Submission Id: 2642

Title

Development of a Fraud Detection Plan in a Survey of Canadian Abortion Providers.

Priority 1 (Research Category)

Research methodology and instrument development

Presenters

Madeleine Ennis, PhD, Wendy Norman, MD, MHSc, Janusz Kaczorowski, PhD, Sheila Dunn, MD, MSc, Regina Renner, MD, MPH, Damien Contandriopoulos, PhD, Edith Guilbert, MD, MSc

Abstract

Context: Prevention and detection of fraudulent online survey responses is challenging for researchers. Objective: Describe the development and implementation of a fraudulent response identification plan in the online Canadian Abortion Provider Survey (CAPS). Study Design: Development and piloting of a fraud detection algorithm in collaboration with the Research Ethics Board (REB), a biostatistician, and CAPS co-investigators. Dataset: The CAPS data was collected as part of a cross-sectional, anonymized, self-administered, online survey using the Research Electronic Data Capture (REDCap) platform between July and December 2020. Population Studied: The survey targeted physicians, nurse practitioners, and administrators who provided abortion care in Canada in 2019. Results: The authors used a two-phase approach to detect fraud. After extensive piloting, phase one of the fraud detection algorithm flagged respondents with: (1) nonsensical or non-probable response combinations that could indicate fraudulence; (2) survey completion date after October 2020 (when the first potential fraudulent responses were identified); and (3) use of a suspicious email address to request remuneration/future contact, including those with alternating letters and numbers and similar name patterns that were not linked to traceable individuals. With REB approval, the separate main survey and remuneration survey were linked using timestamps. Phase one protected respondents who provided: (1) free-text responses specific to abortion provision; and (2) email addresses with academic or medical institution domains. After phase one, all respondents were included, excluded, or marked for further review. During phase two, study personnel reviewed the records for respondents marked for further review, for additional nonsensical, non-probable or protective open-text responses. All respondents (n=933) were either marked for inclusion (n=518) or exclusion (n=415) after team consensus was reached. Conclusions: Prevention and identification of fraud in surveys is key to optimize data integrity. The authors used an iterative process including human based identification of potential indicators of fraudulent / non fraudulent responses to build a score, balancing computer- with human-based analysis of potential

indicators and their weight towards an overall score and final decision on scoring system and removal of flagged respondents. This may inform REB and researchers conducting online surveys.