

# METHODOLOGY



SAMPLING

Consistent with previous years, a two-stage cluster approach was used for the selection of the VASyR sample. The main sample frame used was the number of Syrian refugees known to UNHCR as of January 2019. Sampling was based on a “30 x 7” two-stage cluster scheme initially developed by the World Health Organization. This method outlines a sample size of 30 clusters per geographical area and seven households per cluster which provides a precision of +/- 10 percentage points.<sup>1</sup> Districts were considered as the geographical level within which 30 clusters were selected. There are 26 districts in Lebanon, where Beirut and Akkar each represent a district and a governorate. As such, to ensure representativeness of these two districts as governorates, an additional two cluster samples were considered for each. Additionally, one cluster sample was also added to Baalbek-El Hermel governorate which is composed of only two districts. Adding an additional cluster sample ensured all governorates have at least three cluster samples.

The primary sampling unit was defined as the village level (i.e. cluster) and UNHCR cases served as the secondary sampling unit. A case was defined as a group of people who are identified together as one unit (usually immediate family) under UNHCR databases. Villages were selected using probability proportionate to size where villages with a larger concentration of refugees were more likely to be selected and 30 clusters per village were selected<sup>2</sup> with four replacement clusters, per district.

In order to estimate the sample size needed to generate results that are representative on a district, governorate and national level, the following assumptions were used:

- 50% estimated prevalence
- 10% precision
- 1.5 design effect
- 5% margin of error

Using the above parameters, 165 cases per district was required, leading to a target of 4,950 cases nationally. Due to the known high level of mobility of the Syrian refugee population and based on experience in previous rounds of the VASyR and other household level surveys,

a 40% non-response rate was considered, yielding 8,250 cases as the pool from which cases were targeted. Cases from the pool were selected by the following breakdown:

- 8,250 cases distributed over 30 districts/ 34 clusters per district: 8 cases per cluster

Due to some clusters having less than eight cases, a total of 8,079 cases were used as the sample pool for the survey. Of these, 4,769 were visited.

TRAINING AND FIELD WORK

Separate enumerator trainings were carried out in each operational region (Bekaa, Mount Lebanon, North and South) covering the data collection tool, contextual background, methodology and ethical considerations. The trainings were administered by UNHCR, WFP and UNICEF staff over the course of seven days, including two field test days. Data was collected and entered on electronic tablets by the enumerators during the interviews using Open Data Kit (ODK) software. The data was then sent to UNHCR’s Refugee Assistance Information System (RAIS) Platform. Data collection took place between 8 April and 3 May 2019 through face-to-face interviews at refugee homes by four partners in each region, as shown in Table 1.

Table 1: Partners that conducted data collection

Akkar	Caritas
Baalbek-El Hermel	World Vision International
Beirut	Makhzoumi Foundation
Bekaa	World Vision International
Mount Lebanon	Makhzoumi Foundation
El Nabatieh	SHEILD
North	Caritas
South	SHEILD

<sup>1</sup> World Health Organization. Training for Mid-level Managers: The EPI Coverage Survey. Geneva: WHO Expanded Programme on Immunization, 1991. WHO/EPI/MLM/91.10  
<sup>2</sup> Using the Emergency Nutrition Assessment (ENA) Software.

## QUESTIONNAIRE

The 2019 VASyR questionnaire consisted of 486 questions that collected data at the household level and individual level. Questions included those on demographics, legal documentation, safety and security, shelter, WASH, health, food security, livelihoods, expenditures, food consumption, debt, coping strategies and assistance, as well as questions specifically relating to women, children and people with disabilities. The VASyR questionnaire is a household survey administered with either the head of the household or any other adult household member. The full questionnaire can be downloaded via the following link:

<https://data2.unhcr.org/en/documents/details/71337>



## DATA QUALITY ASSURANCE

The following steps were taken to monitor the quality of collected data:

1. Using a harmonized check list, each VASyR core agency conducted frequent spot checks on each of the data collection teams across Lebanon<sup>3</sup>. Feedback was provided directly after the interview was completed and reports were scanned and shared with the respective area coordinator and Core Group members. No interview was interrupted, unless crucial intervention was needed in events such as violation of the ethical regulations.
2. Agencies conducted follow up phone calls for a randomly selected 5% of the weekly target number of households each week to verify a few questions from the interview and get feedback on the enumerators' performance.
3. At the end of each week, a data collection summary report was shared with all agencies to check on the progress of data collection.
4. A WhatsApp group was created among the enumerators and general feedback was shared on weekly basis.

## DATA PROCESSING

Weighting of the data was necessary to ensure that the geographical distribution of the population was reflected in the analysis and to compensate for the unequal probabilities of a household being included in the sample. The normalized weight was calculated for each district using the following formula:

$$w_n = \frac{(N_s/N)}{(n_s/n)}$$

Where  $w_n$  is the normalized weight,  $N_s$  is total sample frame of the district,  $N$  is the total national sample frame,  $n_s$  is the number of household visited in the district and  $n$  is the total visited households.

The data was cleaned for any significant outliers and consistency checks were applied to spot any data errors. Results were disaggregated by district, governorate, gender of the household head, shelter type, food security and economic vulnerability, were deemed necessary. Data was analysed using SPSS version 20.

## LIMITATIONS

As with any survey, limitations are expected; several main limitations are discussed here:

1. The VASyR relies primarily on self-reported data which may be bias or untrue. To minimize the impact of this bias, enumerators were trained on providing a comprehensive informed consent to reassure confidentiality, purpose, risks and benefits.
2. Sample sizes for specific age groups may be small as the sampling strategy was not conducted for this purpose. Thus, results for such age groups are either not reported (e.g., cases below 25), not segregated by geography (e.g., IYCF) or reported but with caution.
3. The VASyR sampling frame excluded Syrian refugees who have never approached UNHCR (unless within a targeted household). This population is a consistent gap in data on Syrian refugees in Lebanon.
4. The VASyR questionnaire and respective indicators are subjected to adjustment and changes in order to ensure that the most accurate definition or calculation is being

<sup>3</sup> Refer to <http://ialebanon.unhcr.org/vasyr> for a detailed description of the spot checks procedure and tools used.



used. These changes may have caused some results not to be directly comparable with previous years.

5. The VASyR is a household survey and the interview is usually conducted with the head of household or any other adult household member. As such, there are no individual interviews carried out with each family member and obtaining accurate information on particularly sensitive topics is a challenge (i.e. child labour or harassment).

6. The VASyR methodology does not mitigate for gender bias in respondents' answers. Given that an interview is typically conducted with the head of household, and only 18% of households were female-headed, the voices and issues of females may not be accurately captured. As a result, the data is likely to be skewed towards male voices and perspectives.

7. Families that had recently moved to a different governorate and whose address was not updated with UNHCR were not captured through this assessment. This is because households are sampled according to their geographical area of residence (known to UNHCR) and linked to the predetermined clusters.

## FIELD CONSULTATIONS

---

In 2019, field consultations were held in all 4 field offices to validate the VASyR quantitative findings. The consultations complement the data analysis and are summarized in the *Voices from the field* boxes.

Participants included sector leads in the field and key actors implementing activities targeting Syrian refugees.