BMJ Global Health

Effect of airtime incentives on response and cooperation rates in non-communicable disease interactive voice response surveys: randomised controlled trials in Bangladesh and Uganda

Dustin G Gibson,[1] Adaeze C. Wosu,[2] George William Pariyo,[1] Saifuddin Ahmed,[3] Joseph Ali,[1,4] Alain B Labrique,[1] Iqbal Ansary Khan,[5] Elizeus Rutebemberwa,[6] Meerjady Sabrina Flora,[5] Adnan A Hyder[1,7]

ABSTRACT

Background The global proliferation of mobile phones offers opportunity for improved non-communicable disease (NCD) data collection by interviewing participants using interactive voice response (IVR) surveys. We assessed whether airtime incentives can improve cooperation and response rates for an NCD IVR survey in Bangladesh and Uganda.

Methods Participants were randomised to three arms: a) no incentive, b) 1X incentive or c) 2X incentive, where X was set to airtime of 50 Bangladesh Taka (US$0.60) and 5000 Ugandan Shillings (UGX; US$1.35). Adults aged 18 years and older who had a working mobile phone were sampled using random digit dialling. The primary outcomes, cooperation and response rates as defined by the American Association of Public Opinion Research, were analysed using log-binomial regression model.

Results Between 14 June and 14 July 2017, 440 262 phone calls were made in Bangladesh. The cooperation and response rates were, respectively, 28.8% (353/1227) and 19.2% (580/3016) in control, 39.2% (370/945) and 23.9% (507/2120) in 50 Taka and 40.0% (362/906) and 24.8% (532/2148) in 100 Taka incentive groups. Cooperation and response rates, respectively, were significantly higher in both the 50 Taka (risk ratio (RR) 1.36, 95% CI 1.21 to 1.53) and (RR 1.24, 95% CI 1.12 to 1.38), and 100 Taka groups (RR 1.39, 95% CI 1.23 to 1.56) and (RR 1.29, 95% CI 1.16 to 1.43), as compared with the controls. In Uganda, 174 157 phone calls were made from 26 March to 22 April 2017. The cooperation and response rates were, respectively, 44.7% (377/844) and 35.2% (552/1570) in control, 57.6% (404/701) and 39.3% (508/1293) in 5000 UGX and 58.8% (421/716) and 40.3% (535/1328) in 10 000 UGX groups. Cooperation and response rates were significantly higher, respectively in the 5000 UGX (RR 1.29, 95% CI 1.17 to 1.42) and (RR 1.12, 95% CI 1.02 to 1.23), and 10 000 UGX groups (RR 1.32, 95% CI 1.19 to 1.45) and (RR 1.15, 95% CI 1.04 to 1.26), as compared with the control group.

Summary box

What is already known?
► We published a literature review that documented the use of mobile phone surveys to collect population-representative estimates in low-income and middle-income countries (LMICs).
► We identified six LMIC studies that examined the effect of airtime incentives on an interactive voice response (IVR) survey completion and their results were inconclusive.

What are the new findings?
► We employed a random digit dial sampling method and a standardised protocol to evaluate the effect of airtime incentives in Uganda and Bangladesh.
► We found that the small and large incentives similarly improved response and cooperation rates of an IVR survey.
► The provision of an airtime incentive approach cost-neutrality, with respect to the control arm, by reducing the number of incomplete interviews.

What do the new findings imply?
► Our study suggests that small airtime incentives may be a useful tool to improve mobile phone survey participation in LMIC.
► Although we observed a skewed distribution of complete interviews in favour of younger males, future studies may consider employing quota sampling to increase survey representativeness to obtain national estimates.

Conclusion In two diverse settings, the provision of an airtime incentive significantly improved both the cooperation and response rates of an IVR survey, with no significant difference between the two incentive amounts.
Trial registration number NCT03768323.

INTRODUCTION

More than two-thirds of global deaths are now attributable to non-communicable disease (NCD), with most of these deaths occurring in low-income or middle-income countries (LMICs).[1 2] The majority of NCDs are attributed to four behavioural, and largely modifiable, risk factors by WHO: tobacco use, excessive alcohol consumption, unhealthy diet and inadequate physical activity.[3] In order to prevent premature deaths from NCDs, effective surveillance and population-level prevention efforts are vital.[4] Current NCD surveillance activities include the WHO's STEPwise approach to Surveillance (WHO STEPS) survey, which has collected nationally representative population estimates of self-reported behavioural risk factors and physical and biochemical measurements since 2002.[5] Despite WHO recommendations to implement and repeat STEPS surveys every 3–5 years, as of 2015, only 27 countries had conducted at least two STEPS or STEPS-aligned national surveys.[6] The current paucity of timely data hinders efforts to strengthen NCD surveillance and programme implementation.[7]

For several decades, high-income countries have supplemented in-person data collection with telephone surveys, such as the Behavioural Risk Factor Surveillance Survey.[8] These have been validated as unbiased sources of population-level data on behavioural risk factors for a number of chronic conditions.[9–11] With the exponential growth of mobile phone access and ownership in LMICs, from 22.9 mobile subscriptions per 100 people in 2005 to 98.7 in 2016,[12] there is an opportunity to complement traditional household surveys with more frequent surveillance through the use of mobile phone surveys (MPS).

This rapid change opens the possibility of reaching large segments of the populations of LMICs using MPS. These surveys include interactive voice response (IVR), where a participant uses a mobile phone keypad to enter responses to a prerecorded and automated questionnaire. IVR surveys have been used to collect subnational[13 14] and national estimates[15] of health and demographic indicators in LMICs. However, optimal implementation of MPS in these contexts remains poorly understood, particularly in regard to increasing participation rates and ensuring equitable data collection from different population segments.[16] Although monetary incentives have been found to increase response and cooperation rates of mail, telephone and interviewer-mediated surveys in high-income countries,[17 18] the evidence for their use in LMICs is insufficient.[13]

The objective of this study was to assess the effect of airtime incentives on cooperation and response rates of a random digit dial, NCD behavioural risk factor IVR survey in Bangladesh and Uganda using two randomised controlled trials. This data should help inform the future use of MPS in LMIC settings.

METHODS

Study design and participants

Two randomised controlled trials were conducted in Bangladesh and Uganda, where mobile phone subscription rates were 83 and 55 subscribers per 100 people, respectively.[12] Participants were randomised into one of three airtime incentive amounts: 1) no incentive; 2) 1X incentive or 3) 2X incentive, where X was equal to 50 Bangladeshi Taka (US$0.60) or 5000 Ugandan Shillings (UGX; US$1.35 as of 3 April 2018). The conduct, analysis and reporting of results were done in accordance with the Consolidated Standards of Reporting Trials guidelines.[19]

Participants were sampled using random digit dialling (RDD) method.[20] The country codes for Bangladesh (880) and Uganda (256) and the three-digit sequence specific to mobile network operators were used as the phone number's base. The remaining seven digits were then randomly generated to create a 10-digit telephone number. Participants who indicated an age ≥18 years were eligible to participate. Age-eligible participants were asked to 'press 1' on their mobile phones if they agreed to participate, after hearing a consent disclosure statement. We placed no quotas on age or gender categories.

A detailed description of the methods and protocol has been reported.[21] The trials are registered with ClinicalTrials.gov, number NCT03768323.

Randomisation and masking

Participants were randomised to one of the three study arms after selecting a preferred language for survey completions, but before consenting to the survey. This randomisation was automated and done by the IVR platform (Viamo). Due to the nature of the intervention and study design, study participants were not masked to their study group allocation (the assigned airtime amount was described during the survey introduction). Data cleaning and analyses were done by a statistician blinded to the allocation.

Procedures

In each country, IVR surveys were administered between 08:00 and 20:00 hours local time and sent only once to each randomly sampled phone number. The survey was programmed to have a designated local number appear on the respondent's caller ID screen. RDD participants who picked up the phone were first asked to select a survey language (eg, "If you would like to take the survey in English, press 1"). For Bangladesh, the survey was offered in Bangla and English; for Uganda, languages offered were Luganda, Luo, Runyakitara and English. The survey comprised the following parts: 1) language selection, 2) survey introduction, 3) age-screening questions, 4) consent, 5) demographic questions and 6) NCD modules. The five NCD modules were groups of topically similar questions and included: 1) tobacco, 2) alcohol, 3) diet, 4) physical activity and 5) blood pressure and diabetes. The order of the NCD modules was

randomised to minimise attrition bias, but to preserve skip patterns, questions within each NCD module were not randomised. Airtime incentives, as told to participants during the survey introduction, were only given to those who completed the entire IVR survey. Participants did not incur any expenses (ie, airtime used) for taking the survey.

Outcomes

Complete interviews (I) were defined as participants who answered at least four of the five NCD modules and partial interviews (P) as those who answered one, two or three NCD modules. Non-interviews (R) were classified as either refusals or break-offs. Refusals were defined as age-eligible participants who either did not press a button on their mobile phone to indicate consent or who terminated the survey at the consent question. Participants who were age-eligible, consented, but did not complete an NCD module were classified as break-offs. Respondents who initiated the survey but did not answer the age question were defined as unknowns (U). We computed the estimated proportion of age-eligible respondents (*e*) from among those who were screened for age-eligibility but remained of unknown status. Ineligible participants were defined as either respondents who indicated an age <18 years or phone numbers that were dialled, but it could not be ascertained if the phone number was active.

The primary outcomes were response rate and cooperation rate, using equations #4 and #1, respectively, as defined by the American Association for Public Opinion Research (AAPOR).[22] Cooperation rate was defined as the number of complete interviews divided by the sum of complete, partial and non-interviews; in other words, the proportion of complete interviews from those who were age-eligible. Response rate was defined as the number of complete and partial interviews divided by the sum of complete and partial interviews, non-interviews and the estimated proportion of age-eligible unknowns. Secondary outcomes included contact rate #2 and refusal rate #2 from AAPOR. See online supplementary appendix 1 for the AAPOR equations used in primary and secondary outcomes.

Statistical analysis

In both countries, the same assumptions were used to calculate sample sizes. With a control arm cooperation rate of 30%, an alpha of 0.05 (type 1 error) and statistical power of 80%, 376 individuals were needed to complete an IVR survey in each arm in order to detect an absolute 10% difference in survey cooperation rates between two study arms. As each trial contained three arms, the total sample size per trial was 1128 complete surveys. Sample sizes were not inflated for multiple comparisons, which is recommended by Rothman[23] because of the concern that reducing type I error for a null association increases the type II error probabilities for non-null associations, and thus increasing the sample size was not needed.

For each trial, risk ratios (RRs) and corresponding 95% CIs for contact, response, refusal and cooperation rates were calculated for the incentive arms compared with the control group using log-binomial regression.[24] To assess the heterogeneity of incentive effects on cooperation rates by various demographic characteristics, the log-binomial models were extended and interaction terms were tested. We did not assess any possible heterogeneity for response rates given that its equation includes disposition codes of unknowns where respondents do not provide any demographic information. DerSimonian and Laird random-effects meta-analysis was used to calculate pooled RRs for the different incentive amount subgroups and the heterogeneity statistic was estimated by the Mantel-Haenszel method.[25] Demographic characteristics of complete and partial interviews were compared using χ^2 test. We conducted a sensitivity analysis for contact, response and refusal rates where we did not apply estimated proportion of age-eligible to unknown disposition codes. Analyses were done with STATA/SE (V.14.1; StataCorp, College Station, Texas, USA). An alpha of 0.05 was assumed for all tests of statistical significance.

We calculated the direct delivery cost per complete survey, which included the cost of airtime used to deliver the survey and the incentive amount, as applicable. We summed the total call durations by arm and multiplied by the per-minute cost of airtime—US$0.04 in Bangladesh and US$0.10 in Uganda—to produce the estimated cost per completed survey.
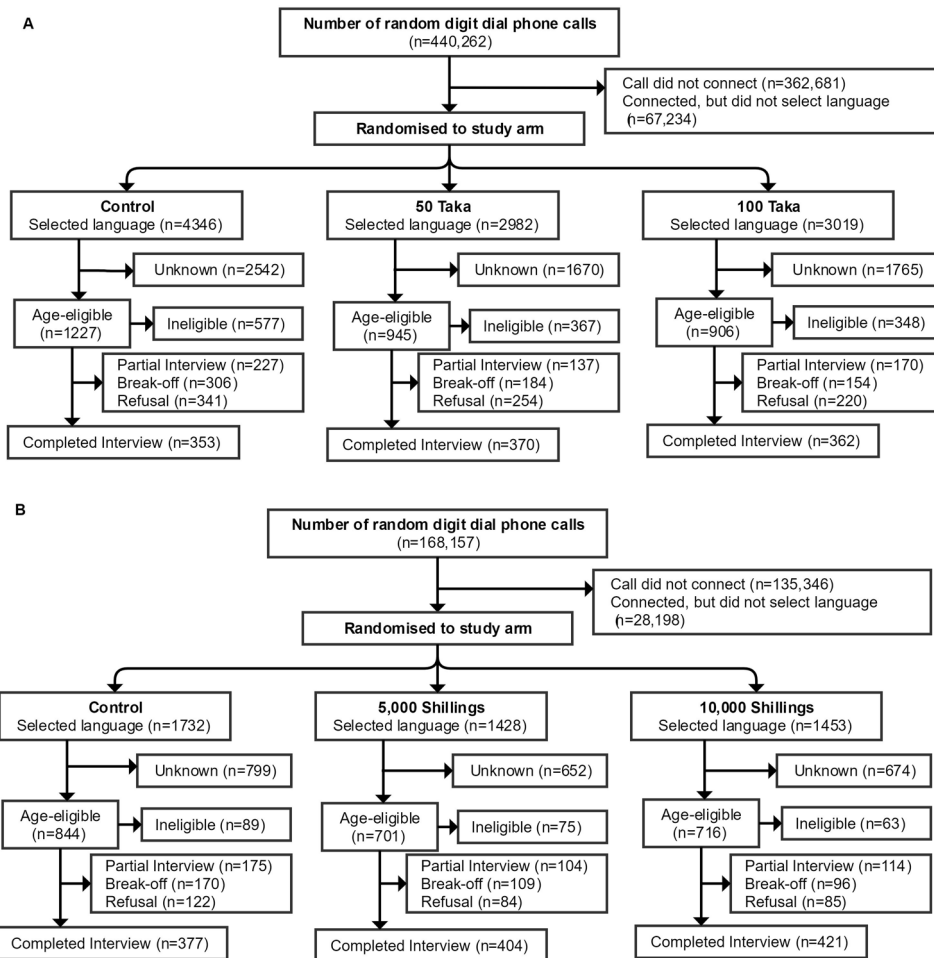
Patient involvement

In 2015, experts in NCDs, mobile health and survey methodology selected questions from standardised surveys such as WHO STEPS and Tobacco Questions for Surveys to be included in the IVR survey.[21] The initial questionnaire was tested at Johns Hopkins University for cognitive understanding and usability of the IVR platform with people who identified as being from an LMIC.[26] Then, in each country, a series of key informant interviews, focus group discussions and user-group testing were conducted to identify appropriate examples to be used in the questions (ie, local fruits and types of physical activity, ensure the questionnaire was comprehensible and that the IVR platform was usable).

RESULTS

From 14 June to 14 July 2017, 440 262 RDD phone calls were made in Bangladesh (figure 1A). Sociodemographic characteristics of the complete interviews were similar across the control (n=353), 50 Taka (n=370) and 100 Taka (n=362) groups (table 1). The majority of complete interviews were between ages of 18 and 29 years, 70.0% (n=759) and male, 87.7% (n=952). Approximately half of the complete interviews reported living in a rural location. There were no differences in sociodemographic characteristics between complete and partial interviews (online supplementary appendix 2). The average

A

Number of random digit dial phone calls
(n=440,262)

Call did not connect (n=362,681)
Connected, but did not select language
(n=67,234)

Randomised to study arm

Control Selected language (n=4346)	50 Taka Selected language (n=2982)	100 Taka Selected language (n=3019)
Unknown (n=2542)	Unknown (n=1670)	Unknown (n=1765)
Age-eligible (n=1227) — Ineligible (n=577)	Age-eligible (n=945) — Ineligible (n=367)	Age-eligible (n=906) — Ineligible (n=348)
Partial Interview (n=227) Break-off (n=306) Refusal (n=341)	Partial Interview (n=137) Break-off (n=184) Refusal (n=254)	Partial Interview (n=170) Break-off (n=154) Refusal (n=220)
Completed Interview (n=353)	Completed Interview (n=370)	Completed Interview (n=362)

B

Number of random digit dial phone calls
(n=168,157)

Call did not connect (n=135,346)
Connected, but did not select language
(n=28,198)

Randomised to study arm

Control Selected language (n=1732)	5,000 Shillings Selected language (n=1428)	10,000 Shillings Selected language (n=1453)
Unknown (n=799)	Unknown (n=652)	Unknown (n=674)
Age-eligible (n=844) — Ineligible (n=89)	Age-eligible (n=701) — Ineligible (n=75)	Age-eligible (n=716) — Ineligible (n=63)
Partial Interview (n=175) Break-off (n=170) Refusal (n=122)	Partial Interview (n=104) Break-off (n=109) Refusal (n=84)	Partial Interview (n=114) Break-off (n=96) Refusal (n=85)
Completed Interview (n=377)	Completed Interview (n=404)	Completed Interview (n=421)

Figure 1 Consolidated Standards of Reporting Trials diagram. (A) Bangladesh. (B) Uganda.

time spent completing the IVR survey was 15 min, 8 s (SD: 1 min, 42 s). The average cost of airtime including incentive (where applicable) per completed interview was US$3.76 in the control, US$3.79 in the 50 Taka and US$4.46 in the 100 Taka groups (table 2).

In Uganda, 174 157 RDD phone calls were made from 26 March to 22 April 2017 (figure 1B). Sociodemographic characteristics of the complete interviews were similar across the control (n=377), 5000 UGX (n=404) and 10 000 UGX (n=421) groups (table 1). Similar to Bangladesh, the majority of Ugandan complete interviews were aged 18–29 years, 76.9% (n=900) and male, 77.1% (n=910). There were no differences in demographic characteristics between complete and partial interviews (online supplementary appendix 2). The average time spent completing the IVR survey was 13 min, 49 s (SD: 1 min, 35 s). The average cost of airtime including incentive (where applicable) per completed interview was US$2.94 in the control, US$3.91 in the 5000 UGX and US$5.15 in the 10 000 UGX groups (table 2).

In Bangladesh, the cooperation and response rates were, respectively, 28.8% (353/1227) and 19.2% (580/3016) in control, 39.2% (370/945) and 23.9% (507/2120) in 50 Taka incentive and 40.0% (362/906) and 24.8% (532/2148) in 100 Taka incentive groups

(table 3). In Uganda, the cooperation and response rates were, respectively, 44.7% (377/844) and 35.2% (552/1570) in control, 57.6% (404/701) and 39.3% (508/1293) in 5000 UGX and 58.8% (421/716) and 40.3% (535/1328) in 10 000 UGX groups. Cooperation rates in Bangladesh were significantly higher in the 50 Taka (RR 1.36, 95% CI 1.21 to 1.53, p<0.0001), and 100 Taka groups (RR 1.39, 95% CI 1.23 to 1.56, p<0.0001), as compared with control. Response rates were also significantly higher in the 50 Taka (RR 1.24, 95% CI 1.12 to 1.38, p=0.0001), and 100 Taka groups (RR 1.29, 95% CI 1.16 to 1.43, p<0.0001), as compared with control. Similar to Bangladesh, cooperation and response rates in Uganda were significantly higher in the 5000 UGX, respectively (RR 1.29, 95% CI 1.17 to 1.42, p<0.0001) and (RR 1.12, 95% CI 1.02 to 1.23, p=0.0225), and 10 000 UGX groups (RR 1.32, 95% CI 1.19 to 1.45, p<0.0001) and (RR 1.15, 95% CI 1.04 to 1.26, p=0.0045), as compared with control. In both countries, there were no differences in cooperation and response rates when the estimated proportion of age-eligible participants were excluded from analysis (online supplementary appendix 3). The response and cooperation rates had minor variations by time of day (online supplementary appendix 4).

Table 1 Demographics of complete interviews by study arm

Demographic	Bangladesh			Uganda		
	Control (n=353)	50 Taka (n=370)	100 Taka (n=362)	Control (n=377)	5000 UGX (n=404)	10 000 UGX (n=421)
Sex						
Male	313 (88.7%)	321 (86.8%)	318 (87.9%)	282 (76.2%)	308 (78.4%)	320 (78.4%)
Female	39 (11.1%)	46 (12.4%)	41 (11.3%)	88 (23.8%)	85 (21.6%)	88 (21.6%)
Other	1 (0.3%)	3 (0.8%)	3 (0.8%)	NA	NA	NA
Missing*	n=0	n=0	n=0	n=7	n=11	n=13
Age group (years)						
18–29	236 (66.9%)	269 (72.7%)	254 (70.2%)	281 (75.7%)	306 (78.3%)	313 (76.7%)
30–49	99 (28.1%)	85 (23.0%)	89 (24.6%)	83 (22.4%)	79 (20.2%)	84 (20.6%)
50–69	8 (2.3%)	10 (2.7%)	7 (1.9%)	6 (1.6%)	5 (1.3%)	8 (2.0%)
70+	10 (2.8%)	6 (1.6%)	12 (3.3%)	1 (0.3%)	1 (0.3%)	3 (0.7%)
Missing*	n=0	n=0	n=0	n=6	n=13	n=13
Education attempted						
None	27 (22.1%)	22 (19.8%)	30 (23.4%)	65 (17.5%)	66 (16.6%)	62 (15.2%)
Primary	93 (76.2%)	89 (80.2%)	98 (76.6%)	114 (30.7%)	92 (23.2%)	97 (23.7%)
Secondary	NA	NA	NA	125 (33.6%)	161 (40.6%)	172 (42.1%)
Tertiary or higher	NA	NA	NA	68 (18.3%)	78 (19.7%)	78 (19.1%)
Refused	2 (1.6%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Missing*†	n=231	n=259	n=234	n=5	n=7	n=12
Location						
Urban	170 (48.2%)	205 (55.4%)	188 (51.9%)	206 (55.5%)	228 (57.6%)	243 (59.1%)
Rural	182 (51.6%)	165 (44.6%)	172 (47.5%)	165 (44.5%)	168 (42.4%)	168 (40.9%)
Refused	1 (0.3%)	0 (0.0%)	2 (0.6%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Missing*	n=0	n=0	n=0	n=6	n=8	n=10
Language						
Bangla	348 (98.6%)	366 (98.9%)	360 (99.5%)	NA	NA	NA
English	5 (1.4%)	4 (1.1%)	2 (0.5%)	69 (18.3%)	67 (16.6%)	65 (15.4%)
Luganda	NA	NA	NA	217 (57.6%)	230 (56.9%)	267 (63.4%)
Luo	NA	NA	NA	33 (8.8%)	35 (8.7%)	35 (8.3%)
Runyakitara	NA	NA	NA	58 (15.4%)	72 (17.8%)	54 (12.8%)

Data are n (%) for participants classified as complete interviews.
*Missing values for Uganda are due to errors in platform that prevented storing data.
†Missing values for Bangladesh are due to incorrect coding of IVR platform.
NA, not applicable; UGX, Ugandan Shillings.

For secondary outcomes, contact rate in Bangladesh only improved in the 50 Taka incentive group (RR 1.10, 95% CI 1.03 to 1.17, p<0.0052) as compared with control group, while there were no differences in contact rates by study group in Uganda (table 3). Refusal rates in Bangladesh were significantly lower in the 100 Taka group (RR 0.81, 95% CI 0.72 to 0.91, p=0.0004), but not the 50 Taka group, as compared with control participants (table 3). In Uganda, both the 5000 UGX (RR 0.80, 95% CI 0.68 to 0.95, p=0.0095) and 10 000 UGX (RR 0.73, 95% CI 0.62 to 0.87, p=0.0003) had lower refusal rates than the control group. In both countries, there were no

significant differences in responses to NCD questions by study arm (online supplementary appendix 5).

Subgroup analysis for Bangladesh found that overall, age significantly modified the effect of incentives on cooperation rate (p=0.015; table 4). Participants aged 26 years and older had lower cooperation rates than those who were younger than 25 years in the 50 Taka group (RR 0.71, 95% CI 0.56 to 1.90, p=0.0055) as compared with controls; with no significant difference by age in the 100 Taka group. Unlike Bangladesh, subgroup analysis for Uganda found no significant effect modifiers of the interventions for cooperation rate (table 5). Although

Table 2 Disposition codes by study arm in Bangladesh and Uganda

	Bangladesh			Uganda		
	Control	50 Taka	100 Taka	Control	5000 UGX	10 000 UGX
Complete interview (I)	353	370	362	377	404	421
Partial interview (P)	227	137	170	175	104	114
Refusal (R)						
Refusal	341	254	220	122	84	85
Breaks-off	306	184	154	170	109	96
Non-contact (NC)	0	0	0	0	0	0
Other (O)	0	0	0	0	0	0
Unknown household (UH)	0	0	0	0	0	0
Unknown other (UO)	2542	1670	1765	799	652	674
e Unknown e(UO)*	1789	1175	1242	726	592	612
Ineligible						
Under age	577	367	348	89	75	63
Call did not connect†	120 893	120 894	120 894	45 115	45 115	45 116
Connected, but no language selection†	22 411	22 411	22 412	9400	9399	9399
Average cost per complete interview (US$)‡	3.76	3.79	4.46	2.94	3.91	5.15

*Estimated proportion of unknown cases that were age-eligible was 70.4% for Bangladesh and 90.9% for Uganda.
†Evenly distributed to each study arm due to randomisation occurring after language selection.
‡Only includes cost of airtime participants spent on survey and airtime incentive, as applicable.

females had lower cooperation rates than males in the 5000 UGX group (RR 0.78, 95% CI 0.62 to 0.97, p=0.0233) as compared with control, the overall interaction for gender was not statistically significant (p=0.066).

The pooling of the different incentive groups in Uganda and Bangladesh found that cooperation rate was significantly higher in the 1X incentive group (figure 2; pooled RR (pRR) 1.32, 95% CI 1.22 to 1.42) and 2X incentive group (pRR 1.34, 95% CI 1.25 to 1.45) and that the results were highly consistent, respectively (I^2 0.0%, p=0.497) and (I^2 0.0%, p=0.493). Overall, providing an incentive of any amount significantly improved the cooperation rate by 33% (pRR 1.33, 95% CI 1.26 to 1.40) as compared with control group. Pooled analysis found that

Table 3 Survey rates by study arm in Bangladesh and Uganda

	Bangladesh			Uganda		
	Control	50 Taka	100 Taka	Control	5000 UGX	10 000 UGX
Contact rate	40.7%	44.6%	42.2%	53.8%	54.2%	53.9%
Risk ratio (95% CI)	*Ref.*	1.10 (1.03 to 1.17)	1.04 (0.97 to 1.11)	*Ref.*	1.01 (0.94 to 1.08)	1.00 (0.94 to 1.07)
P value	*Ref.*	0.0052	0.2810	*Ref.*	0.8070	0.9324
Response rate	19.2%	23.9%	24.8%	35.2%	39.3%	40.3%
Risk ratio (95% CI)	*Ref.*	1.24 (1.12 to 1.38)	1.29 (1.16 to 1.43)	*Ref.*	1.12 (1.02 to 1.23)	1.15 (1.04 to 1.26)
P value	*Ref.*	0.0001	<0.0001	*Ref.*	0.0225	0.0045
Refusal rate	21.5%	20.7%	17.4%	18.6%	14.9%	13.6%
Risk ratio (95% CI)	*Ref.*	0.96 (0.86 to 1.07)	0.81 (0.72 to 0.91)	*Ref.*	0.80 (0.68 to 0.95)	0.73 (0.62 to 0.87)
P value	*Ref.*	0.4941	0.0004	*Ref.*	0.0095	0.0003
Cooperation rate	28.8%	39.2%	40.0%	44.7%	57.6%	58.8%
Risk ratio (95% CI)	*Ref.*	1.36 (1.21 to 1.53)	1.39 (1.23 to 1.56)	*Ref.*	1.29 (1.17 to 1.42)	1.32 (1.19 to 1.45)
P value	*Ref.*	<0.0001	<0.0001	*Ref.*	<0.0001	<0.0001

Ref, reference group; UGX, Ugandan Shilling.

Table 4 Subgroup analyses of cooperation rates in Bangladesh

	Control	50 Taka	Stratum-specific RR	P value*	100 Taka	Stratum-specific RR	P value*
Gender							
Male	313/757 (41%)	321/586 (55%)	1.32 (1.18 to 1.48)	0.20	318/581 (55%)	1.32 (1.18 to 1.48)	0.88
Female	39/89 (44%)	46/65 (71%)	1.61 (1.22 to 2.14)		41/69 (59%)	1.36 (1.00 to 1.84)	
Age (years)							
≤25	196/674 (29%)	219/474 (46%)	1.59 (1.36 to 1.85)	0.006	190/447 (43%)	1.46 (1.25 to 1.71)	0.40
>25	157/553 (28%)	151/471 (32%)	1.13 (0.94 to 1.36)		172/459 (37%)	1.32 (1.11 to 1.58)	
Education							
No school	27/74 (36%)	22/57 (39%)	1.06 (0.68 to 1.65)	0.75	30/53 (57%)	1.55 (1.06 to 2.27)	0.28
Primary	93/199 (47%)	89/166 (54%)	1.15 (0.93 to 1.41)		98/171 (57%)	1.23 (1.01 to 1.49)	
Location							
Urban	170/427 (40%)	205/354 (58%)	1.45 (1.26 to 1.68)	0.036	188/349 (54%)	1.35 (1.16 to 1.57)	0.26
Rural	182/365 (50%)	165/283 (58%)	1.17 (1.01 to 1.35)		172/287 (60%)	1.20 (1.05 to 1.38)	

Data are n/N (%) and RR (95% CI). Overall interaction terms were significant for cooperation rates only in the age subgroup: gender, p=0.31; age, p=0.015; education, p=0.32; location, p=0.11.
*P values obtained from an interaction term between intervention groups and demographic characteristic.
RR, risk ratio.

Table 5 Subgroup analyses of cooperation rates in Uganda

	Control	5000 UGX	Stratum-specific RR	P value*	10 000 UGX	Stratum-specific RR	P value*
Gender							
Male	282/543 (52%)	308/443 (70%)	1.34 (1.21 to 1.48)	0.023	320/463 (69%)	1.33 (1.20 to 1.47)	0.090
Female	88/155 (57%)	85/144 (59%)	1.04 (0.86 to 1.26)		88/140 (63%)	1.11 (0.92 to 1.34)	0.76
Age (years)							
≤25	222/485 (46%)	245/403 (61%)	1.33 (1.17 to 1.50)	0.43	250/413 (61%)	1.32 (1.17 to 1.50)	
>25	149/343 (43%)	146/275 (53%)	1.22 (1.04 to 1.44)		158/284 (56%)	1.28 (1.09 to 1.50)	
Education							
≤Primary	179/308 (58%)	158/234 (68%)	1.16 (1.02 to 1.32)	0.12	159/231 (69%)	1.18 (1.04 to 1.35)	0.18
>Primary	193/371 (52%)	239/344 (69%)	1.34 (1.18 to 1.51)		250/360 (69%)	1.33 (1.18 to 1.50)	
Location							
Urban	206/393 (52%)	228/339 (67%)	1.28 (1.14 to 1.45)	0.60	243/344 (71%)	1.35 (1.20 to 1.51)	0.085
Rural	165/291 (57%)	168/242 (69%)	1.22 (1.07 to 1.40)		168/257 (65%)	1.15 (1.01 to 1.32)	

Data are n/N (%) and RR (95% CI). Overall interaction terms were not significant for cooperation rates in all subgroups: gender, p=0.066; age, p=0.72; education, p=0.27; location, p=0.18.
*P values obtained from an interaction term between intervention groups and demographic characteristic.
RR, risk ratio; UGX, Ugandan Shilling.
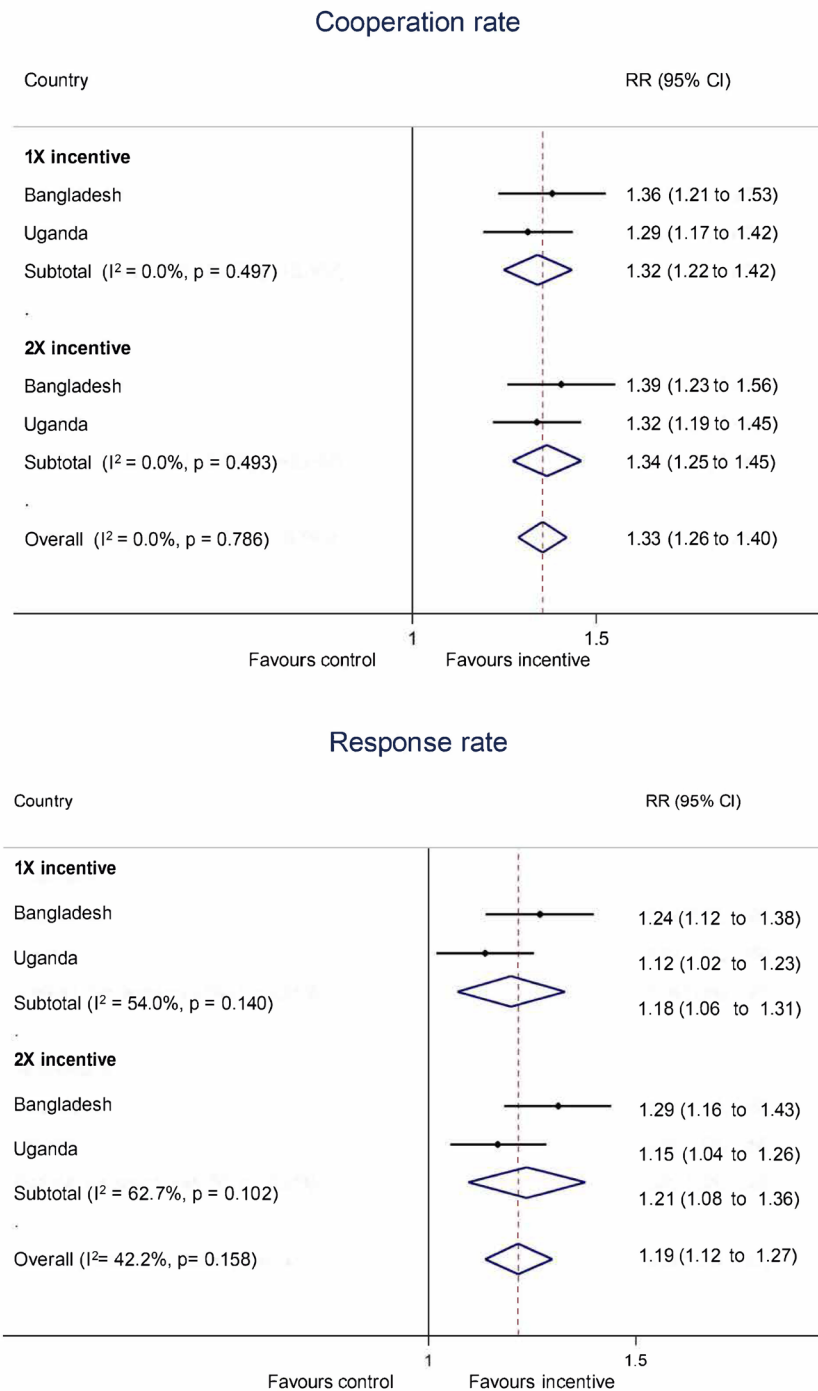
Figure 2 Forest plots for cooperation and response rates in Bangladesh and Uganda. RR, risk ratio.

response rate was significantly higher in the 1X incentive group (pRR 1.18, 95% CI 1.06 to 1.31) and the findings were moderately heterogeneous (I^2 54.0%, p=0.140). Similarly, the 2X incentive group significantly improved response rate (pRR 1.21, 95% CI 1.08 to 1.36) and had moderate heterogeneity (I^2 62.7%, p=0.102).

DISCUSSION
In our study of 2287 participants who completed IVR surveys, the provision of airtime incentives significantly

increased cooperation and response rates compared with no incentive in both Bangladesh and Uganda. The pRRs across two geographically, linguistically and culturally distinct countries found that the effect of 1X and 2X incentives were highly homogenous at improving cooperation rates and moderately heterogeneous for improving response rates, implying that incentives may be a useful tool in improving mobile phone survey participation in other countries.

The lack of difference in cooperation and response rates between the two incentive amounts is consistent with other studies, which found no significant differences in computer-assisted telephone interview response rates with airtime compensation amounts equivalent to 300 or 500 Tanzanian Shillings (or US$0.17 vs US$0.42).[27] However, RDD IVR surveys conducted in Afghanistan and Mozambique found that a 4 min airtime incentive did not significantly improve completion rates as compared with a zero incentive control group.[28] The similarity in cooperation and response rates between airtime incentive arms in our study suggests potential threshold effects of incentives, while the null findings from other incentive studies may imply that a minimum threshold level exists. This phenomenon has been documented for incentive studies of HIV testing, paediatric immunisation and voluntary male medical circumcisions in LMICs.[29–31] and in survey research in high-income countries.[32]

We computed contact, cooperation, refusal and response rates using guidelines established by AAPOR to allow for standardisation of outcomes and comparability with similar studies in LMIC settings.[22] Our observed contact (39%) and response (31%) rates respectively in control arm participants were similar to those observed in a nationally administered RDD IVR survey of persons aged ≥18 years in Ghana where no incentives were provided.[15] However, we observed higher refusal rates in Bangladesh (21%) and Uganda (19%), and lower cooperation rates, 29% and 45%, respectively, than what was observed in Ghana (refusal rate=7%; cooperation rate=59%). Differences in questionnaire design, length and the calculation of disposition codes for break-offs, refusals and partial interviews between the two studies may explain these variations.

A common critique of the use of incentives in LMIC settings are concerns about sustainability. However, our data show that the cost in airtime for a complete survey in the Bangladesh control group, US$3.76, was very similar to the 50 Taka airtime arm, which had a delivery cost in airtime and incentive of US$3.79 per complete survey. In this case, the provision of an incentive was cost-neutral in terms of direct delivery costs and may represent savings in reducing costs due to decreased numbers of partial interviews. Still, even if use of incentives represents an added cost, this may be justified if it increases participation of marginalised subpopulations and increases survey representativeness. As an example, for Bangladeshi females, those in the airtime incentive groups had 15%–21% higher cooperation rates than females from the control group. Given the skewed distribution of IVR participation to males, the use of airtime incentives, regardless of cost, may be required if population-representative estimates are needed. Lastly, although the cost per complete survey was higher in the Uganda intervention groups than in the control groups, it is likely that the incentive would become cost neutral if we used a lower incentive amount, assuming that completion rates remain unchanged.

In regard to possible ethical concerns, the incentives we used were less than one day's average working wage, the amounts used were guided by input from community members and ministry officials, and their application was not paired to a risky or dangerous behaviour. However, it is uncertain if our use of airtime incentives was equitable. Although the subgroup analyses should be interpreted with caution due to small sample size, we found that 50 Taka incentives significantly favoured younger participants for cooperation rates in Bangladesh. Moreover, although not statistically significant (p=0.066), airtime incentives differentially improved cooperation rates of Ugandan males as compared with females. Lastly, although there is no published 'optimal' approach to conducting consent for mobile phone surveys in LMICs, we developed a relatively brief survey informational module and provided respondents the opportunity to opt-in or opt-out from participation. Additional studies that examine alternative ways to consent participants for mobile phone surveys are needed to maximise desired scientific and ethical outcomes.[33]

A larger concern about the use of IVR and other types of mobile phone surveys is the challenge of generating population representative estimates given that certain subpopulations (eg, women, the less-educated and elderly) may have lower levels of phone access and ownership and that in rural areas, the mobile network coverage may be suboptimal. Similar to RDD IVR surveys conducted in Afghanistan, Ethiopia, Ghana, Mozambique and Zimbabwe,[15 28] we observed that males were more likely to participate in our IVR surveys, making up 86.8%–88.7% of participants across the three arms in Bangladesh, and 76.2%–78.4% of participants in Uganda. While this 'digital divide' exists,[34] it appears to be shrinking with time as more investments are made by telecommunications operators and mobile phone penetration rates continue to rise.[35] In the meantime, quota sampling could be used to ensure a more equal distribution among age-sex strata[36] and alternative strategies to reach these under-represented groups, such as more motivational introductory messages and interviewing style to secure participation, lottery incentives, scheduling surveys to meet participants' availability, calling back participants multiple times and the use of computer assisted telephone interviews, need to be tested.

This study has several strengths. First, the randomisation of participants to study groups was automated by the IVR platform, thereby ensuring that our cooperation and response rates were unbiased from any potential deviation in study group allocation. Second, the results in Bangladesh and Uganda are comparable because we employed a standardised protocol in each country and used the same technology platform to administer the IVR survey. Third, our sample included all known mobile network operators in each country, thereby reducing any potential selection bias that could result from differences in subscriber characteristics between mobile network operators. Moreover, we found no differences in demographic characteristics

between complete and partial interviews in both countries. Lastly, we conducted sensitivity analyses for our primary outcomes and found that the cooperation and response rates remained similar when not applying the estimated proportion of age-eligible individuals to the unknown disposition code.

However, this study also has several limitations. First, there was a substantial number of phone calls where we were unable to disaggregate them into working or non-working phone numbers. We chose to classify all of these instances as non-working numbers because they were not randomised to a study group. As a result, our response, contact and refusal rates are likely inflated, with no effect on the cooperation rates. Future studies could consider re-contacting a sample of these unascertainable numbers to estimate the percentage of working and non-working numbers.[37] Second, our survey was available in only three out of six major Ugandan language groups.[38] People who did not speak English or any of these languages had no possibility of being captured within our study, potentially resulting in selection bias. Although this bias is likely minimal and inconsequential for our study, it has larger implications for delivering nationally representative surveys.[28]

In conclusion, the provision of either of two airtime incentives significantly improved the cooperation and response rates of an IVR survey in Bangladesh and Uganda. Given the geographical and cultural diversity, as well as differences in mobile phone subscription rates and number of languages spoken between these two countries, our findings suggest that airtime incentives may be useful to improve IVR survey participation in other resource-constrained settings. As mobile phone ownership continues to increase in LMICs, more research studies are needed to fully optimise the potential of these methods of data collection for health-related research.

Author affiliations
[1]International Health, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA
[2]Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA
[3]Population, Family And Reproductive Health, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA
[4]Johns Hopkins University Berman Institute of Bioethics, Baltimore, Maryland, USA
[5]Institute of Epidemiology Disease Control and Research, Dhaka, , Bangladesh
[6]Makerere University College of Health Sciences, Kampala, , Uganda
[7]George Washington University Milken Institute of Public Health, Washington, District of Columbia, USA

REFERENCES

1. Dicker D, Nguyen G, Abate D, et al. Global, regional, and national age-sex-specific mortality and life expectancy, 1950–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet* 2018;392:1684–735.
2. World Health Organization. *Noncommunicable diseases: progress monitor 2017*. Geneva, 2017.
3. World Health Organization. *Global status report on noncommunicable diseases 2010*. Geneva, 2010.
4. World Health Organization. *Global action plan for the prevention and control of noncommunicable diseases 2013-2020*. Geneva, 2013.
5. World Health Organization. Who stepwise approach to surveillance (steps), 2018. Available: https://www.who.int/ncds/surveillance/steps/en/ [Accessed 8 Nov 2018].
6. Riley L, Guthold R, Cowan M, et al. The world Health organization stepwise approach to noncommunicable disease risk-factor surveillance: methods, challenges, and opportunities. *Am J Public Health* 2016;106:74–8.
7. Pariyo GW, Wosu AC, Gibson DG, et al. Moving the agenda on noncommunicable diseases: policy implications of mobile phone surveys in low and middle-income countries. *J Med Internet Res* 2017;19:e115.
8. Centers for Disease Control and Prevention. Behavioral risk factor surveillance system, 2018. Available: https://www.cdc.gov/brfss/index.html [Accessed 11 Dec 2018].
9. Pierannunzi C, Hu SS, Balluz L. A systematic review of publications assessing reliability and validity of the behavioral risk factor surveillance system (BRFSS), 2004-2011. *BMC Med Res Methodol* 2013;13:49.
10. Jungquist CR, Mund J, Aquilina AT, et al. Validation of the behavioral risk factor surveillance system sleep questions. *J Clin Sleep Med* 2016;12:301–10.
11. Hu SS, Pierannunzi C, Balluz L. Integrating a multimode design into a national random-digit-dialed telephone survey. *Prev Chronic Dis* 2011;8:A145.
12. International Telecommunications Union. Mobile-cellular subscriptions. Available: http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx [Accessed 8 Nov 2018].
13. 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.
14. Greenleaf AR, Gibson DG, Khattar C, et al. Building the evidence base for remote data collection in low- and middle-income countries: comparing reliability and accuracy across survey modalities. *J Med Internet Res* 2017;19:e140.
15. L'Engle K, Sefa E, Adimazoya EA, et al. Survey research with a random digit dial national mobile phone sample in Ghana: methods and sample quality. *PLoS One* 2018;13:e0190902.
16. 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:e133–10.
17. Singer E, Ye C. The use and effects of incentives in surveys. *Ann Am Acad Pol Soc Sci* 2013;645:112–41.
18. Brick J, Montaquila JM, Hagedorn MC, et al. Implications for RDD design from an incentive experiment. *J Off Stat* 2005;21:571–89.

19. Schulz KF, Altman DG, Consort MD. Statement: updated guidelines for reporting parallel group randomised trials. *PLoS Med* 2010;2010:e1000251.
20. Waksberg J. Sampling methods for random digit Dialing. *J Am Stat Assoc* 1978;73:40–6.
21. Gibson DG, Pariyo GW, Wosu AC, *et al*. Evaluation of mechanisms to improve performance of mobile phone surveys in low- and middle-income countries: research protocol. *JMIR Res Protoc* 2017;6:e81.
22. The American Asssociation for Public Opinion Research. *Standard Definitions: Final dispositions of case codes and outcome rates for surveys*. 9th edn. AAPOR, 2016.
23. Rothman KJNo adjustments are needed for multiple comparisons. *Epidemiology* 1990;1:43–6.
24. Wacholder S. Binomial regression in GLIM: estimating risk ratios and risk differences. *Am J Epidemiol* 1986;123:174–84.
25. Deeks J, Altman D, MJ B. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger M, Smith G, DG A, eds. *Systematic reviews in health care: meta-analysis in context*, 2008: 285–312.
26. Gibson DG, Farrenkopf BA, Pereira A, *et al*. The development of an interactive voice response survey for noncommunicable disease risk factor estimation: technical assessment and cognitive testing. *J Med Internet Res* 2017;19:e112.
27. Hoogeveen J, Croke K, Dabalen A, *et al*. Collecting high frequency panel data in Africa using mobile phone interviews. *Canadian Journal of Development Studies / Revue canadienne d'études du développement* 2014;35:186–207.
28. Leo B, Morello R, Mellon J, *et al*. *Do mobile surveys work in poor countries*, 2015.
29. Gibson DG, Ochieng B, Kagucia EW, *et al*. Mobile phone-delivered reminders and incentives to improve childhood immunisation coverage and timeliness in Kenya (M-SIMU): a cluster randomised controlled trial. *Lancet Glob Health* 2017;5:e428–38.
30. Thornton RL. The demand for, and impact of, learning HIV status. *Am Econ Rev* 2008;98:1829–63.
31. Thirumurthy H, Masters SH, Rao S, *et al*. Effect of providing conditional economic compensation on uptake of voluntary medical male circumcision in Kenya: a randomized clinical trial. *JAMA* 2014;312.
32. Singer E, Van Hoewyk J, Maher MP. Experiments with incentives in telephone surveys. *Public Opin Q* 2000;64:171–88.
33. Ali J, Labrique AB, Gionfriddo K, *et al*. Ethics considerations in global mobile Phone-Based surveys of noncommunicable diseases: a conceptual exploration. *J Med Internet Res* 2017;19:e110.
34. GSMA Connected Women. The mobile gender gap report 2018, 2018. Available: https://www.gsma.com/mobilefordevelopment/connected-women/the-mobile-gender-gap-report-2018/ [Accessed 24 Aug 2018].
35. Tran MC, Labrique AB, Mehra S, *et al*. Analyzing the mobile "digital divide": changing determinants of household phone ownership over time in rural bangladesh. *JMIR Mhealth Uhealth* 2015;3:e24.
36. Labrique A, Blynn E, Ahmed S, *et al*. Health surveys using mobile phones in developing countries: automated active strata monitoring and other statistical considerations for improving precision and reducing biases. *J Med Internet Res* 2017;19:e121.
37. Lau C, di Tada N. Identifying Non-Working phone numbers for response rate calculations in Africa. *Surv Pract* 2018;11.
38. Guwatudde D, Mutungi G, Wesonga R, *et al*. The epidemiology of hypertension in Uganda: findings from the National non-communicable diseases risk factor survey. *PLoS One* 2015;10:e0138991.