

Conducting Household Surveys on Reproductive Health in Urban Settings: Lessons From Karachi, Pakistan

Mir Baz Khan

Aga Khan University

Sidrah Nausheen

Aga Khan University

Imtiaz Hussain

Aga Khan University

Kristy Hackett

Harvard University HSPH: Harvard University T H Chan School of Public Health

Zahra Kaneez

Aga Khan University

Khalid Feroze

Aga Khan University

David Canning

Harvard University HSPH: Harvard University T H Chan School of Public Health

Iqbal Shah

Harvard University HSPH: Harvard University T H Chan School of Public Health

Sajid Bashir Soofi (✉ sajid.soofi@aku.edu)

Aga Khan University <https://orcid.org/0000-0003-4192-8406>

Research article

Keywords: Household surveys, challenges, loss-to-follow up, implementation, sampling frame

Posted Date: October 22nd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-94413/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on February 18th, 2021. See the published version at <https://doi.org/10.1186/s12874-021-01216-x>.

Abstract

Background: Data collection is the most critical stage in any population health study and correctly implementing fieldwork enhances the quality of collected information. However, even the most carefully planned large-scale household surveys can encounter many context-specific issues. This paper reflected on our research team's recent experiences of conducting surveys for a quasi-experimental evaluation of a reproductive health program in urban areas of Karachi, Pakistan.

Methods: The study followed a three-stage random sampling design.

Result: This paper has described the issues that were encountered around technical problems related to geographical information system (GIS) usage and computer assisted personal interviews (CAPI), household listing, interviewing respondents on sensitive topics and their expectations, and other field related concerns such as ensuring privacy etc. during the survey.

Conclusion: The papers has also underscored on lessons learned from this process and presented some potential solutions for conducting future household surveys in similar urban environments.

Background

The United Nations (UN) has recognized universal access to reproductive healthcare as a priority for global health. One of the Sustainable Development Goals (SDGs) is to ensure universal access to sexual and reproductive health (SRH) services, including family planning, with a particular focus on the use of modern contraceptive methods (indicator 3.7.1). Women's reproductive health services in low and middle income countries (LMICs) have not fully met the human rights standards [1]. Family planning in several LMICs is fragile due to constrained existing policies and dependence on donor funding. The unmet need for contraception is the highest among the poorest couples [2]. Access to family planning in such settings can help in meeting women's and couples' reproductive goals and improve maternal and child health [2].

For decades, family planning has been a contentious issue in Pakistan with the highest total fertility rate (3.8 children per woman) in South Asia after Afghanistan. Despite 98-99% of universal knowledge about family planning, the overall contraceptive prevalence rate (CPR) is 34% in Pakistan [3, 4]. The major obstacles to contraceptive use are lack of motivation to avoid pregnancy, awareness and knowledge of contraception, perception of the husband's preferences and attitudes, health concerns, and perceived access to services [5]. Moreover, it is also difficult to talk to women/families openly about sexual preferences in Pakistani context. Several issues impede gathering reliable data from marginalized groups [6, 7] that includes the sensitive nature of information gathered in data collection related to Sexual and Reproductive Health (SRH), specifically some past bad experiences related to those [8]. The influence of strong religious identity affects women's willingness to discuss contraception with others. Likewise, strong institutionalized religious doctrines in combination with cultural beliefs in urban communities, and men's role as the overall head of household makes women less likely to share their experiences and practices related to family planning [9]. Data quality on the uptake of family planning therefore remains a

challenge to address maternal mortality and to reduce reproductive risks by discussing unwanted pregnancies in (LMICs) [10].

In any research process, data collection is the first stage of gathering information with respect to respondents' needs, goals, and attitudes toward the problem area. In the process of obtaining information from respondents, researchers may encounter unanticipated challenges. Enumerators working in LMICs often encounter respondents who are reluctant to participate in the interview and that makes it difficult to build rapport [11]. A number of factors may influence a respondent's willingness to participate. For example, the response of the interviewee depends on the types of clothes worn (formal or casual) during the field visit. Enumerators may also face issues persuading a respondent in a joint family system to participate and convincing some respondents to discuss sensitive topics such as sexuality in health research [12]. The respondent literacy in understanding the questions consequently impacts the process of data collection. If respondents perceive a risk of breaches in confidentiality, they may be reluctant to reveal personal information to the enumerator during the interview [13, 14]. Similarly, social and cultural taboos attached with SRH in some religious and cultural groups, do not permit women to discuss health issues particularly sexual activities with someone other than closed family and friends. A study on women from constituencies of highly religious groups of urban Zimbabwe reported women's refusal to participate in interviews related to SRH because of their religious doctrines in the context of Pakistan, strong societal taboos linked with moral values, husband and religious opposition restrict women for open discussion of SRH services in communities of perceived religious origin [15, 16]. There are also cultural myths of not sharing personal information with people from outside who you don't trust [17, 18]. Similarly, studies from LMICs have shown that lack of trust is another challenge, due to previous research that provided no direct benefit or feedback to respondents [19, 20]. Moreover, respondents' busy schedules and lengthy questionnaires may lead to refusal or incomplete surveys. Understanding local norms and developing culturally appropriate strategies to build trust within communities significantly reduces refusals [19, 20]. The data collection issues experienced in South Africa, reported that urban settlements do not have space for interviews, therefore privacy was often a problem [21]. Furthermore, densely populated urban settings in LMICs also pose logistical as well as technical difficulties for researches such as identification of eligible respondents, congestion of infrastructure, poor transportation, and lack of political and administrative willingness to support independent researchers) [18].

Evaluation of the Willows Reproductive Health Program

The present study was a part of large Willows Reproductive Health Program [22]. The project included evaluation of a community-based reproductive health program providing reproductive health information and education to women of reproductive age in Karachi, Pakistan. The program included both prospective and retrospective assessment of separate intervention and control sites. These sites from Karachi included communities with adequate provision of family planning services that typically house women of low socio-economic status. The initial stage of the program involved a registration period in which field workers enrolled all eligible women living in the defined intervention area. The field educators

then carried out home visits informing women about the benefits of family planning, provided information on the range of available contraceptive methods, and referred women to locally-based healthcare services, when required, and coordinated with providers to ensure the delivery of high quality family planning services. The program also applied an algorithm that determined women most in need of family planning counseling and, the field educators visited them to provide education, information, and referrals. The non-users with high risk of unintended pregnancy, ineffective users of traditional methods, and unsatisfied users of modern methods were given highest priority for home visits. The retrospective assessment was carried out between 2013 and 2015 and prospective assessment from 2017 to 2020 in urban areas of Karachi, Pakistan. This paper reflects our research team's recent experiences in carrying out these assessments for a quasi-experimental evaluation of the Willows reproductive health program. We aim to describe the issues encountered and lessons learned from this process of carrying out assessments in urban settlements particularly around SRH topics and recommend some potential solutions for future surveys in similar urban dwellings.

Methodology

Parent Study design

Lessons learned reflect our team's experiences in undertaking a larger impact evaluation, which consisted of two components. Component 1: was a baseline prospective assessment conducted prior to implementation of a reproductive health program, and component 2: was a cross-sectional retrospective assessment in a site where the program has ended.

Study sites and population

The program included both prospective and retrospective assessment of separate intervention and control sites in urban settlements of Karachi, Pakistan. Jamshed Town as an intervention site and Yousuf Goth as a control site for prospective assessment; and Korangi Town as an intervention site and PIB colony and Dalmia/Shanti as control sites for retrospective surveys were selected. The program was implemented from April 2013 through September 2015 in Korangi Town (for retrospective assessment), and from June 2017 to March 2020 in Jamshed Town (for prospective assessment). The control sites were selected to match the intervention areas in terms of socio-demographic characteristics, such as type of area (urban or peri-urban), population size, and ethnicity or language. Women were eligible if they were aged 16-44 years for prospective assessment and 16-49 years for retrospective assessment, married, usual members of the household, and spoke at least one of the four commonly used languages (Urdu, Pushto, Sindhi, or English).

Jamshed town is the most populated (0.73 million) municipality in the East district of Karachi with majority of Muslim population. The town is populated by diversified ethnic groups including Muhajirs, Punjabis, Sindhis, Kashmiris, Seraikis, Pakhtuns, Balochis, Memons, Bohras, Ismailis, etc. and is also home to minority groups such as Christians, Parsis, and Hindus. Yousuf Goth with nearly one million inhabitants is a peri-urban setting in the neighborhood of district Malir with majority of Muslim

population of various ethnic groups. Jamshed town and Yousuf Goth share similar socio-demographic characteristics and consist of large slum areas with some population belonging to the upper-middle class. Both areas have dynamic working-class population and have public/private schools, malls and shopping plazas. The religious festivals, variability in weather pattern influence seasonal migration within and between districts and provinces. For example, families of slum areas move to interior Sindh districts to harvest crops from April to July each year. Some families live temporarily in rented houses in selected settings and frequently shift between nearby communities.

On the other hand, Korangi town with the population size of 2.4 million is a peri-urban neighborhood in the East of Karachi with multi-ethnic population including Muhajirs, Sindhi, Balochi, Pushtuns, and Gilgiti. This is an industrial area and is home to families from Afghanistan and Bangladesh who migrated for the purpose of employment in the garment and leather factories. Population here belong to low and middle class. PIB and Dalmia/Shanti Nagar with the population size of one million and 0.1 million, respectively, are the peri-urban areas located in Gulshan town with majority of Muslim population. In both areas, children's enrollment in schools is low and they work as helpers in the factories. Some women work in garment factories, some work as housekeepers, but many have low-paying private school teaching jobs. Due to ethnic diversity, the seasonal mobility varies in each ethnicity.

Sampling strategy and sample size

Both prospective and retrospective surveys followed a three-stage random sampling design. Initially, GIS technology was used to construct the sampling frame with distinct area mapping and cluster demarcation of the intervention and control sites. GIS technology was used because Pakistan lack reliable and updated statistics related to structure and household number in union councils of urban slums. Secondly, it is cost-effective method to in-person visits required to validate cluster boundaries and searching for a reliable statistics and unnecessary delays for developing sampling frames. Based on the geographical demarcation by GIS, 283 clusters were formed in intervention sites and 200 clusters in control site for the prospective assessment, whereas, 548 clusters in the intervention site and 160 in the control sites for the retrospective assessment. Each cluster was created on cadastral scale consisted of 60-100 structures. The second stage involved random selection of clusters in the intervention and comparison sites. A total of 105 clusters from an intervention and 100 from control site for the prospective assessment, and 110 clusters in each site for the retrospective assessment were randomly selected for inclusion in the study.

The second stage involved a complete household listing in selected clusters and random sampling of eligible women. The data management unit of Aga Khan University (AKU) developed an android application program for the household listing activity. All households were included in the listing, and the questionnaire sought to determine which had any eligible women who were between the ages of 16-44 years for the prospective component and 16-49 for the retrospective component. Field teams also collected pertinent details on the household location, including GPS coordinates, addresses, written directions, and the name of the household head. The application was tested multiple times to fix bugs

and queries prior to household listing. The household listing enrolled 8,179 households in the intervention and 6,406 households in the control sites for prospective assessment, and 9,010 in the intervention and 8,182 in control sites in retrospective assessment to generate a sampling frame for selection of households (secondary sampling units).

The final stage involved random selection of women from these sites. A sample size of 1836 (~2000) women from each site was required to the retrospective and prospective survey. The study team used household listing and identified 2,019 eligible women in intervention and 2,147 eligible women in control sites for prospective survey; and 2,750 eligible women each from intervention and control sites for the retrospective assessments using a computerized process. The calculated sample size was powered to test even at least a 5% (percentage point) difference in critical value of CPR in intervention sites compared to control sites at 0.05 significance with 90% probability of exceeding the critical value on a two-sided test. This process was carried out in STATA using a uniform [0, 1] random number generator with a fixed seed. Women were ranked by the number drawn. The lowest random draws from eligible women in the household were selected to participate in the survey. The randomly selected women were uploaded to the CommCare application for interview by the enumerators assigned to each cluster. The CommCare application randomly selected one if there were more than one eligible woman in a selected household.

Survey questionnaires (additional file 1 & 2)

We finalized the questionnaire for the retrospective and prospective assessments using country-specific standard data collection instruments such as demographics and health surveys. The survey questionnaire covered a range of topics on women's reproductive health, including marital status, contraceptive knowledge and use, childbearing, and abortion and access to safe abortion services. It sought insights on issues related to implementation, uptake, and continuation of a modern method. The assessments used to describe month-by-month history of certain key events i.e. births, pregnancies, termination of pregnancies, and family planning use of the respondents for the five years calendar period (January 2013 to 2018) preceding the date of interview. It also captured the information on change in contraceptive methods, and reasons of discontinuation [4]. The survey questionnaire, originally devised in English, was translated into Urdu (the national language) and was then back translated into English. We used digital data collection process which is useful for on-time data entry and checking, however requires resources such accessories, power bank, and internet connectivity for uploading data from fields. We conducted two pilot-testing exercises prior to data collection and modified the questionnaire based on feedbacks. In the initial week, enumerator was asked to complete one questionnaire a day, which was increased to two completed questionnaires a day in the second week, and after wards three completed questionnaires a day for the rest of the survey period.

Data quality assurance

The study team implemented a number of strategies to ensure quality fieldwork. These included regular quality checks in CommCare application to identify inconsistencies and outliers in the data. The

CommCare is a software developed by Dimagi that allows non-programmers to design and use mobile application for android devices. This is used across multiple sectors with the need to collect data digitally. The application had built-in quality checks to minimize errors and ensure proper skip patterns. The enumerators were trained to sync completed forms for review by the team leader. This means that completed forms on the enumerators' tablet were uploaded to a central server so that the team leader could access them on her tablet before leaving the field so as to clarify any mistakes. The syncing process was done on completion of each form. Each team leader reviewed the collected information on her tablet during and after the field visits. Team leaders also visited randomly selected 5% of the households to verify collected information, however, no such discrepancies were found. We hired all female enumerators and team leaders based upon the local norms where a female is allowed to enter the household and talk to woman. We developed a protocol for re-visiting cases with major errors such as when an enumerator skipped collecting information on pregnancy even when the respondent had a birth in the last 30 months. The data management unit documented the data cleaning notes and ran rounds of STATA cleaning codes to fix case specific errors on discussions with field manager, then re-ran to confirm all issues were resolved. In addition, enumerator refresher trainings were conducted weekly to discuss issues with data and fieldwork and to discuss updated field protocols where appropriate. In order to minimize loss to follow up between the household listing and re-visits for the survey, when a selected woman could not be found at their household, the enumerator made three attempts to contact her on different days. After each interview or attempt to make contact with a selected woman, enumerators completed an 'interviewer contacts' form in CommCare, which documented the status of each case (e.g. completed, not eligible, moved away, re-scheduled, etc.). To improve response rates, enumerators also scheduled interviews for days/times most convenient for participants (for example, on weekends or evenings, outside of working hours).

Ethical Approval

The Ethical Review Committee of the Aga Khan University (4964-Ped-ERC-17) and the National Bioethics Committee of Pakistan approved the study. The enumerator read aloud the consent form before asking survey questions to eligible women and signed the consent form on their behalf. The enumerator also provided a hard copy of consent form to the respondent.

Results

This study documented various issues encountered in data collection with potential solutions. These are broadly categorized into four main themes with further sub-categories (Table 1). The four main themes are technical issues related to GIS usage and computer assisted personal interviews (CAPI), household listing issues, respondent issues, and field issues.

Technical issues related to GIS and CAPI

We found that using GIS technology to demarcate clusters, particularly in urban slums, was an efficient approach to establishing a sampling frame; however, its use is subject to issues around cluster boundary

interpretation and household identification, particularly in narrow and congested streets. This approach required time and effort of extra human resource to physically validating the boundaries through field visits. Efforts were also required to work closely with program's staff to ensure that the boundaries around areas intended for program implementation were accurately demarcated.

Collecting data using electronic devices had many advantages but was also prone to some challenges such as power outages, low signals, and slow uploading speed. We ensured that enough tablets were available as backup in such circumstances. When data issues/queries emerged in the field, waiting to receive a response from the technical support team in AKU office who could access the central server was a time-consuming procedure. In areas of low availability of signals, enumerators struggled with the intermittent connectivity to sync the data, especially when conducting interviews on ground floors of tower buildings. Nevertheless, the application was programmed such that the survey itself could be run without any connectivity, and syncing could occur later (i.e. each evening upon return to the office). The dedicated staff were also on board to deal such queries in the field.

Household listing issues

Karachi is a metropolitan city with fast growing urban dwellings and settlements. Residential apartments and multi-storey buildings are common. Majority of the housing areas do not have formal address of the streets and residential sites, which makes it difficult to track the household.

In a dynamic population like our study sample, having a long gap between the household listing and the household survey can lead to high loss-to-follow-up, and recall errors among enumerators, such as forgetting important sites and other landmarks such as door markings. The household listing team marked unique identification number (ID) at a visible place on the doors during household listing after permission of participants to relocate and confirm the households randomized for data collection during household survey team by identifying the ID of the cluster. We trained two separate teams to conduct the household listing and household survey. Initially, our plan was to complete the entire household listing in all selected clusters before beginning the household survey. However, there was a shift in strategy mid-course when study team identified a need to reduce the time lag between the two activities. In our revised field protocol, the household listing and survey teams worked in tandem so that interviews began immediately after each cluster was listed, rather than listing all clusters before beginning the household surveys.

Respondent issues

The study revealed that respondent issues were the most important barriers to high quality data collection (table 1). In Karachi, people's religious beliefs about SRH, perceived security threats and gender-based household dynamics made them reluctant to participate in surveys [9]. Respondents consider enumerator as a stranger due to the fear of theft and child kidnapping incidents happened in the town. We found that discussing sensitive topics such as women's socioeconomic status (concerned of getting robbed) and sexual and contraceptive history posed issues to fieldwork and led to high refusal rates if not

approached with sensitivity. Likewise, certain community sub-groups e.g. Muhajirs and Pushtuns were also reluctant to participate during household data collection. To counter this challenge and to build good rapport, male field supervisors coordinated with male community leaders personally using a gatekeeper script to initiate the discussion, and we were privileged to have all the female enumerators who were trained to improve mutual trust between enumerators/respondents prior to interview. As a result, refusal rates were low (Jamshed town 9.8%, Yousuf Goth 2.5%, Korangi town 5%, and PIB & Dalmia/Shanti Nagar 4.3%).

In some instances, respondents expressed difficulties understanding Urdu language and had to be assigned to a new enumerator who could communicate in their native language. To maximize inclusion in our study sites, we ensured that our enumerator team included women who spoke the three most common languages in Karachi: English, Urdu, Pushto, or Sindhi.

The survey showed that engaging respondents for more than an hour in a busy personal schedule created problems for successful data collection. This also increased participant discomfort and unwillingness to participate in the data collection process. In these cases, we rescheduled the remaining portion of the interviews as per the availability of the respondents (e.g. on Sundays, outside of working hours, etc.).

Another issue this study revealed was that some respondents demanded material benefits, extra healthcare services or sometimes lodged complaints against the healthcare system when visited for surveys, thus enumerators had to be properly trained to deal with such situations. The enumerators were trained to explain that a report will be generated from the study findings and will be disseminated to concerned stakeholders. This report might inform policies that could benefit respondents in the long run. We also gave a tea mug as a gift to all respondents who completed the interview as a token of appreciation for their time and effort once the interview ended.

Field issues

This study also highlighted the difficulties of conducting long interviews in a dense, urban environment. Our protocol was to conduct the interview privately, in women's homes, but this was sometimes difficult because other family members were present and did not want to leave. Selecting an alternative private interview location was a challenge in some areas since data collection did not take place in a room, particularly in joint family systems, but rather in an open space, usually at the doorsteps with loud background noises outside and warm temperature. In addition to this, ensuring privacy away from other family members and neighbors was a challenge. We addressed this issue by scheduling a revisit after confirming another time when most of the family members were away from the home.

The interview period in both components was subject to some seasonal events e.g. fasting and Eid and field activities were postponed in retrospective sites due to political events and general elections in July in 2018. These data collection protocol adaptations have serious time and budget (additional cost incurred

for GIS and cluster demarcation) implications, but we view them as necessary investments to ensure high quality data.

Discussion

The study demonstrates how data collection in urban Karachi presents unique issues that influence the data collection process. However, appropriate mitigation measures to address the environmental and socio-cultural context enabled successful data collection in this setting. Lessons learned may be usefully applied in similar urban settings in other LMICs.

Using GIS technology to demarcate clusters for our sampling frame is a unique and efficient approach of present study. However, the process is time intensive, as interpretation of accurate cluster boundaries and household identification, particularly within neighborhoods with congested and narrow streets can be challenging. Working with experienced field teams and program staff who know the area well is critical to success. The finding is consistent with other studies conducted in LMICs, where researchers lacked accurate and detailed spatial images, and tracing the exact location of zoning area was difficult [23, 24]. This is strength of our study that we used a unique approach compare to sampling frame used for other studies such as the 2017-18 Pakistan Demographic and Health Survey (PDHS) which used complete list of enumeration blocks (EBs) created for the Pakistan Population and Housing Census 2017. The Pakistan Bureau of Statistics (PBS) has formulated EBs into three categories of income groups i.e. low, middle and high, keeping in view the living standard of the majority of the people.

We demonstrated that the growing interest and efforts in using electronic data collection methods in LMICs has many advantages, however, it takes time when a query emerges in the field and to receive a response from the technical support team who have access to central server. This finding is supported by available evidences from other LMICs [25, 26] which reported that implementation is challenging both administratively and technically. This can be reduced by building technical capacity of enumerators using the training of trainer (ToT) model, where team leaders can be properly train first, which may help enumerator's queries on the spot in the field to reduce errors.

The study revealed considerable issues from respondent's end during the data collection process. Like other LMICs, we also found considerable seasonal migration of respondents, and faced difficulty in identifying listed respondents in a cluster. A study conducted in Bolivia, Kenya, and South Africa showed similar finding of attrition rates between survey rounds and interviewer team tracked, followed and re-interviewed 84% of those women moved [27]. In our study, coordinating two teams to work in tandem and begin interviews immediately once each cluster's household listing was completed helped to minimize attrition between data collection activities.

Some respondents were reluctant to participate in our study because discussing family planning, especially contraceptive methods and abortion services is highly stigmatized among some Muslim religious groups. However, we also found that highly religious Christian families also had reservations about talking to strangers about such topics. Previous studies revealed similar findings that sexual and

reproductive issues are considered sensitive subjects to be rarely discussed in certain Muslim communities and both religious traditions strongly influence the uptake of family planning [28, 29]. This points to the important role of religious leaders as important influencers in conservative societies and suggests the need to engage them in efforts to promote sexual and reproductive health.

Respondents' perceived security threat is also an inherent issue hence suggesting to take into confidence influential gate keepers in the community and district administration. This finding is consistent with a study in South Africa where respondents denied access to houses and perceived enumerator with suspicion and resisted to participate as security risk and complained to local police despite official letters [30]. These issues were exacerbated when enumerators discussed sensitive topics such as family planning and reproductive history with respondents of particular religious groups. Similar findings were reported in another study where attitudes of some Muslims and Christian groups frequently oppose family planning with diverse sexual orientations [31]. Our study recruited female enumerators who were familiar with respondent's language and cultural norms hence respondent developed trust in them and felt comfortable in sharing information on sensitive topics.

Our study demonstrated that some respondents experienced difficulty in understanding Urdu language. We recruited skilled enumerators fluent in multiple native languages. This finding is endorsed by other studies, which encountered language barriers while collecting data and used bilingual enumerators fluent in local languages [18, 30, 32]. This highlights the importance of recruiting and training a diverse team of enumerators to ensure that all potential respondents have the same ability to communicate with the research team. Misinterpretation of risks and biases could be reduced through training and mock exercises of enumerators.

We adopted a good strategy where the training did not stop after the initial two week session; rather, the team adapted to learnings from the field throughout the data collection process, and came up with strategies to address them as they emerged. We discussed weekly progress of every enumerator with errors in the data. The field manager identified discrepancies through data quality checks and discussed strategies to resolve these queries.

The usefulness of household surveys to capture accurate information depends heavily on its content and length [33]. This finding was comparable to a study which indicated that participants rushed through the data collection process and provided inaccurate information. In our study, we rescheduled interviews to avoid incomplete or inaccurate information. Furthermore, lengthy questionnaires can lead to enumerator fatigue, but we did not encounter this difficulty in our study. This might be because we did not overburden our enumerators with difficult daily targets like we given target of three households' complete interviews in a day to each enumerator [12].

An interesting learning from our fieldwork was that some respondents expect monetary benefits or in other way showed sympathy for some direct benefits. The enumerator however explained that participants would receive no direct benefits. This likely reflects the fact that people get assistance from many local and international NGOs involved in charity programs in these urban societies. Studies in

Pakistan and other LMICs [6, 34, 35] reported similar findings, where both rural as well as urban communities had strong impression of some unconditional beneficial packages when enumerators visit homes. The respondents could be appreciated for their time and participation through a small non-monetary gift (e.g. tea mug) at the end of the interview.

The interview environment is a crucial aspect of the data collection process to ensure privacy [1]. Our study found that in certain instances, interviews had to be rescheduled especially in joint families living in limited spaces. The respondents felt uneasiness and nervousness discussing sensitive information such as a sexual activity and contraceptive use. In this environment, securing a private interview room to maintain privacy remained a challenge. Previous literature reports that getting access to participants and managing privacy and obtaining accurate responses to sensitive topics is a challenge in many health studies [6, 35, 36]. This suggests that in some settings, interviews must be conducted at an impartial, private venue that is suitable and secure for both the respondent and enumerator [1].

To our knowledge, this is the first study to lay out fieldwork issues at different stages of the data collection process in urban environments of Karachi, Pakistan. Our study also had some limitations. First, due to competing household tasks and busy schedules, some respondents may have rushed through the survey. Secondly, we did not track/follow respondents who could not be reached after three visits and were thus unable to capture information on this potentially important sub-group.

Finally, we implemented our study following strict standard protocol to ensure high quality data although time and cost substantially increased. This underlines the need to prioritize high quality data over cost and time complications.

Conclusion

The GIS technology may be used which is an innovative and cost-effective method for developing sampling frames in such resource constrained settings. Secondly, dedicated data management team should be on board to monitor electronic data collection in real time that will facilitate efficient detection of errors and inconsistencies and improve data quality. Thirdly, the strategy of interviewing women immediately after the cluster may be applied to make it easier to re-locate selected respondents and to reduce loss-to-follow up. Understanding local norms and developing culturally appropriate strategies should be developed to build trust with communities that may significantly reduce refusals. The trained female enumerators should be hired since they are very useful on how to interview respondents on sensitive reproductive health topics in such urban communities. Ensuring privacy in joint family households may be guaranteed by rescheduling interviews for another time when most family members were not at home. Findings of this study will help improve the quality and efficiency of future household surveys in urban settings.

Abbreviations

GIS (Geographical Information System), SDG (Sustainable Development Goals), LMICs (Low and Middle Income Countries), SRH (Sexual and Reproductive Health), AKU (Aga Khan University), CPR (Contraceptive Prevalence Rate), CAPI (Computer Assisted Personal Interviews) PDHS (Pakistan Demographic and Health Survey), EBs (Enumeration Blocks)

Declarations

Ethics approval and consent to participate

Necessary ethical clearance was granted by ethical review committee of Aga Khan University, Karachi, Pakistan (4964-Ped-ERC-17). Written consent was obtained from the participants prior to enrollment.

Consent to publish

Not applicable

Availability of data and materials

The datasets generated and/or analysed during the current study are available available from the corresponding author on reasonable request.

Competing interests

The authors declared that they have no competing interests.

Funding

Funding for this study received by a grant from an anonymous donor to Harvard T.H. Chan School of Public Health.

Authors Contributions

SBS was the principal investigator of the trial for Pakistan. MBK drafted the manuscript with input from all co-authors. SN, IH, KH, ZK, KF, DC, IS, SBS critically reviewed the manuscript. All authors have read, contributed to, and approved the final version of the manuscript.

Acknowledgments

The authors would like to thank the Department of Paediatrics and Child Health, Aga Khan University, Pakistan, for providing support for this research. We would also like to acknowledge Harvard T.H Chan School of Public Health. Authors express their sincere gratitude to all study participants for cooperation during the study.

References

1. Gill P, Stewart K, Treasure E, Chadwick B. Methods of data collection in qualitative research: interviews and focus groups. *British dental journal*. 2008;204(6):291–5.
2. Prata N. Making family planning accessible in resource-poor settings. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2009;364(1532):3093–9.
3. Ali M, Azmat SK, Hamza HB. Assessment of modern contraceptives continuation, switching and discontinuation among clients in Pakistan: study protocol of 24-months post family planning voucher intervention follow up. *BMC Health Serv Res*. 2018;18(1):359.
4. National Institute of Population Studies, P. Pakistan demographic and health survey 2017-18. Pakistan: In: NIPS/Pakistan and ICF Islamabad; 2019.
5. Casterline JB, Sathar ZA, ul Haque M. Obstacles to contraceptive use in Pakistan: A study in Punjab. *Stud Fam Plann*. 2001;32(2):95–110.
6. Cheema AR, Mehmood A, Khan FA. Challenges of research in rural poverty: lessons from large field surveys. *Development in Practice*. 2018;28(5):714–9.
7. Nori-Sarma A, Gurung A, Azhar GS, Rajiva A, Mavalankar D, Sheffield P, Bell ML. Opportunities and challenges in public health data collection in Southern Asia: examples from Western India and Kathmandu Valley, Nepal. *Sustainability*. 2017;9(7):1106.
8. Dickinson FM, Pyone T, van den Broek N. Experiences from the field: maternal, reproductive and child health data collection in humanitarian and emergency situations. *International Health*. 2016;8(2):83–8.
9. Sensoy N, Korkut Y, Akturan S, Yilmaz M, Tuz C, Tuncel B. (2018). Factors Affecting the Attitudes of Women toward Family Planning. *Family Planning*, 33.
10. Tlale LB, Morake B, Lesetedi O, Maribe L, Masweu M, Faye C, Asiki G. (2019). Data quality self-assessment of child health and sexual reproductive health indicators in Botswana, 2016–2017. *PloS one*, 14(8).
11. Bell K, Fahmy E, Gordon D. Quantitative conversations: the importance of developing rapport in standardised interviewing. *Quality quantity*. 2016;50(1):193–212.
12. Rimando M, Brace AM, Namageyo-Funa A, Parr TL, Sealy D, Davis TL, Martinez LM, Christiana RW. (2015). Data Collection Challenges and Recommendations for Early Career Researchers. *The Qualitative Report*, 20(12), 2025–2036. Retrieved from <https://nsuworks.nova.edu/tqr/vol20/iss12/8>.
13. Bonevski B, Randell M, Paul C, Chapman K, Twyman L, Bryant J, . . Hughes C. Reaching the hard-to-reach: a systematic review of strategies for improving health and medical research with socially disadvantaged groups. *BMC medical research methodology*. 2014;14(1):42.
14. Dearnley C. (2005). A reflection on the use of semi-structured interviews. *Nurse researcher*, 13(1).
15. Iqbal S, Zakar R, Zakar MZ, Fischer F. Perceptions of adolescents' sexual and reproductive health and rights: a cross-sectional study in Lahore District, Pakistan. *BMC international health human rights*. 2017;17(1):5.

16. Stephenson R, Hennink M. (2004). Barriers to family planning use amongst the urban poor in Pakistan.
17. Machingura F. (2014). The Martyring of people over radical beliefs. *Multiplying in the spirit*, 175.
18. Munyoro I. Research data collection in challenging environments: Barriers to studying the performance of Zimbabwe's Parliamentary Constituency Information Centres (PCICs). *African Journal of Information Communication*. 2018;21:81–95.
19. Nel E, Rich E, Morojele N, Burnhams H, Williams NPetersen, P., & Parry C. Data collection challenges experienced while conducting the International alcohol control study (IAC) in Tshwane, South Africa. *Drugs: Education Prevention Policy*. 2017;24(5):376–83.
20. Schmeets H. Increasing response rates and the consequences in the Dutch Parliamentary Election Study 2006. *Field Methods*. 2010;22(4):391–412.
21. Harvard University. (2020, July 15). Willows impact evaluation: Reproductive Health Study. Retrieved from <https://projects.iq.harvard.edu/willowsimpacteval>.
22. Fadaei H, Bayazidi M. a Comparative Study of Utilizing Mobile-Gis Technology to Collect Online Crime. *The International Archives of Photogrammetry Remote Sensing Spatial Information Sciences*. 2019;42:355–61.
23. Mennecke BE, West LA Jr. Geographic Information Systems in developing countries: issues in data collection, implementation and management. *Journal of Global Information Management (JGIM)*. 2001;9(4):44–54.
24. Agarwal S, Perry HB, Long LA, Labrique AB. Evidence on feasibility and effective use of mHealth strategies by frontline health workers in developing countries: systematic review. *Tropical medicine international health*. 2015;20(8):1003–14.
25. Pujari SJ, Palipudi KM, Morton J, Levinsohn J, Litavec S, Green M. (2012). Electronic data collection and management system for global adult tobacco survey. *Online Journal of Public Health Informatics*, 4(2).
26. Alderman H, Behrman JR, Kohler H-P, Maluccio JA, Watkins CS. (1999). *Attrition in longitudinal household survey data: some tests for three developing-country samples*: The World Bank.
27. Alomair N, Alageel S, Davies N, Bailey JV. Factors influencing sexual and reproductive health of Muslim women: a systematic review. *Reproductive health*. 2020;17(1):1–15.
28. Sundararajan R, Yoder LM, Kihunrwa A, Aristide C, Kalluvya SE, Downs DJ, Downs JA. How gender and religion impact uptake of family planning: results from a qualitative study in Northwestern Tanzania. *BMC Women's Health*. 2019;19(1):99.
29. Mathee A, Harpham T, Naicker N, Barnes B, Plagerson S, Feit M,.. . Naidoo S. Overcoming fieldwork challenges in urban health research in developing countries: a research note. *Int J Soc Res Methodol*. 2010;13(2):171–8.
30. Cense M, de Neef M, Visscher W. (2018). Culture, Religion and Sexual and Reproductive Health & Rights.

31. Smith HJ, Chen J, Liu X. Language and rigour in qualitative research: problems and principles in analyzing data collected in Mandarin. *BMC medical research methodology*. 2008;8(1):44.
32. Glewwe P. (2005). An overview of questionnaire design for household surveys in developing countries. *Household sample surveys in developing and transition countries*.
33. Darling J. Emotions, encounters and expectations: The uncertain ethics of 'the field'. *Journal of Human Rights Practice*. 2014;6(2):201–12.
34. Holden RJ, Scott AMM, Hoonakker PL, Hundt AS, Carayon P. Data collection challenges in community settings: Insights from two field studies of patients with chronic disease. *Qual Life Res*. 2015;24(5):1043–55.
35. Ashton S. (2014). Researcher or nurse? Difficulties of undertaking semi-structured interviews on sensitive topics. *Nurse researcher*, 22(1).

Tables

Table 1. Issues and mitigation measures for conducting a study in community of urban population

Challenges encountered	Mitigation measures
1. Technical issues: GIS usage and computer assisted personal interviews (CAPI)	
<p>i. GIS Mapping: Difficulty in cluster boundary interpretation and household identification at each survey point, especially where dense structures are present.</p>	<ul style="list-style-type: none"> · Physically validating the boundaries through field visits.
<p>ii. Low internet signals in ground floors of tower buildings sometimes made it difficult to sync data.</p>	<ul style="list-style-type: none"> · Having dedicated staff available during field activities to deal with unexpected tablet issues in real time.
<p>iii. Reporting a technical query and waiting for a solution from staff with access to central server was time consuming.</p>	<ul style="list-style-type: none"> · Being able to download data later upon returning to the office.
2. Household Listing issues	
<p>i. Long duration between household listing and survey. Enumerators forgot site details and landmarks, respondent migration was higher, door markings were more difficult to identify.</p>	<ul style="list-style-type: none"> · Started data collection soon after household listing.
3. Respondent issues	
<p>i. Sensitive information: Respondents' perceptions and beliefs make it difficult to discuss some topics. Asking sensitive questions about sexuality can be controversial.</p>	<ul style="list-style-type: none"> · Refresher training sessions conducted on how to discuss sensitive topics with respondents.
<p>ii. Security issues: Respondents' fear about child kidnapping and theft and linked it with previous such incidents.</p>	<ul style="list-style-type: none"> · Took into confidence influential gate keepers in the community and district administration.
<p>iii. Unwillingness to participation: Certain community sub-groups i.e. Urdu speaking and Pushto were reluctant to participate in the study.</p>	<ul style="list-style-type: none"> · Coordinated with community leaders (especially men) personally; used gatekeeper script.
<p>iv. Language barriers: Some respondents expressed difficulties understanding Urdu language.</p>	<ul style="list-style-type: none"> · Assigned new enumerators who could communicate in their native language.
<p>v. Length of questionnaire and participants' schedules: Engaging the respondent for an hour or more in a busy personal schedule created barriers to successful data collection.</p>	<ul style="list-style-type: none"> · Rescheduled such cases as per the availability of the respondents, including weekends and after working hours.
<p>vi. Participant expectations: Many participants expected extra healthcare services, or another way requested material benefits from enumerators.</p>	<ul style="list-style-type: none"> · Enumerators were trained to explain indirect benefits e.g. sharing research findings with key stakeholders and informing policies that may benefit the respondents in long run.

4. Field issues

i. *Environmental issues:* Interviews conducted in dark rooms with no electricity, bad odors, dirty streets, in extreme temperatures; enumerators reported feelings of isolation.

· Worked in groups in neighborhoods where enumerators felt uncomfortable.

· Supervisors waited nearby and were available to support if needed.

ii. Presence or nearness of family members

· Scheduled a revisit after confirming another time when most of the family members were away from the home.

iii. Loud background noises outside

iv. *Ensuring privacy:* Some interviews couldn't be done inside the house (especially in joint family systems).

· Revisit/reschedule cases as per respondent's availability.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [RetrospectiveSurveyQuestionnaire.pdf](#)
- [ProspectiveSurveyQuestionnaire.pdf](#)