



Original Investigation | Global Health

# Performance and Resource Requirements of In-Person, Voice Call, and Automated Telephone-Based Socioeconomic Data Collection Modalities for Community-Based Health Programs

## A Systematic Review

Luke N. Allen, MPH, MBChB; Shona Mackinnon, MSc; Iris Gordon, PhD; David Blane, PhD; Ana Patricia Marques, PhD; Stephen Gichuhi, PhD; Alice Mwangi, MSc; Matthew J. Burton, PhD; Nigel Bolster, PhD; David Macleod, PhD; Min Kim, MSc; Jacqueline Ramke, PhD; Andrew Bastawrous, PhD

### Abstract

**IMPORTANCE** Gathering data on socioeconomic status (SES) is a prerequisite for health programs that aim to improve equity. There is a lack of evidence on which approaches offer the best combination of reliability, cost, and acceptability.

**OBJECTIVE** To compare the performance of different approaches to gathering data on SES in community health programs.

**DATA SOURCES** A search of the Cochrane Library, MEDLINE, Embase, Global Health, ClinicalTrials.gov, the World Health Organization International Clinical Trials Registry Platform, and OpenGrey from 1999 to June 29, 2021, was conducted, with no language limits. Google Scholar was also searched and the reference lists of included articles were checked to identify further studies. The search was performed on June 29, 2021.

**STUDY SELECTION** Any empirical study design was eligible if it compared 2 or more modalities to elicit SES data from the following 3 categories: in-person, voice call, or automated telephone-based systems.

**DATA EXTRACTION AND SYNTHESIS** Two reviewers independently screened titles, abstracts, and full-text articles and extracted data. They also assessed the risk of bias using Cochrane tools and assessed the certainty of the evidence using the Grading of Recommendations, Assessment, Development and Evaluation approach. Findings were synthesized thematically without meta-analysis.

**MAIN OUTCOMES AND MEASURES** Response rate, equivalence, time, costs, and acceptability to patients and health care professionals.

**RESULTS** The searches returned 3943 records. The 11 included studies reported data on 14 036 individuals from 7 countries, collecting data on 11 socioeconomic domains using 2 or more of the following modes: in-person surveys, computer-assisted telephone interviews (CATIs), and 2 types of automated data collection: interactive voice response calls (IVRs) and web surveys. Response rates were greater than 80% for all modes except IVRs. Equivalence was high across all modes (Cohen  $\kappa > 0.5$ ). There were insufficient data to make robust time and cost comparisons. Patients reported high levels of acceptability providing data via IVRs, web surveys, and CATIs.

(continued)

### Key Points

**Question** What are the relative strengths and weaknesses of different socioeconomic data collection modes?

**Findings** In this systematic review of 11 studies with 14 036 individuals, high levels of equivalence and acceptability were found across in-person surveys, computer-assisted telephone interviews, and 2 types of automated data collection: interactive voice response calls and web surveys; cost and time comparisons were rarely performed. Response rates were greater than 80% for all modes except interactive voice response.

**Meaning** This systematic review identified no substantial evidence that remote and automated data collection modes are any worse than in-person approaches, and there was no compelling evidence that these approaches are faster or cost less.

### + Supplemental content

Author affiliations and article information are listed at the end of this article.

**Open Access.** This is an open access article distributed under the terms of the CC-BY License.

Abstract (continued)

**CONCLUSIONS AND RELEVANCE** Selecting an appropriate and cost-effective modality to elicit SES data is an important first step toward advancing equitable effective service coverage. This systematic review did not identify evidence that remote and automated data collection modes differed from human-led and in-person approaches in terms of reliability, cost, or acceptability.

*JAMA Network Open.* 2022;5(11):e2243883. doi:10.1001/jamanetworkopen.2022.43883

---

## Introduction

### Rationale

Inequalities in health are pervasive and persistent. Women and girls, individuals living in rural areas, and persons with lower levels of income, education, and social status all tend to experience higher barriers to accessing care than other groups.<sup>1-4</sup> To understand and redress socioeconomic inequalities, international development partners are increasingly calling for socioeconomic status (SES) data to be routinely collected and analyzed by all health systems and programs.<sup>5,6</sup>

Previous work has reported that SES data can be collected using a variety of modalities in the community setting, including in-person interviews, telephone calls, and automated telephone-based systems.<sup>7</sup> There is growing interest in using mobile phones to collect data for global health programs on the basis that this modality is lower cost, faster, and more flexible than in-person approaches.<sup>8,9</sup>

Croke et al<sup>10</sup> have argued that telephone-based data collection is acceptable in settings where mobile phone ownership rates exceed 80%. While this percentage is an arbitrary threshold, we note that the share of the population that has access to a telephone exceeds the proportion of those who own a telephone. Mobile phone ownership has increased sharply in the past decade such that there are now approximately 100 mobile phone subscriptions for every 100 people in low- and middle-income countries (LMICs).<sup>11,12</sup> Across Sub-Saharan Africa, where telephone ownership is lowest, telephones have been used for a wide range of applications including surveillance, surveys, behavior change interventions, monitoring and evaluation, and training.<sup>12-17</sup>

It is well known that the mode of data collection (eg, in-person, telephone interview, or short message service [SMS]) can influence survey response rates and other performance characteristics, especially when the questions are of a sensitive nature.<sup>18,19</sup> Previous research suggests that telephone-based data collection approaches may reduce social desirability bias—where responders provide what they perceive to be socially acceptable answers even if they are not accurate—compared with in-person approaches.<sup>20</sup> However, telephone-based approaches also tend to have lower response rates and have historically presented under-coverage biases due to lower penetration among less-educated and low-income groups.<sup>21</sup>

Pariyo and colleagues<sup>22</sup> have noted the dearth of research comparing different modalities of SES data collection in LMICs. Given the increasing feasibility and potential efficiency gains of using telephones for SES data collection, we aimed to systematically review the findings of empirical studies that have compared in-person vs voice call vs telephone-based modalities for gathering SES data for community-based health programs in terms of their performance characteristics, resource requirements, and acceptability to participants and service professionals.

---

## Methods

This registered review followed a published protocol.<sup>23</sup> It also followed the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline and Cochrane guidelines.<sup>24,25</sup>

## Eligibility Criteria

### Population

In this systematic review, the population was composed of studies rather than people, namely, those that sought to compare 2 or more modalities for SES data collection from individuals enrolled in community-based health programs. Studies that reported on only 1 mode of data collection were excluded.

For the purpose of this review, *health programs* were defined as organized activities to improve 1 or more health outcomes in a defined population. *Community-based* encompasses all settings except hospitals. Some researchers exclude primary care facilities from definitions of community-based care<sup>26</sup>; however, these facilities were included in this review, along with outreach and mobile clinics, community centers, schools, workplaces, and people's own homes.

*Socioeconomic status* is a critically important but nebulous concept that pertains to social and economic standing within society.<sup>27</sup> It determines exposure to the social determinants of health; "the conditions in which people are born, grow, live, work, and age"<sup>2</sup>; and relates to issues of privilege, power, and control.<sup>28</sup> Almost all health outcomes are patterned according to SES, with the most disadvantaged populations experiencing the worst health outcomes.<sup>2,28,29</sup> Socioeconomic status is commonly measured using income, educational level, occupation, and other metrics, such as wealth, caste, and place of residence. We included all of these domains, as well as any other proxies that are identified by researchers as capturing SES.

### Interventions

The interventions being studied are 3 different groups of modalities for collecting SES data (**Box**). The focus is on the modality of data collection (eg, in-person vs voice call vs automated) rather than the content of the wording that is used to elicit information.

We excluded approaches that used a blend of modes to elicit SES data. We excluded studies in which the SES questions and wording were not kept constant across modes. Studies that gather SES

#### Box. Definitions of the 3 Data Collection Approaches Used in This Review

**In-person data collection** included any form of exchange between a program implementer and a participant or their responsible guardian where the program implementer asks predefined questions to ascertain the participants' socioeconomic status and a synchronous response is received, ie, both parties occupy the same time and space, and the response is recorded by the implementer before the encounter is terminated. Any recording modality used by the program implementer will be included, such as pen and paper or completion of an electronic form. For this review we will also include self-administered questionnaires as a subtype of in-person data collection, provided that the data collection instrument was provided when the participant presented to a program implementer in person, the participant was asked to complete the data entry form, and the participant submitted their responses before departing. Any nonhospital location was accepted.

**Voice call data collection** includes real-time, telephone-based verbal exchanges between program implementers and participants whereby SES data are elicited and recorded by the program implementer using predefined questions. This category included computer-assisted telephone interviews—where the interviewer follows prompts on a computer screen, usually in a call center—as

well as non-computer-assisted telephone interviews. Videocalls were included as another subtype of voice calls.

**Automated telephone-based data collection** included any mobile telephone-based asynchronous exchange of information whereby participants are sent a standardized text message (also known as a short message service [SMS]), multimedia message (MMS), or automated phone call (sometimes called interactive voice response or IVR) and asked to provide SES data. Interactive voice response calls use prerecorded messages that prompt respondents to provide answers using speech, eg, state your age in years or by entering numbers on the keypad eg, press 1 for yes and 2 for no. We allowed responses to be provided using the same modality or any other digital form, eg, entering details on a webpage/web survey. Interventions that required participants to engage with human program implementers (eg, human-led SMS exchanges) were excluded from this modality. All forms of phrasing of the requests and responses were included. Reasoning that all smartphones come with a preloaded browser, we included web surveys that can be accessed by a hyperlink, as long as the link was sent via SMS or MMS. We excluded data collection approaches that required the download of third-party software, including email.

data at the household or community level were only included if these data were used to make assumptions about the SES of identifiable individual participants enrolled or due to be enrolled in the service delivery program of interest.

### Comparator

Studies that examined any 2 or more modalities were eligible. We excluded studies that only reported outcomes for 1 modality, that is, in which comparisons were not possible between modes. There was no index or gold standard data collection modality. Interventions that bundled requests for SES data with requests for other data (eg, broader demographic data) were included, as long as separate results were reported for the SES data collection element.

### Outcomes

Our 2 primary outcomes were performance characteristics and resource requirements. We reported these outcomes at the level of the following individual SES items.

#### Performance Characteristics

- Response rate: number of completed SES items divided by the total number of elicitation attempts. This outcome was calculated at the level of each SES item.
- Equivalence: agreement between the responses obtained from 2 or more different modalities. Recognizing that equivalence can vary by question, we report equivalence for each SES item. We report equivalence figures that aggregated multiple SES questions in a secondary analysis; however, we do not report aggregate equivalence figures that mixed SES items with non-SES items. Following Marcano Belisario et al<sup>30</sup> and Gwaltney et al,<sup>31</sup> we used comparisons of mean scores between modalities and/or correlations and/or measures of agreement, including intraclass correlation coefficients, Pearson product-moment correlations, Spearman  $\rho$ , and weighted  $\kappa$  coefficients.

#### Resource Requirements

- Time: the time taken to gather SES data using each approach (range and mean).
- Costs: any financial data on the costs of operating the data collection approach. These approaches include fixed costs (equipment, software, insurance, and personnel required to set up a given data elicitation modality) and ongoing support costs. We aimed to calculate the fixed and per-person costs to purchasers per completed survey.

Our secondary outcome was acceptability to participants and service professionals, based on survey or interview results reporting on how program implementers and participants perceived the collection modality in terms of intrusiveness, ease of use, time requirement, and general acceptability, as well as perceived advantages, barriers, disadvantages, and additional costs presented by the beneficiaries, data collectors, or study authors.

#### Measures of Effect

For each outcome we present raw values and risks ratios. We used the most commonly studied modality (computer-assisted telephone interview [CATI]) as the reference group.

### Study Types to Be Included

All empirical study designs that compared 2 or more data collection modalities were included. Studies were only included if they compared modalities that had been used to gather data from participants. Studies that used simulated data or data obtained from populations other than the intended beneficiaries were excluded. Both quantitative and qualitative study designs were included as long as they reported 1 or more of the outcomes of interest. Review articles were excluded, but the primary studies they discussed were screened for potential inclusion.

## Information Sources

We searched the following information resources: the Cochrane Library, MEDLINE, Embase, Global Health, ClinicalTrials.gov, and the World Health Organization International Clinical Trials Registry Platform for current and ongoing trials. We searched OpenGrey for gray literature and the first 20 pages of Google Scholar. We checked the reference lists of included studies and relevant systematic reviews to identify any additional potentially relevant reports of studies. We contacted key authors to uncover additional or upcoming studies.

## Search Strategy

The search strategy was built around 3 blocks: data collection modalities, SES concepts, and study design and setting terms (eMethods in the [Supplement](#) provides the full strategy). The search was limited to human studies published since 1999 (the year that it first became possible to send cross-network SMS messages). We searched for full-text studies published in any language. We did not include reports of studies published as conference abstracts. The search was performed on June 29, 2021.

## Study Selection

Two of us (L.N.A. and S.M.) independently screened all titles and abstracts and full texts using online software (Covidence). Studies that did not meet the inclusion criteria were excluded. Disagreements were resolved through consensus-based discussion and discussion with a third reviewer (D.B.) when necessary. We recorded reasons for exclusion at the full-text screening stage.

## Data Extraction and Management

Two of us (L.N.A. and S.M.) independently extracted study characteristics and data from the included studies using a custom data extraction form that was based on the Cochrane template.<sup>25</sup> We emailed study authors to request additional information and primary data if any aspect of their article precluded the assessment of eligibility or inclusion in the data synthesis.

## Data Items

We extracted the following items from each study:

- Article details
- Study design, population, and setting
- Questions used to assess SES (SES domains and individual response options)
- Number of times SES data were collected from each participant (eg, cross-sectional or serial)
- Modalities used to collect SES data:
  - Modality name and definition
  - Who gathered the SES data
  - When data were gathered in the patient journey/program
  - Equipment used
  - Who provided the data
  - Synchronous or asynchronous data collection
- Types of comparison and outcome measures
- Outcomes: response rate, completeness, equivalence, time, costs, and all qualitative text provided on acceptability

## Risk of Bias Assessment for Included Studies

Two of us (L.N.A. and S.M.) independently assessed risk of bias using the Cochrane RoB2 tool for randomized studies<sup>32</sup> and ROBINS-I<sup>33</sup> for nonrandomized studies. Disagreements were resolved by consensus and discussion with a third reviewer (D.B.) if necessary. The risk of bias for each outcome across individual studies was summarized by risk of bias tables. We also produced a review-level narrative summary of the risk of bias.

### Principal Summary Measures

We used ratios to present principal differences between modalities as we considered the relative level of agreement, cost, or acceptability between each approach for a given SES item to be more important than the absolute level.

### Strategy for Data Synthesis

Had data been available, we planned to pool effect estimates using a random-effects model.<sup>34</sup> Given the heterogeneity in study design, interventions, and outcomes of the included studies, we used a narrative synthesis without a meta-analysis approach, following reporting guidelines from Campbell and colleagues.<sup>35</sup> We stratified the synthesis by modality, SES domain, and outcome. We assessed heterogeneity by considering study design, interventions, and outcomes. To assess the risk of bias across studies we assessed selective outcome reporting by comparing protocols (when available) with published reports.

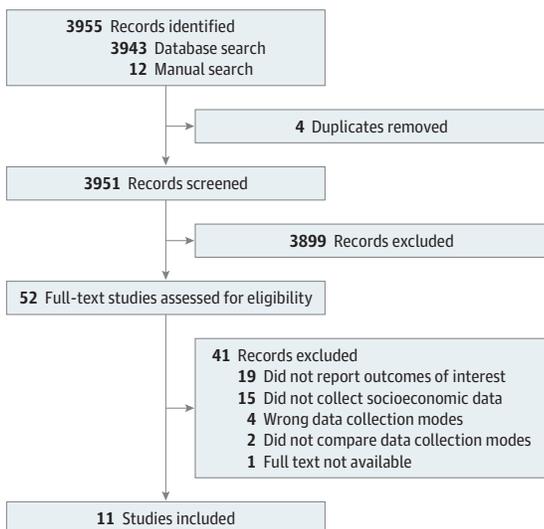
### Additional Analyses

We planned to exclude studies at high risk of bias from the synthesis and primary analysis. We planned to perform a secondary analysis that included all studies irrespective of their risk of bias. We also planned to perform a secondary analysis assessing whether findings differed between high-income and LMICs.

### Assessment of Certainty of Evidence

We used the Grading of Recommendations, Assessment, Development and Evaluations criteria to assess the certainty of the primary outcomes.<sup>36,37</sup> One of us (L.N.A.) collated the evidence for each primary outcome and suggested initial ratings that were discussed with another of us (S.M.) and agreed on by joint decision. For randomized clinical trials, evidence was assumed to be of high certainty and then downgraded due to risk of bias, inconsistency of results, indirectness of evidence, imprecision, or publication bias. For observational studies, evidence started at low certainty but was upgraded for a large effect size, dose-response, gradient, or plausible confounding that decreases the magnitude of effect.

Figure 1. Study Flow Diagram



## Results

### Search Results

Our search returned 3943 records and additional searches returned a further 11 studies (Figure 1). We contacted 24 study authors for full texts or missing data. Only 1 study<sup>38</sup> was excluded because we could not obtain the full text.

### Study Characteristics

The 11 included studies reported data on 14 036 individuals from 7 countries: 5 from the US,<sup>39-43</sup> 2 from Australia,<sup>44,45</sup> and 1 each from Bangladesh and Tanzania,<sup>22</sup> Burkina Faso,<sup>46</sup> Kenya,<sup>47</sup> and the Netherlands<sup>48</sup> (Figure 2). As such, 3 studies (27.3%) reported data from 4 LMICs. All studies were published in English. Table 1 summarizes the included studies' designs, modes used, SES domains, and outcomes.

### Study Designs

One study used a randomized crossover survey design.<sup>22</sup> Parallel 2-arm<sup>39,40,48</sup> and 4-arm<sup>44</sup> surveys were more prevalent, with participants randomly allocated to different survey instruments and comparisons made between the instruments. Gagliardi et al<sup>41</sup> used a nonrandomized parallel 2-arm approach. Greenleaf et al<sup>46</sup> randomized participants to CATIs or interactive voice response calls (IVRs) and compared response rates between arms, but also compared both arms with findings from an in-person survey completed 11 months previously to calculate equivalence. The 4 remaining studies used test-retest approaches.<sup>42,43,45,47</sup> The vast majority of studies collected SES data as part of broader surveys. Only Chittleborough et al<sup>45</sup> had the primary aim of comparing different modalities for collecting SES data.

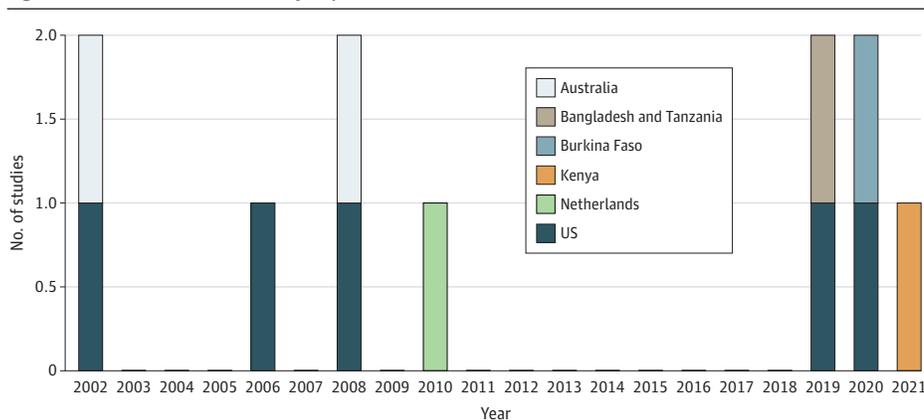
### Risk of Bias

eFigure 1 and eFigure 2 in the Supplement summarize the risk of bias for each study. Overall, 7 studies were found to be at low risk of bias; we had some concerns regarding 4 studies, and none were found to be at high risk of bias. The risk of bias across studies (including selective outcome reporting) was low to moderate.

### Data Collection Modalities

None of the included studies used SMS, multimedia message (MMS), or non-CATI approaches. CATIs were used in all 11 studies: this approach entails conducting real-time telephone calls and leading participants through a series of questions read from a computer screen, with responses usually entered using the same program. Three studies used data collected as part of an existing national

Figure 2. Publication Year and Study Population Location of Included Studies



survey,<sup>44,45,48</sup> 1 study used data collected by primary care administrative staff as part of an implementation and comparative effectiveness study,<sup>41</sup> and the remaining 7 studies used members of the research team to collect the data; in 3 of these studies data were collected as part of a larger parent study.<sup>42,43,47</sup>

In-person data collection,<sup>40,45-47</sup> web surveys,<sup>39,42,43,48</sup> and IVRs<sup>22,41,44,46</sup> were each assessed by 4 studies. Two studies<sup>44,46</sup> included hybrid IVR arms when a researcher called the participant at the beginning or end of the IVR data collection activity. We included these studies because all SES data were collected during the IVR phase; however, we have singled these studies out in the ensuing analyses because we might expect this approach to achieve a higher response rate than IVR approaches with no associated human interaction. All of the studies directly compared CATIs against one other approach except for Greenleaf et al,<sup>46</sup> who compared CATIs against IVRs for response rate, time, and costs, and they compared CATI and IVR approaches against in-person survey for equivalence. eFigure 3 in the Supplement illustrates the comparisons made between each modality.

### Socioeconomic Domains

Eleven different SES domains were reported across the 11 included studies (eTable 1 in the Supplement). More than one-third of the studies collected data on educational level, marital status, household income, and employment; however, multiple different response options were used, and

Table 1. Study Characteristics of Included Studies Reporting the Performance Characteristic of 2 or More SES Data Collection Modes

Source	Design	Population	Study focus	Modes used to collect SES data	SES domains	Outcome domains
Corkrey and Parkinson, <sup>44</sup> 2002, Australia	Parallel, randomized, 4-arm survey	2880 Adults with fixed telephone connections, nationally representative sample	Drugs and alcohol use survey	CATI, IVR, hybrid CATI/IVR	Educational level, marital status, country of birth, employment	Costs, acceptability
Ellen et al, <sup>39</sup> 2002, US	Randomized, parallel 2-arm survey	223 African American adolescents living in San Francisco	Teen sexual behavior data collection	CATI, web survey	Household structure, school enrollment	Costs, acceptability
Graham et al, <sup>43</sup> 2006, US	Test-retest: CATI followed by web survey 2 d later	213 Internet users who searched for stop smoking and navigated to the intervention site	Smoking habits survey, nested within a RCT testing a smoking cessation intervention	CATI, web survey	Household income	Equivalence (Cohen κ)
Graham and Papandonatos, <sup>42</sup> 2008, US	Test-retest: CATI then web 2 d later	422 Internet users who searched for stop smoking and navigated to the intervention site	Smoking habits survey, nested within an RCT testing a smoking cessation intervention	CATI, web survey	Household income	Equivalence (Cohen κ)
Chittleborough et al, <sup>45</sup> 2008 Australia	Test-retest: in-person then CATI 6 mo later	2206 South Australian adults living in metropolitan areas and listed in the electronic white pages	SES data collection	CATI, in-person	Parental educational level, occupation, employment status, household income, educational level, urban/rural, country of birth, marital status	Response rate
Nagelhout et al, <sup>48</sup> 2010, the Netherlands	Randomized, parallel, 2-arm survey	2072 Adult smokers registered with an online survey database	Tobacco use data collection	CATI, web survey	Educational level, marital status	Response rate, time, costs
English et al, <sup>40</sup> 2019, US	Parallel, randomized, 2-arm survey	900 Adults from rural American Indian communities in New Mexico	General public health survey	CATI, in-person	Educational level; household income, employment status	Response rate, time, costs
Pariyo et al, <sup>22</sup> 2019, Bangladesh and Tanzania	Randomized crossover survey	2196 Adults with mobile phone access in Bangladesh and Tanzania	Noncommunicable diseases data collection	CATI, IVR	Education, urban/rural	Equivalence (Cohen κ)
Gagliardi et al, <sup>41</sup> 2020, US	Parallel, nonrandomized, 2-arm study	1008 Women overdue for cancer screening in a US health system	Primary care cancer screening outreach	CATI, IVR	Primary care registration	Costs
Greenleaf et al, <sup>46</sup> 2020, Burkina Faso	Randomized, parallel 2-arm survey	1766 Women aged 15-49 y who own a mobile phone	Family planning data collection	CATI, IVR <sup>a</sup>	Educational level, marital status, urban/rural	Response rate, equivalence (Cohen κ), costs, time
Ashgibie et al, <sup>47</sup> 2021, Kenya	Test-retest CATI then in person <24 h later for a 10% subsample	130 Adults registered with Kenyan health facilities	Access to medicines survey	CATI, in person	Educational level, wealth	Time, costs

Abbreviations: CATI, computer-assisted telephone interview; IVR, interactive voice response; RCT, randomized clinical trial; SES, socioeconomic status.

<sup>a</sup> Greenleaf et al<sup>46</sup> used hybrid-IVR: participants were first called by a researcher to set up the process and take consent, and the participant was then transferred to an IVR system for data collection.

no 2 studies used exactly the same wording or response options. eTable 2 in the [Supplement](#) provides the survey items and response options used for each SES domain within each study.

### Response Rate

Four studies presented data on the response rates for individual questions, defined as the number of completed SES responses divided by the total number of elicitation attempts (**Table 2**). Not every study provided sufficient data to permit the calculation of 95% CIs.

Socioeconomic status data collection using CATIs was found to have either superior or equivalent response rates compared with IVRs. The response rates were found to be similarly high in each domain by Gagliardi et al<sup>41</sup> and 100% in all the SES domains collected by Greenleaf et al,<sup>9</sup> whereas response rates using IVRs ranged from 68% to 73%. Nagelhout et al<sup>48</sup> found response rates using CATIs and web-based data collection to be similarly high.

Chittleborough et al,<sup>45</sup> the only study to report response rates for individual SES domain-level questions, compared CATIs and in-person data collection and found similar response rates between the 2 methods, although English et al<sup>40</sup> reported an overall survey completion rate of 35.7% using CATIs compared with 68.9% in-person—a ratio of 0.52. English et al<sup>40</sup> also reported that this lower rate was noted despite the fact that the CATI was significantly shorter (25 vs 45 minutes). A potential confounding factor was that a nominal incentive was offered to individuals who completed the in-person survey, but this was not logistically possible to offer those completing CATIs, although the English et al<sup>40</sup> highlighted that the interviewers were trained not to mention the incentive until after the survey had been completed to reduce the risk of bias.

### Equivalence

Six studies assessed the level of agreement between the SES responses obtained from 2 or more different modalities. All used weighted  $\kappa$  coefficients. eTable 3 in the [Supplement](#) presents findings by SES domain. In a crossover design, Pariyo et al<sup>22</sup> presented 2 sets of coefficients for each indicator depending on which modality was used first. The authors provided no interpretation for the very low agreement between IVRs and CATIs for education in Tanzania. They noted that the higher levels of agreement observed with performing IVRs first for other domains (which extend beyond the 2 SES domains presented herein) may be due to a form of selection bias where less-educated people may drop out of IVRs.<sup>22</sup> Apart from this domain, all other  $\kappa$  values were greater than 0.51, which Cohen<sup>49</sup> suggested interpreting as moderate agreement, with many exceeding 0.8: almost perfect agreement.

**Table 2. Response Rates**

Source	Domains <sup>a</sup>	CATI response rate, %	Response rate, comparator, %	Ratio CATI/comparator
Chittleborough et al, <sup>45</sup> 2008	Highest level of education	100	100 In person	1.00
	Occupation (6 categories)	100	100 In person	1.00
	Employment status (7 categories)	99	98.1 In person	1.01
	Household income (4 categories)	89.2	88.4 In person	1.01
	Area of residence (metropolitan/country)	100	100 In person	1.00
	Marital status	100	100 In person	1.00
	Country of birth	100	100 In person	1.00
Pariyo et al, <sup>22</sup> 2019	Residential area	100	68 IVR	1.47
	Ever attended school	100	71 IVR	1.41
	Marital status	100	73 IVR	1.37
Nagelhout et al, <sup>48</sup> 2010	Educational level	96.8	99.2 Web survey	0.98
	Marital status	99.5	99.7 Web survey	1.00
Gagliardi et al, <sup>41</sup> 2020	Insurance	99.3	99.5 IVR	1.00

Abbreviations: CATI, computer-assisted telephone interview; IVR, interactive voice response.

<sup>a</sup> The denominator for each domain is the entire population for each study listed in Table 1.

## Time

Three studies quantified the time taken to gather SES data using different approaches (eTable 4 in the [Supplement](#)). None presented ranges and Nagelhout et al<sup>48</sup> and English et al<sup>40</sup> did not present times for both of the approaches that they used. All 3 studies presented the time taken to complete the entire survey—not just the SES instruments. Ellen et al<sup>39</sup> and Nagelhout et al<sup>48</sup> used the same number and wording of questions irrespective of modality. Ashigbie et al<sup>47</sup> found that CATIs were 1.48 times slower than in-person surveying, but crucially, this did not include the time taken to travel to each household.

## Costs

Seven studies presented cost data<sup>39-41,44,46-48</sup>; however, there was little consistency in the cost items included in the estimations for each modality and, in some cases, specific details of costs included were not provided. All studies that reported cost data compared CATIs with another mode of data collation, and there was notable variability in the cost-effectiveness, measured as cost per completed interview, of the different modalities between the studies related to response rates, interviewer costs, and participant reimbursement. We present the ratio of CATIs to other modes in eTable 5 in the [Supplement](#).

Two studies compared CATIs with in-person interviewing: English et al<sup>40</sup> found that both methods incurred high costs, but in-person interviewing was more cost-effective than telephone per completed survey due to the low response rate of telephone administration among American Indian or Alaska Native rural populations. Conversely, Ashigbie et al<sup>47</sup> found telephone interviewing to be less expensive than in-person interviewing in semiurban and rural communities in Kenya. Although the interviews took longer, the process was less time-consuming because data collectors did not have to travel, often via poor road networks, to houses that may not be close to each other, incurring further cost. Nagelhout et al<sup>48</sup> found web surveys to be more cost-effective than CATIs due to lower fieldwork costs and slightly lower participant reimbursements required, while Ellen et al<sup>39</sup> found web surveys to be more expensive when combining actual costs for interviewers, mailing, telephones, travel, incentives, and supplies.

One study found IVRs to be more cost-effective than CATIs owing to reduced personnel costs,<sup>41</sup> but 2 studies<sup>44,46</sup> found IVRs to be less cost-effective due to the costs associated with recording the automated survey in multiple languages, additional airtime costs to complete the survey, and lower completion rates.

## Acceptability

None of the studies explored acceptability to providers. Two studies presented data on acceptability to participants: Ellen et al<sup>39</sup> found no statistically significant differences ( $P > .05$ ) in perceived comfort, honesty, and accuracy in answering full surveys delivered by CATIs vs web survey. We note that Ellen et al<sup>39</sup> did not single out acceptability of the SES-specific questions. Corkrey and Parkinson<sup>44</sup> assessed participants' perception of ease, enjoyment, stress, and likability using IVRs and CATIs. Both methods scored equally highly for all 4 domains. eTable 6 in the [Supplement](#) presents the GRADE level of certainty for each of the key findings from the review's primary outcomes.

## Secondary Analyses

None of the studies had high risk of bias, so none were excluded from the primary analyses. When we repeated the analyses comparing studies conducted in high-income vs LMIC settings we found that there were insufficient data to compare equivalence or time requirements for different modes. Greenleaf et al<sup>46</sup> found a lower response rate with IVRs in Burkina Faso (72%) than Nagelhout et al<sup>48</sup> found with the same modality in the Netherlands (99%); however, participants in the latter study were financially reimbursed, so this example is not a fair comparison. Ashigbie et al<sup>47</sup> and Greenleaf

et al<sup>46</sup> both obtained very high CATI response rates (>95%) in LMICs; however, response rates were similarly high for the same items asked in high-income settings.

The cost per completed CATIs ranged from AU \$6 to US \$211 (approximately AU \$7 and US \$240 in 2022) depending on accounting practices. Heterogeneity in the application of each method and accounting practices precludes any firm conclusions, but data collection modes used in LMICs do not appear to be systematically more or less expensive than those used in high-income countries.

---

## Discussion

### Summary of Main Findings

Our systematic review included 11 studies that collected data on 11 different SES domains using 4 different modalities under the 3 overarching categories of in-person, voice call, and automated approaches. All studies used CATIs, 4 used web surveys, 4 used in-person approaches, and 3 studies used IVR methods. None of the included studies used SMS data collection, and all of the in-person approaches involved home visits. Despite an overall low risk of bias across the studies, comparisons were limited by marked heterogeneity in the SES items used.

There is not enough evidence to say whether automated approaches are less costly than nonautomated data collection modalities. This lack of evidence is mainly due to differences in costing approaches used, as well as heterogeneity in how each modality was used. Only Ashigbie et al<sup>47</sup> compared the time taken to complete surveys, finding that CATI was 1.48 times slower than in-person elicitation; however, their figure did not include the travel time involved for home visits so the level of certainty for this finding is very low. Two studies compared the acceptability of CATIs vs IVR<sup>44</sup> and CATIs vs web survey,<sup>39</sup> finding no statistically significant differences in reported comfort, honesty, accuracy, ease, enjoyment, stress, or likability, which were assessed at the level of the whole survey rather than isolating the SES questions.

We can be moderately certain that response rate is equally high for SES questions asked via CATI, web survey, and in-person interview. Response rates may be slightly lower for IVR than for other modes, which may be largely related to incomplete responses. Greenleaf et al<sup>46</sup> found high rates of break-off, where 19.7% of individuals ( $n = 174$ ) consented but answered less than 50% of the relevant questions using this method. We postulate that human-led interactions exert a stronger social pressure not to terminate the call partway through the interview.

Equivalence between answers elicited using automated vs nonautomated approaches was moderate to substantial for all comparisons made. Responses provided by CATIs seem to be equivalent to those provided by web survey and in-person interviews.

Equivalence was also generally moderate to high between CATI and IVR, with the marked exception of eliciting educational attainment in Tanzania ( $\kappa = 0.03$ ), where there appeared to be systematic underreporting at initial IVR compared with CATI follow-up. This finding suggests that there may have been a systematic issue in understanding this prerecorded question. The authors also noted that if a respondent accidentally entered an incorrect option on IVR, there was no facility to change their answer.<sup>22</sup>

In sum, CATI, web surveys, and in-person approaches can all attain very high response rates and appear to collect equivalent data. Our review found a slightly lower response rate with IVRs than the other modes, although this finding is based on 2 studies. We did not find sufficient evidence to suggest that time requirements, costs, or acceptability vary meaningfully between modes. Automated approaches (ie, web surveys and IVRs) have comparable response rates and similarly high perceived levels of acceptability compared with surveys conducted in person or with the telephone, although there are very few studies contributing evidence.

The time and costs for each mode seem to depend on the baseline telephone response rate for the population of interest and the distances involved in home visits: sometimes it may be more cost-effective to visit households than to repeatedly call. The length of telephone calls can also be a material factor when airtime is expensive, and there is low-quality evidence to suggest that IVRs may

take longer than human-led calls. However, we note that IVRs do not involve personnel costs beyond setting up and managing the software.

### Comparisons With the Wider Literature

The World Health Organization recommends that health programs and researchers should routinely gather socioeconomic data on a wide range of domains.<sup>5,6</sup> We note that none of our included studies collected data on religion, sexuality, or disability.

We found that respondents using IVRs and CATIs felt they were honest with their answers, even when answering sensitive questions. The wider evidence suggests that automated approaches, such as IVRs and web surveys, may obtain more honest answers than CATIs or in-person interviews<sup>15,50-52</sup> due to reduced social distance and desirability bias.<sup>22</sup> Automated approaches may also reduce bias that can arise from the social dynamics of interacting with a human, such as acquiescence<sup>18,53</sup> and nonuniform questions, because a computer presents the same question in the same way every time, whereas a person does not.<sup>54</sup> Social dynamics involved in providing answers to a real person may reduce the risk of satisficing (ie, providing the first/easiest option to complete the survey quickly).<sup>55</sup> We did not find evidence to support or refute this hypothesis. Self-administered approaches, such as web surveys, may place a higher cognitive burden on respondents that can lead to disengagement<sup>53</sup> and satisficing.<sup>55</sup> Coupled with our findings that web surveys tended to achieve low response rates and were not much less costly than other options, we recommend that researchers consider using alternative options. One final important source of difference between automated and nonautomated modes is the measurement error that can stem from the fact that respondents can ask for clarifications and amend their answers, whereas these options are often not available for IVR and some web survey modes.<sup>22</sup>

We did not find enough data to make robust comparisons between the use of different modes in LMICs vs high-income countries. Reviews conducted by Gibson et al<sup>12</sup> and Greenleaf et al<sup>9</sup> suggest that more research is required to understand the reliability and accuracy of different modes in low-income settings.

In 2015, Ballivian and colleagues<sup>8</sup> argued that telephone-based data collection approaches can introduce selection bias. This argument is less of a problem now that telephone ownership is so high around the world; however, low-income groups may be the least likely to own mobile phones and this is a material consideration for work seeking to obtain representative SES data for a given population. Remote and rural communities may also have unreliable network coverage. A further issue raised by Ballivian and colleagues<sup>8</sup> is the lower response rates from telephone-based approaches vs face-to-face data collection modes; however, we did not find this factor to be an issue in the included studies.

None of our included studies examined SMS/MMS or clinic-based data collection. A 2008 study from a California ambulatory care service found that collecting race and ethnicity and language data using a paper questionnaire at the front desk yielded an 88% response rate at a cost of \$0.21 per completed survey.<sup>56</sup> West and colleagues<sup>57</sup> found that CATIs were faster and less costly than manual SMS data collection for a 15-item survey of Nepalese adults. These studies were excluded from our review because they did not use comparators.

### Strengths and Limitations

Our study had a number of strengths: our search was designed by a Cochrane information specialist (I.G.), and we included a wide range of databases and other sources. We used independent dual screening, data extraction, and quality scoring, and followed best practice guidelines throughout the study. We included a wide range of outcomes to maximize the utility of the review for program managers faced with difficult decisions about which modality to use.

This study has limitations. The performance of individual SES items in a given questionnaire is likely to be influenced by the preceding items, the focus of the overall survey, and broader contextual factors. To minimize bias, we calculated and reported intermodal comparison rates rather than reporting absolute levels. Although this approach is methodologically robust, decision-makers are

unlikely to select a mode on the basis of how it performs for individual survey items. We did not search for or extract data on sample frame errors and nonresponse errors.<sup>58</sup> We excluded articles that were published before 1999, which may have excluded useful studies. We note that not all telephones can be used to access web surveys.

---

## Conclusions

Our review reinforces the message that the choice of survey mode should be guided by the type of questions being asked, the population, and the resources available.<sup>8,10</sup> We found that CATIs, IVRs, web surveys, and in-person interviews have all been used to attain high response rates with comparable answers in a range of settings. Marked heterogeneity in their deployment makes it very difficult to reach conclusions about their relative costs and benefits, and future work should aim to align accounting practices with those used by major reviews. Given the absence of evidence that automated and telephone-based systems deliver inferior data, we recommend that decision-makers try approaches that are likely to offer cost savings; however, it is important to review response rates early on and consider the extent to which selection bias is influencing the findings.

---

## ARTICLE INFORMATION

**Accepted for Publication:** October 13, 2022.

**Published:** November 28, 2022. doi:10.1001/jamanetworkopen.2022.43883

**Open Access:** This is an open access article distributed under the terms of the [CC-BY License](#). © 2022 Allen LN et al. *JAMA Network Open*.

**Corresponding Author:** Luke N. Allen, MPH, MBChB, International Centre for Eye Health, Department of Clinical Research, London School of Hygiene & Tropical Medicine, Keppel St, London WC1E 7HT, United Kingdom ([luke.allen@lshtm.ac.uk](mailto:luke.allen@lshtm.ac.uk)).

**Author Affiliations:** International Centre for Eye Health, Department of Clinical Research, London School of Hygiene & Tropical Medicine, London, United Kingdom (Allen, Gordon, Marques, Burton, Kim, Ramke, Bastawrous); NHS Education for Scotland, Glasgow, United Kingdom (Mackinnon); Institute of Health and Wellbeing, University of Glasgow, Glasgow, United Kingdom (Blane); Department of Ophthalmology, University of Nairobi, Nairobi, Kenya (Gichuhi); Operation Eyesight, Nairobi, Kenya (Mwangi); Peek Vision, Berkhamsted, United Kingdom (Bolster); London School of Hygiene & Tropical Medicine, London, United Kingdom (Bolster); International Statistics & Epidemiology Group, Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, United Kingdom (Macleod); School of Optometry and Vision Science, University of Auckland, Auckland, New Zealand (Ramke).

**Author Contributions:** Dr Allen had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

*Concept and design:* Allen, Marques, Gichuhi, Mwangi, Burton, Bolster, Kim, Bastawrous.

*Acquisition, analysis, or interpretation of data:* Allen, Mackinnon, Gordon, Blane, Macleod, Ramke.

*Drafting of the manuscript:* Allen, Gordon, Gichuhi.

*Critical revision of the manuscript for important intellectual content:* Allen, Mackinnon, Blane, Marques, Gichuhi, Mwangi, Burton, Bolster, Macleod, Kim, Ramke, Bastawrous.

*Statistical analysis:* Allen, Macleod, Kim.

*Obtained funding:* Burton, Bolster, Bastawrous.

*Administrative, technical, or material support:* Gordon, Marques, Gichuhi, Bolster, Bastawrous.

*Supervision:* Allen, Blane, Burton, Bastawrous.

**Conflict of Interest Disclosures:** Dr Bolster reported receiving grants from Wellcome Trust during the conduct of the study; salary from Peek Vision outside the submitted work. Dr Macleod reported receiving grants from Wellcome Trust during the conduct of the study. Dr Bastawrous reported being founder and chief executive officer of the nonprofit social enterprise Peek Vision, which powers eye health program providers to strengthen systems and service delivery with a software and data intelligence platform. No other disclosures were reported.

**Funding/Support:** This work was supported by the National Institute for Health Research (NIHR) using the UK's Official Development Assistance Funding and Wellcome (grant 215633/Z/19/Z) under the NIHR-Wellcome Partnership for Global Health Research.

**Role of the Funder/Sponsor:** The funding organization had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

**Disclaimer:** The views expressed are those of the authors and not necessarily those of Wellcome, the NIHR, or the Department of Health and Social Care.

## REFERENCES

1. Marmot M, Allen J, Goldblatt P, et al. *Fair Society Healthy Lives: The Marmot Review*. Institute of Health Equity; 2010. Accessed November 11, 2021. <https://www.instituteofhealthequity.org/resources-reports/fair-society-healthy-lives-the-marmot-review>
2. World Health Organization. Closing the gap in a generation: health equity through action on the social determinants of health—final report of the Commission on Social Determinants of Health. August 27, 2008. Accessed November 11, 2021. <https://www.who.int/publications-detail-redirect/WHO-IER-CSDH-08.1>
3. Burton MJ, Ramke J, Marques AP, et al. The *Lancet* Global Health Commission on Global Eye Health: vision beyond 2020. *Lancet Glob Health*. 2021;9(4):e489-e551. doi:10.1016/S2214-109X(20)30488-5
4. Hart JT. The inverse care law. *Lancet*. 1971;1(7696):405-412. doi:10.1016/S0140-6736(71)92410-X
5. World Health Organization. World Bank. Monitoring progress towards universal health coverage at country and global levels: framework, measures and targets. 2014. Accessed November 11, 2021. <https://apps.who.int/iris/handle/10665/112824>
6. World Health Organization. Making fair choices on the path to universal health coverage: final report of the WHO consultative group on equity and universal health coverage. 2014. Accessed October 15, 2021. <https://apps.who.int/iris/handle/10665/112671>
7. Allen LN, Smith RW, Simmons-Jones F, Roberts N, Honney R, Currie J. Addressing social determinants of noncommunicable diseases in primary care: a systematic review. *Bull World Health Organ*. 2020;98(11):754-765B. doi:10.2471/BLT.19.248278
8. Ballivian A, Azevedo J, Durbin W. Using mobile phones for high-frequency data collection. In: Toninelli D, Pinter R, de Pedraza P, eds. *Mobile Research Methods: Opportunities and Challenges of Mobile Research Methodologies*. Ubiquity Press; 2015:21-39.
9. Greenleaf AR, Gibson DG, Khattar C, Labrique AB, Pariyo GW. 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(5):e140. doi:10.2196/jmir.7331
10. Croke K, Dabalen A, Demombybes G, Giugale M, Hoogeveen J. Collecting high frequency panel data in Africa using mobile phone interviews. World Bank. June 2012. Accessed July 7, 2021. <https://openknowledge.worldbank.org/bitstream/handle/10986/9313/WPS6097.pdf?sequence=1&isAllowed=y>
11. World Bank. World development indicators 2021. Accessed July 7, 2021. <https://databank.worldbank.org/source/world-development-indicators>
12. Gibson DG, Pereira A, Farrenkopf BA, Labrique AB, Pariyo GW, Hyder AA. Mobile phone surveys for collecting population-level estimates in low- and middle-income countries: a literature review. *J Med internet Res*. 2017;19(5):e139. doi:10.2196/jmir.7428
13. Diedhiou A, Gilroy KE, Cox CM, et al. Successful mLearning pilot in Senegal: delivering family planning refresher training using interactive voice response and SMS. *Glob Health Sci Pract*. 2015;3(2):305-321. doi:10.9745/GHSP-D-14-00220
14. Long LA, Pariyo G, Kallander K. Digital technologies for health workforce development in low- and middle-income countries: a scoping review. *Glob Health Sci Pract*. 2018;6(suppl 1):S41-S48. doi:10.9745/GHSP-D-18-00167
15. 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(5):e110. doi:10.2196/jmir.7326
16. Larmarange J, Kassoum O, Kakou É, Fradier Y, Sika L, Danel C. Faisabilité et représentativité d'une enquête téléphonique avec échantillonnage aléatoire de lignes mobiles en Côte d'Ivoire. *Population*. 2016;71(1):121-134. doi:10.3917/popu.1601.0121
17. Leo B, Morello R, Mellon J, Peixoto T, Davenport ST. Do mobile phone surveys work in poor countries? April 24, 2015. Accessed April 12, 2022. <https://papers.ssrn.com/abstract=2597885>

18. Tourangeau R, Rips LJ, Rasinski K. *The Psychology of Survey Response*. Cambridge University Press; 2000. doi:10.1017/CBO9780511819322
19. Christian L, Foster K. Multi-mode surveys. In: Lavrakas PJ, ed. *Encyclopedia of Survey Research Methods*. SAGE Publications; 2008.
20. Kreuter F, Presser S, Tourangeau R. Social desirability bias in CATI, IVR, and web surveys: the effects of mode and question sensitivity. *Public Opin Q*. 2008;72(5):847-865. doi:10.1093/poq/nfn063
21. Mahfoud Z, Ghandour L, Ghandour B, Mokdad AH, Sibai AM. Cell phone and face-to-face interview responses in population-based surveys: how do they compare? *Field Methods*. 2015;27(1):39-54. doi:10.1177/1525822X14540084
22. Pariyo GW, Greenleaf AR, Gibson DG, et al. Does mobile phone survey method matter? reliability of computer-assisted telephone interviews and interactive voice response non-communicable diseases risk factor surveys in low and middle income countries. *PLoS One*. 2019;14(4):e0214450. doi:10.1371/journal.pone.0214450
23. Allen LN, Mackinnon S, Gordon I, et al. Performance and resource requirements of in-person versus voice call versus automated telephone-based socioeconomic data collection modalities for community-based health programmes: a systematic review protocol. *BMJ Open*. 2022;12(4):e057410. doi:10.1136/bmjopen-2021-057410
24. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372(71):n71. doi:10.1136/bmj.n71
25. Higgins J, Altman D, Sterne J. Assessing risk of bias in included studies. 2020. Accessed March 16, 2021. [https://handbook-5-1.cochrane.org/chapter\\_8/8\\_assessing\\_risk\\_of\\_bias\\_in\\_included\\_studies.htm](https://handbook-5-1.cochrane.org/chapter_8/8_assessing_risk_of_bias_in_included_studies.htm)
26. The King's Fund. Community health services explained. January 14, 2019. Accessed April 28, 2022. <https://www.kingsfund.org.uk/publications/community-health-services-explained>
27. American Psychological Association. Socioeconomic status. Accessed April 28, 2022. <https://www.apa.org/topics/socioeconomic-status>
28. World Health Organization. Social determinants of health: key concepts. May 7, 2013. Accessed April 28, 2022. <https://www.who.int/news-room/questions-and-answers/item/social-determinants-of-health-key-concepts>
29. World Health Organization. Social determinants of health. The solid facts. Second edition. Accessed April 28, 2022. <https://apps.who.int/iris/handle/10665/326568>
30. Marcano Belisario JS, Jamsek J, Huckvale K, O'Donoghue J, Morrison CP, Car J. Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods. *Cochrane Database Syst Rev*. 2015;(7):MR000042. doi:10.1002/14651858.MR000042.pub2
31. Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value Health*. 2008;11(2):322-333. doi:10.1111/j.1524-4733.2007.00231.x
32. Cochrane. RoB 2: a revised Cochrane risk-of-bias tool for randomized trials. Accessed April 13, 2022. <https://methods.cochrane.org/bias/resources/rob-2-revised-cochrane-risk-bias-tool-randomized-trials>
33. Cochrane. ROBINS-I tool. Accessed April 13, 2022. <https://methods.cochrane.org/methods-cochrane/robins-i-tool>
34. Deeks JJ, Higgins JPT, Altman DG. Chapter 10: analysing data and undertaking meta-analyses. Accessed April 13, 2022. <https://training.cochrane.org/handbook/current/chapter-10>
35. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWIM) in systematic reviews: reporting guideline. *BMJ*. 2020;368:l6890. doi:10.1136/bmj.l6890
36. BMJ Best Practice. What is GRADE? Accessed March 19, 2021. <https://bestpractice.bmj.com/info/toolkit/learn-ebm/what-is-grade/>
37. Murad MH, Mustafa RA, Schünemann HJ, Sultan S, Santesso N. Rating the certainty in evidence in the absence of a single estimate of effect. *Evid Based Med*. 2017;22(3):85-87. doi:10.1136/ebmed-2017-110668
38. Midanik LT, Greenfield TK, Rogers JD. Reports of alcohol-related harm: telephone versus face-to-face interviews. *J Stud Alcohol*. 2001;62(1):74-78. doi:10.15288/jsa.2001.62.74
39. Ellen JM, Gurvey JE, Pasch L, Tschann J, Nanda JP, Catania J. A randomized comparison of A-CASI and phone interviews to assess STD/HIV-related risk behaviors in teens. *J Adolesc Health*. 2002;31(1):26-30. doi:10.1016/S1054-139X(01)00404-9
40. English KC, Espinoza J, Pete D, Tjemslund A. A Comparative analysis of telephone and in-person survey administration for public health surveillance in rural American Indian communities. *J Public Health Manag Pract*. 2019;25(suppl 5):S70-S76. doi:10.1097/PHH.0000000000001007

41. Gagliardi KS, Coleman S, Intinarelli G, et al. An automated telephone call system improves the reach and cost-effectiveness of panel management outreach for cancer screening. *J Ambul Care Manage*. 2020;43(2):148-156. doi:10.1097/JAC.0000000000000322
42. Graham AL, Papandonatos GD. Reliability of internet- versus telephone-administered questionnaires in a diverse sample of smokers. *J Med internet Res*. 2008;10(1):e8. doi:10.2196/jmir.987
43. Graham AL, Papandonatos GD, Bock BC, et al. Internet- vs telephone-administered questionnaires in a randomized trial of smoking cessation. *Nicotine Tob Res*. 2006;8(suppl 1):S49-S57. doi:10.1080/14622200601045367
44. Corkrey R, Parkinson L. A comparison of four computer-based telephone interviewing methods: getting answers to sensitive questions. *Behav Res Methods Instrum Comput*. 2002;34(3):354-363. doi:10.3758/BF03195463
45. Chittleborough CR, Taylor AW, Baum FE, Hiller JE. Non-response to a life course socioeconomic position indicator in surveillance: comparison of telephone and face-to-face modes. *BMC Med Res Methodol*. 2008; 8(1):54. doi:10.1186/1471-2288-8-54
46. Greenleaf AR, Gadiaga A, Choi Y, et al. Automated and interviewer-administered mobile phone surveys in Burkina Faso: sociodemographic differences among female mobile phone survey respondents and nonrespondents. *JMIR Mhealth Uhealth*. 2020;8(7):e17891. doi:10.2196/17891
47. Ashgibie PG, Rockers PC, Laing RO, et al. Phone-based monitoring to evaluate health policy and program implementation in Kenya. *Health Policy Plan*. 2021;36(4):444-453. doi:10.1093/heapol/czab029
48. Nagelhout GE, Willemsen MC, Thompson ME, Fong GT, van den Putte B, de Vries H. Is web interviewing a good alternative to telephone interviewing? findings from the International Tobacco Control (ITC) Netherlands survey. *BMC Public Health*. 2010;10:351. doi:10.1186/1471-2458-10-351
49. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37-46. doi:10.1177/001316446002000104
50. Ali J, DiStefano MJ, Coates McCall I, et al. Ethics of mobile phone surveys to monitor non-communicable disease risk factors in low- and middle-income countries: a global stakeholder survey. *Glob Public Health*. 2019;14(8):1167-1181. doi:10.1080/17441692.2019.1566482
51. Midanik LT, Greenfield TK. Interactive voice response versus computer-assisted telephone interviewing (CATI) surveys and sensitive questions: the 2005 National Alcohol Survey. *J Stud Alcohol Drugs*. 2008;69(4):580-588. doi:10.15288/jsad.2008.69.580
52. Keeter S. From telephone to the web: the challenge of mode of interview effects in public opinion polls. Pew Research Center Methods. May 13, 2015. Accessed April 12, 2022. <https://www.pewresearch.org/methods/2015/05/13/from-telephone-to-the-web-the-challenge-of-mode-of-interview-effects-in-public-opinion-polls/>
53. Ongena YP, Dijkstra W. A model of cognitive processes and conversational principles in survey interview interaction. *Appl Cogn Psychol*. 2007;21(2):145-163. doi:10.1002/acp.1334
54. Groves RM. Theories and methods of telephone surveys. *Annu Rev Sociol*. 1990;16(1):221-240. doi:10.1146/annurev.so.16.080190.001253
55. Krosnick JA, Narayan S, Smith WR. Satisficing in surveys: initial evidence. *New Dir Eval*. 1996;1996(70):29-44. doi:10.1002/ev.1033
56. Palaniappan LP, Wong EC, Shin JJ, Moreno MR, Otero-Sabogal R. Collecting patient race/ethnicity and primary language data in ambulatory care settings: a case study in methodology. *Health Serv Res*. 2009;44(5, pt 1):1750-1761. doi:10.1111/j.1475-6773.2009.00992.x
57. West BT, Ghimire D, Axinn WG. Evaluating a modular design approach to collecting survey data using text messages. *Surv Res Methods*. 2015;9(2):111-123.
58. Biemer PP. Total survey error: design, implementation, and evaluation. *Public Opin Q*. 2010;74(5):817-848. doi:10.1093/poq/nfq058

#### SUPPLEMENT.

**eMethods.** Search Strategy

**eFigure 1.** RoB-2 Risk of Bias Table for Randomized Studies

**eFigure 2.** ROBINS-I Risk of Bias Table for Nonrandomized Studies

**eFigure 3.** Comparisons Between Data Collection Modes

**eTable 1.** Socioeconomic Domains Used in the Included Studies

**eTable 2.** Response Options Within Each SES Domain

**eTable 3.** Equivalence Between Modalities of 6 Studies Reporting Equivalence Between Data Collection Modes

**eTable 4.** Time Requirement for Each Modality

**eTable 5.** Cost per Completed Interview for Different Modes

**eTable 6.** Summary of Key Findings

**eReferences**