhogar” + “id_woman” + “lb_num”
IHME_SMI_PAN_HHS_2013_MOD2C_EMBERA_Y2019M02D27 [CSV, DTA]	PAN Emberá-Wounaan household module 2c: 1 interview with the mother/caregiver of each eligible child aged 0-59 months in the household	Each row represents one child, unique identifier “seg” + “nhogar” + “id_kid”
IHME_SMI_PAN_HHS_2013_MOD2C_GUNA_Y2019M02D27 [CSV, DTA]	PAN Guna Yala household module 2c: 1 interview with the mother/caregiver of each eligible child aged 0-59 months in the household	Each row represents one child, unique identifier “seg” + “nhogar” + “id_kid”

File Name	Description	Data structure
IHME_SMI_PAN_HHS_2013_MOD3_EMBERA_Y2019M02D27 [CSV, DTA]	PAN Emberá-Wounaan household module 3: 1 anthropometric measurement module per eligible child aged 0-59 months in the household	Each row represents one child, unique identifier “seg” + “nhogar” + “id_kid”
IHME_SMI_PAN_HHS_2013_MOD3_GUNA_Y2019M02D27 [CSV, DTA]	PAN Guna Yala household module 3: 1 anthropometric measurement module per eligible child aged 0-59 months in the household	Each row represents one child, unique identifier “seg” + “nhogar” + “id_kid”

Additional File Information

Inventory

A file inventory (**IHME_SMI_PAN_HHS_2013_FILE_INVENTORY_Y2019M02D27.XLSX**) contains a list of all files and information on each. This information includes file name, format (CSV, PDF, etc.), type (data, codebook, questionnaire, or documentation), description, and version date.

Language

The questionnaires and codebooks for all Belize surveys are in English. The questionnaires and codebooks for the El Salvador Baseline Health Facility Survey are in English. The questionnaire for the El Salvador Baseline Census and Household Survey is in both English and Spanish, while the codebooks are in Spanish. Questionnaires and codebooks for all other surveys are in Spanish. Some questionnaires and codebooks also include indigenous languages.

The comarca of Guna Yala in Panama is also referred to as Kuna Yala and both names may be seen over time throughout this study.

Codebooks

These contain variable names, questions (variable labels), numeric values and labels for coded values, and question types. Accents have been removed in SMI codebooks, which affects Spanish and indigenous language translations. Some codebooks, however, contain other special characters that do not display properly if a CSV is opened in Excel. Therefore, codebooks are provided in both the machine-actionable CSV format and, for reference, the human-readable XLSX format.

There are nine major types of questions found in the codebooks: calculated, checkAllItem, checkAllSummary, comment, hidden, poplist, preload, radioGroup, and text. These question types

are determined by the survey software program and can be grouped into the following categories:

- Check all that apply: The check all that apply questions are labeled as either checkAllItem or checkAllSummary
- Single response option: The single response options are labeled as either poplist, preload, or radioGroup. These are defined by the type of table that was used in the DatStat survey.
- Text response: The text response is labeled as either text or comment
- Calculated or pre-populated variables: Any variable that is created by the survey software system, such as time it takes to complete the survey, is labeled as either calculated or hidden. For the purposes of this study, no birth dates or identifiable information is collected, but internal survey calculations were done based on the date of birth to determine age in years. Other examples of pre-populated variables include geographic information piped into the survey from external census data.

The codebooks contain the most accurate list of variables asked in the survey. The PDF questionnaires produced do not reflect questions that were hidden from participants and interviewers after the initial survey was published for testing and piloting purposes.

Methodological Statement

Data Collection

Data collection for the SMI-Panama baseline measurement was conducted by Fundación para la Educación y el Desarrollo Social (Fundación FES). All surveys were conducted using a computer-assisted personal interview (CAPI). The CAPI was programmed using DatStat Illume and installed onto computer netbooks. CAPI supports skip patterns, inter-question answer consistency, and data entry ranges. The aim of introducing CAPI to the field was to reduce survey time by prompting only relevant questions, maintain a logical answering pattern across different questions, decrease data entry errors, and permit rapid data verification.

Portable scales and stadiometers were used for the anthropometric measurements, and hemoglobin levels were assessed in the field using a portable HemoCue machine. Medically trained personnel (i.e., professional nurses) performed all physical measurements.

The SMI-Panama baseline household census, which captures basic demographic characteristics of all usual household occupants, was carried out between April 2, 2013 and June 27, 2013.

Data collection for the SMI-Panama baseline household survey began on April 23, 2013 and was completed on August 31, 2013. To assure completeness of the sample, field staff were instructed to return to selected households up to three times (on different days, and at least once on a weekend) in an attempt to complete the Household Characteristics Questionnaire, the Maternal and Child Health Questionnaire, and the Physical Measurements Module. Households that refused to participate or were absent at all three visits were substituted with randomly selected alternates.

Data collection teams, consisting of one supervisor and three to five interviewers, were deployed

to conduct the SMI household census and the SMI household survey. Supervisors were responsible for reviewing questionnaires for quality and consistency prior to departing to each segment. There were five supervisors overseeing the SMI household census and SMI household survey at baseline.

Data were collected using computer netbooks equipped with CAPI software. Field team leaders monitored the implementation of the survey and report feedback. Data collection using CAPI allowed data to be transferred instantaneously once a survey was completed via a secure connection to IHME. IHME monitored collected data on a continuous basis and provided feedback. Suggestions, surveyor feedback, and any modifications were incorporated into the instruments and readily transmitted to the field.

The research protocol was approved by the Internal Review Board of the University of Washington. All data collection instruments and procedures were approved by the National Ethics Committee of the Ministry of Health of Panama.

Sampling/Population

The study design for the SMI-Panama household survey provides representative estimates of the coverage of key health interventions and indicators for a geographic area that approximates the lowest wealth quintile of the population of Panama.

Panama is divided into ten provinces and three indigenous comarcas. Two comarcas, Emberá and Guna Yala, were purposefully selected for SMI-Panama on the basis of their high concentration of residents in the country's lowest wealth quintile. From these two comarcas, a two-stage clustered random sample of eligible households was selected.

First-stage sample selection: census segments

The household survey uses a two-stage random sampling design in order to balance survey administration costs with the ability to make estimates representative of the population in the study area. For the SMI-Panama household census, the primary sampling unit (PSU) is the *lugar poblado* (populated place) from the 2010 Panama Population Census. A representative sample of these clusters ("segments") was randomly selected from a sampling frame of all segments in SMI municipalities with probability proportional to size, where size is measured by the number of occupied households. Samples for intervention strata (61 segments) were selected independently.

A set of alternate segments was selected using identical methodology, to be surveyed in the event that any of the selected segments could not be surveyed and needed to be replaced due to security concerns, community rejection of the study, or a high proportion of absent households. Four segments were replaced due to logistical reasons. In each case, a randomly selected alternate from the same district was used.

Second-stage sample selection

The SMI-Panama household census is conducted in each of the randomly selected segments prior to the SMI-Panama household survey in order to identify all eligible women and children for second-stage

sampling. Interviewers visit every household in the segment and create a household roster capturing the age and sex distribution of household members.

Eligible households are systematically selected from the complete census listing for participation in the SMI-Panama Household Survey. Thirty households with eligible participants are selected for participation. In order to ensure at least 30 complete interviews per segment, 10 backup households are selected at random in case of refusals or absent households.

All women aged 15-49 years who are members of the selected household are eligible to be interviewed, and all children aged 0-59 months who are members of the selected household are eligible for the physical measurement module. Any household head or other individual knowledgeable about household characteristics and expenditures is permitted to respond to the household characteristics module, while any primary caregiver of a child 0-59 months is eligible to inform for the child health interview module, regardless of sex or age.

Weighting

Survey weights reflect the three-stage cluster sampling design of the study. The primary sampling unit is referred to as the "segment." The segment is censused, and 30 households with eligible participants selected at random. Within selected households, all women 15-49 years of age and all children 0-59 months of age are selected for participation in the survey. Design weights for households, women and children were calculated according to the inverse probability of selection of the unit. No post-stratification adjustments were made to the weights. Although cluster sampling can improve efficiency when the target population is spread out over a large area, the resultant sample consists of observations that are not completely independent of one another. Estimation using SMI survey data should apply sampling weights in order to represent the population of the study area, and should account for intra-class correlation by specifying clusters and strata when calculating variance.

Imputed Variables and/or Constructed Variables – What was Imputed/Constructed and How

There are no imputed variables in the data. The constructed variables are labeled in the dictionary as "IHME Generated" and consist only of the weights constructed by the research team.

Known Data Quality Issues

Sampling errors & Design Effects for Key Indicators

- Visit disposition codes and visit dates in the Dates module are entered by the interviewer and some conflicting responses may exist between individual household visits and final disposition codes.
- Age is registered at the time of the household census, so age in months may represent a lag of up to a month at the day of interview/physical measurements.
- Additional household members added in Module 1 are sometimes duplicates of household members who were censused, then re-added by mistake.
- Continuous data (e.g., household expenditures) have not been cleaned or trimmed for outliers.

- There are data entry errors, specifically with the child roster IDs in Modules 2A and 2B, that were entered in the field. IHME had a thorough data verification process and communication system with the field team, although not all originally reported data may align throughout the census to the household.
- Ages in livebirth roster have not been reconciled or corrected except in some cases where the child is under age 5.
- Module 3 had a separate consent process for child anthropometric measurements and was sometimes refused even for children who have an interview module 2C. Refusals of module 3 are not always captured accurately in the dates module.
- During the cleaning and measurement process for country indicators, some don't know or decline to respond answer values were irreversibly replaced with missing values.
- Anemia screening by anthropometrists was according to a field conversion table for hemoglobin concentration by altitude, and in some cases recorded anemia status may vary from a standard clinical definition or may have been mis-recorded. Height measurement position may also be mis-recorded in some cases.
- Some households may be missing one or more modules due to refusal or tracking errors in the field.

Public Use Dataset Notes

This is a public use dataset. The data have been de-identified. Variables determined to contain identifiable private information, or potentially identifiable private information, for health facilities, health workers, and/or other individuals have been removed in accordance with IHME's microdata release protocol. The protocol's determination for variables that constitute identifiable private information is based primarily on HIPAA'S De-identification Standard.

No personally identifiable information was collected for this study; however, these data were stripped of comments and information on who conducted the interview. Some variables in the dataset do not contain data, such as date of birth, because this information was not stored on the survey or sent to IHME. The date of birth was entered into the survey and an internal calculation was done to provide age.

Additional Information

Terms and Conditions

http://www.healthdata.org/about/terms-and-conditions

Contact Information

To request further information about the Salud Mesoamérica Initiative (SMI), please contact IHME:

Institute for Health Metrics and Evaluation
2301 Fifth Ave., Suite 600
Seattle, WA 98121
USA
Telephone: +1-206-897-2800
Fax: +1-206-897-2899
Email: data@healthdata.org
www.healthdata.org

These files may be updated periodically, so we appreciate hearing feedback or additional information about how these data are being used.