

Collecting mortality data via mobile phone surveys: a non-inferiority randomized trial in Malawi

Chasukwa, Michael^{1,2}
Choko, Augustine T.³
Muthema, Funny¹
Nkahlamba, Mathero M.⁴
Saikolo, Jacob¹
Tlhajoane, Malebogo⁶
Reniers, Georges⁵
Dulani, Boniface^{1,2}
Helleringer, Stéphane^{6,7,*}

¹ Institute of Public Opinion and Research, Zomba (Malawi)

² Chancellor College, Department of Political and Administrative Studies, University of Malawi, Zomba (Malawi)

³ Malawi-Liverpool Wellcome Trust Clinical Research Programme, Blantyre (Malawi)

⁴ Chancellor College, Department of Psychology, University of Malawi, Zomba (Malawi)

⁵ Department of Population Health, London School of Hygiene and Tropical Medicine, London (United Kingdom)

⁶ Program in Social Research and Public Policy, Division of Social Science, New York University, Abu Dhabi (United Arab Emirates)

⁷ Department of Sociology, New York University, New York (United States)

* To whom correspondence should be addressed:

Social Research and Public Policy Program

New York University (NYU) | Abu Dhabi

P.O. Box 129188

Abu Dhabi, United Arab Emirates

Email: sh199@nyu.edu

Introduction: Despite the urgent need for timely mortality data in low-income and lower-middle-income countries, mobile phone surveys rarely include questions about recent deaths. There are concerns that such questions might a) be too sensitive, b) prompt negative/adverse reactions among respondents, c) take too long to ask and/or d) generate unreliable data. We assessed the feasibility of mortality data collection during mobile phone surveys.

Methods: We conducted a non-inferiority trial among a random sample of mobile phone users in Malawi. Participants were allocated 3:1 to an interview about recent deaths in their family (treatment group) or about their economic activity (control group). In the treatment group, half of the respondents completed a short mortality questionnaire, focused on information necessary to calculate recent mortality rates, whereas the other half completed an extended questionnaire that also included questions about symptoms and healthcare use. The primary trial outcome was the cooperation rate. Secondary outcomes included the completion rate, self-reports of negative feelings and stated intentions to participate in future interviews. We also documented the amount of time required to collect mortality data, and we explored the quality of death reports.

Results: The difference in cooperation rates between treatment and control groups was 0.9 percentage points (95% CI = -2.3, 4.1), which satisfied the non-inferiority criterion. Similarly, the mortality questionnaire was non-inferior to the control questionnaire on all secondary outcomes. Collecting mortality data required approximately 2 to 4 additional minutes per reported death, depending on the inclusion of questions about symptoms and healthcare use. More than half of recent deaths elicited during mobile phone interviews had not been reported to the national civil registration system.

Conclusion: Including mortality-related questions in mobile phone surveys appears acceptable and feasible. It might help strengthen the surveillance of mortality trends in low-income and lower-middle-income countries with limited civil registration systems.

KEY QUESTIONS

What is already known?

- In many low-income and lower-middle-income countries, civil registration systems only record a fraction of all deaths. The excess mortality associated with health crises is thus not known in near real-time.
- Mobile phone surveys are increasingly common in low-income and lower-middle-income countries. They could help fill mortality-related data gaps, but there are concerns that asking questions about recent deaths over the phone might be too sensitive, might take too long, and/or might generate unreliable data.

What are the new findings?

- In a randomized trial conducted with mobile phone users in Malawi, asking questions about recent deaths was not less acceptable than asking questions about economic activity and household livelihoods.
- Few participants reported experiencing negative feelings during the interview, and these feelings were temporary.
- More than half of the deaths reported during mobile phone interviews had not been previously registered with the national civil registration system.

What do the new findings imply?

- Including questions about recent deaths in mobile phone surveys appears feasible and acceptable.
- It might help strengthen the surveillance of mortality trends in low-income and lower-middle-income countries with limited civil registration systems.

INTRODUCTION

In many low-income and lower-middle-income countries (LLMICs), few deaths are registered with competent administrative authorities in a timely manner [1–3]. In such settings, mortality statistics are only updated every few years, after periodic household surveys or censuses are conducted. The retrospective data these inquiries generate allow estimating mortality levels for periods stretching a few months to a few years prior to data collection [5], but they do not allow monitoring mortality in near real-time. Household surveys and censuses are also often postponed or canceled during epidemics or other crises, due to elevated risks of disease transmission or heightened safety concerns [6].

In most LLMICs, the data available to track the mortality impact of health crises are obtained from more partial data collection systems. For example, the counts of deaths routinely reported during epidemics such as Ebola or COVID-19 often only include the deaths that occur among those who have been diagnosed with the disease. Yet, the coverage of testing services is very limited in LLMICs [7]. Some COVID-19 cases may be lost to follow-up after diagnosis, and surveillance systems predominantly record deaths that occur at health facilities [8], even though many deaths occur at home [9,10]. Epidemics also indirectly affect mortality, for example by disrupting health services [11,12]. As a result, reported counts of deaths only include a fraction of the excess mortality caused by an epidemic [15]. This might foster perceptions that global health crises (e.g., COVID-19) have “spared” LLMICs, and it might preclude countries from effectively advocating for resources required to mitigate the impact of such crises [16].

In the medium to long-term, strengthening civil registration systems is the key intervention required to address this data gap [17,18]. Increasing death registration is thus a key indicator of

progress towards the sustainable development goals (SDGs). In the short term however, interim data sources are needed to better understand how ongoing crises affect population health in LLMICs. The World Health Organization (WHO) has recommended setting up rapid mortality surveillance systems to monitor the death toll of the COVID-19 pandemic [19], for example by extracting data about deaths from registers maintained by health facilities, or by tallying the number of burials taking place at local cemeteries. New data sources (e.g., satellite images, social media) might also help keep track of increasing numbers of deaths [20,21]. Rapid mortality surveillance systems have documented the previously unknown extent of excess mortality caused by the COVID-19 pandemic in several LLMICs [16].

Unfortunately, setting up such surveillance is complex. It requires significant investments in data acquisition and management. Baseline data that pre-date health crises might be difficult to obtain, for example if historical records were not properly kept at a health facility or cemetery. Behaviors related to the management of deaths (e.g., where and when to carry out a burial) might also change during epidemics and other crises, due to lockdowns and restrictions on mobility or access to health facilities. In such contexts, short-term fluctuations in mortality documented by surveillance systems might be due to behavioral changes as well as excess mortality. Finally, rapid mortality surveillance systems have been predominantly established in large urban areas [22]. Small towns and rural areas are less often included in such initiatives.

Additional tools are needed to better document the total impact of health crises on mortality in LLMICs with limited death registration. We investigated the use of mobile phone surveys (MPS) to collect mortality data. MPS are surveys in which participants are recruited and interviewed entirely by phone [23]. They are increasingly conducted in LLMICs, for example to monitor risk factors for non-communicable diseases [24] or to document the effects short-term fluctuations in economic activity, schooling or healthcare use [25,26]. They present several advantages over

other modes of data collection. Because they are implemented remotely, they remove the need for physical interactions between data collectors and participants, and can be sustained during epidemics [27]. They can also be repeated more frequently than household surveys or censuses, since they require less complicated logistics. Finally, owing to the rapid penetration of mobile phones in rural areas of LLMICs, they might allow monitoring mortality trends outside of large cities more conveniently than other rapid surveillance systems.

Several MPS have already included questions about mortality. During the Ebola epidemic in west Africa, a survey in Monrovia (Liberia) asked randomly selected phone owners about recent deaths in their households [6]. Since the beginning of the COVID-19 pandemic, surveys in India and Bangladesh have monitored recent household deaths [28,29]. In particular, a question about recent COVID deaths was included in a large phone survey, which allowed estimating that the death toll related to the COVID-19 pandemic in India might be 6-7 times larger than officially reported [30].

Such examples of mortality-related MPS remain however isolated. There are concerns that questions about recent deaths might be too sensitive to ask during phone interviews, thus prompting high levels of refusals to participate, or to answer specific questions. In some instances, questions about deaths might upset respondents or trigger distress. Whereas such reactions can be addressed when interviews occur in person (e.g., through signs of empathy), they may be more difficult to mitigate remotely. Questions about deaths might also take too long to administer. There are strict recommendations to keep the duration of MPS short [31], and eliciting mortality data about deaths with sufficient detail requires time. Existing mortality-related MPS have thus only included limited ascertainties of deaths and their circumstances. Finally, MPS interviewers might not be able to probe and cross-check answers provided by respondents

as thoroughly as during in-person interviews. As a result, the mortality data generated by MPS might be of lesser quality.

We report results from a randomized trial of the collection of mortality data via MPS recently conducted in Malawi.

METHODS

Trial design: we compared responses to mobile phone interviews about recent deaths (treatment group), to those obtained in mobile phone interviews about economic activity (control group). The control questionnaire has been used in multiple LLMICs since the beginning of the COVID-19 pandemic, with high levels of reported acceptability [25]. We conducted a non-inferiority randomized trial [32], in which we tested whether the treatment questionnaire did not lead to unacceptably worse response patterns than the control questionnaire. Within the treatment group, participants were randomly assigned to either a short or an extended questionnaire on deaths. The short mortality questionnaire only ascertained information required to measure and triangulate recent mortality rates. The extended mortality questionnaire also included questions about symptoms and circumstances of reported deaths.

Study setting: the trial was conducted in Malawi, a low-income country in southeastern Africa, with a population of approximately 18 million inhabitants. Despite a legal obligation to register vital events, and the possibility to register events in decentralized offices and locations, few deaths are reported to the National Registration Bureau, i.e., the authority in charge of vital records. As a result, data series based on death registration, typically used to monitor short-term fluctuations in mortality in high-income countries, are not available in Malawi [33]. The most recent household surveys that collected mortality data in Malawi occurred between 2015

and 2017, whereas the national census was conducted in 2018. Since the beginning of the COVID-19 pandemic, several mobile phone surveys have been conducted in Malawi [34–36], but they have not investigated recent mortality trends.

Participants: Study participants were mobile phone users aged 18-64 years old at the time of the study. Mobile phone users younger than 18 years old were excluded due to practical difficulties in obtaining parental consent during MPS. Mobile phone users aged 65 years and older were excluded in part due to difficulties in reaching such age groups during MPS. In addition, people aged 65 years and older are typically excluded from household surveys that collect data on mortality. For example, the demographic and health surveys (DHS) only include women aged 15-49 years old, and men aged up to 54 years old [37]. The population-based HIV impact assessments (PHIA) include individuals aged up to 64 years old [38].

We recruited study participants among users of Malawi’s two major mobile networks through random digit dialing (RDD). We worked with Sample Solutions, a firm specializing in the provision of RDD samples. Sample Solutions first generated a list of phone numbers at random using the country’s numbering scheme. They then matched this list to a global registry of authorized network subscribers, and excluded numbers that could not be located. Finally, a team of 15 interviewers contacted the selected numbers to introduce the study, assess the eligibility of mobile phone users who were reached, and ask for their consent to participate in interviews.

We implemented sampling quotas based on age, gender and regional residence. We formed 18 sampling strata based on these characteristics, and enrollment continued in each stratum until the quota was filled or until progress towards this quota stopped. All interviews were conducted in local languages (Chichewa, Chitumbuka, ChiYao) or in English, depending on mobile users’

preferences. If a mobile user did not speak any of these languages, he/she was not included in the trial.

Randomization: after interviewers successfully contacted a mobile phone number, they introduced themselves to the user and stated that they were conducting a survey about the impact of the COVID-19 pandemic on several aspects of the lives of Malawians. They asked the phone user if they were interested in learning more about the study. If so, they assessed the user's eligibility by asking 4 questions about gender, age and residence. Mobile users who did not meet the age criterion and those whose sampling stratum was already filled were told that they were not eligible for the study. Eligible mobile users were randomized into the treatment and control groups in a 3:1 ratio.

At that time, interviewers read consent scripts and sought oral consent from mobile users. In both study groups, the consent scripts included similar explanations of how phone numbers were selected, and statements about that the interview would take approximately 15 minutes to complete. Consent scripts in both groups also mentioned that participants would receive a small amount of airtime (1,200 Malawian Kwachas, or 1.5 US Dollar) as a token of appreciation upon completion of the interview. Since the main themes of the interview differed between treatment and control groups, the description of other study procedures included differed between the two groups (see supporting documents).

Study interviewers were unaware of the mobile users' assignment to the treatment or control group until the users' sampling stratum had been determined and their eligibility had been confirmed. Randomization was stratified by sampling stratum. It was conducted using random numbers generated in Stata 15.1. Within the treatment group, randomized allocation to the short vs. extended mortality questionnaires was conducted in a similar manner in a 1:1 ratio.

Outcomes: The primary outcome of the trial was the cooperation rate [39]. It was calculated among mobile users randomized to either the treatment or control group, and it was defined as the number of completed interviews divided by the sum of completed interviews, call-backs, partial interviews and refusals. Secondary outcomes included a) the completion rate, i.e. the number of completed interviews divided by the sum of completed interviews and partial interviews and b) the proportion of respondents who stated their willingness to participate in future interviews of varying durations. We also measured c) the proportion of respondents who self-reported that some of the survey questions made them upset.

Sample size: We used standard formulae for the determination of sample size in non-inferiority trials [40]. Based on preliminary pilot data, we assumed that a cooperation rate of 86 per cent in the control group and we set a non-inferiority margin of 5 percentage points. Our null hypothesis was that the cooperation rate in the treatment group is lower than in the control group by more than 5 percentage points. Our alternative hypothesis was that the cooperation rate in the treatment group is lower than in the control group by less than 5 percentage points. To test this hypothesis with 80% power and $\alpha = 0.05$, we required at least 1,194 mobile subscribers in the treatment group and 398 mobile subscribers in the control group. Sample size calculations were performed using the `SampleSize4ClinicalTrials` package in R.

Study instruments and procedures: in each study group, questionnaires elicited respondents' socio-economic characteristics such as their age, marital status and educational level, as well as their household's access to water and electricity. In the treatment group, respondents were asked to list deaths that had occurred among members of their households since the beginning of 2021, and to indicate the survival status of their parents and (maternal) siblings.

Additional questions were included for respondents who had reported that one of their parents or siblings had died since 2019. Respondents allocated to the short mortality questionnaire were solely asked to state the age at death, date of death (month and year) and status of their relative's death in the national civil registration system. Respondents allocated to the long mortality questionnaire were also asked if their deceased relative(s) had experienced symptoms commonly observed in COVID-19 patients (e.g., cough, fatigue, loss of taste/smell), and whether they had sought healthcare in the weeks prior to death. Finally, they were asked to indicate the place of death (e.g., at the hospital or at home) and place of burial of their deceased relatives. In the control group, instead of questions about deaths, respondents were asked to report their recent economic activities, to list the sources of livelihood of their household, and to describe how they manage their finance (e.g., ownership of bank accounts).

In both study groups, we asked respondents about their reactions to the interview. We adapted debriefing questions used in post-disaster surveys (e.g., terror attacks) to identify respondents who might have experienced negative feelings during the interview [41]. Participants who self-reported negative feelings were asked to classify these feelings on a 3-point scale ranging from "a little upset" to "very upset". They were also asked if they were still upset by these questions at the end of the interview, or if they were "okay now". Finally, we asked interviewers to indicate if they noticed signs of emotional distress during the interview (e.g., crying, long silences, voice alterations).

We offered psychological support to all respondents who self-reported ongoing negative feelings, or who were identified by interviewers as having displayed signs of emotional distress. If a participant indicated being interested in such a service, the study interviewer transmitted the participant's phone number to an on-call clinical psychologist. This practitioner then called the

referred participant to a) assess their psychological well-being, b) provide phone-based counseling, and c) if needed, provide required referrals for further follow-up.

Finally, respondents in both groups were provided details about who to contact in case of questions, and were asked if they wanted to receive information about SARS-CoV-2. If so, study interviewers read a short script that included details about modes of transmission, symptoms and strategies to prevent infection/transmission. They also indicated how to get additional information through resources provided by the Ministry of Health (e.g., hotline, social media).

Data collection: Interviewers were trained for six days. This included a review of recruitment, screening and consent procedures, and explanations of survey questions in each study group. Training also included sessions about the use of tablets for mobile data collection, and the development of skills to conduct sensitive interviews and detect potential signs of distress among mobile respondents. After mock interviews, interviewers conducted a 3-day pilot. The trial was then conducted between September 21st, 2021 and October 12th, 2021.

All data were collected on tablets using surveyCTO. We recorded how interviewers administered key parts of a randomly selected subset of interviews, including the consent statement, the ascertainment of deaths in the treatment group, and the assessment of negative reactions to the interview. Study supervisors listened to these audio-files to monitor compliance with study instructions, and provided feedback to study interviewers on the basis of these recordings. Study supervisors also placed follow-up calls to a 1 in 15 sample, which included a) all participants who experienced negative feelings during the interview, and b) a random sample of respondents who did not experience such reactions. During these follow-up calls, supervisors re-assessed participants' reactions to the interview. They also offered psychological support to

participants who indicated experiencing ongoing negative feelings, and gathered other feedback participants had about their interview.

Ethical clearance: the study was approved by the institutional review boards of New York University – Abu Dhabi (HRPP-2021-93) and the University of Malawi (UNIMAREC, P .07/21/76). The protocol was registered prospectively in the American Economic Association's Registry of randomized Controlled Trials (#0008065).

Patient and public involvement: participants in the 3-day pilot provided feedback that informed study procedures. Other members of the study population were not directly involved in the design of the study. Dissemination of study results is ongoing with local stakeholders in Malawi.

Statistical methods: We conducted pre-specified intent-to-treat analyses of trial data. We measured primary and secondary outcomes according to participants' assigned study groups, then we computed the differences in those outcomes between treatment and control groups. We calculated two-sided confidence intervals around these differences [42], and we assessed the position of these confidence intervals relative to the pre-specified non-inferiority margin (-5 percentage points). In computing confidence intervals, we assumed unequal variances in estimates of proportions between the two study groups.

We explored whether respondents with recent deaths among their relatives experienced negative feelings more frequently. We reported the number of participants who asked to obtain counseling and support from our on-call psychologist. In the treatment group, we described the amount of time required to collect data on mortality during mobile phone interviews. We estimated a linear regression model in which interview duration (in minutes) was the dependent variable, and predictors included a binary variable denoting the type of mortality questionnaire

(short vs. long), a categorical variable indicating how many recent deaths were reported by the respondent (none, one or two or more deaths) and an interaction term between those two variables.

We investigated the quality of data on the characteristics of deaths reported during the study. We measured the proportion of reported deaths with missing data on age at death and month of death. We assessed whether these proportions varied by source of the death report (e.g., household deaths, parental deaths or siblings' deaths). Due to small sample sizes in some categories, we used exact methods to calculate confidence intervals [43]. We described the time series of deaths reported during the survey, by source of the death report. Finally, we measured the proportion of reported deaths that were also registered in the national civil registration system, by year of death and source of the death report.

RESULTS

Enrollment, participant flow and baseline characteristics

Study interviewers dialed more than 7,000 unique mobile numbers (figure 1). They reached 3,054 mobile users. However, 698 mobile users (22.8%) immediately indicated that they were not interested in the study. Only 5 mobile users (0.2%) were excluded due to language-related reasons. Among mobile users whose eligibility was assessed ($n = 2,318$), 69 did not meet the age-related inclusion criteria (2.9%), and 566 were excluded because their quota had already been filled (24.4%). In total, 1,683 mobile users were randomized to the treatment or control groups.

[FIGURE 1 ABOUT HERE]

Due to preset quotas, the study sample included similar numbers of men and women (table 1). More than 1 in 4 participants was aged 18-24 years old, whereas 1 in 7 participants was aged 45 years and older. Approximately 20% of the study sample was located in the northern region of Malawi. The southern and central regions each accounted for approximately 40% of the sample. Twenty-eight per cent of the mobile users who were randomized to the treatment or control groups resided in a city, whereas 49.3% of the mobile users described their place of residence as a town or trading center, and 22.7% were residents of villages. There were no important differences in background characteristics between mobile users allocated to the control vs. treatment groups.

[TABLE 1 ABOUT HERE]

The distribution of interview results among randomized mobile users is presented in table A1. The overall cooperation rate was 92.2% (1552/1683). Seventy-eight mobile users refused to provide consent (4.6%). Thirty-two mobile users (1.9%) consented but opted to be called at another time to complete the interview (“call-backs”). Study interviewers could not reach them again despite multiple attempts. Finally, 21 mobile users discontinued their interview (1.2%).

Primary outcome: In the treatment group, the cooperation rate was 92.4% (1,174/1,270) vs. 91.5% in the control group (378/413). This difference of 0.9 percentage points (95% CI = -2.3 to 4.1) meets our criterion for non-inferiority (figure 2).

[FIGURE 2 ABOUT HERE]

Secondary outcomes: In the treatment group, 39 out of 1,213 mobile users who consented to participate did not complete their interview (96.8% completion rate, table A1) vs. 14 out of 392 mobile users in the control group (96.4% completion rate). Similarly, 1,134 of the 1,174 members of the treatment who completed the interview did not self-report that they were upset by some of the interview questions (96.6%) vs. 367 out of 378 in the control group (97.1%). Finally, 94.7% of participants in the treatment group and 93.7% in the control group stated their intention to participate again in (hypothetical) similar interviews in the near future. The non-inferiority criterion was met for all pre-specified secondary outcomes (figure 2).

Referrals to psychological counseling: Fifty-one participants self-reported being upset by (some of) the interview questions. In the treatment group, 12 out of 40 participants who self-reported negative feelings indicated that they were “very upset” by some of the interview questions. Participants with a recent death among their relatives reported being “very upset” more frequently than other participants (figure A1). In the control group, all participants who self-reported negative feelings (n = 11) indicated being either “a little upset” or “moderately upset”.

Among respondents who self-reported negative feelings, 3 stated that they were still upset at the end of the interview (all in the treatment group). Study interviewers identified 6 additional participants who displayed signs of distress during the interview, but who did not self-report being upset (3 in the treatment group, and 3 in the control group). In follow-up calls with study supervisors that took place 1-2 days after the interview, a few respondents who did not initially self-report negative feelings indicated that some of the questions asked during the interview had made them upset (table A2, n=7). None of these respondents however reported that these negative feelings persisted at the time of their call with study supervisors. In total, 9 respondents were informed about the possibility to talk to an on-call clinical psychologist who would provide support and information. Three participants (2 in the treatment group, and 1 in the control group)

accepted this offer and were contacted. After the call with the clinical psychologist, none of those 3 participants required additional referrals and follow-up.

Interview duration: On average, study participants in the treatment group required 18 minutes to complete all study procedures, including screening, consent and debriefing (figure A2).

Participants who did not report any recent death required on average 16 minutes to complete study procedures (figure 3). Among participants allocated to the short mortality questionnaire, those who reported one recent death among their household members or relatives required an additional 1.9 minute to complete the interview, whereas those who reported 2 or more such deaths required 4.4 additional minutes. Among those who were assigned to the extended mortality questionnaire, similar estimates of additional time required were 4.0 minutes and 12.3 minutes, respectively.

[FIGURE 3 ABOUT HERE]

Data quality: The characteristics of recent deaths reported by mobile respondents are described in figure 4. The likelihood of missing data on age at death and month of death varied by source of death report, with lowest levels observed among household deaths (0-3%) and highest levels observed among deaths of siblings (>10%, panel A). Counts of reported siblings' deaths varied between 2 and 12 per quarter, whereas reported parental deaths varied between 4 and 14 per quarter (figure 4, panel B). Questions on household deaths generated between 9 and 21 reports of deaths per quarter in 2021. Between half and two thirds of recent deaths reported during the survey were not registered with the national registration bureau, depending on the source of the report and the year of death (figure 4, panel C).

[FIGURE 4 ABOUT HERE]

DISCUSSION

In this randomized trial in Malawi, a mortality-related questionnaire administered by mobile phone was non-inferior to a questionnaire about economic activity previously administered in several LLMICs, and recognized as highly acceptable. Cooperation and completion rates were high among respondents asked to answer questions about recent deaths in their households and families. The frequency of self-reported negative feelings was low (<3%), and it was not heightened in respondents who were asked mortality-related questions. Respondents who reported a recent family or household death during the interview experienced stronger negative feelings, but these feelings were transitory. The few respondents who opted to talk to an on-call clinical psychologist did not require further follow-up. In this trial, concerns that questions related to mortality might be too sensitive to ask during MPS were thus not realized.

We included questions about mortality that emulated those asked during household surveys and censuses. Our questionnaires were thus more detailed than instruments included in previous MPS, which focused on specific types of deaths (e.g., Ebola or COVID deaths), or only ascertained deaths among members of the respondent's household. Asking questions about mortality required an additional 2-4 minutes per reported death, depending on the inclusion of questions about circumstances, symptoms and healthcare use around the time of death. This time investment concerns only a small group of survey participants, since >80% of respondents did not report any recent death in our survey.

Mortality-related phone interviews generated data that appeared of good quality. Missing data on age at death, and date of death, were limited. They were more frequent in reported deaths of siblings, and least frequent in reported deaths of household members. The levels of missing data we observed were comparable to those observed in many household surveys conducted

in-person [44]. In this sample, the mortality questions generated reports of approximately 8.4 deaths per quarter over the past 3 years, with marked increases in the most recent time period (i.e., the first 3 quarters of 2021) when questions about household deaths were applicable. Future work should investigate the sample sizes required in mobile surveys to allow detecting short-term fluctuations in mortality.

Our data indicate that 50-65% of the deaths that were reported during the mobile interview had not been previously registered with the National Registration Bureau, i.e., the administrative unit in charge of civil registration in Malawi. Conducting mortality-related MPS might thus supplement existing data collection systems. It might also document short-term fluctuations in mortality among population groups that might otherwise be frequently excluded from data collection (e.g., small towns and rural villages).

Our trial has several limitations. First, our sample size was too small to assess data quality in more detail. We could not measure heaping in reported ages of siblings [45], nor could we investigate the reporting of deaths on shorter time scales (e.g., weeks). Second, some of the outcomes we considered (e.g., negative feelings, death registration) were based on self-reported data. They might have been affected by social desirability biases. If the extent of such biases did not differ by trial group, our assessments of the non-inferiority of the mortality questionnaire are however unaffected. In addition, we implemented robustness checks (e.g., audio-recordings, supervisor follow-ups) to enhance the reliability of these data. Finally, we did not explore the acceptability of more extensive mortality questionnaires that might enable the attribution of causes of deaths [46], nor did we investigate the measurement of child mortality through the collection of birth or pregnancy histories [47].

The development of MPS as a reliable tool for mortality surveillance in LLMICs also requires investigating the selectivity of samples recruited by mobile phone. These samples exclude population members who do not have access to a phone. They disproportionately include younger and more urban respondents. In some settings, they predominantly include males, who are more likely to be phone owners. Patterns of phone ownership and utilization might also change over the course of an epidemic, for example if poor households are forced to sell mobile phones due to hardship, or are less able to afford charging their phone (and may not be reachable). Population in the rural areas may not be reachable because their phones would be off as a result of not charging the battery associated with erratic power supply. Statistical models that account for such selectivity should be developed.

Despite limitations and ongoing research needs, our work in Malawi suggests that MPS are a potentially useful tool for mortality surveillance in LLMICs with limited civil registration systems.

SOURCES OF FUNDING: This study was supported by funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R01HD088516, PI: Helleringer), the National Institute on Aging (R03AG070660, PI: Helleringer), the Bill and Melinda Gates Foundation (INV-023211, PI: Reniers) and New York University-Abu Dhabi.

REFERENCES

- 1 Haider MM, Alam N, Ibn Bashir M, *et al.* Adult death registration in Matlab, rural Bangladesh: completeness, correlates, and obstacles. *Genus* 2021;**77**:13. doi:10.1186/s41118-021-00125-7
- 2 Mikkelsen L, Phillips DE, AbouZahr C, *et al.* A global assessment of civil registration and vital statistics systems: monitoring data quality and progress. *Lancet* 2015;**386**:1395–406. doi:10.1016/s0140-6736(15)60171-4
- 3 Setel PW, Macfarlane SB, Szreter S, *et al.* A scandal of invisibility: making everyone count by counting everyone. *Lancet* 2007;**370**:1569–77. doi:10.1016/S0140-6736(07)61307-5
- 4 Hill K, Choi Y, Timæus I. Unconventional approaches to mortality estimation. *Demographic Research* 2005;**S4**:281–300. doi:10.4054/DemRes.2005.13.12
- 5 Masquelier B, Reniers G, Pison G. Divergences in trends in child and adult mortality in sub-Saharan Africa: survey evidence on the survival of children and siblings. *Popul Stud (Camb)* 2014;**68**:161–77. doi:10.1080/00324728.2013.856458
- 6 Kuehne A, Lynch E, Marshall E, *et al.* Mortality, Morbidity and Health-Seeking Behaviour during the Ebola Epidemic 2014-2015 in Monrovia Results from a Mobile Phone Survey. *PLoS Negl Trop Dis* 2016;**10**:e0004899. doi:10.1371/journal.pntd.0004899
- 7 Hasell J, Mathieu E, Beltekian D, *et al.* A cross-country database of COVID-19 testing. *Sci Data* 2020;**7**:345. doi:10.1038/s41597-020-00688-8
- 8 Riffe T, Acosta E, the COVERAGE-DB team. Data Resource Profile: COVERAGE-DB: a global demographic database of COVID-19 cases and deaths. *International Journal of Epidemiology* 2021;**50**:390–390f. doi:10.1093/ije/dyab027
- 9 Chisumpa VH, Odimegwu CO, De Wet N. Adult mortality in sub-saharan Africa, Zambia: Where do adults die? *SSM Popul Health* 2017;**3**:227–35. doi:10.1016/j.ssmph.2017.02.001
- 10 Lankoandé B, Duthé G, Soura A, *et al.* Returning home to die or leaving home to seek health care? Location of death of urban and rural residents in Burkina Faso and Senegal. *Global Health Action* 2018;**11**:1475040. doi:10.1080/16549716.2018.1475040

- 11 Wilhelm JA, Helleringer S. Utilization of non-Ebola health care services during Ebola outbreaks: a systematic review and meta-analysis. *J Glob Health* 2019;**9**:010406. doi:10.7189/jogh.09.010406
- 12 Delamou A, Delvaux T, El Ayadi AM, *et al.* Public health impact of the 2014-2015 Ebola outbreak in West Africa: seizing opportunities for the future. *BMJ Glob Health* 2017;**2**:e000202. doi:10.1136/bmjgh-2016-000202
- 13 Connolly E, Boley EJ, Fejfar DL, *et al.* Childhood immunization during the COVID-19 pandemic: experiences in Haiti, Lesotho, Liberia and Malawi. *Bull World Health Organ* 2022;**100**:115-126C. doi:10.2471/BLT.21.286774
- 14 Walker PGT, White MT, Griffin JT, *et al.* Malaria morbidity and mortality in Ebola-affected countries caused by decreased health-care capacity, and the potential effect of mitigation strategies: a modelling analysis. *The Lancet Infectious Diseases* 2015;**15**:825–32. doi:10.1016/S1473-3099(15)70124-6
- 15 Dorrington RE, Moultrie TA, Laubscher R, *et al.* Rapid mortality surveillance using a national population register to monitor excess deaths during SARS-CoV-2 pandemic in South Africa. *Genus* 2021;**77**:19. doi:10.1186/s41118-021-00134-6
- 16 Whittaker C, Walker PGT, Alhaffar M, *et al.* Under-reporting of deaths limits our understanding of true burden of covid-19. *BMJ* 2021;**375**:n2239. doi:10.1136/bmj.n2239
- 17 AbouZahr C, Bratschi MW, Cercone E, *et al.* The COVID-19 Pandemic: Effects on Civil Registration of Births and Deaths and on Availability and Utility of Vital Events Data. *Am J Public Health* 2021;**111**:1123–31. doi:10.2105/AJPH.2021.306203
- 18 Rao C, John AJ, Yadav AK, *et al.* Subnational mortality estimates for India in 2019: a baseline for evaluating excess deaths due to the COVID-19 pandemic. *BMJ Global Health* 2021;**6**:e007399. doi:10.1136/bmjgh-2021-007399
- 19 Setel P, AbouZahr C, Atuheire EB, *et al.* Mortality surveillance during the COVID-19 pandemic. *Bull World Health Organ* 2020;**98**:374. doi:10.2471/BLT.20.263194
- 20 Koum Besson ES, Norris A, Bin Ghouth AS, *et al.* Excess mortality during the COVID-19 pandemic: a geospatial and statistical analysis in Aden governorate, Yemen. *BMJ Glob Health* 2021;**6**:e004564. doi:10.1136/bmjgh-2020-004564
- 21 Warsame A, Bashiir F, Freemantle T, *et al.* Excess mortality during the COVID-19 pandemic: a geospatial and statistical analysis in Mogadishu, Somalia. *International Journal of Infectious Diseases* 2021;**113**:190–9. doi:10.1016/j.ijid.2021.09.049
- 22 Mwananyanda L, Gill CJ, MacLeod W, *et al.* Covid-19 deaths in Africa: prospective systematic postmortem surveillance study. *BMJ* 2021;**372**:n334. doi:10.1136/bmj.n334
- 23 Gibson DG, Pereira A, Farrenkopf BA, *et al.* Mobile Phone Surveys for Collecting Population-Level Estimates in Low- and Middle-Income Countries: A Literature Review. *J Med Internet Res* 2017;**19**:e139. doi:10.2196/jmir.7428

- 24 Hyder AA, Wosu AC, Gibson DG, *et al.* Noncommunicable Disease Risk Factors and Mobile Phones: A Proposed Research Agenda. *J Med Internet Res* 2017;**19**:e133. doi:10.2196/jmir.7246
- 25 Egger D, Miguel E, Warren SS, *et al.* Falling living standards during the COVID-19 crisis: Quantitative evidence from nine developing countries. *Science Advances* 2021;**7**:eabe0997. doi:10.1126/sciadv.abe0997
- 26 Morse B, Grépin KA, Blair RA, *et al.* Patterns of demand for non-Ebola health services during and after the Ebola outbreak: panel survey evidence from Monrovia, Liberia. *BMJ Global Health* 2016;**1**:e000007. doi:10.1136/bmjgh-2015-000007
- 27 Hensen B, Mackworth-Young CRS, Simwinga M, *et al.* Remote data collection for public health research in a COVID-19 era: ethical implications, challenges and opportunities. *Health Policy Plan* 2021;**36**:360–8. doi:10.1093/heapol/czaa158
- 28 Bamezai A, Banaji M, Gupta A, *et al.* Survey evidence of excess mortality in Bihar in the second COVID-19 surge. 2021. doi:10.31235/osf.io/zxq97
- 29 Barnwal P, Yao Y, Wang Y, *et al.* No excess mortality detected in rural Bangladesh in 2020 from repeated surveys of a population of 81,000. 2021. doi:10.1101/2021.05.07.21256865
- 30 Jha P, Deshmukh Y, Tumbe C, *et al.* COVID mortality in India: National survey data and health facility deaths. *Science* 2022;**eabm5154**. doi:10.1126/science.abm5154
- 31 Gourlay S, Kilic T, Martuscelli A, *et al.* Viewpoint: High-frequency phone surveys on COVID-19: Good practices, open questions. *Food Policy* 2021;**105**:102153. doi:10.1016/j.foodpol.2021.102153
- 32 Kaji AH, Lewis RJ. Noninferiority Trials: Is a New Treatment Almost as Effective as Another? *JAMA* 2015;**313**:2371–2. doi:10.1001/jama.2015.6645
- 33 Kanyuka M, Ndawala J, Mleme T, *et al.* Malawi and Millennium Development Goal 4: a Countdown to 2015 country case study. *The Lancet Global Health* 2016;**4**:e201–14. doi:10.1016/S2214-109X(15)00294-6
- 34 Kao K, Lust E, Dulani B, *et al.* The ABCs of Covid-19 prevention in Malawi: Authority, benefits, and costs of compliance. *World Development* 2021;**137**:105167. doi:10.1016/j.worlddev.2020.105167
- 35 Banda J, Dube A, Brumfield S, *et al.* Knowledge, risk perceptions, and behaviors related to the COVID-19 pandemic in Malawi. *Demographic Research* 2021;**44**:459–80. doi:10.4054/DemRes.2021.44.20
- 36 Chikoti L, Vundru WD, Fuje HN, *et al.* Monitoring COVID-19 Impacts on Households in Malawi: Findings from the First Round of the High-Frequency Phone Survey. Washington, DC: : World Bank Group 2020. <http://documents.worldbank.org/curated/en/591551597706342578/Findings-from-the-First-Round-of-the-High-Frequency-Phone-Survey> (accessed 30 Sep 2020).

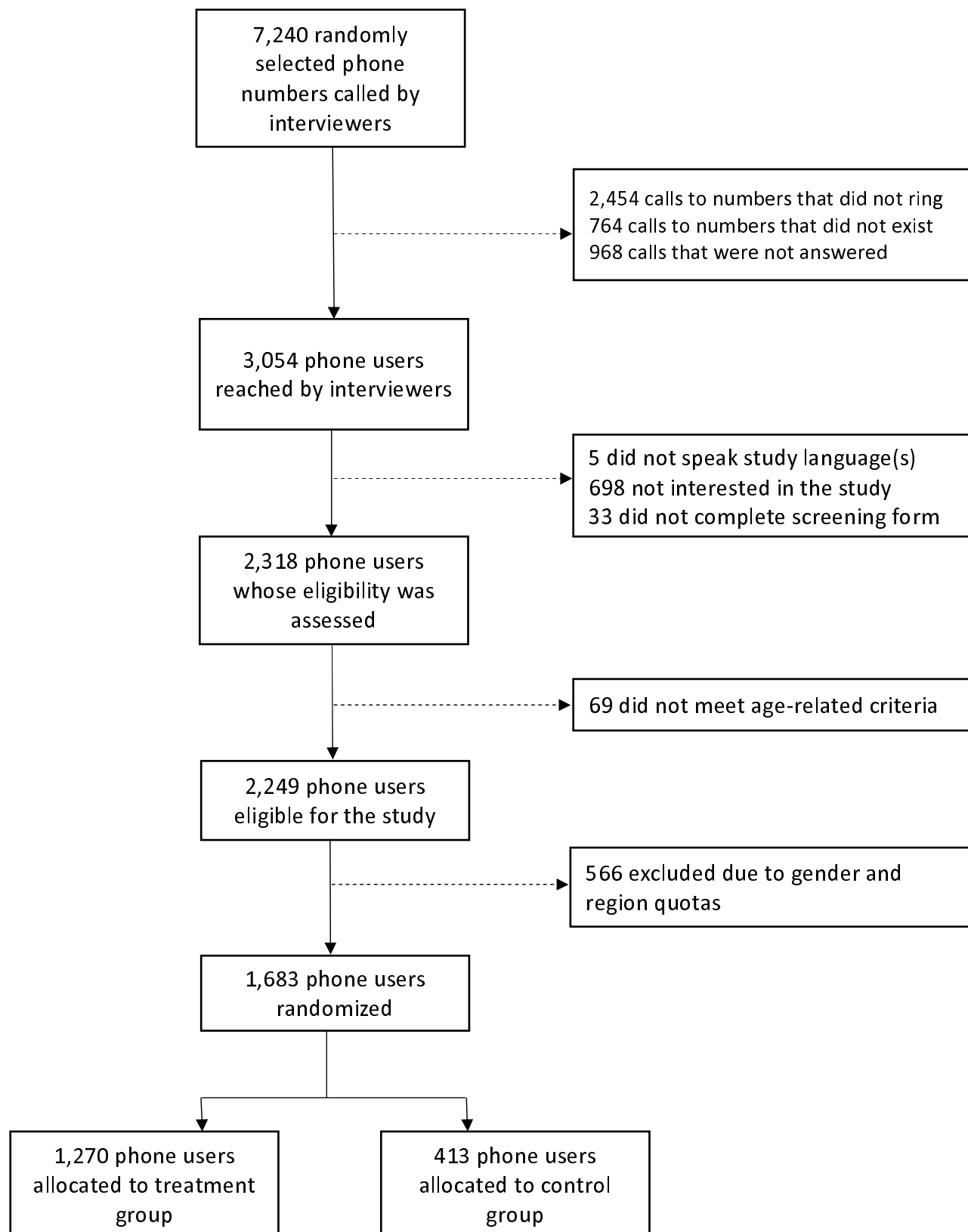
- 37 Corsi DJ, Neuman M, Finlay JE, *et al.* Demographic and health surveys: a profile. *Int J Epidemiol* 2012;**41**:1602–13. doi:10.1093/ije/dys184
- 38 Sachathep K, Radin E, Hladik W, *et al.* Population-Based HIV Impact Assessments Survey Methods, Response, and Quality in Zimbabwe, Malawi, and Zambia. *JAIDS Journal of Acquired Immune Deficiency Syndromes* 2021;**87**:S6. doi:10.1097/QAI.0000000000002710
- 39 Gibson DG, Wosu AC, Pariyo GW, *et al.* Effect of airtime incentives on response and cooperation rates in non-communicable disease interactive voice response surveys: randomised controlled trials in Bangladesh and Uganda. *BMJ Global Health* 2019;**4**:e001604. doi:10.1136/bmjgh-2019-001604
- 40 Schumi J, Wittes JT. Through the looking glass: understanding non-inferiority. *Trials* 2011;**12**:106. doi:10.1186/1745-6215-12-106
- 41 Galea S, Nandi A, Stuber J, *et al.* Participant reactions to survey research in the general population after terrorist attacks. *Journal of Traumatic Stress* 2005;**18**:461–5. doi:10.1002/jts.20053
- 42 Dunn DT, Copas AJ, Brocklehurst P. Superiority and non-inferiority: two sides of the same coin? *Trials* 2018;**19**:499. doi:10.1186/s13063-018-2885-z
- 43 Vollset SE. Confidence intervals for a binomial proportion. *Statistics in Medicine* 1993;**12**:809–24. doi:10.1002/sim.4780120902
- 44 Ahmed S, Li Q, Scraftford C, *et al.* An assessment of DHS maternal mortality data and estimates. Published Online First: 1 September 2014. <https://dhsprogram.com/publications/publication-mr13-methodological-reports.cfm> (accessed 12 Jan 2022).
- 45 Lyons-Amos M, Stones T. Trends in Demographic and Health Survey data quality: an analysis of age heaping over time in 34 countries in Sub Saharan Africa between 1987 and 2015. *BMC Research Notes* 2017;**10**:760. doi:10.1186/s13104-017-3091-x
- 46 Duarte-Neto AN, Marinho M de F, Barroso LP, *et al.* Rapid Mortality Surveillance of COVID-19 Using Verbal Autopsy. *Int J Public Health* 2021;**66**:1604249. doi:10.3389/ijph.2021.1604249
- 47 Akuze J, Blencowe H, Waiswa P, *et al.* Randomised comparison of two household survey modules for measuring stillbirths and neonatal deaths in five countries: the Every Newborn-INDEPTH study. *The Lancet Global Health* 2020;**8**:e555–66. doi:10.1016/S2214-109X(20)30044-9

	Treatment (N=1270)	Control (N=413)	Overall (N=1683)
Gender			
Male	650 (51.2%)	209 (50.6%)	859 (51.0%)
Female	620 (48.8%)	204 (49.4%)	824 (49.0%)
Age group			
18-24y	317 (25.0%)	107 (25.9%)	424 (25.2%)
25-34y	479 (37.7%)	170 (41.2%)	649 (38.6%)
35-44y	290 (22.8%)	80 (19.4%)	370 (22.0%)
45-54y	134 (10.6%)	41 (9.9%)	175 (10.4%)
55-64y	50 (3.9%)	15 (3.6%)	65 (3.9%)
Region			
Northern	260 (20.5%)	72 (17.4%)	332 (19.7%)
Central	481 (37.9%)	180 (43.6%)	661 (39.3%)
Southern	529 (41.7%)	161 (39.0%)	690 (41.0%)
Area of residence			
City	356 (28.0%)	115 (27.8%)	471 (28.0%)
Town/Trading Centre	621 (48.9%)	209 (50.6%)	830 (49.3%)
Villages	293 (23.1%)	89 (21.5%)	382 (22.7%)

Table 1: characteristics of study participants, by study group

Notes: figures in parentheses are column percentages. The variables included in this table are extracted from the screening form, which was completed by all randomized participants.

Figure 1: Flow chart of study enrollment



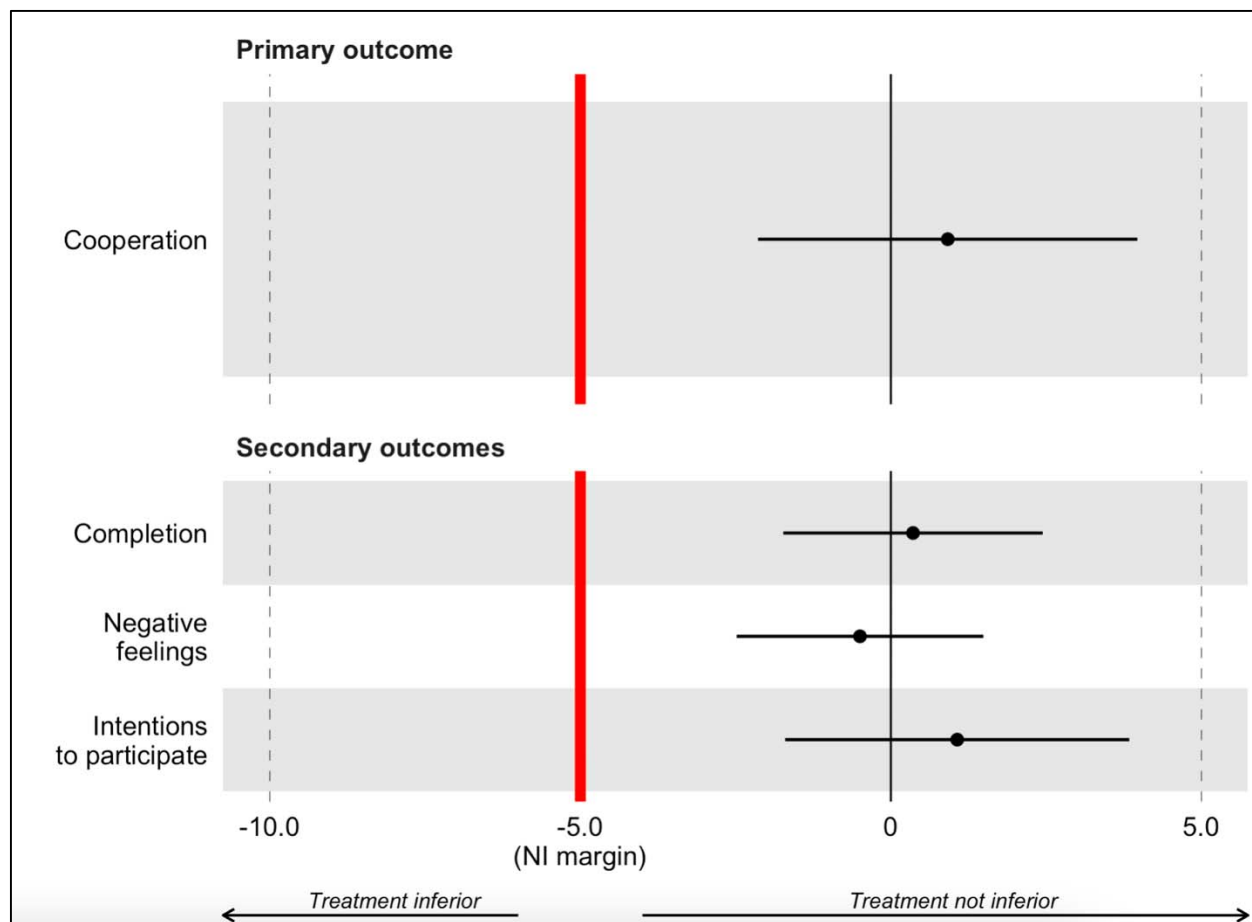


Figure 2: Differences in study outcomes between treatment and control groups

Notes: Values of the x-axis are expressed in percentage points. They are calculated as % treatment group minus % control group. Error bars represent two-sided 95% confidence intervals around the difference in proportions between study groups. The non-inferiority criterion is met when the confidence interval remains to the right of the non-inferiority margin (red vertical line). Similar results were obtained when calculating one-sided confidence intervals. NI = Non-inferiority.

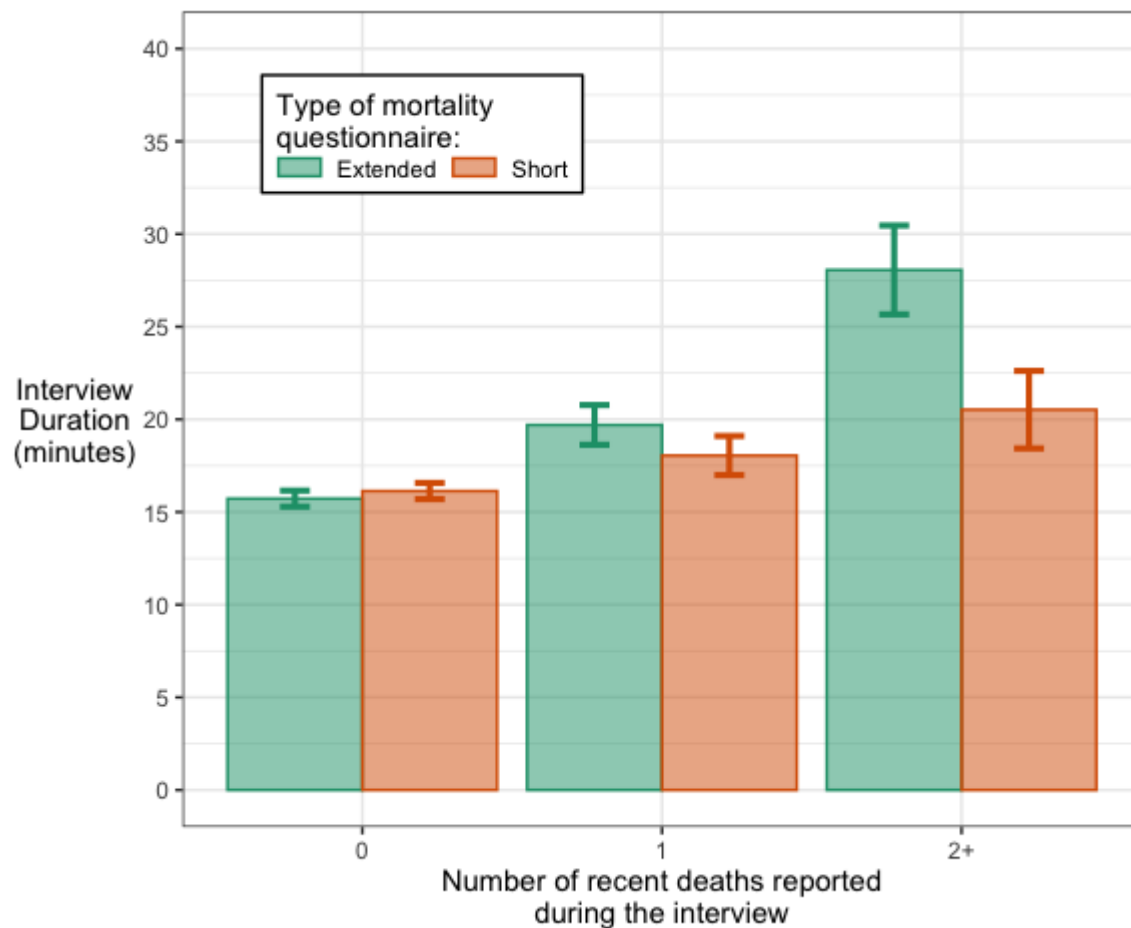


Figure 3: Duration of treatment group interviews, by type of mortality questionnaire and recent deaths among respondents' relatives.

Notes: error bars represent 95% confidence intervals obtained from linear regression models.

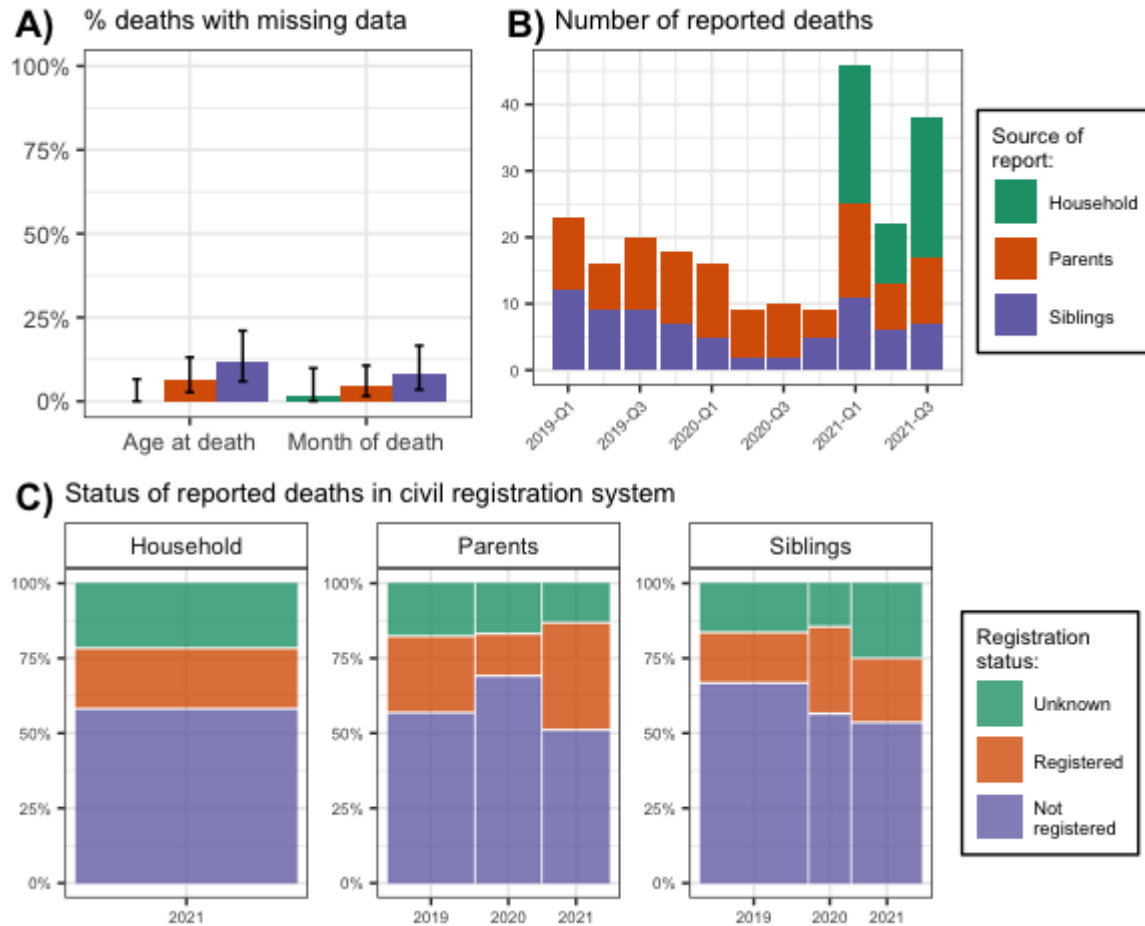


Figure 4: Exploratory assessments of the quality of reported data on recent deaths.

Notes: In Panel A), error bars represent 95% exact confidence intervals. The difference in the likelihood of missing data on age at death by source of report was significant at the $p < 0.05$ level.

In Panel B), Q1 refers to the first quarter of a year, i.e., January, February and March. Two deaths reported to have occurred in October 2021 by respondents interviewed after October 1st, 2021 were omitted from the plot. Household deaths were only elicited for the period between the start of 2021 and the survey data.

In Panel C), the width of each bar is proportional to the number of deaths reported in each year.

	Treatment (N=1270)	Control (N=413)	Overall (N=1683)
Study result			
Completed interview	1174 (92.4%)	378 (91.5%)	1552 (92.2%)
Refusal	57 (4.5%)	21 (5.1%)	78 (4.6%)
Busy/call-back	24 (1.9%)	8 (1.9%)	32 (1.9%)
Interview discontinued	15 (1.2%)	6 (1.5%)	21 (1.2%)

Table A1: distribution of study results, by study group

Notes: “Busy/call-back” refers to respondents who consented to being interviewed, indicated that they would prefer being called-back at a later time, and could not be reached again before the completion of the study.

		Interviewer data	
		No negative feelings	Some negative feelings
Supervisor follow-up	Treatment group		
	No negative feelings	57 (91.9)	4 (22.2)
	Some negative feelings	5 (8.1)	14 (78.8)
	Total	62 (100.0)	18 (100.0)
	Control group		
	No negative feelings	15 (88.2)	0 (0.0)
	Some negative feelings	2 (11.8)	7 (100.0)
Total	17 (100.0)	7 (100.0)	

Table A2: concordance of self-reported data on negative feeling

Notes: supervisor follow-ups were conducted 1-2 days after the initial interview. The supervisors were not aware of the answers provided about self-reported feelings during the initial interview; nor were they aware of the respondents’ assignment to the control or treatment groups.

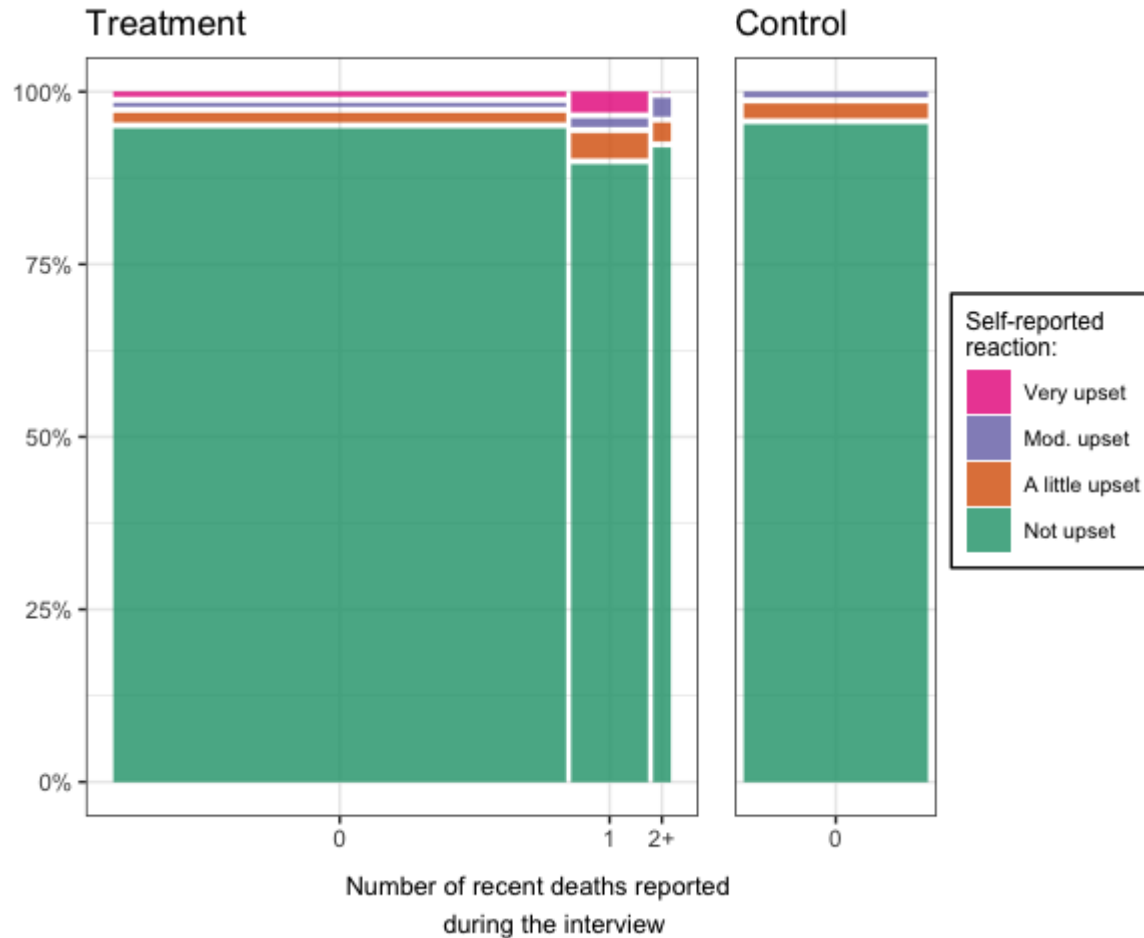


Figure A1: Strength of self-reported negative feelings experienced during the interview, by reports of recent deaths and study group.

Notes: respondents were first asked if any of the interview questions made them upset, and if so, they were asked to indicate the severity of their negative feelings. In the treatment group, the strength of negative feelings was associated with recent reports of deaths ($p=0.02$). In the treatment group, the width of each bar is proportional to the number of respondents reporting different numbers of deaths during the interview.

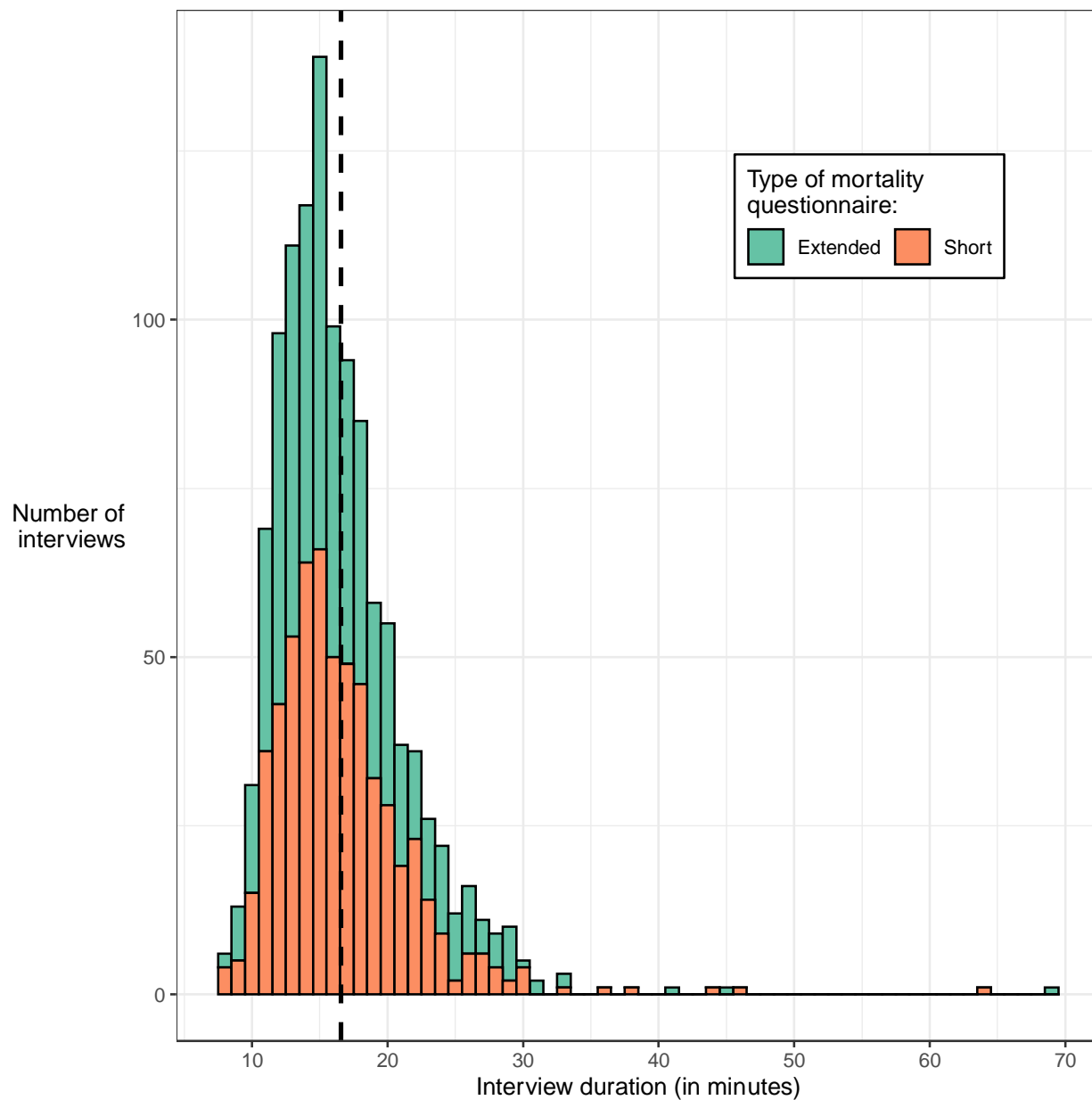


Figure A2: Interview duration in the treatment group, by type of mortality questionnaire administered

Notes: the dashed vertical line represent the mean duration of interviews in the treatment group, regardless of questionnaire type.