Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Detecting inattentive responding in commercial-panel surveys of healthcare professionals: a multi-indicator framework with empirical demonstration in a rural Appalachian sample

Temitope S Ajagbe[1]

temitope.ajagbe1@waldenu.edu

Walden University

Research Article

Keywords: survey methodology, data quality, inattentive responding, satisficing, commercial online panel, healthcare professionals, sensitivity analysis, reporting checklist, CHERRIES, rural healthcare

Posted Date: May 21st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9647512/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Commercial online research panels are increasingly used to recruit healthcare professionals for survey research, but data-quality screening procedures for clinician-panel samples are not standardized, and the sensitivity of substantive effect estimates to analytical choices about respondent inclusion is rarely reported. Existing reporting checklists such as CHERRIES address survey administration but do not cover post-collection data-quality screening or sensitivity-analysis design.

Methods

We propose a multi-indicator framework for detecting inattentive responding in commercial-panel surveys of healthcare professionals. The framework combines three indicator families: completion-time indicators (median, percentile thresholds, sub-threshold completions), response-variance indicators (unique Likert responses, modal-response proportion, full straight-lining), and combined indicators (fast-and-modal). We empirically demonstrate the framework using data from a cross-sectional online survey of self-identified rural Appalachian healthcare professionals (N = 100) recruited through a commercial panel between September and December 2025, with a 40-item Likert instrument adapted from validated Diffusion of Innovations (DOI) measures. We report 6 sensitivity analyses comparing the focal effect estimate (the association between a structural-readiness composite and current mobile health [mHealth] adoption) across alternative respondent-inclusion definitions.

Results

In the demonstration sample, completion-time indicators flagged 18% (18/100) of respondents as fast (under 2 minutes for a 40-item instrument) and 47% (47/100) as completing in under 3 minutes. Response-variance indicators flagged 15% (15/100) as full straight-liners (1 unique Likert response across 34 attitudinal items) and 23% (23/100) with modal-response proportion at or above 80%. Sensitivity analyses across 6 respondent-inclusion definitions (full sample [n = 100]; excluding straight-liners [n = 85]; excluding fast and modal [n = 66]; excluding modal at or above 80% [n = 77]; verified target-state respondents only [n = 91]; named-setting respondents [n = 78]) yielded focal odds ratios of 0.18 to 0.25 (all P<.05), demonstrating that the substantive effect estimate moved within a 39% range across analytically defensible inclusion choices. We synthesize these results into a 9-item reporting checklist (DQ-CSC) intended to complement CHERRIES for clinician-panel surveys.

Conclusions

Data-quality screening choices in clinician-panel surveys meaningfully affect substantive effect estimates and warrant transparent reporting. The proposed framework and DQ-CSC reporting checklist provide a reusable procedure for detecting and reporting inattentive responding that is calibrated to clinician-panel samples and complementary to existing internet-survey reporting guidelines. The empirical demonstration is restricted to a single rural healthcare professional sample; further calibration across specialty, urban-rural, and international clinician panels is needed.

Background

Modern health services and digital health research frequently utilize online platforms for participant recruitment and survey administration. Commercial online research panels provide a practical and efficient means of reaching healthcare professionals across diverse regions and specialties, serving as a robust alternative to traditional sampling methods like state licensure registries or professional society lists. However, as established in survey methodology literature, these commercial panels often face significant data-quality challenges, including "satisficing," inattentive responding, and potential respondent fraud [1–3]. Recent studies indicate that nonprobability online panels typically show higher rates of repetitive response patterns (straight-lining) compared to probability-based panels[4], and these behaviors tend to persist over time [5]. More importantly, such inattentive or fraudulent responses can distort statistical findings in ways that are not easily identified through standard analysis[6, 7].

Surveys of healthcare professionals face unique data-quality hurdles. Clinicians are costly to recruit, resulting in smaller samples where even a few poor-quality responses can significantly alter results that influence policy or practice. Despite this, guidance for vetting these data remains scattered. While the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) is the standard for reporting how web surveys are run, it does not detail how to clean the data once collected [8]. Recent studies have introduced methods to spot fraudulent responses or shared cleaning experiences in health behavior research [9, 10], but these ideas have not yet formed a unified system for clinician research. Specifically, they lack clear instructions on how to report how much final results might change based on which respondents are included or excluded.

Three main gaps in current research drive this study. First, common quality checks—such as how fast someone finishes a survey or if they give the same answer to every question—are usually reported in isolation. They are rarely used together as a structured screening system with specific "pass/fail" rules. Second, researchers seldom report how their final findings vary depending on their data-cleaning choices. Without this, readers cannot know if a reported result is stable or if it would look different under a slightly different set of inclusion rules. Third, existing reporting standards like CHERRIES cover the "who and how" of survey administration but offer no checklist for the "what and why" of post-collection data screening.

To address these gaps, we introduce a multi-indicator framework to identify inattentive responding in commercial-panel surveys of healthcare professionals. We test this framework using a survey of 100

rural Appalachian clinicians. We focused on this group because rural healthcare workers are often recruited via commercial panels to save costs, making rigorous data screening vital. While our specific thresholds (like completion time) may need adjustment for different surveys, we argue that the overall structure of our framework and reporting checklist provides a valuable starting point for the field.

This work makes three contributions. First, we provide a three-part framework for screening clinician survey data. Second, we show through our example that a key result shifted by 39% across six different, defensible ways of defining the sample, highlighting how screening choices impact outcomes. Finally, we provide a new 9-item checklist, the DQ-CSC (Data Quality reporting for Clinician Surveys via Commercial panels), designed to work alongside CHERRIES to improve transparency in reporting.

Methods

Framework structure

The proposed framework is organized into 3 indicator families, each addressing a distinct mechanism by which inattentive responding manifests in survey data.

Family 1: Completion-time indicators

Completion-time indicators capture the simplest signal of inattentive responding: respondents who complete the instrument implausibly quickly. The framework specifies 5 reportable completion-time indicators: mean completion time; median completion time; 25th–75th percentile range; full range; and the proportion of respondents completing the instrument at or below 2 fixed thresholds (typically 60 seconds, 120 seconds, 180 seconds, and 300 seconds). Threshold selection is instrument-dependent. We recommend that researchers calibrate thresholds against an a priori estimate of the minimum plausible completion time, computed as approximately 3 to 5 seconds per Likert item plus an allowance for introduction, transitions, and demographic items. For a 40-item instrument with brief demographic items, this yields a minimum plausible completion time of approximately 2 minutes. Completion below this threshold is high-risk; completion below half this threshold (60 seconds for the example instrument) is implausible and warrants exclusion.

Family 2: Response-variance indicators

Response-variance indicators capture inattentive responding that manifests as low variability in answers across attitudinal items. The framework specifies 4 reportable response-variance indicators computed across the attitudinal Likert items only (excluding demographic and adoption-status items, which legitimately have low variance for many respondents): the proportion of respondents giving a single unique Likert response across all attitudinal items (“full straight-lining”); the proportion giving 2 or fewer unique Likert responses; the proportion with modal-response proportion at or above 90%; and the

proportion with modal-response proportion at or above 80%. Modal-response proportion is the share of attitudinal items that received the most-frequent response by that respondent. Full straight-lining (unique = 1) is a strong signal of non-engagement; modal-response proportions of 80% to 90% capture respondents whose engagement was variable but predominantly stereotyped. Both signal types are well-established in the satisficing literature [4, 11, 12].

Family 3: Combined indicators

Combined indicators address respondents who exhibit moderate patterns on both speed and variance dimensions but might not be flagged by either family alone. The framework specifies 1 primary combined indicator: the proportion of respondents who completed the instrument in less than 3 minutes AND had a modal-response proportion at or above 50% ("fast-and-modal"). This indicator captures respondents whose individual signals are subthreshold but whose joint pattern suggests low engagement. Researchers may define additional combined indicators relevant to their instrument (e.g., fast completion AND skipping open-ended items where present).

Sensitivity-analysis design

The framework requires that researchers report the sensitivity of their focal substantive estimates to alternative respondent-inclusion choices. We recommend a 6-restriction design that maps directly onto the 3 indicator families plus 2 sampling-frame validation restrictions: (A) full sample [primary analysis]; (B) excluding full straight-liners; (C) excluding the joint fast-and-modal subgroup; (D) excluding all respondents with modal-response proportion at or above 80%; (E) restricted to respondents whose self-reported sampling-frame characteristics (e.g., licensure state, role) are consistent with the planned frame; (F) restricted to respondents whose practice setting falls within a named (non-residual) category. Each restriction is analytically defensible: each would be a reasonable choice for a researcher concerned about the corresponding threat. The 6-restriction design enables readers to see whether the focal estimate is robust across the analytical-choice space rather than contingent on a single inclusion definition.

Researchers reporting sensitivity analyses should report, for each restriction: (1) the n retained; (2) the focal effect estimate with 95% confidence interval and P value; (3) the effect-size difference from the primary analysis; and (4) a brief substantive interpretation of whether the restriction materially changes the study's conclusions. Where the focal estimate moves substantially across restrictions, the primary specification should be reconsidered.

Demonstration sample and instrument

We demonstrate the framework using data from a cross-sectional online survey of self-identified rural Appalachian healthcare professionals (N = 100) recruited through a commercial online research panel

between September and December 2025. The panel screened for self-identified practice in 7 Central or Southern Appalachian US states (Kentucky, Tennessee, West Virginia, Virginia, North Carolina, South Carolina, Georgia). Respondents completed a 40-item instrument including demographic items, an adoption-status item, and 34 attitudinal Likert items adapted from validated Diffusion of Innovations measures. Items were scored on 5-point Likert scales (1 = strongly disagree to 5 = strongly agree). The substantive analysis focused on the association between a 4-item structural-readiness composite (financial resources, leadership support, technical support, infrastructure adequacy; Cronbach α = .79) and current mobile health (mHealth) adoption, estimated using logistic regression adjusted for relative advantage, complexity, provider-estimated patient technology access, and years in practice.

The substantive findings of the survey are reported in 2 prior outputs from the same dataset: the corresponding author's doctoral dissertation [13] and a companion manuscript currently under consideration at JMIR mHealth and uHealth [14]. Both are cited transparently and the data overlap is fully disclosed. The present paper uses the same dataset to demonstrate a methodological framework, with focal substantive estimates reported only as the dependent variable in the sensitivity-analysis demonstration. Readers interested in the substantive findings are directed to the JMIR submission [14]; readers interested in the methodological framework should treat the substantive findings here as illustrative rather than primary.

Ethical considerations

The underlying survey was approved by the Walden University Institutional Review Board (approval number 08-28-25-1124176). Informed consent was obtained through a survey-based implied consent procedure approved by the IRB. No personally identifiable information was collected; survey responses were stored on the panel provider's secure platform. The present methodological analysis used de-identified data only and required no additional IRB review. Reporting follows the STROBE guideline for observational cross-sectional research where applicable; the proposed DQ-CSC checklist (Table 4) extends rather than replaces existing reporting standards.

Statistical analysis

Data-quality indicators were computed as described in the framework structure section. Sensitivity analyses replicated the primary substantive specification across 6 respondent-inclusion restrictions. All analyses were conducted in Python 3.11 (Python Software Foundation) using the pandas, scipy, and statsmodels libraries. P values are reported to 3 decimal places; values below .001 are reported as P<.001.

Results

Demonstration sample composition

The sample (N = 100) reported the following primary roles: physicians (11/100, 11%), nurse practitioners (28/100, 28%), physician assistants (14/100, 14%), and other healthcare professionals (47/100, 47%). Practice settings included hospital-owned clinics (39/100, 39%), independent or small group practices (25/100, 25%), Federally Qualified Health Centers and Rural Health Clinics (14/100, 14%), and respondents who selected “none of the above” (22/100, 22%). State licensure was distributed across the 7 target states, with 9% (9/100) reporting licensure in none of the 7 target states despite self-identifying as Appalachian healthcare professionals at panel screening. The mismatch between screening-stage self-identification and item-level licensure response provides one example of the verification gap that motivates Family 3 of the framework.

Family 1: Completion-time indicators

Survey completion duration (Table 1) was variable: median 186 seconds (IQR 138–297; range 51–1,930). Eighteen percent (18/100) of respondents completed the 40-item instrument in under 120 seconds, and 47% (47/100) completed in under 180 seconds. The minimum-plausible-completion estimate for this instrument was approximately 2 minutes (3 seconds per Likert item across 34 attitudinal items, plus 30 seconds for demographic and transition items); 18% of respondents fell below this threshold. One respondent (1/100) completed the instrument in 51 seconds, which is implausible for any plausible engagement model.

Table 1
Completion-time indicators for the demonstration sample (N = 100).

Indicator	Value	n (% of sample)
Mean completion time (seconds)	284	—
Median completion time (seconds)	186 (3 min 6 sec)	—
25th–75th percentile (seconds)	138–297	—
Range (seconds)	51–1,930	—
Completed in < 60 seconds	—	1 (1%)
Completed in < 120 seconds	—	18 (18%)
Completed in < 180 seconds	—	47 (47%)
Completed in < 300 seconds	—	75 (75%)

Indicators are computed across the full 40-item instrument. The 120-second threshold corresponds to an a priori minimum plausible completion time of approximately 2 minutes for the demonstration instrument (3 seconds per attitudinal item across 34 items, plus 30 seconds for demographic and transition items).

Family 2: Response-variance indicators

Across the 34 attitudinal Likert items (Table 2), 15% (15/100) of respondents gave a single unique Likert response to every item (“full straight-lining”), and 26% (26/100) gave 2 or fewer unique responses.

Modal-response proportion was at or above 90% for 18% (18/100) of respondents and at or above 80% for 23% (23/100). The full-straight-lining rate of 15% in this sample is consistent with the upper end of rates reported in nonprobability panel research (Cornesse and Blom [4] reported straight-lining rates of 8% to 17% across 7 nonprobability online panels in Germany; Goldammer et al. [11] reported careless-responding prevalence between 12% and 26% across 4 questionnaire-based studies). The 23% modal-response-proportion-at-or-above-80% rate suggests that response variance concerns extend beyond the full-straight-lining subgroup.

Table 2
Response-variance indicators for the demonstration sample (N = 100).

Indicator	n (% of sample)
Full straight-lining (1 unique Likert response across 34 items)	15 (15%)
≤ 2 unique Likert responses across 34 items	26 (26%)
Modal-response proportion ≥ 90% (across 34 items)	18 (18%)
Modal-response proportion ≥ 80% (across 34 items)	23 (23%)
Combined indicator: <3 min completion AND modal-response ≥ 50%	34 (34%)

Indicators are computed across the 34 attitudinal Likert items only (excluding demographic and adoption-status items). Modal-response proportion is the share of items receiving the most-frequent response by that respondent.

Family 3: Combined indicators

The combined fast-and-modal indicator (completion in less than 3 minutes AND modal-response proportion at or above 50%) flagged 34% (34/100) of respondents. This is substantially larger than either single-indicator subgroup (18% completed in under 2 minutes; 23% had modal-response proportion at or above 80%). The combined indicator captures respondents who are subthreshold on any single indicator but whose joint pattern suggests low engagement: a respondent who completes a 40-item instrument in 2 minutes 30 seconds with 60% modal-response proportion would not be flagged on any individual indicator but is captured by the combined family. The proportion flagged by the combined indicator is the most policy-relevant statistic for researchers deciding whether to retain the full sample or to apply restrictions — a third of the demonstration sample exhibited some combination of fast and stereotyped responding.

Sensitivity-analysis design and effect-estimate variation

We applied the proposed 6-restriction sensitivity-analysis design to the focal substantive estimate from the underlying survey: the association between a 4-item structural-readiness composite and current mHealth adoption, estimated using logistic regression adjusted for relative advantage, complexity, provider-estimated patient technology access, and years in practice. Across the 6 restrictions (Table 3), the focal odds ratio ranged from 0.18 (95% CI 0.06–0.56; P=.003) in the full sample to 0.25 (95% CI

0.07–0.88; P=.031) when respondents reporting “none of the above” for practice setting were excluded. The focal estimate remained statistically significant (P<.05) in all 6 restrictions, but the point estimate moved across a 39% range — from 0.18 to 0.25 — across analytically defensible inclusion choices. The widest confidence interval appeared in restriction (F) (CI width 0.81), and the narrowest in the full sample (CI width 0.50).

The sensitivity-analysis output provides 2 distinct types of information for the reader. First, it confirms that the focal substantive conclusion (a strong negative association between structural constraints and adoption) does not depend on a single inclusion choice. Second, it quantifies the analytical-choice-induced uncertainty in the point estimate: a reader interpreting the primary OR of 0.18 should know that defensible alternative inclusion choices yield estimates as high as 0.25, and that the true uncertainty in the point estimate is therefore wider than the primary 95% CI alone implies.

Table 3
Sensitivity analysis: focal effect estimate (structural-readiness composite as predictor of current mHealth adoption) across 6 respondent-inclusion restrictions.

Restriction	Inclusion definition	n	OR (95% CI)	P value
A	Full sample (primary analysis)	100	0.18 (0.06–0.56)	.003
B	Excluding full straight-liners	85	0.22 (0.07–0.67)	.008
C	Excluding fast (< 3 min) AND mode ≥ 50%	66	0.23 (0.07–0.79)	.020
D	Excluding modal-response ≥ 80%	77	0.19 (0.06–0.64)	.007
E	Verified target-state respondents only	91	0.23 (0.07–0.72)	.012
F	Excluding “None of the above” practice setting	78	0.25 (0.07–0.88)	.031

All 6 restrictions replicate the primary specification (logistic regression of current mHealth adoption on structural-constraints composite, relative advantage, complexity, provider-estimated patient technology access, and years in practice), restricted to the indicated subsample. Across the 6 restrictions, the focal odds ratio ranged from 0.18 to 0.25 (39% range) and remained statistically significant (P<.05) in all restrictions.

Reporting checklist: DQ-CSC

We synthesize the framework into a 9-item reporting checklist (Table 4) intended to complement, not replace, CHERRIES [8] and related internet-survey reporting guidelines. CHERRIES focuses on what was administered (sample frame, response rates, instrument design); DQ-CSC focuses on what was kept (data-quality screening, inclusion choices, sensitivity to those choices). The 2 are complementary and we recommend they be used together. The 9 items are intended to be addressable in approximately 1 paragraph of Methods and 1 paragraph of Results, plus 1 supplementary table; DQ-CSC reporting should not impose substantial additional length on a typical research article.

Table 4
DQ-CSC: Data Quality reporting checklist for Clinician Surveys via Commercial panels.

#	Item	Description
1	Panel provider and screening	Name the commercial panel provider, screening criteria for clinician status, and any verification of self-identified professional role beyond panel-stage screening.
2	Completion time — distribution	Report mean, median, IQR, and full range of completion times in seconds.
3	Completion time — thresholds	Report the proportion of respondents completing the instrument under at least 2 thresholds, including an a priori minimum-plausible-completion threshold calibrated to the instrument length.
4	Response variance — straight-lining	Report the proportion of respondents giving a single unique Likert response across attitudinal items ("full straight-lining").
5	Response variance — modal proportion	Report the proportion of respondents with modal-response proportion at or above 80% across attitudinal items.
6	Combined indicators	Report at least 1 combined indicator capturing respondents subthreshold on individual signals but exhibiting joint fast-and-stereotyped patterns (e.g., fast completion AND modal-response proportion at or above 50%).
7	Sampling-frame validation	Report the proportion of respondents whose item-level self-reported sampling-frame characteristics (state, role, specialty) are consistent with panel-stage screening.
8	Sensitivity analyses	Report focal effect estimates across at least 4 alternative inclusion definitions, including straight-liner exclusion, modal-response exclusion, sampling-frame validation, and the full sample.
9	Pre-registration of thresholds	Where feasible, pre-register data-quality thresholds and exclusion rules. Where post hoc, disclose this explicitly.

DQ-CSC is intended to complement CHERRIES (Eysenbach 2004 [8]*) and STROBE for observational cross-sectional research. CHERRIES addresses survey administration; DQ-CSC addresses post-collection data-quality screening and sensitivity-analysis reporting. The 9 items are reportable in approximately 1 paragraph of Methods, 1 paragraph of Results, and 1 supplementary table.*

Discussion

Principal results

We propose a 3-family multi-indicator framework for detecting inattentive responding in commercial-panel surveys of healthcare professionals, demonstrate it empirically using a rural healthcare professional sample (N = 100), and synthesize the framework into a 9-item reporting checklist (DQ-CSC) intended to complement CHERRIES. Three findings warrant emphasis. First, in the demonstration sample, individual indicators flagged response-quality concerns at rates consistent with the upper end of

published nonprobability-panel rates: 18% completed the instrument in under 2 minutes; 15% gave a single Likert response to every attitudinal item; 23% had modal-response proportion at or above 80%. Second, the combined fast-and-modal indicator flagged 34% of the sample — substantially larger than any single indicator — demonstrating that single-indicator screening underestimates the size of the engagement-concern subgroup. Third, the 6-restriction sensitivity-analysis design demonstrated that the focal substantive effect estimate moved within a 39% range across analytically defensible inclusion choices (OR 0.18 to 0.25; all P<.05). The substantive conclusion did not change, but the analytical-choice-induced uncertainty in the point estimate was non-trivial.

Comparison with prior work

Existing methodological work on online survey data quality has established the prevalence of satisficing and inattentive responding [1–7], the differential vulnerability of nonprobability versus probability-based panels [4], the persistence of satisficing across panel waves [5], and dynamic protocols for detecting fraudulent responses in health surveys [9, 10]. Reporting standards for internet surveys are dominated by CHERRIES [8] and have been extended in recent work on dynamic fraud-detection protocols [9]. The present framework is additive to this prior work in 3 specific ways. First, it provides a coordinated 3-family structure rather than a list of indicators, enabling researchers to address speed, variance, and combined-pattern threats in a single screening procedure. Second, it operationalizes a 6-restriction sensitivity-analysis design that empirically demonstrates the impact of inclusion choices on substantive estimates — precisely the type of empirical methodology study BMC Medical Research Methodology invites. Third, the DQ-CSC checklist (Table 4) is, to our knowledge, the first reporting checklist focused specifically on post-collection data-quality screening for clinician-panel surveys. Where CHERRIES asks what was administered, DQ-CSC asks what was kept and how that choice affected results.

The framework's thresholds are deliberately conservative and are calibrated against this single demonstration instrument and sample. The 2-minute completion-time threshold corresponds to approximately 3 seconds per Likert item, which is at the lower bound of plausible cognitive engagement based on response-time research [12]. The 80% modal-response-proportion threshold is consistent with thresholds used in published satisficing research [11] but is not universal; researchers using shorter or longer instruments, or instruments with more complex item structures (e.g., reverse-scored items, branching), should calibrate the threshold accordingly. We do not propose universal thresholds; we propose universal indicators, with calibration by instrument and context.

Practical implications

For survey-based health services researchers planning clinician-panel data collection, 3 implications follow directly from the framework. First, data-quality indicators should be planned pre-specification rather than reported post hoc. Pre-registration of completion-time thresholds and modal-response exclusion rules reduces the risk that screening choices appear opportunistic to reviewers. Second, sensitivity analyses across 4 or more respondent-inclusion definitions should be reported routinely, not only when reviewers request them. The cost of running an additional 5 logistic regressions is negligible

relative to the interpretive value for readers. Third, the DQ-CSC checklist (Table 4) provides a 9-item template that addresses these recommendations in approximately 2 paragraphs of additional manuscript text.

For survey-based researchers planning recruitment specifically through commercial panels, the framework also implies a budgetary consideration. The verification gap that motivates Family 3 of the framework — 9% of respondents in the demonstration sample reported licensure outside the planned target states despite panel-stage screening for those states — cannot be eliminated post hoc; it must be addressed by either tighter screening (which increases panel costs) or supplementary recruitment from a verified frame (which increases costs proportionally to the supplement size). Even a small verified subsample (e.g., 10% to 15% of the total target n recruited from a state licensure registry) provides comparative leverage for assessing whether panel-frame and verified-frame estimates diverge. Researchers without budget for a verified supplement should at minimum report restriction (E) of the framework (verified target-state respondents only) so that the reader can assess the gap.

Limitations

The empirical demonstration is based on a single sample of 100 self-identified rural Appalachian healthcare professionals. We foreground 3 specific limitations of this demonstration. First, the framework's thresholds are calibrated to a 40-item Likert instrument administered to clinicians, and we do not claim that these specific thresholds (2-minute completion-time floor, 80% modal-response ceiling) generalize to instruments of substantially different length or to non-clinician populations. The framework's 3-family structure and the 6-restriction sensitivity-analysis design generalize; the specific threshold values require recalibration. Second, the demonstration sample is drawn from a rural healthcare professional population in a specific US region (Central and Southern Appalachia). The data-quality patterns observed here may differ from those in urban, international, or specialty-specific clinician panels. We chose this sample deliberately because rural healthcare professional samples are precisely where commercial-panel recruitment is most frequently used and where verified sampling frames are most expensive to access; the data-quality concerns documented here are therefore most acute in exactly the population where the framework is most needed. But the calibration question — whether thresholds calibrated to this sample apply to urban specialty-clinician panels, or to international samples — remains open. Third, the proposed DQ-CSC checklist has not been validated through Delphi consensus or formal stakeholder review, as CHERRIES was [8]. We propose it as a starting point for discussion in the field rather than as a finalized standard. Future methodological work should refine and validate the checklist through structured consensus methods.

A further limitation worth naming explicitly: the focal substantive estimate used to demonstrate sensitivity-analysis variation is itself a single estimate from a single regression specification. Demonstrating that the focal estimate moves within a 39% range across 6 inclusion definitions does not establish that all focal estimates in clinician-panel research will move within similarly bounded ranges. Estimates near the null, or estimates with much wider primary confidence intervals, or estimates from non-linear specifications, may exhibit larger analytical-choice-induced variation. The 39% range we report

is informative for one well-behaved logistic regression with a moderately strong focal signal; it should not be read as a universal expectation.

Conclusions

Data-quality screening choices in commercial-panel surveys of healthcare professionals meaningfully affect substantive effect estimates and warrant transparent reporting. The 3-family multi-indicator framework, the 6-restriction sensitivity-analysis design, and the 9-item DQ-CSC reporting checklist proposed here provide a reusable procedure that complements CHERRIES and other internet-survey reporting standards. In the rural healthcare professional demonstration sample, the focal substantive estimate moved within a 39% range across analytically defensible inclusion choices, illustrating both the importance of routine sensitivity-analysis reporting and the value of a coordinated framework for thinking about which inclusion choices are defensible. Further calibration of the framework's threshold values across specialty, urban-rural, and international clinician panels is a natural next step, and Delphi-based validation of the DQ-CSC checklist would convert it from a starting-point proposal into a field standard.

Abbreviations

CHERRIES
Checklist for Reporting Results of Internet E-Surveys
CFIR
Consolidated Framework for Implementation Research
CI
confidence interval
DOI
Diffusion of Innovations
DQ-CSC
Data Quality reporting for Clinician Surveys via Commercial panels
IQR
interquartile range
IRB
Institutional Review Board
mHealth
mobile health
OR
odds ratio
STROBE
Strengthening the Reporting of Observational Studies in Epidemiology
TAM
Technology Acceptance Model

UTAUT
Unified Theory of Acceptance and Use of Technology

Declarations

Ethics approval and consent to participate

The underlying survey was approved by the Walden University Institutional Review Board (approval number 08-28-25-1124176). Informed consent was obtained through a survey-based implied consent procedure approved by the IRB. The present methodological analysis used de-identified data only and required no additional IRB review.

Consent for publication

Not applicable.

Availability of data and materials

The de-identified dataset and analysis code used for the demonstration in this paper are available from the corresponding author on reasonable request. The substantive analyses of the underlying survey are reported in the corresponding author's doctoral dissertation [13] and in a companion manuscript currently under consideration at JMIR mHealth and uHealth [14].

Competing interests

The author is the founder and Chief Executive Officer of SafeguardMedix, a healthcare technology company that operated during the underlying doctoral research period. SafeguardMedix does not market or sell mobile health applications of the kind discussed in the underlying substantive analyses. The methodological framework proposed in this paper has no commercial relevance to SafeguardMedix's products or services. The full disclosure is provided in the underlying dissertation [13] and in the companion JMIR submission [14].

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The underlying survey data were collected as part of doctoral research at Walden University.

Authors' contributions

TSA conceived and designed the methodological framework, conducted the empirical demonstration analyses, designed the DQ-CSC checklist, and drafted and revised the manuscript.

Acknowledgements

The author thanks the dissertation committee at Walden University for guidance during the underlying research, and the survey respondents for their time.

Prior publication and data overlap

The dataset used for the empirical demonstration in this paper was collected for and is reported substantively in the corresponding author's doctoral dissertation [13] and in a companion manuscript currently under consideration at JMIR mHealth and uHealth [14]. The present paper addresses a distinct methodological research question (the framework, sensitivity-analysis design, and reporting checklist) rather than the substantive findings reported in those outputs. The substantive estimate (structural-readiness composite as predictor of mHealth adoption) is reported here only as the dependent variable for demonstrating sensitivity-analysis variation, not as a primary finding.

References

1. Krosnick JA. Response strategies for coping with the cognitive demands of attitude measures in surveys. Appl Cogn Psychol. 1991;5(3):213–36. 10.1002/acp.2350050305.
2. Meade AW, Craig SB. Identifying careless responses in survey data. Psychol Methods. 2012;17(3):437–55. 10.1037/a0028085. PMID: 22506584.
3. Bowling NA, Huang JL, Bragg CB, Khazon S, Liu M, Blackmore CE. Who cares and who is careless? Insufficient effort responding as a reflection of respondent personality. J Pers Soc Psychol. 2016;111(2):218–29. 10.1037/pspp0000085. PMID: 26594821.
4. Cornesse C, Blom AG. Response quality in nonprobability and probability-based online panels. Sociol Methods Res. 2023;52(2):879–908. 10.1177/0049124120914940.
5. Kraemer F, Silber H, Struminskaya B, Weiß B, Bosnjak M, Koßmann J, Sand M. Satisficing response behavior across time: assessing negative panel conditioning using an experimental design with six repetitions. Surv Res Methods. 2023;17(3):269–300. 10.18148/srm/2023.v17i3.7986.
6. Berry K, Rana R, Lockwood A, Fletcher L, Pratt D. Factors associated with inattentive responding in online survey research. Pers Individ Differ. 2019;149:157–9. 10.1016/j.paid.2019.05.043.
7. Goldammer P, Annen H, Stöckli PL, Jonas K. Careless responding in questionnaire measures: detection, impact, and remedies. Leadersh Q. 2020;31(4):101384. 10.1016/j.leaqua.2020.101384.
8. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res. 2004;6(3):e34. 10.2196/jmir.6.3.e34. PMID: 15471760.

9. Tan KKB, Lim WC, Ang ZQ, Lim BHY, Lee YJ, Lee TC, Tan MZE, Tan SHX, Lim ZJ, Sng GGR. Increasing rigor in online health surveys through the reduction of fraudulent data. J Med Internet Res. 2025;27:e64252. 10.2196/64252.
10. Storozuk A, Ashley M, Delage V, Maloney EA. Got bots? Practical recommendations to protect online survey data from bot attacks. Quant Methods Psychol. 2020;16(5):472–81. 10.20982/tqmp.16.5.p472.
11. Dykema J, Schaeffer NC, Garbarski D, Nordheim EV, Banghart P, Cyffka K. The impact of parental involvement and other interventions on response rates and respondent burden in mailed and internet-delivered surveys of adolescents. Surv Pract. 2021;14(1):1–11. 10.29115/SP-2021-0008.
12. Anduiza E, Galais C. Answering without reading: IMCs and strong satisficing in online surveys. Int J Public Opin Res. 2017;29(3):497–519. 10.1093/ijpor/edw007.
13. Ajagbe TS. Barriers and incentives for smartphone mHealth adoption among health care providers in Central and Southern Appalachia [doctoral dissertation]. Minneapolis, MN: Walden University; 2026. Walden Dissertations and Doctoral Studies, paper 19241. URL: https://scholarworks.waldenu.edu/dissertations/19241
14. Ajagbe TS. Structural readiness and diffusion of innovations attributes in rural Appalachian mobile health adoption: cross-sectional online survey [manuscript under consideration]. JMIR mHealth Uhealth; 2026.