



WETHOLOGY

Sampling

Sampling for the VASyR followed a two-stage cluster approach, keeping with the methodology of previous years. The UNHCR database of known Syrian refugees as of May 2021 served as the sample frame. Cases with missing addresses were excluded. 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.¹ 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 similar representativeness with other governorates, an additional two strata samples were considered for each, yielding 90 cluster selections for each. The governorate of Baalbek-El Hermel is made up of only two districts, and thus to ensure an adequate sample in that governorate, one additional cluster sample was considered.

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/household) under UNHCR databases. Using the Emergency Nutrition Assessment (ENA) software, villages were selected with a probability proportionate to size where villages with a larger concentration of refugees were more likely to be selected and 30 clusters/villages were selected with four replacement clusters per district.

In order to determine the sample size needed to generate results representative at 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/cluster selection were required, leading to a target of 5,115 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. In the final sample, 8,662 cases were targeted across all districts of which 5,035 households were visited.

Training and field work

Like the previous year, due to the COVID-19 pandemic, enumerator training took place remotely. Separate enumerator trainings were carried out online for each operational region (Bekaa, Mount Lebanon, the North,

and the South) covering the data collection tool, contextual background, methodology, and ethical considerations. Additionally, enumerators were required to attend a two-hour online COVID-19 training, provided by the Lebanese Red Cross, which covered key information about the virus, transmission, and precautionary methods. The bulk of the training was administered by UNHCR, WFP, and UNICEF staff. Training on the Washington Group Question Set of Functioning was provided by Humanity and Inclusion.

Data was collected and entered on electronic tablets by the enumerators during the interviews using KoBo toolbox software. The data was then sent to the UNHCR Refugee Assistance Information System (RAIS) platform.

Data collection took place between June 7 and July 7 2021 through face-to-face interviews at refugee homes by four partners in each region, as shown in the table below.

Table 1: Partners that conducted interviews for the VASyR

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

COVID-19 safety measures during data collection

With the support of the Lebanon Crisis Response Plan Health Working Group and WHO, the detailed guidelines, put in place in 2020, were updated to ensure the safety of enumerators and refugee families during the face-to-face data collection. Firstly, prior to the visit, households were screened over the phone to ensure that no member was exhibiting COVID-19 related symptoms and to inform respondents of the measures to be taken during the interview. UNHCR field offices and partners liaised closely with local authorities to inform them of the exercise and the measures taken to ensure access to specific areas. During the data collection activity, enumerators were provided with personal protective equipment such as masks and sanitizing equipment, which were also provided to refugees participating in the interviews. Enumerators were equipped with digital thermometers to measure body temperature of respondents prior to beginning the interview. Interviews took place with one household member at a safe social distance and in an outdoor or a well-ventilated area. If these conditions were not met or if any household member was showing COVID-19 related symptoms, the interview was called off.

¹ Using the Emergency Nutrition Assessment (ENA) Software.

Questionnaire

The 2021 VASyR questionnaire consisted of around 513 questions that collected data at the household and individual level including demographics, legal documentation, safety and security, shelter, WASH, energy, 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/88944>

Data quality assurance

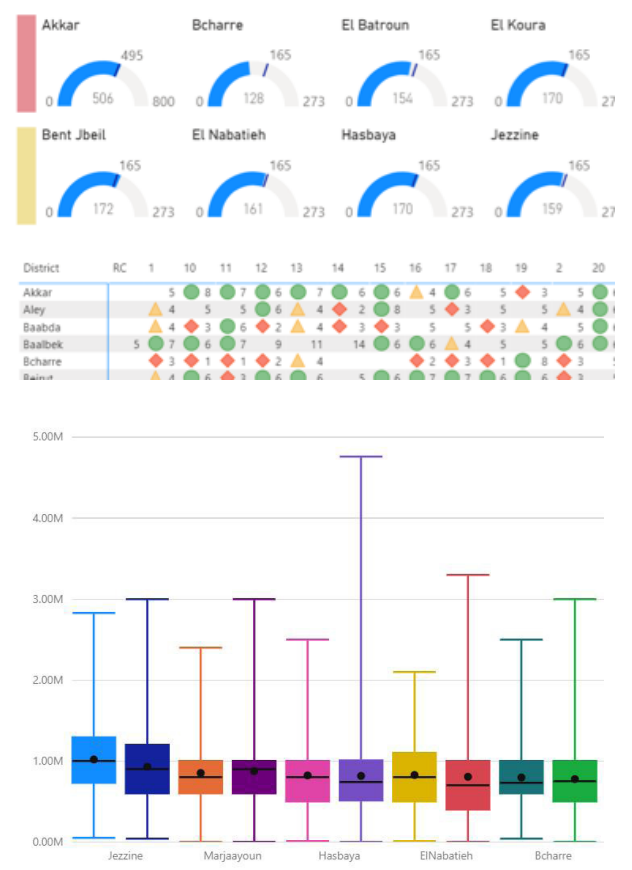
The following steps were taken to monitor the quality of collected data:

1. Using a harmonized check list, each VASyR core agency conducted frequent spot checks on each of the data collection teams across Lebanon.² Feedback was provided to enumerators directly after the interview was completed, and reports were drafted 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. Important feedback was shared with all enumerators through field coordinators via WhatsApp groups.
2. Agencies conducted follow up phone calls for randomly selected 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 daily.
5. A dynamic dashboard was created to monitor the progress of data collection in real-time. The dashboard included tools to monitor accuracy of data collected and to identify outliers during data collection.

² Refer to <http://ialebanon.unhcr.org/vasyr> for a detailed description of the spot checks procedure and tools used.

METHODOLOGY

Sample charts from the dashboard to monitor progress and identify outliers in real time



Data processing

Data weighting 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 the total sample frame of the district, N is the total national sample frame, n_s is the number of households visited in the district, and n is the total visited households.

The data was cleaned from 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 when deemed necessary. Data was analyzed using the IBM SPSS Statistics 20 software.

Limitations and constraints

As in any survey, limitations were expected. Key limitations of the VASyR were as follows:

1. In 2021, the data collection was in June-July as opposed to August-September in 2020 and May-June of previous years (2019 and before). The changes in timeline were due to the COVID-19 outbreak and finding the right time to conduct the survey. This may have had implications on indicators that concern behaviors with eventual seasonal variations.

2. The VASyR relies primarily on self-reported data which may give rise to bias. To minimize the impact of this bias, enumerators were trained in providing comprehensive informed consent to reassure confidentiality, purpose, risks, and benefits.

3. The VASyR sampling frame excluded Syrian refugees who have never approached UNHCR (unless within a targeted household). It is worth noting that this population is a consistent gap in data on Syrian refugees in Lebanon.

4. The VASyR questionnaire and respective indicators were subjected to adjustments and changes in order to ensure that the most accurate definition or calculation was being used. This 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, individual interviews are not carried out with each family member and obtaining accurate information on particularly sensitive topics is a challenge (i.e. child labor or harassment).