Methodological, contextual and ethical challenges in studying early marriage among Syrian refugees in Bekaa, Lebanon

Introduction

According to the United Nations High Commissioner for Refugees (UNHCR), there were more than 65 million forcibly displaced people globally in 2016. Among those, 20 million are refugees (UNHCR 2016). The vulnerabilities of refugee populations and their growing needs have generated extensive interest among researchers and policy makers. Research in refugee settings is vital to accurately describe their situation, or assess their needs in order to respond to such needs appropriately. Nevertheless, engaging in research with refugees entails challenges. Researchers must be able to generate knowledge that would be useful to protect and improve the lives of refugees while at the same time they must abide by high standards of academic rigor (Jacobsen & Landau, 2003). However, there are some instances in displacement and refugee settings where abiding by rigid scientific standards prove to be challenging. Despite such difficulties, researchers should strive to build the best evidence in such settings.

Refugee research employs various quantitative and qualitative methods. In either case, good research practices require that researchers reveal and explain their methods (King, 1994; Ragin, 1994). Being transparent about the research approach permits judgement of the validity of the findings, makes assessment of the quality and limitation of the data possible, allows replicating the study when needed and comparing results with studies in similar settings.

This paper draws on a cross-sectional survey that was conducted in the summer of 2016 (July to September) in Bekaa, Lebanon. The study explores the prevalence and determinants of early marriage among 9-24 Syrian refugee girls/women in three areas in Bekaa, Lebanon; Bar Elias, Qab Elias and Marj. We discuss three sets of key challenges encountered while conducting research with Syrian refugees in Lebanon: methodological challenges, contextual challenges and ethical challenges. We also discuss the approaches we adopted to address such concerns and the implications of encountered challenges on our study.
Background:

A brief description of the situation of Syrian refugees in Lebanon follows.

The Syrian war, which started in 2011, has led to the displacement of more than 5 million refugees (UNHCR 2017). Lebanon hosts more than one million Syrian refugees (UNHCR 2017) and has the largest number of refugees in relation to its own population size. Lebanon is not a signatory of the 1951 convention in relation to the status of refugees (Stevens 2014), and did not establish official camps for Syrian refugees as other neighboring countries such as Jordan and Turkey did (Knudsen 2016). The majority of Syrian refugees in Lebanon live in informal tented settlements (ITS) often run by local government entities and UN agencies or in rented accommodation that is often shared with other refugee families (UNHCR 2014).

Many Syrian refugees in Lebanon face legal and residency difficulties, which have implications on their access to services and on their relationship with the host community. Only 21% of surveyed households in a vulnerability assessment survey in 2015 reported that all of their members had legal residency permit (UNHCR et al 2015). Additionally, the employment laws for Syrian refugees in Lebanon are strict. Although Syrians are allowed to work in Lebanon, their work opportunities are limited to "third sector jobs", such as agriculture, domestic services and construction (Errighi & Griesse 2016). Syrians working in Lebanon do not have full coverage under the National Social Security Fund and require a work permit to be employed (ILO 2014). Work permits are granted on an annual basis according to specific conditions, necessitating having a Lebanese sponsor, and ensuring that there is no competition between the applicant and Lebanese workers. The work permit of any foreigner can be cancelled when there are incorrect documents, or whenever the interest of the Lebanese workforce is considered to be threatened (ILO 2014). Given the difficulties encountered with renewing residency permits and the challenges in obtaining work permit, many Syrians in Lebanon work illegally and for minimum wages (ILO 2013). Moreover, the displacement of Syrian refugees to mostly poor areas in Lebanon has resulted in competition over employment (ILO 2013), and limited resources (UNHCR 2015).
In terms of the focus of the study on early marriage, which was a common practice in Syria before the war, anecdotal evidence from neighboring displacement countries indicated that the practice was increasing (UNICEF 2014; Save the Children 2014). Yet, existing studies on early marriage among Syrian refugees in Lebanon were limited, and not representative of the Syrian population residing in Lebanon (UNFPA et al 2014; USJ 2014). One survey was restricted to refugees registered with UNHCR (UNFPA et al 2014), another was not explicit about its sampling strategy (USJ 2014), and the remaining studies, though informative, were qualitative and not generalizable (Mourtada et al 2017; Cherri et al, 2017). The study at hand is a first attempt to provide an estimate of early marriage among Syrian refugees in Lebanon based on a representative sample of Syrian refugees residing in selected areas (Abdulrahim et al 2016).

Our study took place in Bekaa, a region that hosts 35% of the overall Syrian refugee population in Lebanon (UNHCR 2017). We carried out a community based cross-sectional survey, which collected data from a sample of Syrian refugee girls/women in three regions in West Bekaa that have a large number of Syrian refugees: Bar Elias (31,510), Qab Elias (22,539) and Marj (14,252). Two out of these three regions where our survey took place, Bar Elias and Qab Elias, were considered among the most vulnerable localities in Bekaa (UNHCR 2015). Eligible selected households included a Syrian family who came to Lebanon since the Syrian conflict began in 2011 and that had at least one girl/woman aged 9-24 years old living in the same household. Data on girls under 18 years old were collected mostly from mothers or any other adult family member in the same household.

**Challenges**

The challenges we faced in our study include methodological related challenges, contextual related challenges and ethical related challenges. We will proceed with describing each of those challenges and how we dealt with them.

**Methodological related challenges**

Sampling is one of the main challenges faced by researchers wanting to carry out surveys in refugee settings. Generally, random sampling is the gold standard of survey sampling
(Bostoen et al 2007). However, such sampling technique can be achieved only in settings where the target communities are often stable, transport and logistics are not problematic, and most importantly, sampling frames of such stable populations are comprehensive and more readily available, allowing researchers to select a representative random sample (Bostoen et al 2007). Regrettably, such conditions are often not met in humanitarian settings where sampling frames are lacking. Refugees are generally scattered and households are often hard to reach. In such situations, and with the lack of agreement on the most appropriate methodologies to use for sampling (Suleiman-Hill and Thompson 2011), researchers usually adopt the best possible sampling strategies and justify the choice of their approaches.

Since our study is descriptive and only sought to provide estimates on prevalence of early marriage and its determinants, we assumed that 2000 households would provide a large enough sample for estimation (e.g. comparison between married versus unmarried girls/women inside versus outside the settlements). Our aim was to sample 2000 households drawn equally from inside and outside the settlements and across the three selected areas; 333 HHs from inside the ITS and 333 HHs from outside the ITS in each of the three areas.

Generally, probability techniques are used to ensure representativeness of and generalizability to the population from which refugees are drawn from (Bloch 2007). These can be simple random, stratified or cluster random samples but a key aspect of probability sampling is that there is a complete and accurate frame of sampling units (Bloch 2007). In our study, and since Syrian refugees live inside and outside ITS, and those are two different settings, we had to adopt two sampling strategies to ensure representativeness of Syrian refugees residing in each of those settings.

**Refugees inside settlements:**

Camps around the world differ in terms of their accessibility, with some that are easy to access and others that are not (Vogler 2007). There are different control mechanisms that rule camps or settlements imposed by governments, non-governmental organizations (NGOs) and
sometimes by refugees themselves. Such bureaucratic control mechanisms can shape the fieldwork process (Vogler 2007). In the case Lebanon, the government’s approach to the Syrian refugee crisis, which resulted in not establishing official camps that would have been strictly run by the government, has led to the rise of local Syrian men known as the “Shawishs” within the ITS (Ghaddar 2017). The absence of direct governmental control over the ITS permitted informal security actors/settlement leaders, such as the “Shawishs”, to prosper in Lebanon’s informal refugee camps. Additionally, the control granted to those “Shawishs” by the local governments resulted in them exploiting camp residents and bullying international aid organizations. This had implications on our research strategy and sampling refugees inside the ITS. Establishing contacts with such gatekeepers, or the “Shawishs” was a crucial and inevitable step to proceed with our research activities within the ITS. To simplify access to the ITS, we obtained official letters from the municipalities in each of the three areas that grant an official permission for data collectors to access those ITS.

For refugees living inside the settlements, it was possible to undertake a probability sampling strategy since we were able to obtain a list of households (HHs)/tents in all the camps/settlements from the municipalities in each of the three areas. Those lists constituted the sampling frame for selecting HHs inside the camps/settlements. We assumed, and based on previous experience in those refugee settings, that people living in the same camp are more likely to share the same socio-economic characteristics, are more likely to come from the same governorate in Syria, and are more likely to receive the same type of aid or support from NGOs and UN agencies operating in those areas. Therefore, to ensure an appropriate representation of each camp, we adopted stratified sampling in each area. This technique ensures that the sample profile is comparable to the population in important elements (Lynn 2016).

Our aim was to collect data from 333 tents/HHs from inside all of the camps/ITS in each area. There was a total of 53 camps in Bar Elias, 42 camps in Qab Elias and 13 camps in Marj. We sampled 400 tents/HHs from inside all of the camps/settlements from each area to adjust for non-response, and non-eligible tents, using an internal stratified systematic sampling (ISS).
technique, which produced a list of tents/HHs with predetermined assigned numbers in each camp. ISS involves ranking the elements following some ordering principle (which was the size of the camp/settlement in our case that was assessed based on the number of tents/HHs in the camp) and then applying systematic sampling. The sampling interval was calculated based on the total size of the tents/HHs in those lists and the desired sample size. The first tent/HH was selected randomly, followed by a systematic selection of the rest of the tents/HHs (based on the sampling interval). The number of selected households in each camp/settlement was proportional to the size of the camp/settlement.

Data collectors had to check the eligibility of each of the sampled tents/HHs before proceeding with the interview. Ideally, each tent in the ITS should have had a unique number, assigned by the municipality, written on the tent close to the main entrance, which should have been easily identified by data collectors. Moreover, the total number of tents in the ITS should have matched the one in the latest ITS’ lists obtained from the municipalities. However, realities on the ground were different.

Tents within the ITS were numbered by municipalities as well as UNHCR, and sometimes more than once. Consequently, each tent had at least two different numbers in different colors. Sometimes, those numbers were unclear and many times tents were covered by an additional protective cloth, which made it difficult for data collectors to verify the number of a tent. Data collectors were able to overcome this challenge by using the help of the ITS leader or the "Shawish", as known locally, who had a complete list of the tents and was able to help data collectors with identifying the sampled tents.

Due to the dynamic nature in refugee settings, there were constant changes within those ITS with new tents being installed and old tents deserted resulting in a number of ineligible tents and the omission of many tents from the obtained sampling frame. These dynamics were highlighted in a report assessing the ITS in Lebanon. The report (UNHCR and REACH 2014) described the challenges that are faced by humanitarian actors when coordinating effective interventions in such ITS. For instance, new ITS may be formed but only exist for short periods
of time, or ITS that have existed for comparatively long periods of time may face sudden eviction by landlords or municipalities, and disagreements within refugees living in those ITS may divide communities and drive some members to relocate. Despite having access to the most updated ITS' lists from the municipalities that were closest to realities on the ground, we were unable to deal with all the dynamic changes described above, which had implications on our sample. We achieved only 65% of our intended sample drawn from inside the ITS. The final sample was weighted and adjusted for for non-response and non-eligible cases (Yansaneh 2013).

Although we faced some inevitable challenges, conducting formative research about the organization and the dynamics within the ITS enabled us to overcome some of the potential hindrances that would have occurred had the research team entered those ITS without prior knowledge or preparations. Instead, the team established good relations with the municipalities in each of those areas months prior to conducting the survey. Maintaining good relations with the municipalities helped us acquire the latest ITS lists, and obtain official approval letters to access each of the ITS. Obtaining such letters, not only helped data collectors gain easy access to the camps but it also helped them deal with some of the challenges encountered on the ground, such as the difficulties with identifying the sampled tents or even identifying which tents were installed recently, as some tents were numbered randomly by illegal refugees. The official letters, issued by the municipalities, encouraged “the Shawishs” who had good relations with most municipalities to assist data collectors with dealing with some of the challenges they encountered. The use of gatekeepers in research with refugees was also described by Bloch when discussing methodological considerations and guidelines when carrying out a survey with refugees (Bloch 1999). Bloch describes how using gatekeepers can facilitate research in refugee setting. Bloch argues that using gatekeepers is not normally a feature of survey research and requires a long process to identify the appropriate gatekeepers and discuss research activities with them. However, it is necessary to identify local channels in such settings to facilitate research activities (Bloch 1999).

Refugees outside the settlements
Generally, aid agencies and programs for refugees focus on refugees residing into camps or on refugees registered with UNHCR disregarding the large number of self-settled refugees who mostly settle in urban areas with the host communities (Bakewell 2008; Schenker et al 2014). Ninety percent of the nearly 5 million Syrian refugees live in urban settings in host countries. Yet, the literature on urban refugees is thin compared to that on camp refugees (Kelberer 2016). This is understandable as it is particularly challenging to work with such refugees; they are often dispersed in several geographical areas, they have no registration with UNHCR (Spiegel 2010), and it is not as easy to identify them as those living within the camps (Schenker et al 2014).

Spiegel (2010) highlighted the difficulties in conducting population-based surveys in non-camp settings, and discussed that some of the employed sampling methods still require validation. Sampling from outside the ITS constituted a particular challenge for us with the lack of a defined sampling frame, a common problem in refugee settings (Tilbury 2006; Reichel and Morals 2017). Even if we wanted to use UNHCR records as a sampling frame, which are difficult to access, many refugees in Lebanon were not registered with UNHCR, which would have resulted in an unrepresentative sample. Given this reality, we wanted to obtain a sample as representative as possible of the Syrian refugee population living outside the ITS in each of the three areas.

We recognized that with the absence of clear sampling frames in urban settings, neighborhoods or blocks often provide a suitable aggregation of households and can be used when constructing a sampling frame (Collins 2004). Consequently, and for each area, we constructed such sampling frames using Google Earth and the help of locals who knew those areas well. In each area, we purposively identified ten clusters and created maps for each of those clusters. Each cluster comprised a number of neighboring streets with the highest concentration or density of Syrian refugees. Although it was hard to assess the exact number of refugees residing in each cluster, we attempted to construct clusters that would comprise at least 50 HHs with Syrian refugee families, to ensure that we will find at least 33 eligible HHs who agree to participate, which was the required sample size in each cluster to achieve a total of 333 HHs outside the ITS. Households were selected systematically from those clusters with
the first HH selected randomly from one of the cluster's four corners. Data collectors would knock on the door of those HHs one by one and when they identify an eligible Syrian HH agreeing to participate, they would skip five houses in the same building to decrease the likelihood of selecting members of the same family. The data collectors would stop collecting data from a cluster when they reached 33 HHs or when there were no more eligible HHs in that cluster.

The process of constructing the clusters, identifying their four boundaries and preparing the corresponding maps was challenging with the absence of clear street names or landmarks both on the ground and on Google Maps. One of the main limitations of using Google Earth is that it depends on the accuracy of the database (Lefer et al 2018). In our case, the available downloaded Google maps were not up to date. Therefore, to simplify the sampling process for data collectors, the study coordinator first documented the street names (as known locally) on the maps using the help of two locals in each area who knew the areas well and who knew where Syrian refugees resided. Secondly, the study coordinator visited each of those areas and identified the four borders for each cluster based on the clearest landmarks available (e.g. schools, shops, or hospitals), and marked them on Google Maps as well. A similar strategy of identifying clusters to simplify the sampling process for data collectors was adopted when conducting a population-based mortality survey in Iraq (Galway et al 2012).

Suleiman and Thompson, (2011) recognized the difficulties in obtaining a truly representative random sample of the refugee population they were interested in, due to the cautiousness and invisibility of refugees, therefore, they attempted to sample a cross section of participants that is as wide as possible. Similarly, in her paper on overcoming complexities when sampling refugees from an urban environment, Vearey argues that conducting surveys in complex settings should strive to be "representative enough" of the complexities of the urban environment (Vearey 2012). Although this is what we aimed to achieve in our study, we believe that our sampling strategy, which focused on refugees who live in clusters in residential areas, may have missed less visible refugees who live in non-residential or less-densely populated area such as in garages or warehouses.
Contextual related challenges

Animosity between refugees and the host population constituted another challenge for us when sampling refugees residing outside the ITS in urban areas with the host communities. Urban refugees are not only less accessible than those living inside the ITS, they are not as readily distinguishable (Schenker et al 2014), resulting in greater risk that the Syrian data collectors are exposed to. Given the reported resentment toward Syrian refugees from host communities due to many factors, including perceived competition in the labor market, many Syrian refugee feel apprehensive about being in direct contact with the Lebanese community. This reality had implications on our strategy of recruiting and pairing our data collectors. All of the data collectors working inside the ITS were recruited from the Syrian refugee community as they had easier access to the ITS that had only Syrian refugees. However, since sampling outside the ITS involved knocking on the doors of random houses and it involved direct interaction with the Lebanese community, Syrian data collectors working outside the ITS were paired up with Lebanese data collectors who were less intimidated and faced less threat by approaching random households looking for Syrian families. This strategy for choice of data collectors may have had ethical implications, which we discuss in the next section.

In our study and on some occasions, mothers, mainly those coming from rural areas, were unable to accurately identify the age of their daughters who were less than 18 years old. Thus, data collector asked for the help of other family members whenever possible or asked to review other documents such as UNHCR registration documents to corroborate the age of the girls. Lack of documentation is often a problem for the urban internally displaced. Official documentation is frequently lost or destroyed fleeing emergency situations, or during subsequent displacements (Fielden 2008). In Sri Lanka it is estimated that more than 70 percent of survivors of the tsunami of December 2004 lost their documentation (Fielden 2008).

The conflict and forced displacement caused many Syrian families to lose their identification documents. A research study by the Norwegian Refugee Council that interviewed 580 Syrian refugee households in Lebanon, Jordan and Iraq concluded that 70% of refugees lack basic
identity documents (NRC 2017). A number of reports discussed the implication of losing ID documents on Syrian refugees in their host countries as well as in Syria, but there was no discussion of its implication on research studies, especially when research involves illiterate refugees.

**Ethical related challenges**

Participation in research influences the life of participants. It should be a positive and beneficial experience that involves minimal risks (Block et al 2012). Therefore, researchers must carefully consider the ethical implications of their research activities. Furthermore, when researchers fail to follow the highest ethical standards, they will jeopardize losing the trust of the refugee community, which can result in either failing to obtain honest responses to questions or failing to engage participants altogether (Ellis et al 2007). Considering and reflecting upon ethical considerations is a process that should start once a research project is conceived and end with the presentation of its findings publicly (Block et al 2012).

The dimensions of research ethics are complex and one should differentiate between 'procedural ethics', which involves gaining formal consent through ethics committees and 'ethics in practice', which involves recognizing and responding to context related circumstances than can arise through the course of research (Guillemin and Gillam 2004). In addition to applying to our institutional review board (IRB) for ethical approval, which ensured that the implementation of our research is aligned with the core ethical principles, we also reflected on the ethical implications of our research activities and modified them constantly to minimize the potential risks to the studied refugee population.

The recommendations of our institutional review board influenced the fieldwork process and had implications on our final sample size. Since the survey was on early marriage, we were only interested to collect data on 9-24 year old girls/women. Mothers or other adult family members were required to answer on behalf of girls who were under 18 years old, as commended by our institutional review board, which also required the presence of those girls to give their assent to be part of the survey. Since many young girls, particularly those who lived in the ITS, worked in agriculture in the day time, they were often not present at the time of the interview and it was often difficult for data collectors to return to those camps due to
their distant location in addition to limited time and resources. Although we tried to avoid
such a problem by inquiring from the Shawishs about the best time of the day where most
girls would be available at the camps, girls were not present on many occasions and data
collectors were unable to conduct the survey with the adult family member without the girls’
assent. We believe that this might have created a bias because families with working
daughters are more likely to marry them early due to poverty and such families may have
been under-represented in our sample due to the absence of the girls at the time of the
interview. Difficulties with accessing refugees who work was also highlighted by Elena and Liu
(1986) who explained that interviewing refugees who work was not possible unless data
collectors were willing to conduct the interviews at 11 in the evening or 7 in the morning.

It has been reported that using data collectors from the same community can have ethical
implications as there is a high chance that respondents may lose their privacy, and
consequently may be less forthcoming to members of their own community (JIPS 2014;
Birman 2005), especially about sensitive issues like income levels, well-being, and
employment experiences (Tilbury 2016). To present a positive image, there is a high
probability of having what is known as a ‘social desirability response’ effect as a result of using
insider data collectors collecting data face-to-face (Babbie 2005). Although there was a
chance that some of the refugees may have hesitated about disclosing personal information
or provided biased answers to appear good in front of their community, we believe that the
risk was minimal as our survey did not include sensitive questions. Furthermore, early
marriage is a practice that is accepted culturally and practiced even by some of the data
collectors themselves, so it is unlikely that asking about such practice might have caused
embarrassment to respondents. Jacobson and Landau (2013) suggest that it may be best to
hire data collectors who are not from the same refugee community but speak the same
language. In such case, the other alternative to hiring Syrian data collectors from the same
community, would have been hiring only Lebanese data collectors, but we believe that this
would have triggered such biases, especially with the diminished trust Syrian refugees have
in Lebanese service providers and the fear of being stigmatized (Mourtada et al, 2017). For
instance, some of the information refugees revealed, especially about their current work
status, could be used against them, especially in a country that does not allow Syrian refugees
to work freely. We believe that pairing Lebanese and Syrian data collectors to simplify access to HHs outside the ITS also contributed to reducing such biases, as having a Syrian data collector in addition to the Lebanese one may have made interviewees feel less threatened when disclosing information on matters that might have been sensitive to them such as their work status. This issue was highlighted by Jacobsen & Landau when discussing the challenges when conducting social research in refugee settings. The authors also highlighted the additional positive aspects for hiring refugees, such as providing short-term employment and training for marginalized communities and this is valuable particularly in such challenging settings (Jacobsen & Landau 2003).

**Conclusion and recommendations**

Despite all of the described challenges in this paper, our final sample size was 1513 HHs or families out of the required 2000 HHs. We were able to sample 646 households from inside the ITS and 867 households from outside the ITS, from which we collected data on 2400 9-24 year old girls/women.

Conducting research studies in refugee settings includes a number of difficulties. In this paper, we describe the different challenges we encountered, which we categorized into three sets of challenges: methodological related challenges which involved mainly difficulties with sampling in such dynamic settings and the lack of clear or updated sampling frames, contextual related challenges, which included the animosity between refugees and the host population and its implication on the choice of our fieldworkers, and refugees’ loss of identification documents, and ethical related challenges in which we describe the implications of the recommendations of our institutional review board on our fieldwork and the ethical implications of the choice of data collectors.

The above-described difficulties necessitated coming up with alternative strategies in an attempt to obtain a representative sample and collect valid data but this proved to be a difficult process in such dynamic settings. Spending time in the field, becoming familiar with the research setting and building rapport with the main gatekeepers contributed to our ability to deal with some of the challenges we encountered in an effective and timely manner.
However, we faced unexpected conditions that had implications on our final sample and the results of the study.

The aim of this paper was not only to describe the challenges we encountered while conducting this study but also to be transparent about the choice of our research methods, its implications on our results, and how we aimed to adopt rigorous research practices despite all the challenges to abide with good research practice. We believe that researchers working in dynamic refugee settings need to adopt a ‘continuous process of critical scrutiny and interpretation’ with respect to themselves and the research situation as described by Guillemin and Gillam, (2004), when recommending how to achieve ethics in practice. Jacobson and Landau (2003) argue that “from an academic and advocacy perspective, the benefits of rigorous methods in refugee research outweigh the costs” because a research that is done well can be a powerful tool for policy makers (Tilbury 2016). This paper is a contribution towards encouraging development of strategies for addressing methodological problems in refugee settings while maintaining rigorous methods in research to the best degree possible, and an accounting for problems that could not be avoided and their potential implications on our study.
References


36. Tilbury, F. (2006, December). Everything is excellent” : methodological issues in studying refugee settlement. In TASA Conference, University of Western Australia and Murdoch University (pp. 4-7).


