ISSN 2282-0930 • Epidemiology Biostatistics and Public Health - XIII^th SISMEC Congress - Vol. 20 Suppl. 1

A Framework to Improve Data Quality and Manage Dropout in Web-Based Medical Surveys: Insights from an AI Awareness Study among Italian Physicians

Cofini Vincenza[1], Benvenuti Eugenio[1], Mancinelli Martina[2], Di Pangrazio Ginevra[1], Piccardi Laura[3,4], Muselli Mario[1], Cimino Eleonora[1], Tobia Loreta[1], Fabiani Leila[1], Necozione Stefano[1]

(1) Department of Life, Health and Environmental Sciences, University of L'Aquila, L'Aquila, Italy
(2) Department of Anatomical and Histological Sciences, Legal Medicine and Orthopedics, Sapienza University, Rome, Italy
(3) Psychology Department, Sapienza University, Rome, Italy
(4) San Raffaele Cassino Hospital, Cassino (FR), Italy

CORRESPONDING AUTHOR: *Cofini Vincenza, vincenza.cofini@univaq.it*

BACKGROUND

Ensuring data quality in self-reported online surveys remains a critical challenge in digital health research, particularly when targeting healthcare professionals [1,2]. Self-reported data are susceptible to multiple biases, including careless responding, social desirability bias, and drop-out-related attrition, all of which may compromise the validity of findings [3,4]. In web-based surveys where researcher oversight is limited, structured quality control measures are essential to detect low-quality responses, minimise sampling bias, and enhance data reliability [5]. Previous studies have demonstrated that inadequate quality checks can lead to inflated error rates, reduced statistical power, and misleading conclusions [6]. Objective

This study presents a comprehensive methodological framework for optimising data quality in web-based medical surveys, applied to a national study on AI awareness among Italian physicians. Integrating pre-survey validation, real-time dashboards, response-time filtering, and post-hoc careless responding detection would address key challenges in digital research, while providing a replicable model for future studies.

METHODS

We conducted a national web-based survey using a validated instrument (doi:10.1101/2025.04.11.25325592) via the LimeSurvey platform. The survey incorporated two main sections: (1) a core module assessing knowledge, attitudes and practices regarding AI in medicine; (2) clinical scenarios evaluating diagnostic agreement with AI-generated proposals. Multiple quality control strategies were implemented throughout the survey lifecycle. In terms of survey design and logic, the questionnaire employed an adaptive flow structure, whereby respondents were routed through clinical scenarios relevant to their medical speciality. To reduce the incidence of partial completions and missing data, key questions were marked as mandatory, and completion status was actively tracked. In the monitoring and recruitment phase, a real-time dashboard monitored participant distribution (gender/geographical areas/speciality); referral links were rotated to minimise snowball bias [7]. Time-based data quality checks excluded outliers (completion time <1st or >99th percentile) [8]. Completion time for the first section was analysed for all completers to assess correlations between response speed and quality indicators. Dropout patterns were analysed using Kaplan-Meier survival analysis and logistic regression, to identify systematic attrition predictors. Data quality assessments were performed on the outlier-cleaned dataset (n=587). Response quality was assessed using complementary careless responding indicators applied specifically to opinion scale items (Likert 1-5). Two detection methods were used: low response variance analysis, identifying respondents with insufficient variability (SD < 0.5), and excessive same-response detection, flagging participants using identical responses for >75% of items. Internal consistency analysis (Cronbach's α) evaluated scale reliability across different quality levels.

RESULTS

A total of 736 accesses were recorded on the survey platform. As an initial inclusion criterion, only participants who indicated current registration with the Italian Medical Council were considered eligible: 79 (10.7%) were excluded, yielding

DOI: 10.54103/2282-0930/29202

a sample of 657 eligible participants (89.3%). Among eligible respondents, 597 completed the first section, yielding a dropout rate of 9.1% (n=60). A Kaplan-Meier survival analysis using total survey time revealed that most dropouts occurred early, with critical points at 45% after demographic, 51% after personal AI knowledge items, 71% after opinion items, and 100% before clinical scenarios. Logistic regression showed no significant predictors of completion (LR $\chi^2(6)=3.46$, p= 0.7497; pseudo-R^2=0.014; AUC=0.60, 95%CI: 0.50–0.70). Completion time showed no correlation with response quality (Spearman's ρ = -0.019, p = 0.645). Following outlier removal, data quality assessment among 587 who completed the first section revealed two complementary patterns of careless responding: 8.52% (n=50) exhibited low response variance, while 32 (5.45%) demonstrated excessive same response patterns. Cross-classification analysis showed 23 participants (3.92%) flagged by both indicators, with 71.88% of excessive same responders also showing low variance. Overall, 50 participants (10.05%, 95% CI: 7.9%- 12.8%) exhibited careless responding detectable by at least one indicator. Internal consistency analysis showed robust scale reliability (Cronbach's α = 0.754) that remained stable across quality levels.

CONCLUSION

The integration of real-time monitoring, adaptive design, time-based validation, and systematic careless responding detection provides a robust methodological framework for web-based medical surveys, particularly for complex topics like AI adoption. Comprehensive data quality assessment revealed a 10.05% careless responding rate among completers, which aligns with the literature. The absence of correlation between completion time and response quality shows that careless responding could reflect attentional rather than temporal factors. Our findings suggest that both phenomena likely reflect situational or contextual factors rather than systematic participant characteristics or survey design flaws. This supports the validity and generalizability of the final dataset while providing a replicable quality control framework for future web-based medical research.

REFERENCES

1. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res. 2004 Jun;6(3):e34.
2. Hochheimer C.J., Sabo R.T., Krist A.H. et al. Methods for Evaluating Respondent Attrition in Web-Based Surveys. J Med Internet Res. 2016 Nov;18(11):e301.
3. Meade A.W., Craig S.B. Identifying careless responses in survey data. Psychol Methods. 2012 Sep;17(3):437-55.
4. Bethlehem J. Selection bias in web surveys. Int Stat Rev. 2010 Aug;78(2):161-88.
5. Bosch O.J., Revilla M. The Quality of Survey Questions in Spain: A Cross-National Comparison. Rev Esp Investig Sociol. 2021 Jul-Sep;175:3-26.
6. Schell C., Godinho A., Cunningham J.A. Using a consistency check during data collection to identify invalid responding in an online cannabis screening survey. BMC Med Res Methodol. 2022 Mar;22(1):67.
7. Sauermann H., Roach M. Increasing web survey response rates in innovation research: An experimental study of static and dynamic contact design features. Res Policy. 2013 Feb;42(1):273-86.
8. Matjašič M., Vehovar V., Lozar Manfreda K. Web survey paradata on response time outliers: A systematic literature review. Metodološki Zvezki. 2018;15(1):23-41.

DOI: 10.54103/2282-0930/29202