OXFORD

Remote data collection for public health research in a COVID-19 era: ethical implications, challenges and opportunities

B Hensen [1,*,†], CRS Mackworth-Young [2,†], M Simwinga [3],
N Abdelmagid [4], J Banda[5], C Mavodza [6,7], AM Doyle [8],
C Bonell [7,‡] and HA Weiss [8,‡]

[1]Department of Clinical Research, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK
[2]Department of Global Health and Development, Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK
[3]Zambart, Nationalist Road, Lusaka, Zambia
[4]Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK
[5]Malawi Epidemiology and Intervention Research Unit (MEIRU), Lilongwe, Malawi
[6]Biomedical Research and Training Institute, Seagrave Rd, Avondale, Harare, Zimbabwe
[7]Department of Public Health, Environments and Society, Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK
[8]Medical Research Council Tropical Epidemiology Group, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

*Corresponding author. Department of Clinical Research, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London, UK. E-mail: Bernadette.hensen@lshtm.ac.uk
†These authors are joint first authors.
‡These authors are joint senior authors.

Introduction

The coronavirus disease 2019 (COVID-19) pandemic, caused by the SARS-CoV-2 virus, has had unprecedented impacts on health systems, public health, societies and individuals globally (The Lancet Public Health, 2020). In response to outbreaks, physical distancing measures, national lockdowns and travel restrictions to control the spread of COVID-19 have been implemented in many countries (Chu *et al.*, 2020). In response to these measures, many public health researchers are choosing to switch from standard face-to-face data collection methods to remote data collection in support of continued research. Remote data collection is defined here as the collection of data via the phone, online or other virtual platforms, with study participants and researchers physically distanced.

The aim of this commentary is to summarize methods, key challenges and opportunities of remote qualitative and quantitative data collection for public health research in low- and middle-income countries (LMIC). The framework we use to structure our discussion is the research process, starting from sampling and culminating in analysis. Within this, we draw out the steps in research most

likely to be affected by the pandemic and attendant need to cease face-to-face interactions with research participants. We identify which steps are most affected and what are potential alternatives based on interviews and discussions, held between May and June 2020, with ~30 researchers from the London School of Hygiene and Tropical Medicine and collaborating partners, representing a range of disciplines. Interviewees were selected or volunteered themselves, based on their experience and expertise in designing and conducting remote data collection. These consultations identified the following as the steps in research most likely to require attention: sampling and recruitment; informed consent; response rates; rapport with participants; privacy and safety; and analysis. Whilst the focus of this commentary is on LMIC, many of the lessons learnt are relevant to remote data collection in high-income countries.

What remote data collection methods can I use?

Remote qualitative methods include online or phone-based interviews and focus group discussions (FGDs), audio-diary methods

(Mupambireyi and Bernays, 2019), photovoice (use of photography to capture lived experiences) (Copes *et al.*, 2018), video documenting, documentary analysis of social media (e.g. Facebook and WhatsApp groups, YouTube comments or podcasts) and auto-ethnography (ethnographic study on self) (Ellis and Bochner, 2000; Lupton, 2020). Remote quantitative methods include mobile phone surveys implemented using: interactive voice response (IVR), short messaging service (SMS) or computer-assisted telephone interviews (CATI) and self-completed online questionnaires, shared via email or social media platforms. These methods are not new (Gibson *et al.*, 2017) with telephone and postal surveys used in higher-income countries; yet their use has become essential during the COVID-19 era to support the collection of data directly from individuals and populations.

Each remote data collection method has advantages and disadvantages, which affect their feasibility and acceptability in specific settings (Table 1). For example, when considering a mobile phone survey, although IVR and SMS surveys are cheaper than CATI, they require participants to have high levels of literacy; CATI allows for the inclusion of individuals regardless of literacy and provides opportunities for researchers to encourage participation and study participants to clarify questions (Gibson *et al.*, 2017). With widespread ownership of mobile phones in LMIC, but lower access to smartphones and the Internet, mobile phone methods are more commonly used than online methods and are a key focus of this commentary. Few of the experts interviewed had implemented or were planning online methods due, in part, to their limited reach in certain LMIC. Exceptions include online surveys planned with specific target groups, e.g. members of an established association of professionals and university students.

In the following sections, we describe the specific challenges of remote data collection throughout the design, conduct and analysis of a research study, and discuss the implications for: ethics, sampling and recruiting study participants, obtaining informed consent, maximizing response, protecting participants' privacy and confidentiality and data analysis and interpretation.

Is it ethically appropriate to conduct my research study during the COVID-19 pandemic?

Individuals, communities and societies face heightened social, physical and emotional challenges during the COVID-19 pandemic. Decisions on whether to conduct research using remote methods need to consider the research burden and COVID-19-related risks to study participants. For example, remote collection of data may require greater effort on the part of the study participant, who may be required to use their own phone, their own resources to charge this phone, and to identify a private space to participate in the study. On the other hand, remote methods may be more preferable to study participants, removing the time and opportunity cost associated with travel to study sites. As with any research, potential risks need to be weighed against benefits and the ethical imperative to continue with research to generate the evidence of benefit to public health.

How do I sample and recruit study participants?

Key challenges in remote data collection include garnering diverse experiences (qualitative research), obtaining a sampling frame representative of the population of interest (quantitative research) and contacting 'harder to reach' populations (Tran *et al.*, 2015). Whilst some of these challenges are present in face-to-face research, the limited ability to recruit participants in person, either at home, in a clinic or other venue, alongside the reliance on mobile phones for recruitment, heightens these challenges and creates the need for alternative sampling methods. For qualitative research, sampling approaches include purposive sampling, snowball and convenience sampling. Purposive sampling aims to ensure diversity according to key factors theorized to influence experience. Recruitment can be facilitated via community-based organizations and leaders, neighbourhood health committees or established networks (Sudan case study, Box 1). Snowball sampling can be effective for qualitative research, although drawing from multiple initial participants (who then recruit others from within their networks) is important to achieve diversity (Shaghaghi *et al.*, 2011; Kirchherr and Charles, 2018). These sampling methods can also be used in quantitative research; snowball sampling may be useful for online surveys shared via email or social media platforms (Roy *et al.*, 2020), and a convenience sample can be recruited through online social networking platforms.

For quantitative research, representative samples from the population of interest are either important to maximize internal validity (descriptive research) or useful to maximize external validity (aetiological/evaluation research). In countries where mobile phone ownership is high, a sampling frame of the general population can be obtained by contacting mobile phone network operators or mobile phone survey companies who maintain lists of phone numbers. A sample can then be randomly selected using these lists. Alternatively, random digit dialling could be used to generate a study sample. These methods, however, have limitations. Network operators may be unwilling to provide phone numbers and random digit dialling is unlikely to yield a representative random sample of the population. For a descriptive, population-based survey, lack of representativeness limits the validity of this approach.

As with qualitative research, established relationships, e.g. with participants recruited to a cohort study (Malawi case study, Box 2), can be leveraged to facilitate continued or new research. Where the target population is a specific group, e.g. female sex workers or adolescents, respondent-driven sampling (where individuals representative of the target population are provided a fixed number of coded coupons to incentivize recruitment of their peers to the study) (Heckathorn, 1997; Johnston and Sabin, 2010), is an established method that can be implemented using mobile phones or online to, in principle, obtain a representative sample. Depending on the target population, existing lists that are representative of the population, e.g. registers of school students or email addresses/phone numbers for members of a professional association, can be leveraged. However, data protection and ethical issues around sharing personal details need to be considered; lists should be anonymized to maintain confidentiality and the owners of these lists should inform potential study participants about the research prior to recruitment. Where the target population is individuals attending particular spaces, e.g. bars and sport facilities, or indeed geographical areas, open source maps can be used to generate a sampling frame and existing social networks leveraged to initiate data collection.

In practice, a combination of approaches may be necessary to recruit study participants. However, limitations related to the diversity of experience and representativeness are likely to persist as is restricted participation of more vulnerable populations, including individuals with vision or hearing impairments, low literacy, and older populations. Where a mobile phone survey or interview is planned, one strategy to reach individuals without a phone is contacting, or even interviewing, a phone-owning friend or relative; however, this may not be appropriate for sensitive research topics.

Table 1 Available qualitative and quantitative remote data collection methods, their strengths, limitations and strategies to improve the data quality of different methods

Method	Description	Strengths	Limitations	Strategies to improve data quality and navigate challenges
Qualitative methods				
Phone interviews	In-depth and semi-structured interviews can be conducted by phone. Interviews can also be via WhatsApp calls or online platforms (e.g. Skype or Zoom)	• Can facilitate the collection of high-quality data on personal experiences • Real-time interviews allow for the interviewer to probe, check understanding and follow the direction of conversation	• Challenge to develop rapport and trust with the participant • Inability to see visual cues reduces understanding and appropriate prompting • Technological challenges with network, airtime, batteries • Cost of phone technologies • Disturbance by other noises and activities if participants cannot find a private space • Responsibility of privacy placed on participant	• Training is key to developing rapport with participants over the phone, including role-play practice of interviews, especially how to set the tone at the beginning of the interview with informal conversation • Phone interviews should be shorter than face-to-face interviews, to account for participant fatigue • Perseverance is required to repeatedly call participants at different times of day and days of the week and call back if the interview is disturbed or cut off
Phone or online FGDs	Online platforms or group phone calls can be used to facilitate group discussions remotely. These real-time discussions can be through writing, speaking or with video.	• Facilitates interaction to understand socially normative perceptions • Can provide some peer support	• Moderation can be challenging, and requires a skilled facilitator • Data security of online platforms need to be considered, including end-to-end encryption as some platforms (e.g. Zoom) are less secure • If not participating anonymously, ensuring confidentiality is challenging	• If desired, participants can join anonymously, through providing a pseudonym and not using a video • Group phone calls may be most appropriate in lower-income settings, where access to online is lower, but incurs additional airtime costs
Self-collection of data (including diaries, photovoice, video documenting and auto-ethnography)	Participants record elements of their lived experiences themselves. Diaries or journals can be handwritten, voice memos or through online platforms or applications. Photovoice or video documenting involves participants taking photos or videos about their everyday practices and interactions that they can share with researchers. Auto-ethnography situates the researcher as the participant, documenting their own lives, experiences and perceptions.	• Enables participants to generate data at a time and a place that is convenient for them • Facilitates tailoring data collection to participants' personal experiences	• Attrition to continue data collection can be a problem, especially for longer term studies • Data recorded may deviate from desired areas of enquiry or research questions • Challenges in transferring self-collected data to researchers • Participants using and keeping photo and video technology may not be appropriate in lower-income settings	• Asking participants what method of self-collection of data suits them can better tailor the method to the participant • Providing relevant prompts and guiding questions can help direct participants' documentation • Keeping in touch with participants maintains participation • Self-collection of data can be used as prompts for further discussion in combination with interviews and other methods
Documentary analysis	Analysis of naturally occurring online data, e.g. analysing YouTube	• Data are already existing and available in the public domain	• Lack of depth, or ability to probe, compared to interviews or FGDs	• Can capture and transfer text, videos and

(continued)

Table 1 (continued)

Method	Description	Strengths	Limitations	Strategies to improve data quality and navigate challenges
	comments, or Facebook group discussions	• Data are not produced for the purpose of research, and not directed by the researcher	• Capturing and analysing emojis is a challenge, but important for understanding meaning	images to analytical software (e.g. Nvivo) • Analytics (e.g. through YouTube) can provide data on demographics of those interacting (watching and commenting)
Quantitative methods				
SMS survey	Individual questions sent via SMS to phone number Participants respond with a number corresponding to a response list	• Relative to other phone surveys, may be more representative of women (Lau et al., 2019) • Appropriate for short surveys, quick to implement • More inclusive for individuals who are hard of hearing • May be appropriate for sensitive questions • Relative to CATI, cheaper and may provide more language options • Allows individuals to respond at times most convenient to them • Least expensive relative to CATI and IVR (Lau et al., 2019)	• Low response rates compared to CATI • Potential for question-level breakoff, which can increase with each subsequent question • Under-represents participants with lower literacy levels • Questions need to be highly specific, as there is no opportunity for participants to clarify questions and limits in character numbers • Navigating keyboard can be a challenge (Leo et al., 2015), particularly for older populations (who are under-represented in SMS surveys) (Broich, 2015) • Participants may have concerns regarding data charges for submitting responses	• Send reminders to increase response rates • Combine with other mobile phone survey method to increase reach • Provide an incentive to minimize concerns regarding data charges
IVR survey	Automated phone survey, with individuals asked to key in or state their response to the questions	• Some evidence to suggest this is more inclusive of individuals with lower literacy levels (Gibson et al., 2017; Lau et al., 2019) • More representative of rural populations • Higher response rates than SMS survey, although not necessarily more representative overall (Lau et al., 2019) • Relative to CATI, provides more language options (Lau et al., 2019)	• More expensive than SMS surveys, but cheaper than CATI • Lacks personal touch of CATI surveys, no opportunity for rapport development • Not inclusive for individuals who are hard of hearing • No opportunity for individuals to clarify questions • Individuals may not answer unknown phone number	• Send survey at different day/time combinations • Combine with SMS reminder to increase response rates • Combine with other mobile phone survey method to increase reach
CATI	Participants are phoned by an individual and asked to complete a survey over the phone	• Higher response rates relative to IVR and SMS surveys (Lau et al., 2019) • Higher survey completion rates, as interviewer able to build rapport and explain the purpose of the study	• Costlier than IVR and SMS surveys • Requires more quality control and training • Potential for interviewer bias introduced (Gibson et al., 2017)	• Make phone calls on different days/times of day • Schedule a time to interview the study participant • Send SMS prior to the phone call to introduce the study

(continued)

Table 1 (continued)

Method	Description	Strengths	Limitations	Strategies to improve data quality and navigate challenges
		• Increased reach of populations with lower literacy • Potential for less measurement error, as there is an opportunity for clarification of questions • Can complete longer, more complex/sensitive surveys		
Online surveys	A link to a self-completed questionnaire sent to potential study participants via email, WhatsApp or other social media or networking platform	• Cheaper and easier to execute than mobile phone survey methods • More efficient than mobile phone surveys, reaching a high sample size in a short timeframe (particularly within a specific target population, e.g. university students or members of a professional association) • Can be shared via multiple platforms (email, social media) • Open-ended, more qualitative questions, can be embedded within the survey, which can provide a rich data source	• Under-represents individuals without an email address, smartphone or data on their phone (for survey completion) (Roy et al., 2020) • If survey completion relies on snowball 'sharing' of the survey, the findings may be further subject to bias (Ball, 2019). • Limited reach of individuals with lower literacy • No opportunity for study participants to clarify questions	• Provide an incentive for survey completion and for inviting eligible individuals to complete the survey • Conduct extensive questionnaire testing to minimize any ambiguity in eligibility criteria and questions • Use with specific target group(s) rather than general population (Ball, 2019).

SMS - short message service
IVR - interactive voice response
CATI - computer-assisted telephone interview

How can I obtain informed consent remotely?

Oral consent (over the phone or via a voice note) or written consent (via email, WhatsApp or SMS) is being accepted by some ethics committees as written informed consent becomes challenging, or impossible. For mobile phone-based research with adolescents, which requires parental/guardian consent, additional challenges emerge in confirming the age of the participant to establish whether parental/ guardian consent is needed and in ensuring consent is being provided by the parent/guardian rather than the respondent themselves, a friend or other relative. For these reasons, oral consent, which can be recorded or conducted in combination with written consent where feasible, may be preferable to written consent only. Concise and simple language is required to convey complete information remotely, whilst maintaining the rigorous ethical standards of face-to-face research. Consent should always be appropriately documented, whilst protecting patient data and confidentiality. Documentation could be in the form of a list of participants, stored on a password-protected computer, who consented to participate in different study components, which could also serve as a record for audit purposes.

How do I navigate technological challenges in recruitment to maximize response rates?

Researchers should anticipate higher non-response than face-to-face methods in sample size calculations. For mobile phone surveys, response rates are influenced by factors including phone ownership and autonomy to use phones. In some settings, this means rural women and elderly populations are under-represented. Even where mobile phone ownership is high, low response rates threaten study validity as how representative study participants are of the broader, target population would remain unclear. Among individuals with a phone, response rates are affected by distrust of unknown phone numbers, phone-based harassment (Lamanna *et al.*, 2019), time required to complete the survey, poor network coverage and inadequate access to electricity to charge phones (Malawi case study, Box 2). Online surveys can achieve high participation yet they over-represent higher-income, urban populations with higher literacy and access to smartphones and/or the Internet (Roy *et al.*, 2020).

To improve response rates to mobile phone surveys, researchers can use established relationships with participants or community-

---

Box 1 A remote collaboration with youth networks for research during the COVID-19 pandemic: a case study from Sudan

In April 2020, a study to explore the acceptability and feasibility of strategies to shield high-risk individuals from COVID-19 was launched in six communities in Sudan. Researchers partnered with a Sudanese network of youth volunteers, aged 20–30 years, trained in promoting health and youth participation.

Volunteers were trained using social media; pre-recorded training sessions were shared via WhatsApp along with interview guides. A virtual chat meeting was held to answer questions and receive feedback on the interview guide. Volunteers identified 60 eligible study participants purposively, by calling existing community contacts, and conducted phone-based interviews. Eligible participants were any adult household member in households with a member at high risk of COVID-19 (∼38% of respondents were female). To summarize observations of emerging themes, volunteers were given a reporting template. Conference calls facilitated sharing insights from the reports and volunteers' intimate knowledge of the data. Interview recordings and transcripts were uploaded to a secure cloud platform for further thematic analysis by researchers. Poor connectivity prevented live training, delayed uploading of interview recordings, and disrupted interviews and group discussions. With volunteers using their own phones to conduct interviews, data security concerns also emerged. The volunteer's lack of prior research experience delayed the original study timeline, as frequent support by researchers was needed, e.g. to ensure post-interview clean-up of identifying information about study participants.

Despite challenges, the partnership leveraged the expertise of researchers and the volunteers' existing community links. The study provided an opportunity to invest in an established community-based network, with the prospect of acquiring research skills and adapting their COVID-19 prevention messaging, both of which were key motivators for the volunteers. Despite a lockdown, and without access to a sampling frame, volunteers were able to remotely identify participants and conduct interviews efficiently and with limited resources.

Box 2 Conducting telephone interviews during COVID-19: a case study from Malawi[1]

To document changes in COVID-19-related knowledge, attitudes and behaviours in Malawi, a cohort study of four rounds of mobile phone surveys was initiated in April 2020, with follow-ups due for completion in November 2020. Study participants were primarily adult residents of Karonga district, Northern Malawi, who had previously participated in epidemiological studies, led by MEIRU, on feasibility of measuring mortality. During these pre-COVID-19 studies (December 2019–March 2020), 1 036 individuals were asked for their phone number for recruitment and/or follow-up purposes. Among these individuals, 257 (24.8%) did not have a phone number or refused to provide one.

Interviewers, working from their homes, called these phone numbers and obtained consent to participate after verifying participants' identity. On average, three calls were required to complete an interview. Respondents received airtime credit of $1.50 upon completion of the interview. Of 779 potential respondents, 620 (79.6%; 77.8% of males and 80.9% of females) completed the first interview.

Factors contributing to successful contact with participants included calling at times when they were likely to be free (late afternoon) and at times suggested by participants, making additional calls even when previous attempts were unsuccessful, and attempts at different times and days. Key challenges were that phone numbers did not exist or were disconnected from the network, and calls went unanswered throughout the study (15% overall). The median interview duration was 30 minutes, with significant variation between interviewers despite receiving the same training, practice sessions and having similar previous interviewing experiences. This variation was attributable to the time required by individual interviewers to develop rapport, obtain informed consent and navigate the survey questionnaire. Some calls lasted more than one hour due to multi-tasking on the part of study participants or calls disconnecting because of poor network and limited battery life. Despite challenges, once contacted, non-consent was low (<1%).

[1]This work was funded by the National Institutes of Health R01HD088516 (PI: Helleringer).

---

based organizations or send an SMS, prior to the phone call, to introduce the study and inform individuals that they should anticipate a call. In the absence of transport refunds, the provision of airtime to compensate for participants' time and their own resources needed to charge their phones is important from an ethical standpoint. Airtime incentives to participate in the study and to refer friends to the study can achieve higher response (Gibson *et al.*, 2017). However, issues of joint phone ownership need to be navigated, in which case other compensation, such as vouchers redeemable at local shops, could be considered. Perseverance (i.e.

repeatedly contacting participants at different time and day combinations) is also required, which can be facilitated through protocols detailing the frequency and timing of contacts (Malawi case study, Box 2). To increase survey completion rates, questionnaires and interview guides need to be short (lasting no longer than 30 minutes) (Dabalen *et al.*, 2016). Placing the most pertinent questions near the start of a survey is of greater importance in remote data collection, as technological challenges may occur, participants may be more likely to experience fatigue, be distracted by other activities or have their privacy compromised.

Box 3 Phone interviews with healthcare providers to understand the perceptions and experiences of lockdown measures: a case study from Zimbabwe

Between March and April 2020, a process evaluation nested within an existing cluster randomized trial of a community-based integrated HIV and sexual and reproductive health service for youth in Zimbabwe was adapted to explore healthcare providers' perceptions and experiences of national lockdown measures. In the first week of the lockdown, 15 phone-based interviews were conducted. Written informed consent was obtained at a face-to-face meeting prior to the lockdown with the providers, who were purposively selected to provide diverse experiences across location, role, age and gender and whose phone numbers were already known.

For participants who had existing relationships with the interviewer, rapport was easily established, although lack of visual cues obstructed the ability to probe. To work around more formal and formulaic responses, particularly, for those the interviewer had not met before, the interviewer built informal conversation into the interview, particularly during the first few minutes of discussion. Some participants were, in fact, more open over the phone: the interview offered them a rare chance to express their feelings and concerns during lockdown, knowing that they would not see the interviewer in the foreseeable future.

Logistical and technological challenges were faced. Network issues interrupted interview flow, forcing the interviewer to be flexible with re-scheduling interviews. Many participants could not find a quiet and private space to participate in the interview, with children and other conversations disrupting the interview. Perseverance and flexibility were required, such as allowing participants to reschedule the interview at a time convenient to them.

Despite challenges, conducting the interviews by phone circumvented the need to travel, enabling the rapid collection of data which the researchers considered to be of high quality. Importantly, participants expressed gratitude at having the opportunity to talk to someone and share the challenges they were facing as a result of the lockdown.

How do I build rapport with participants?

Intensive training of interviewers, including role play for phone-based interviews, is critical for developing strategies to build rapport. Rapport should be established in the first few minutes of a call, with informal conversations incorporated in the consent process (Zimbabwe case study, Box 3). Phone-based in-depth interviews and CATI enable researchers to develop rapport with study participants, which can improve response rates and be more appropriate for asking complex and sensitive questions (Gibson *et al.*, 2017; Lau *et al.*, 2019). To increase response to sensitive questions, e.g. sexual behaviours, and the validity of these data, researchers should consider combined approaches, providing individuals the opportunity to respond via SMS or IVR. This is similar to the use of audio computer-assisted survey instruments within face-to-face surveys, which can reduce reporting bias (Langhaug *et al.*, 2010). However, combining methods may have implications on the cost, time and technical expertise required to complete the study.

How do I protect participants' privacy and safety?

When research is face to face, the researcher is responsible for establishing privacy and halting data collection when privacy is compromised. Remote research places this onus on the study participant. Yet, establishing privacy can be difficult where participants share homes and have limited private space or time (Zimbabwe case study, Box 3). Privacy is particularly important for studies exploring sensitive topics, such as gender-based violence, where the consequences of compromised privacy could be harmful (Peterman *et al.*, 2020). At the start of data collection, participants should be advised of the potentially sensitive nature of the study and that they should seek a private space. To mitigate risk, strategies include using 'code words' or an 'exit button' that participants can say or press when their privacy is compromised (Peterman *et al.*, 2020). IVR and online surveys enable participants to complete surveys at a time and place of their choosing, offering more flexibility for participants to establish privacy. These surveys could include a question on whether the respondent completed the survey in private, or in the presence of, e.g. their child, parent/guardian or friend.

Data protection, including end-to-end encryption of phone calls and security of platforms used to deliver online surveys and interview transcripts, is an additional issue relevant to privacy and confidentiality that requires consideration (Eynon *et al.*, 2017). In addition, researchers have a duty of care and need to carefully consider safeguarding issues, especially where COVID-19 has impacted the availability of support services. Information on online or phone-based services should be made available during the consent process. Specific protocols need to be developed for referrals, interviewers need to be informed if particular responses may trigger automatic referrals, and follow-up is required where safeguarding issues emerge. As a part of this protocol, researchers need to establish a system to regularly check that these services have remained operational.

How do I analyse and interpret data collected remotely?

Remotely collected quantitative data will likely be affected by response bias (Labrique *et al.*, 2017). Weighting results using existing data from a census or population-based survey known to be representative of the population of interest can been used to reduce this bias (Lau *et al.*, 2019). However, the use of weights in data analysis reduces precision and may have little effect on estimates (Lau *et al.*, 2019). As with face-to-face data collection, transparency regarding limitations is essential, including reporting response rates and other potential sources of bias (Greenleaf *et al.*, 2017). Data on whether the respondent was alone whilst completing a mobile phone or online survey can be used in a sensitivity analysis to assess whether having another person present compromised responses. Analysis of remote qualitative data needs to account for issues around rapport;

triangulation of data from different methods can help provide depth. Findings emerging from remote methods should be interpreted in light of these limitations and the implications on generalizability discussed.

What opportunities do remote data collection methods present?

Remote data collection presents opportunities and challenges. The methods enable data collection in contexts where face-to-face data collection is less feasible, e.g. during violence and unrest, when travel restrictions are in place, a natural disaster and during other disease outbreaks. The methods may provide greater autonomy and privacy, e.g. through use of a pseudonym during online FGDs and surveys. Self-collected remote qualitative methods, such as audio diaries, photovoice, video documenting and auto-ethnography enable more participant-centred data collection. The engagement of members from the population of interest in the research activities demonstrates to the public the value placed on their perspectives and lived experiences and can be used to inform and strengthen activities already being implemented by communities (Sudan case study, Box 1). Remote data collection also provides an opportunity for more efficient data collection, being less expensive and time consuming than face-to-face data collection. The methods may be preferred by some study participants who may also have more time for participation, particularly during lockdowns. This efficiency, particularly with automated phone surveys, facilitates data collection from a large number of study participants over a short timeframe, providing critical information to inform the response to COVID-19 or similar crises. The benefits may be greatest for follow-up surveys among cohorts already engaged in research. Leveraging the widespread use of mobile phones among younger adult men, often under-represented in face-to-face population-based surveys, provides opportunities to reach broader cross-sections of a population (Lau *et al.*, 2019; L'Engle *et al.*, 2018).

Concluding remarks

In a COVID-19 era, remote data collection is needed to inform the response to the pandemic and other public health issues. The remote collection of data presents key ethical challenges and particular challenges related to identifying and recruiting study participants. With high and increasing ownership, remote data collection is likely to continue to rely on mobile phones, which remains easiest when building on existing relationships, where contact details are known, rapport is developed and trust established. A key challenge requiring further research and navigation is how to involve individuals who do not own mobile phones and have limited access to the Internet. Furthermore, available approaches to remote data collection are restricted in their ability to establish personal connections. Personal connections are more easily developed through face-to-face interaction and can be critical to public health research, e.g. in the case of qualitative research or to quantitative research particularly on sensitive topics. Despite limitations, remote methods can be more efficient than face-to-face data collection and provide platforms to empower individuals to engage in generating and analysing data. Lessons learnt in designing and implementing remote data collection methods in a COVID-19 era are critical to inform future execution of these methods, which are likely to become fundamental to continued research in public health.

References

Ball HL. 2019. Conducting Online Surveys. *Journal of Human Lactation* **35**(3):413–17.

Broich C. 2015. Offline data collection in sub-Sahran Africa using SMS surveys: Lessons learned. Sample Solutions.

Chu DK, Akl EA, Duda S *et al.* 2020. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *The Lancet* **395**: 1973–87.

Copes H, Tchoula W, Brookman F, Ragland J. 2018. Photo-elicitation interviews with vulnerable populations: practical and ethical considerations. *Deviant Behavior* **39**: 475–94.

Dabalen A, Etang A, Hoogeveen A *et al.* 2016. *Mobile Phone Panel Surveys in Developing Countries: A Practical Guide for Microdata Collection.Directions in Development–Poverty*. Washington, DC: World Bank.

Ellis C, Bochner A. 2000. Autoethnography, personal narrative, reflexivity: researcher as subject. In: Denzin N, Lincoln Y (eds). *Handbook of Qualitative Research*. 2nd edn. Thousand Oaks, CA: Sage Publications, 733–68.

Eynon R, Fry J, Schroeder R. 2017. *The SAGE Handbook of Online Research Methods*. London: SAGE Publications Ltd. https://methods.sage pub.com/book/the-sage-handbook-of-online-research-methods-second-edition.

Gibson DG, Pereira A, Farrenkopf BA *et al.* 2017. Mobile phone surveys for collecting population-level estimates in low- and middle-income countries: a literature review. *Journal of Medical Internet Research* **19**: e139.

Greenleaf AR, Gibson DG, Khattar C, Labrique AB, Pariyo GW. 2017. Building the evidence base for remote data collection in low- and middle-income countries: comparing reliability and accuracy across survey modalities. *Journal of Medical Internet Research* **19**: e140.

Heckathorn D. 1997. Respondent-driven sampling: a new approach to the study of hidden populations. *Social Problems* **44**: 174–199, www.jstor.org/stable/3096941, accessed 15 November 2020.

Johnston LG, Sabin K. 2010. Sampling hard-to-reach populations with respondent driven sampling. *Methodological Innovations* **5**: 38–48.

Kirchherr J, Charles K. 2018. Enhancing the sample diversity of snowball samples: recommendations from a research project on anti-dam movements in Southeast Asia. *PLoS One* **13**: e0201710.

L'Engle K, Sefa E, Adimazoya EA, Yartey E *et al.* 2018. Survey research with a random digit dial national mobile phone sample in Ghana: methods and sample quality. *PLoS One* **13**: e0190902.

Labrique A, Blynn E, Ahmed S *et al.* 2017. Health surveys using mobile phones in developing countries: automated active strata monitoring and other statistical considerations for improving precision and reducing biases. *Journal of Medical Internet Research* **19**: e121.

Lamanna C, Hachhethu K, Chesterman S *et al.* 2019. Strengths and limitations of computer assisted telephone interviews (CATI) for nutrition data collection in rural Kenya. *PLoS One* **14**: e0210050.

Langhaug LF, Sherr L, Cowan FM. 2010. How to improve the validity of sexual behaviour reporting: systematic review of questionnaire delivery modes in developing countries. *Tropical Medicine & International Health* **15**: 362–81.

Lau CQ, Cronberg A, Marks L, Amaya A. 2019. In search of the optimal mode for mobile phone surveys in developing countries. A comparison of IVR, SMS, and CATI in Nigeria. *Survey Research Methods* **13**. https://ojs.ub.uni-konstanz.de/srm/article/view/7375, accessed 3 August 2020.

Leo B, Morrel R, Mellon J, Peixoto T, Davenport S. 2015. Do mobile phone surveys work in poor countries? Center for Global Development Working Paper 398. Washington, DC. https://www.cgdev.org/publication/do-mobile-phone-surveys-work-poor-countries-working-paper-398, accessed 29 July 2020 .

Lupton D. 2020. *Doing Fieldwork in a Pandemic: Crowd-Sourced Document.* https://docs.google.com/document/d/1clGjGABB2h2qbduTgfqribHmog9B6P0NvMgVuiHZCl8/edit?ts=5e88ae0a#, accessed 1 July 2020.

Mupambireyi Z, Bernays S. 2019. Reflections on the use of audio diaries to access young people's lived experiences of HIV in Zimbabwe. *Qualitative Health Research* **29**: 680–92.

Peterman A, Bhatia A, Guedes A. 2020. *Remote Data Collection on Violence against Women during COVID-19: A Conversation with Experts on Ethics, Measurement & Research Priorities (Part 1)*. https://www.unicef-irc.org/article/1997-remote-data-collection-on-violence-against-women-during-covid-19-a-conversation-with.html, accessed 5 August 5 2020.

Roy D, Tripathy S, Kar SK, Sharma N *et al.* 2020. Study of knowledge, attitude, anxiety & perceived mental healthcare need in Indian population during COVID-19 pandemic. *Asian Journal of Psychiatry* **51**: 102083. doi:10.1016/j.ajp.2020.102083.

Shaghaghi A, Bhopal RS, Sheikh A. 2011. Approaches to recruiting "hard-to-reach" populations into re-search: a review of the literature. *Heal Promotion Perspectives* **1**: 86–94.

The Lancet Public Health. 2020. COVID-19 puts societies to the test. *The Lancet Public Health* **5**: e235.

Tran MC, Labrique AB, Mehra S *et al.* 2015. Analyzing the mobile "digital divide": changing determinants of household phone ownership over time in rural Bangladesh. *JMIR mHealth and uHealth* **3**: e24.