Research Article

Comparison of Telephone and In-Person Interview Modalities: Duration, Richness, and Costs in the Context of Exploring Determinants of Equitable Access to Community Health Services in Meru, Kenya

Luke N. Allen[1] , Sarah Karanja[2], Malebogo Tlhajoane[1,3] , John Tlhakenelo[4], David Macleod[5], and Andrew Bastawrous[1]

Abstract

Background: Our research team conducted phenomenological interviews in Kenya with people who were not able to access community eye health services, aiming to explore the barriers and ideas for potential service modifications. We conducted an embedded study that compared in-person and telephone interview modalities in terms of time requirements, costs, and data richness. **Methods:** A team of six interviewers conducted 31 in-person interviews and 31 telephone interviews using the same recruitment strategy, topic guide, and analytic matrix for each interview. We compared the mean duration; mean number of themes reported by each participant; total number of themes reported; interviewer rating of perceived richness; interviewer rating of perceived ease of building rapport; number of days taken by the team to complete all interviews; and all costs associated with conducting the interviews in each modality. **Results:** In-person interviews were 44% more expensive and took 60% longer to complete than our telephone interviews (requiring 5 days and 3 days respectively). The average in-person interview lasted 110 seconds longer than the average telephone interview ($p = .05$) and generated more words and themes. However, the full set of interviews from both approaches identified similar numbers of barriers ($p = .14$) and the same number of solutions ($p = .03$). Interviewers universally felt that the in-person approach was associated with better rapport and higher quality data ($p = .01$). Triangulation of themes revealed good agreement, with 88% of all solutions occurring in both sets of interviews, and no areas of thematic dissonance. **Conclusions:** The in-person approach required more time and financial resources, but generated more words and themes per person, and was perceived to afford richer data by interviewers. However, this additional richness did not translate into a greater number of themes that our team can act upon to improve services.

Keywords
qualitative research, in-person interviews, telephone interviews, mode comparison, methods research

[1]Department of Clinical Research, London School of Hygiene and Tropical Medicine, UK
[2]Kenya Medical Research Institute, Kenya
[3]Department for Population Health, London School of Hygiene and Tropical Medicine, UK
[4]Department of Family Medicine and Public Health, University of Botswana, Botswana
[5]Department of Infectious Disease Epidemiology and International Health, London School of Hygiene and Tropical Medicine, UK

Corresponding Author:
Luke N. Allen, Department of Clinical Research, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT, UK.
Email: luke.allen@lshtm.ac.uk

Introduction

Background

Qualitative interviews – especially those grounded in the phenomenological approach - are designed to elicit rich data about participants' lived experiences and perceptions of a given phenomenon. (Pope & Mays, 2020) Various methods can be used to gather these data. In-person interviews are commonly perceived as the 'gold standard' for obtaining rich phenomenological data due to the fact that the interviewer can observe visual cues and quickly build rapport. (Azad, 2021; Novick, 2008; Rahman, 2023; Rubin & Rubin, 2011) However, telephone interviews offer unique advantages: the increased social distance can make it easier for participants to discuss sensitive topics; travel time and interviewer safety concerns are eliminated; power imbalances are partially concealed; and overall costs can potentially be reduced – depending on the specific study design and population. (Novick, 2008; Sturges & Hanrahan, 2004; Vogl, 2013) As access to mobile phones continues to grow across many low and middle-income countries, it would be useful to better understand how telephone interviews compare to those conducted face-to-face.

Our team is conducting a broader, underlying project to explore access to community-based eye services in Kenya, where participants can be spread across vast distances, meaning that the risks, costs, and time-requirements for in-person interviews are likely to compare poorly with telephone interviewing. This project is formed of multiple components, and is still ongoing. In this particular analysis we aimed to assess which interview modality offers the best balance of richness, duration, and costs in the context of our work to explore barriers to access and potential solutions in Meru County, Kenya. Further information on our underlying project is available in the published protocol (Allen et al., 2024).

Mode Comparison

A number of previous studies have sought to compare telephone and in-person interview modalities. (Francis, 2010; Irvine, Drew, & Sainsbury, 2013; Johnson, Scheitle, & Ecklund, 2021; Krouwel, Jolly, & Greenfield, 2019; Rahman, 2023; Sturges & Hanrahan, 2004; Vogl, 2013) In qualitative research, quality is conceptually linked to the 'richness' of the data obtained, described by Charmaz in terms of revealing participants' true feelings, intentions and actions, and accessing their "otherwise inaccessible thoughts". (Charmaz & Henwood, 2003) Many different proxies have been used to approximate richness in mode comparison studies. A crude but relatively common approach is to measure the duration or wordcount of each interview, working from the assumption that longer interviews, with more words spoken, are more likely to provide deeper insights into people's lived experiences. (Sturges & Hanrahan, 2004; Irvine et al., 2013;

Vogl, 2013; Johnson et al., 2021) Interview duration is often used in the same way, based on the assumption that longer interviews generate richer data, with some studies also reporting 'interviewer dominance'; the proportion of the talking that is done by the interviewer as opposed to the participant. (Johnson et al., 2021) Surprisingly few qualitative mode effect studies compare the actual content of the interviews, despite the fact that this is a more nuanced way of assessing the amount of topic-related data that are generated. (Krouwel et al., 2019; Johnson et al., 2021) This approach is also relatively straightforward, requiring the reporting of the total number or unique themes that arise from each set of interviews and/or the mean number of themes identified by each interview.

A further approach entails having researchers subjectively rate their experience of each interview in terms of the perceived richness of the data obtained, as done by Abrams et al. using a simple three-point Likert scale. (Abrams, 2015) Other reported measures include quantifying the word count of associated field notes for each interview and counting the amount of conversational turn-taking that occurs in each interview. (Irvine et al., 2013; Johnson et al., 2021)

Research Objectives

In this study, we aimed to compare the data richness obtained from two sets of in-person and telephone interviews, electing to use a broad range of proxies: interview duration and wordcount; number of themes identified; and subjective interviewer rating of richness and rapport. We aimed to gather additional data on the time taken to complete each set of interviews, and the associated costs. We hypothesised that telephone interviews would be less time-consuming and less expensive to complete than in-person interviews, but offer less-rich data across all metrics of comparison.

Methods

Participant Selection

This study was nested within a broader programme of work to explore barriers and potential solutions to improve equitable access to community-based eye care. (Allen, Nkomazana et al., 2023) In Meru county, Kenya we had previously found that younger adults (aged 18–44 years old) were the least likely to access care. (L. Allen, 2024) We obtained a full list of all of the younger adults who did not receive care from Peek Vision, a partner organisation that provides the screening and patient flow management software for the programme. (Peek Vision, 2018) Peek Vision also provided contact numbers for all participants, under a pre-existing data sharing agreement.

Once we had obtained a full list of all those in the target population who had not accessed care, we used computer-generated random numbers to determine the order in which

participants would be interviewed. After the first 15 interviews had been completed we switched to a maximum variation sampling approach to ensure that we spoke with people from a range of different backgrounds. We also used computer-generated random numbers to assign participants to either in-person or telephone interviews.

Topic Guide for the Underlying Study

We used the same semi-structured topic guide for both interview modalities, (Allen et al., 2024) exploring the factors that prevented each participant from accessing community-based eye care services, and their perceptions of potential solutions or changes we could make to the programme to improve access.

Data Collection for the Underlying Study

Both sets of interviews were conducted by a team of six Kenyan research assistants with training in qualitative methods and previous research experience. These research assistants came from a range of academic research and clinical backgrounds. All were early-to-mid-career researchers. Interviews were conducted in Meru, Kiswahili, or English, depending on participant preference. Interviews were audio recorded and direct quotes were entered into a deductive analytic matrix in English. Further details on the analytical approach are available elsewhere (Allen et al., 2024).

The same team of six research assistants conducted all interviews using the same semi-structured interview guide. The same process for audio recording data and directly transcribing quotes into the analytic matrix was used for both modalities, and the same process of iterative review and analysis across all cases within each modality was used to generate the final themes. Research assistants received two days of training before conducting the interviews in September 2023. Transcriptions were double-checked by local and external research supervisors at daily debrief sessions.

Sample Size

We used thematic saturation to determine our sample size, following Guest and colleagues' approach. (Guest, Namey, & Chen, 2020) After conducting an initial 'base' of at least 12 interviews in each modality, we continued recruiting participants until we had two (consecutive) interviews in a row where no new themes (barriers or potential solutions) were identified. We aimed to compare equal numbers of telephone versus in-person interviews. All of the telephone interviews were completed before any of the in-person interviews were started.

Comparison Domains

Following our protocol, we gathered data on six different domains:

(1) Interview duration: We measured the duration of each interview in minutes from the start of the consenting process until the researcher concluded the interview by thanking the participant for answering all of their questions. In line with previous studies discussed above, this metric was used as a proxy for richness, based on the assumption that longer interviews capture richer data than shorter interviews. Note that we did not use interviewer dominance measures since this is only possible with typed transcripts, and our approach is based around direct-from-audio entry of verbatim quotes.

(2) Matrix wordcount: We counted the total number of words entered into the analytic matrix for each set of interviews. These were verbatim quotes directly transcribed from the audio by the research assistants. In line with previous research, we assumed that a higher wordcount was associated with richer data.

(3) Total number of themes: We counted the total number of unique themes for barriers and solutions that were reported across all interviews with each modality. We assumed that the modality that captured the largest number of unique themes was capturing richer data. From an operational standpoint, our underlying study is primarily concerned with generating potential solutions that will improve equitable access, so the number of unique solutions that emerged from each set of interviews is a particularly important metric.

(4) Number of themes reported by each participant: We also reported the range and mean number of unique themes (barriers and solutions) identified by each participant for each modality. This was to hedge against a situation where one modality generated a greater number of themes than the other, but only because of one or two prolific interviews.

(5) Interviewer subjective rating of richness: After all of the interviews were complete, each of the six research assistants were asked to provide a single global summary rating of the perceived richness obtained from all in-person and all telephone interviews. Following the approach used by previous researchers, we used a simple Likert scale: low = 1, moderate = 2, high = 3. Each of the research assistants was asked to provide their rating based on the prompt: "How would you rate the richness of the data that you were able to gather via telephone?" and "How would you rate the richness of the data that you were able to gather in-person?"

(6) Interviewer subjective rating of rapport: We supplemented the subjective rating of richness with a second question that asked research assistants to provide a global summary rating of the perceived ease of building rapport across all in-person and all telephone interviews. Again, we used a simple Likert scale: low = 1, moderate = 2, high = 3. Each of the research

assistants was asked to provide their rating based on the prompt: "How would you rate the ease of building rapport via telephone?" and "How would you rate the ease of building rapport in-person?"

(7) Time taken to plan and complete all interviews: We documented the total amount of time taken to plan and complete all interviews in each modality to the nearest half-day. This was recorded by the Kenyan research manager in charge of scheduling, supervision, and logistics for the local research activities.

(8) Costs: Working with a health economist, we recorded costs from the payer's perspective. Both modalities use the same sampling and analytic approach, so we only compared costs that were unique to each approach that is those associated with data collection. For telephone interviews these included airtime and staff daily salaries multiplied by the number of days required to complete data collection, starting with the first phone call to recruit the first participant, and ending with the conclusion of the final interview.

For in-person interviews we included the costs of printing consent forms, transport for researchers, transport reimbursement offered to participants; payments for local Community Health Promoters and sub-county health officials to assist with setting up the interviews (mobilisation/sensitisation), and staff daily salaries multiplied by the number of days taken for data collection. The costs of voice recorders were not included in the comparison because they were used for both sets of interviews. Similarly, the same two-day training covered skills required for both interview modalities so this was not included in the comparison. We did not compare overhead costs unless they differed for the modalities. The local research manager also recorded any unforeseen additional costs associated with each modality.

Data Analysis Procedures

We used sign tests for the paired interviewer global ratings of richness and ease of building rapport. We chose the sign test as each interviewer performed two ratings and we were interested in whether there was evidence that they systematically preferred one mode over the other. As the sign test is non-parametric, we didn't need to make any distributional assumptions and given the small sample size and only 3-point scoring scale we felt that this test was more appropriate than a paired t test. For the unpaired mean testing comparisons for (i) interview duration (ii) mean number of barriers identified by a participant, (iii) mean number of solutions identified by a participant, we used histograms to check the data for normality and then used t-tests (for parametric data) or Mann-Whitney-U tests (for non-parametric data), as appropriate, to provide evidence as to whether the two modalities differed across the domains using a = 0.05. We note that hypothesis tests on small

sample sizes are sensitive to bias caused by measurement error.

Methodological Triangulation of Themes

Finally, we compared the barriers and solutions that emerged from both modalities using methodological triangulation; (Arias Valencia, 2022; Denzin, 1978; Kimchi, Polivka, & Stevenson, 1991) a means of assessing agreement between two different approaches that have been used to study the same phenomenon.

We generated a bespoke convergence coding matrix that listed all of the themes identified by both sets of interviews. We then identified themes that had emerged from; both modalities, just the in-person modality, and just the telephone modality. We performed this assessment for all barrier themes, and separately for all solution themes. We presented our findings in terms of 'agreement' (themes that were identified in both sets of interviews, 'silence' (themes that emerged from one set of interviews but not the other), and 'dissonance (where themes from one set conflicted with those from the other).

Findings

We compared data obtained from 31 telephone interviews and 31 in-person by our team of six researchers in September 2023 in Meru county. Table 1 summarises our main findings.

Richness

The average in-person interview lasted 110 seconds longer than the average telephone interview ($p = .05$) and generated 33% more words in the analytic matrix. On average, face-to-face interviews identified a greater number of barriers ($p = .14$) and solutions ($p = .03$), however the entire in-person interview set only identified one additional barrier and the same number of unique solutions as telephone interviews. All six research assistants were unanimous in their ratings of data richness and ease of building rapport, rating in-person interviews as 'high' and telephone interviews as 'moderate' for both measures ($p = .01$).

Time Requirements

It took two days to prepare for the in-person interviews and then three days to complete them. Preparation time included phoning potential participants to invite them to participate, and then scheduling meeting times and places, and organising transport and local logistics. This included working with local Community Health Promoters (CHPs) and sub-county health officials to sensitise and locate interviewees. This is a vital element in building trust and legitimising our work with participants: the CHPs visited each person to discuss the project and answer any questions, and then

Table 1. Performance Characteristics of Each Modality.

Comparison domain	Metric	Modality		p	Ratio
		In-person	Telephone		
Data richness	Mean interview duration in minutes and seconds (range)	10.20 (4.19–15.20)	8.30 (3.10–30.10)	0.005^	1.11
	Analytic matrix wordcount – barriers	4453	2674	N/A	1.67
	Analytic matrix wordcount – solutions	2638	2094	N/A	1.26
	Total number of barriers identified	15	14	N/A	1.07
	Total number of solutions identified	22	22	N/A	1.00
	Mean number of barriers mentioned by each participant (range)	1.94 (0–4)	1.58 (0–3)	0.142^	1.23
	Mean number of solutions mentioned by each participant (range)	2.23 (0–5)	1.61 (0–4)	0.029^	1.39
	Mean interviewer global rating of richness (1–3)	3.0	2.0	0.014*	1.5
	Interviewer global rating of ease of building rapport (1–3)	3.0	2.0	0.014*	1.5
Time requirement	Time taken to organise and complete all interviews	5 days	3 days	N/A	1.67
Costs	Cost to complete all interviews	USD668.29	USD375.71	N/A	1.78

*p-value derived from sign test.
^p-value derived from Mann-Whitney-U tests.

supported the researchers to connect them with the interviewees in the field.

Costs

Telephone interviews required three days of our research assistants' time, plus the airtime used to complete the calls. Higher airtime costs for the telephone modality reflects the fact that phone calls were used for recruitment, consenting, and data collection, whereas the in-person approach only used calls for recruitment. Spending on research assistants' salaries was the same for the in-person interviews – which were also completed in three days - however this modality incurred a number of additional costs. We paid two local county officials to assist with scheduling the in-person interviews, and for sensitization and mobilisation on the days of the interviews. We paid nine Community Health Promoters to build trust, explain the project, and physically locate interviewees. We printed physical consent forms, reimbursed travel to a convenient interview location for our interviewees and paid to transport our research assistants to the same location. Whilst the VIP programme operates across the entire county, at the time that our study was running the programme was operating in Meru town, meaning that all of the interviews were conducted within 15–30 minutes away from our offices. As such, we estimate that the transport costs for research assistants could easily rise by a factor of ten or more for in-person interviews conducted in other parts of the county (Table 2).

Triangulation of Themes

Table 3 presents the 21 unique barriers that were identified across all 62 interviews. There was agreement between in-person and telephone modalities on nine of these barriers (42.3%). There was silence on the remaining 12 (57.7%) with

each modality identifying six unique barriers that did not emerge from the other set of interviews. We found no evidence of thematic dissonance.

Table 4 presents the 25 unique solutions that were identified across all 62 interviews. There was agreement between in-person and telephone modalities on 22 of these barriers (88.0%). There was silence on the remaining three (12.0%) with each modality identifying three unique barriers that did not emerge from the other set of interviews. We found no evidence of thematic dissonance.

Discussion

In this study we examined the quality, costs, and time-requirements of in-person versus telephone modes, based on 62 interviews conducted with young adults who had not been able to access eye care services in Meru, Kenya. Even with serendipitously low transport costs, in-person interviews were almost twice as expensive as telephone interviews and took 1.7 times longer to complete. They delivered longer interviews with more words transcribed into the analytic matrix and more themes identified per interview. Our research assistants universally ascribed higher ratings of richness and ease of building rapport to in-person interviews. However, across both modalities, exactly the same number of unique solutions were identified.

Our findings align with the wider literature. Irvine et al. found that telephone interviews tended to be shorter than in-person interviews, although their study only included 11 interviews in total. (Irvine et al., 2013) Novick's review of the literature found evidence that telephone interviews are generally less expensive and shorter than in-person interviews. (Novick, 2008) In their retrospective mode-effect analysis of 300 interviews, Johnson et al. found that in-person interviews produced longer transcripts and more word-dense field notes,

Table 2. Cost of Telephone and Physical Interviews in US dollar.

Line item	In-person	Telephone
Salaries for two sub-county health officials to schedule in-person interviews	50.10	N/A
Payments to nine community health promoters for sensitization and mobilization activities	103.32	N/A
Printing consent forms	23.29	N/A
Transport for research assistants	50.10	N/A
Transport reimbursement for interviewees	97.06	N/A
Airtime	12.52	37.57
Data collector salaries to complete the interviews	338.15	338.15
Total	674.55	375.73

1 USD = 159.691 KES.

Table 3. Thematic Overlap Across In-Person and Telephone Modalities: Barriers.

	Number of interviews where this barrier was raised	
Barriers	In-person	Telephone
Conflicting work engagement	14	15
Long queue	11	9
Other conflicting engagement	7	3
Transport costs	7	1
Clinic not open at stated times	4	0
Fear	3	0
Perceived cost of eye drops/spectacles	2	2
Distance to the clinic	2	6
Insufficient numbers of staff	2	2
Lack of clear information on clinic opening times	2	0
Lack of clear information on services available	2	0
Insufficient counselling at the point of referral	1	0
Lack of clear information on the clinic appointment date	1	0
Opportunity costs from loss of wages/income	1	2
Sought services elsewhere	1	2
Forgot	0	3
Did not receive the SMS reminder message	0	2
Assumption that their eye problem wouldn't be addressed	0	1
Dislike of crowded places	0	1
Male health seeking behaviour	0	1
Mixed genders in the queue	0	1

but generated the same themes as telephone and videocall-based interviews. (Johnson et al., 2021) Interestingly, subjective interviewer ratings were also similar across the approaches. Krouwel and colleagues also compared in-person interviews to those conducted using video-calling software. They found that in-person interviews generated more data but the overall number of themes derived from each approach was similar. (Krouwel et al., 2019) Vogl's triangulation of the themes that emerged from two sets of interviews with 56 children found negligible differences. (Vogl, 2013) Finally, in his systematic review comparing telephone and in-person approaches, Rahman concluded that both telephone and in-person modalities can generate comparably rich data, with telephone interviews tending to be less time consuming and less expensive. (Rahman, 2023)

Given that empirical mode comparisons consistently find that remote interviews are able to generate similar qualitative themes at lower costs and in shorter time periods than in-person interviews, irrespective of research question and population studied, Rahman has argued that the in-person modality should only be used if the specific research question demands it. (Rahman, 2023) The relationship between depth of detail, number of themes, and agreement between themes is intriguing. Participants tend to provide much more detail about a given phenomena during in-person interviews, as indicated by longer transcripts, interview durations, and

Table 4. Thematic Overlap Across In-Person and Telephone Modalities: Solutions.

Solutions	Number of interviews where this barrier was raised	
	In-person	Telephone
Hold the clinics on additional days	10	9
Add more staff to each clinic	9	4
Hold the clinics in different locations	6	4
Add a greater number of clinics	6	5
Add more drugs and essential supplies	4	2
Introduce phone call reminders	4	2
Explain the costs and services available	3	0
Extend clinic opening hours	3	2
Specify the clinic dates and times	3	0
Provide a door-to-door service	2	1
Provide transport fare	2	1
Provide transport to the clinic	2	1
Issue public reminders	2	1
Schedule fewer people to attend each clinic every day	2	2
Use SMS reminders	2	2
Have sperate clinic queues for men, women, young & elderly	2	1
Improve staff punctuality	2	1
Hold "mop-up" clinics for those who miss their appointment	1	3
Explain clinic importance at point of referral	1	1
Pay people to attend	1	1
Subsidise treatment	1	0
Hold weekend outreach clinics	1	3
Allow people to choose their appointment day	0	2
Provide specific appointment slots	0	1
Refer non-attenders to the next available clinic via SMS	0	1

analytic matrix wordcounts. However, this additional detail does not seem to translate into identification of novel codes or themes when compared to remote approaches.

We found several differences in the themes that emerged from both sets of interviews. Whilst the differences between the solutions was fairly minor, several of the barriers that were raised during the telephone interviews were potentially more candid than those derived from in-person interviews. A form of social desirability bias might have been at play, with interviewees feeling more comfortable disclosing potentially embarrassing or taboo issues when the interviewer was not physically sat in front of them. (Kreuter, Presser, & Tourangeau, 2008; Bispo Júnior, 2022) Some of the barriers that emerged exclusively from the phone interviews included forgetting about the appointment, assuming that the service would not meet their needs, and perceiving the mixing of men and women in a single queue as 'shameful'.

In terms of strengths and limitations, whereas most research in this field tends to employ one or two metrics, our study compared eight different dimensions of performance, including proxies for richness (duration and wordcount), mean and aggregate themes, and subjective interviewer ratings, supplemented with an assessment of costs and time requirements. We conducted a relatively large number of interviews on a topic that is central to global efforts to extend Universal Health Coverage as part of the Sustainable Development Goals. (UN General Assembly, 2015; World Health Organization, 2019, 2021)

The research assistants' subjective ratings of rapport and richness were necessarily subjective and may have been biased towards the in-person modality. The generalisability of our findings is limited by our relatively focused research question and the homogeneity of our population. Ultimately, whilst our study presents multiple measures, we are not able to definitely say which approach is best, as there is no single 'right' way to balance differences in richness, costs, and time requirements. We focused on the telephone modality rather than videocalls because access rates to internet-enabled devices are relatively low in semi-rural Kenya, as they are across the sub-Saharan region, especially when contrasted with access to basic telephone services. (GSMA, 2023a; 2023b) Whilst telephone-based interviews remain an important tool in the qualitative researcher's belt, they eliminate visual cues, can potentially make it harder to build rapport, can lead to sampling bias (as not everyone owns a phone), can make it impossible to safeguard confidentiality if others are in the

room, and can compromise data quality through poor connections and background noise. Obviously the telephone modality cannot be used in the place of site visits or to make observations of the participant's environment. (Burnard, 1994; Vogl, 2013; Wilson, Roe, & Wright, 1998)

Previous research has documented that the impact of qualitative research findings on real-world programmes is influenced by the timeliness of the findings, (Johnson & Vindrola-Padros, 2017; Allen, Azab et al., 2023) and our broader embedded qualitative work places a premium on rapidly identifying barriers and potential solutions to improve equitable access to care within a live, ongoing screening programme. Given our focus on identifying solutions and service modifications that can be rapidly tested, the lack of dissonance between the modes, lower costs, lower time requirements, and additional researcher safety benefits associated with telephone interviews means that we are very likely to continue using this approach.

Our project was conducted in Meru county, Kenya, and focused on access to community-based eye services. Moreover we exclusively engaged with younger adults who needed eye services but had not been able to attend. Our findings are likely to be transferable to other similar cultural, demographic, and geographic settings, but care must be taken not to generalise our findings wholesale. We feel comfortable transferring these findings to other work we are conducting in Kenya's Kwale county, and areas of Botswana, India, and Nepal where we are using exactly the same approach to study the same issues in groups that are struggling to access community-based eye services. We note that our findings align with studies from a range of different contexts, suggesting that the time and cost savings associated with the telephone modality are likely to apply across a broad range of populations and geographies. We are less confident that the themes identified by both approaches will remain comparable across different contexts.

Conclusions

Our set of 31 telephone interviews was completed in less time and at less expense than the 31 in-person interviews. Whilst the in-person modality generated longer interviews and more data, the ultimate number of themes that derived from both sets was nearly identical.

Author Contributions

LA conceived the study with SK. LA and SK performed the data collection and analyses. LA wrote the initial draft. SK and MT provided additional written content. JT provided health economics input and reviewed the manuscript. DM provided statistical input and reviewed the manuscript. AB reviewed the protocol and provided additional methodological comments. All authors reviewed and approved the final version of the manuscript.

ORCID iDs

Luke N. Allen ⓘ https://orcid.org/0000-0003-2750-3575
Malebogo Tlhajoane ⓘ https://orcid.org/0000-0002-0601-5516

References

Abrams, K. M. (2015). Data richness trade-offs between face-to-face, online audiovisual, and online text-only focus groups. *Social Science Computer Review*, *33*(1), 80–96. https://doi.org/10.1177/0894439313519733

Allen, L., Karanja, S., Gichangi, M., Bunywera, C., & Muturi, E. (2024). Identifying barriers and potential solutions to improve equitable access to community eye services in central Kenya: A rapid exploratory sequential mixed methods study. *medRxiv*, *1*(1), 8. https://doi.org/10.1101/2024.03.13.24304156

Allen, L., Karanja, S., Gichangi, M., Bunywera, C., & Rono, H. (2024). *Access to community-based eye services in Meru, Kenya: A cross-sectional equity analysis*. medRxiv. Available at: https://doi.org/10.1101/2024.02.23.24303185

Allen, L. N., & Azab, H. (2023). *Rapid methods for identifying barriers and solutions to improve access to community health services: A scoping review*. BJGP Open. Available at: https://doi.org/10.3399/BJGPO.2023.0047

Allen, L. N. (2024). A protocol for the comparison of telephone and in-person interview modalities: Duration, richness, and costs in the context of exploring determinants of equitable access to community health services in Meru, Kenya. *International Journal of Qualitative Methods*, *23*(1), 16094069241245352. https://doi.org/10.1177/16094069241245352

Allen, L. N., & Nkomazana, O. (2023). Improvement studies for equitable and evidence-based innovation: An overview of the "IM-SEEN" model. *International Journal for Equity in Health*, *22*(1), 116. https://doi.org/10.1186/s12939-023-01915-5

Arias Valencia, M. M. (2022). Principles, scope, and limitations of the methodological triangulation. *Investigación y Educación en Enfermería*, *40*(2), Article e03. https://doi.org/10.17533/udea.iee.v40n2e03

Azad, A. (2021). Conducting in-depth interviews via mobile phone with persons with common mental disorders and multimorbidity: The challenges and advantages as experienced by participants and researchers. *International Journal of Environmental Research and Public Health*, *18*(22), 11828. https://doi.org/10.3390/ijerph182211828

Bispo Júnior, J. P. (2022). Social desirability bias in qualitative health research. *Revista de Saúde Pública*, *56*(1), 101. https://doi.org/10.11606/s1518-8787.2022056004164

Burnard, P. (1994). The telephone interview as a data collection method. *Nurse Education Today*, *14*(1), 67–72. https://doi.org/10.1016/0260-6917(94)90060-4

Charmaz, K., & Henwood, K. (2003). Grounded theory. In J. A. Smith (Ed.), *Qualitative psychology: A practical guide to research methods*. Sage Publications Ltd. Available at: https://doi.org/10.4135/9781848607927

Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods*. McGraw-Hill.

Francis, J. J. (2010). What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychology and Health*, *25*(10), 1229–1245. https://doi.org/10.1080/08870440903194015

GSMA. (2023a). *The mobile economy sub-saharan Africa 2023*. GSMA. Available at: https://event-assets.gsma.com/pdf/20231017-GSMA-Mobile-Economy-Sub-Saharan-Africa-report.pdf

GSMA. (2023b). *The state of mobile internet connectivity report 2023 - mobile for development*. GSMA. Available at: https://www.gsma.com/r/somic/

Guest, G., Namey, E., & Chen, M. (2020). A simple method to assess and report thematic saturation in qualitative research. *PLoS One*, *15*(5), Article e0232076. https://doi.org/10.1371/journal.pone.0232076

Irvine, A., Drew, P., & Sainsbury, R. (2013). "Am I not answering your questions properly?" Clarification, adequacy and responsiveness in semi-structured telephone and face-to-face interviews. *Qualitative Research*, *13*(1), 87–106. https://doi.org/10.1177/1468794112439086

Johnson, D. R., Scheitle, C. P., & Ecklund, E. H. (2021). Beyond the in-person interview? How interview quality varies across in-person, telephone, and skype interviews. *Social Science Computer Review*, *39*(6), 1142–1158. https://doi.org/10.1177/0894439319893612

Johnson, G. A., & Vindrola-Padros, C. (2017). Rapid qualitative research methods during complex health emergencies: A systematic review of the literature. *Social Science & Medicine*, *189*(1), 63–75. https://doi.org/10.1016/j.socscimed.2017.07.029

Kimchi, J., Polivka, B., & Stevenson, J. S. (1991). Triangulation: Operational definitions. *Nursing Research*, *40*(6), 364–366.

Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys: The effects of mode and question sensitivity. *Public Opinion Quarterly*, *72*(5), 847–865. https://doi.org/10.1093/poq/nfn063

Krouwel, M., Jolly, K., & Greenfield, S. (2019). Comparing Skype (video calling) and in-person qualitative interview modes in a study of people with irritable bowel syndrome – an exploratory comparative analysis. *BMC Medical Research Methodology*, *19*(1), 219. https://doi.org/10.1186/s12874-019-0867-9

Novick, G. (2008). Is there a bias against telephone interviews in qualitative research? *Research in Nursing & Health*, *31*(4), 391–398. https://doi.org/10.1002/nur.20259

Peek Vision. (2018). *News, Kenya's Vision Impact Project reaches one million people in its first year*. Peek Vision. Available at: https://peekvision.org/en_GB/news/ (Accessed 13 October 2023).

Pope, C., & Mays, N. (2020) *Qualitative research in health care*, 4th ed. Wiley Blackwell. Available at: https://www.wiley.com/en-gb/Qualitative+Research+in+Health+Care%2C+4th+Edition-p-9781119410836 (Accessed: 5 January 2023).

Rahman, R. B. A. (2023). Comparison of telephone and in-person interviews for data collection in qualitative human research. In *Interdisciplinary Undergraduate Research Journal [Preprint]*. Available at: https://doi.org/10.25417/uic.22217215.v1

Rubin, H. J., & Rubin, I. S. (2011). *Qualitative interviewing: The art of hearing data*. Sage.

Sturges, J. E., & Hanrahan, K. J. (2004). Comparing telephone and face-to-face qualitative interviewing: A research note. *Qualitative Research*, *4*(1), 107–118. https://doi.org/10.1177/1468794104041110

UN General Assembly. (2015). *A/RES/70/1: Transforming our world: The 2030 agenda for sustainable development*. UN General Assembly. Available at: https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E

Vogl, S. (2013). Telephone versus face-toface interviews: Mode effect on semistructured interviews with children. *Sociological Methodology*, *43*(1), 133–177.

Wilson, K., Roe, B., & Wright, L. (1998). Telephone or face-to-face interviews? A decision made on the basis of a pilot study. *International Journal of Nursing Studies*, *35*(6), 314–321. https://doi.org/10.1016/S0020-7489(98)00044-3

World Health Organization (2019) *Thirteenth general programme of work, 2019–2023: Promote health, keep the world safe, serve the vulnerable*. WHO/PRP/18.1. World Health Organization. Available at: https://apps.who.int/iris/handle/10665/324775

World Health Organization (2021) *Universal health coverage UHC*. World Health Organization.