

Original Paper

# Exploring Methods to Mitigate Fraud in Web-Based Surveys: Multicase Study Analysis

Madeleine Ennis<sup>1\*</sup>, PhD; Regina-Maria Renner<sup>1\*</sup>, MD, MPH; Claudia Morando-Stokoe<sup>2</sup>, BHSc; Sharon James<sup>2</sup>, BN, Grad.Dip.OHSN, MPH, PhD; Patricia A Janssen<sup>3</sup>, RN, BSN, MPH, PhD; Sara Leckie<sup>3</sup>, BSc, MSc; Sheila Dunn<sup>4,5</sup>, MD, MSc; Danielle Mazza<sup>2</sup>, MD; Wendy V Norman<sup>6</sup>, MD, MHSc

<sup>1</sup>Department of Obstetrics and Gynaecology, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada

<sup>2</sup>SPHERE Centre of Research Excellence in Women's Sexual and Reproductive Health in Primary Care, Department of General Practice, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia

<sup>3</sup>School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada

<sup>4</sup>Department of Family and Community Medicine, University of Toronto, Toronto, ON, Canada

<sup>5</sup>Women's College Research Institute, Toronto, ON, Canada

<sup>6</sup>Department of Family Practice, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada

\*these authors contributed equally

## Corresponding Author:

Madeleine Ennis, PhD

Department of Obstetrics and Gynaecology

Faculty of Medicine

University of British Columbia

2194 Health Sciences Mall

Vancouver, BC, V6T 1Z3

Canada

Phone: 1 (604) 806 9966

Email: [madeleine.ennis@cw.bc.ca](mailto:madeleine.ennis@cw.bc.ca)

## Abstract

**Background:** Web-based surveys are a cost-effective technique to engage a large population of participants in research projects, including those who were previously difficult to reach due to geographic location, safety, and vulnerability. While web-based surveys have many advantages, they can be more susceptible to fraud, especially when a generic invitation link or a financial incentive is offered. There is a paucity of literature presenting experiences for mitigating this type of fraudulent study response, yet this important foundation is needed to inform the work of researchers and institutional review boards (IRBs) to support the collection of high-quality, appropriate data.

**Objective:** This study aims to analyze, compare, and contrast the range of strategies used to prevent, detect, and remove fraudulent responses by investigating 4 web-based surveys in Australia and Canada, each of which experienced fraudulent responses.

**Methods:** Our descriptive multiple case study presents 4 research projects from Australia and Canada that experienced survey fraud. These web-based surveys recruited patients of, or clinicians providing, family planning services. We describe each study's approach to preventing fraud (primary prevention; eg, CAPTCHA) and a screening protocol to detect fraudulent responses during data collection (secondary prevention). Once fraud was detected, each study team developed strategies to protect data integrity, in consultation with coinvestigators, ethics committees/ IRBs, and biostatisticians, to remove fraudulent respondents from the dataset (tertiary prevention).

**Results:** All studies recruited via a generic survey link and provided remuneration, which are common risk factors for fraud. Several studies also relied on social media for recruitment. All 4 studies implemented tertiary fraud detection strategies to identify and remove fraudulent responses and maintain data integrity (removing between 16% and 45% of respondents). Including personal identifiers during data collection provided 3 of the studies with a more robust option to identify and remove fraudulent respondents. Where personal identifiers could not be used (eg, to protect the identity of a vulnerable study population), investigators relied on a complex fraud detection algorithm verified by manual team review.

**Conclusions:** Commonly used web-based anonymized survey methods, particularly those offering incentives for participation, are at substantial risk for fraud. Across these 4 studies, robust fraud detection methods were essential to ensure data reliability,

with varying strategies, such as using personal identifiers, applied based on specific survey contexts. Fraud mitigation criteria explored in this multicase analysis can be adapted to other web-based surveys, survey topics, and populations. Implementing the fraud prevention and detection methods within survey design will assist researchers and IRBs in protecting data integrity.

**Trial Registration:** Australian New Zealand Clinical Trials Registry ACTRN12622000655741; <https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=383919> and ClinicalTrials.gov NCT05793944; <https://clinicaltrials.gov/study/NCT05793944>

(*J Med Internet Res* 2025;27:e78671) doi: [10.2196/78671](https://doi.org/10.2196/78671)

## KEYWORDS

abortion; data collection; fraud; long-acting reversible contraception; methods; pregnancy; reimbursement incentive; surveys and questionnaires; women's health

## Introduction

Increasingly, as society adapts digital technologies, academic research has moved from in-person, mail, fax, and phone surveys to web-based data collection [1-4]. Benefits of web-based surveys include cost-effective recruitment over a wide geographic area and access to unknown or equity-deserving populations that have previously been difficult to reach [3,5-8]. The internet facilitates anonymous research with fewer barriers for respondents who may take the survey at their convenience [3,5-7].

However, web-based surveys, especially when anonymous and using a generic link for recruitment, increase the risk of fraudulent responses that threaten data integrity [3,6]. Offering financial incentives as well as conducting research on socially or politically sensitive or stigmatized topics are additional risk factors for fraud [3,9]. Fraudulent web-based survey responses can be the result of human and computer-programmed (ie, bot) activity [3]. These can be broadly categorized as (1) those who misrepresent themselves in order to suit the survey's eligibility criteria, (2) eligible respondents who duplicate their responses to receive additional financial incentive, or (3) those with malicious intent to invalidate the survey results [3].

Researchers have developed approaches and algorithms to decrease the risk for survey fraud with careful choices of their study design, and monitoring incoming survey records for responses that suggest fraud [3,6]. Nonetheless, preventing, identifying, and managing fraudulent responses still pose major challenges for both investigators and ethics committees or institutional review boards (IRBs) [10]. The increase in web-based surveys and associated fraudulent activity has necessitated research on how to reduce the risk of fraud when using anonymized surveys, particularly when financial incentives are offered [8,11-13]. However, little has been written about survey fraud mitigation, with a particular lack of papers that discuss detection strategies and their sensitivity or specificity, nor those that provide recommendations on how to diminish the impact of fraudulent responses on the validity of survey results.

To address this gap, we analyze, compare, and contrast a range of strategies used to prevent, detect, and remove fraudulent responses in 4 web-based surveys in Australia and Canada. Our analysis and implications may assist research teams and IRBs/ethics committees planning web-based surveys by highlighting important tools to protect data integrity as well as respondents' identities.

## Methods

### Overview

We present a multicase study approach [14] with 4 web-based survey studies included in our descriptive analysis. We documented the study aims, design, setting, population, and data collection dates in Table 1. Each author was a coinvestigator in at least one of the presented case studies. For our analysis, we defined 3 stages of fraud prevention analogous to primary, secondary, and tertiary disease prevention [15]. Primary fraud prevention included survey and recruitment design to minimize the risk of experiencing fraud. Secondary prevention consisted of monitoring incoming responses for early fraud detection. We describe the study designs, recruitment, and primary and secondary prevention strategies in the Methods section. In Table S1 in Multimedia Appendix 1, we map detailed prevention strategies to apply during the stages of study development, recruitment, data cleaning, and analysis. Some of the strategies functioned at multiple levels of primary, secondary, and tertiary prevention. Once studies detected fraud, each developed tertiary prevention approaches to minimize the impact of fraud on the final results by removing fraudulent respondents. This required the development of fraud removal strategies and algorithms, as described in the Results section. Each case study team worked with their local IRB/ethics committee upon detecting fraud in their sample, with approval to conduct the analysis to identify and remove fraudulent respondents as part of their data cleaning. We followed the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) reporting checklist (Table S2 in Multimedia Appendix 1). Each survey is available in Multimedia Appendix 2.

**Table 1.** Characteristics of the 4 case studies of web-based surveys that experienced fraud.

	Study aim	Study design	Setting and population	Timeframe
Case Study 1: CAPS2019 <sup>a</sup>	<ul style="list-style-type: none"><li>To describe demographics of abortion providers, clinical care characteristics of abortion care services, and provider experiences with stigma and harassment.</li></ul>	Self-administered, cross-sectional, anonymized, bilingual (French and English), web-based survey	Pan-Canadian study including physicians, nurse practitioners, and administrators who provided abortion care in 2019.	Recruitment between July and December 2020.
Case Study 2: AusCAPPS <sup>b</sup>	<ul style="list-style-type: none"><li>To examine the knowledge, attitudes, and practices of general practitioners, general practice nurses, and community pharmacists about medication abortion in Australia.</li></ul>	Cross-sectional web-based survey	Australian general practitioners, general practice nurses, and community pharmacists.	Data collection between July and December 2021.
Case Study 3: EX-TEND-PREFER	<ul style="list-style-type: none"><li>Primary aim: To assess whether a web-based educational video, describing all available contraceptive methods, developed for and with stakeholders and young Australian women living in rural and remote areas and for women from CALDc backgrounds speaking one of 4 languages (Mandarin, Cantonese, Hindi or Arabic), would increase their contraceptive knowledge, preference for, and uptake of LARCd.</li><li>Secondary aims: To examine participant-reported likelihood of changing contraceptive use, including the use of LARC, after intervention; and to examine the cost and effectiveness of using social media to increase access to reliable contraceptive information.</li></ul>	Single group pre-post intervention design in which a web-based self-report questionnaire was completed at 3 timepoints	Australia based study including women aged 16-25 years, living in rural or remote Australia or from a CALD background who were not currently pregnant, did not wish to become pregnant within a year of the study taking place, had been sexually active with a male partner within the last 6 months or anticipated near-future sexual activity with a male partner and have not undergone any permanent form of contraception (including tubal ligation, hysterectomy or partner vasectomy).	Recruitment between February and April 2023.
Case Study 4: SmartMom RCT <sup>e</sup>	<ul style="list-style-type: none"><li>To evaluate the effectiveness of the SmartMom prenatal education texting program on knowledge of healthy pregnancy and childbirth, adoption of healthy behaviors, and maternal, fetal, and newborn health outcomes.</li></ul>	Two-arm, single-blind (data analysis) RCT	Pan-Canadian study of healthy pregnant people.	Recruitment began in November 2023 and is to be completed in 2026.

<sup>a</sup>CAPS2019: The 2019 Canadian Abortion Provider Survey.  
<sup>b</sup>AusCAPPS: Australian Contraception and Abortion Primary Care Practitioner Support Network.  
<sup>c</sup>CALD: culturally and linguistically diverse.  
<sup>d</sup>LARC: long-acting reversible contraception.  
<sup>e</sup>RCT: randomized controlled trial.

**Case Study 1: The 2019 Canadian Abortion Provider Survey**

**Survey Design**

Between July and December 2020, this study aimed to document the Canadian abortion care workforce characteristics and distribution, clinical care characteristics, and experiences with stigma and harassment via a survey of physicians, nurse

practitioners, and abortion clinic administrators who provided abortion care in 2019 [16]. The Canadian Abortion Provider Survey was available in French and English on a Research Electronic Data Capture (REDCap) platform [17] and included sections on: demographics, clinical care characteristics of abortion care services, care of diverse populations, provider experience with stigma and harassment, and remuneration/contact for study results or future research (Multimedia Appendix 2). Complex skip pattern logic allowed

respondents to see only relevant questions based on their prior answers. The survey took 30 to 80 minutes, depending on the range of abortion services respondents provided. As per the IRB's request, the study team maximized protection of respondents' identities and collected anonymized data. This study did not collect the IP addresses of respondents.

### **Recruitment**

In Canada, abortion providers could include obstetrician/gynecologists, family physicians, and nurse practitioners. The identities and numbers of these individuals within their larger groups are unknown. The researchers partnered with various health care professional organizations, such as the Society of Obstetricians and Gynaecologists of Canada and the College of Family Physicians of Canada, to maximize recruitment of all eligible participants in the target populations. The partnering organizations distributed a generic survey link to their membership via email and newsletters. Additionally, we recruited via relevant institutions such as hospital departments of OB-GYN and family medicine, abortion clinics, and our web-based community of abortion practice platform [18]. The survey invitation listed the inclusion criteria. Initial mandatory survey questions verified respondents' eligibility, and some questions required open-ended typed responses. Upon survey completion, respondents were directed to a separate elective survey, in which they could include contact details for remuneration, future research, and study results. This was the only survey section that asked respondents to provide a personal identifier in the form of an email. The researchers did not link the data from the 2 surveys in order to maintain the anonymity of respondents.

### **Ethical Considerations**

This study received ethics approval (H18-03313) from BC Children's and Women's Research Ethics Board (REB) for both original data collection and subsequent fraud detection analysis. Participants reviewed and completed an anonymous consent statement through REDCap prior to the start of the survey. Respondents were offered a CAD \$50 (equal to US \$37.30) gift certificate as financial compensation, as detailed in the consent form, but not advertised in recruitment materials. All surveys were conducted and data stored using REDCap, hosted by BC Children's Hospital Research Institute, aligned with Canadian privacy regulations.

### **Primary Fraud Prevention Strategies**

This case implemented several primary fraud prevention strategies. The researchers recruited through professional organizations and university medical departments rather than via social media. Recruitment material did not indicate remuneration. Respondents could not request remuneration until they had completed all survey sections relevant to their designation. The beginning of the survey included a CAPTCHA to detect "bots." The initial survey question asked respondents to confirm that they had not taken the survey before.

### **Secondary Fraud Prevention and Initial Fraud Detection**

The research team monitored incoming responses every other week. Based on the literature, this consisted of confirming sensical combinations between select questions in the

demographics [19]. The team identified suspicious patterns (answer combinations not possible within a regulatory or medical context) starting October 6, 2020, and determined that a more rigorous approach was required to identify and remove all fraudulent data prior to analysis.

The researchers identified the approaches of Ballard et al [3] and Teitcher et al [6], an evaluation and review of fraud detection techniques in web-based surveys, as the most applicable for informing the development of their tertiary fraud detection methods. They used a variety of criteria, such as questionnaire/instrument level data, software for detecting bots such as CAPTCHA, tracking respondents' nonquestionnaire data, including personal information such as email addresses, and study design level decisions, such as consent and compensation.

### **Case Study 2: The Knowledge, Attitudes and Practices of Australian Primary Care Clinicians in the Provision of Long-Acting Reversible Contraception and Medication Abortion**

#### **Survey Design**

Launched on July 19, 2021, three national web-based surveys aimed to describe the knowledge, attitudes, and practices of Australian general practitioners, practice nurses, and community pharmacists regarding long-acting reversible contraception (LARC) and medication abortion. Based on previous LARC and medication abortion surveys we conducted this as part of the Australian Contraception and Abortion Primary Care Practitioner Support Network study (ACTRN12622000655741) [20]. An ethics committee approved respondent explanatory statement was embedded on the survey landing page that gave details about the research purpose, inclusion criteria (eg, Australian Health Practitioner Registration), and investigator team. From piloting, the survey took an estimated 10 to 15 minutes to complete and consisted of a maximum of 50 questions over 9 pages, written in English language. Questions could be revisited by respondents prior to completion. Questions were open and closed in design and included demographics about age, gender, postcode, and number of years they had been practicing as a clinician. The survey used timestamps and complex branching logic. Answer options included Likert scales and free-text (eg, "other" responses).

#### **Recruitment**

Using a generic survey link, the research team used convenience sampling methods, such as through research team contacts, a purchased professional address list from the Australasian Medical Publishing Company and trial partners from industry and professional organization newsletters relating to the 3 clinician groups (eg, Royal Australian College of General Practitioners, Australian Primary Health Care Nurses Association, Pharmaceutical Society of Australia). The social media platforms used for recruitment included Facebook, X, and LinkedIn; tagging these clinician groups and women's health organizations (eg, Sexual Health Victoria, Family Planning New South Wales). Recruitment concluded 6 months following launch, once enrolment reached the desired sample



target of approximately 500 verified responses for each professional group.

### **Ethical Considerations**

The Monash University Human Research Ethics Committee approved the Australian Contraception and Abortion Primary Care Practitioner Support Network knowledge, attitudes, and practices (#28002) survey documents as well as any changes to respondent verification to improve the fraud detection process. Participation was voluntary, and consent was implied by survey completion. Recruiting materials advertised an Aus \$40 (equal to US \$30.05) e-gift card incentive and the time required to complete the surveys. All surveys were conducted using REDCap, hosted and managed by Helix (Monash University) [21,22]. Once collected, the data were exported to Microsoft Excel for cleaning and saved in line with Monash University data storage requirements.

### **Primary Fraud Prevention Strategies**

Given the reimbursement offer for survey completion, response inclusion was first determined through manual matching of the respondent's name and postcode to their listing on the publicly available Australian Health Practitioner Registration Agency website. A combination of human and computer-based strategies were then applied, including: duplicate removal, ensuring only those who met the inclusion criteria could complete the survey (eg, Australian Health Practitioner Registration, used Australian postcodes and worked in general practice as a primary or secondary place of employment), a minimum set time of 5 minutes for survey completion (determined at piloting), timestamp review, review for nonsensical email addresses (eg, mostly numbers or no relationship with the respondent's name) and responses (eg, phrases given instead of a name), and respondent behaviors such as survey completion in "batches."

### **Secondary Fraud Prevention and Initial Fraud Detection**

The researchers recorded an influx of general practitioner and practice nurse surveys in which respondent names did not match email addresses over a period of 4 days. The practice nurse survey received the majority of unverified responses over the course of survey recruitment. The fraudulent survey activity was detected approximately 2 weeks after launch. At this time, the team learned that the surveys had been advertised on private Facebook groups and through one organizational newsletter. The researchers then adopted a more rigorous approach by trying to match recruitment methods (eg, the identification of particular social media posts) to legitimate or fraudulent survey response activity.

## **Case Study 3: The EXTEND-PREFER Project**

### **Study Design**

The EXTEND-PREFER project aimed to assess whether a co-designed web-based educational video could increase contraceptive knowledge, preference for, and uptake of LARC in young women from priority populations [23]. From February 2023 to April 2023, web-based advertising recruited young culturally and linguistically diverse (CALD) women and women living in rural and remote areas aged 16-25 years. Respondents completed a prevideo survey (S1; 60 questions across 5 pages);

watched a 13-minute co-designed video, and completed a follow-up survey (S2; 35 questions across 3 pages) immediately after; and another survey (S3; 40 questions across 3 pages) 6 months later. The prevideo survey (S1) collected information relating to demographic characteristics and provided a baseline for contraceptive knowledge, preference for, and uptake of LARC (using a Likert scale). The prevideo survey (S1) also collected personal identifiers, including participants' first name, last name, email address, phone number, postcode, and IP address for verification purposes and so that they could be contacted to complete the follow-up survey (S2). The follow-up survey (S2) addressed respondents' contraceptive knowledge and preference after watching a 13-minute co-designed video detailing all contraceptive methods available in Australia. The 6-month follow-up survey (S3) collected data on contraceptive methods currently used to assess whether the video changed participants' preference for, knowledge of, or uptake of LARC. All surveys were completed using Qualtrics (Qualtrics International Inc) [23], and the data were exported to Microsoft Excel for cleaning and analysis. After completion of S1 and S2, respondents were offered an Aus \$20 (equal to US \$15.025) e-gift card for their time. Respondents who completed S3 were offered an additional Aus \$20 (equal to US \$15.025) e-gift card resulting in a maximum of Aus \$40 (equal to US \$30.05) for participants' time.

### **Recruitment**

The project used posts on LinkedIn and X and paid advertising on Facebook and Instagram to recruit participants. In order to appropriately target respondents, the research team developed 5 sets of advertising assets, which included one ad set targeted at rural communities and 4 ad sets targeted at CALD communities speaking Arabic, Cantonese Mandarin, or Hindi. The ad assets were in English, Arabic, Cantonese, Mandarin, and Hindi and each ad asset had a generic survey link for the language being used.

### **Ethical Considerations**

This study received ethical approval (#24907) from the Monash University Human Research Ethics Committee for data collection and analysis. Consent was implied by survey completion. Respondents were provided a maximum of Aus \$40 (equal to US \$30.05) (Aus \$20 for completing S1 and S2, Aus \$20 for S3 [Aus \$20 equal to US \$15.025]). Survey incentives were advertised on recruitment materials. All surveys were conducted and data stored using Qualtrics, hosted and managed by Monash University.

### **Primary Fraud Prevention Strategies**

In order to take part in the study, prospective respondents were required to complete an initial eligibility survey (first half of S1) that included questions about the respondents' age, relationship status, pregnancy status, pregnancy intention, ethnicity, rurality and sexual activity during the 6-month period prior to commencing the survey. Respondents were unable to proceed if they failed to meet the study inclusion criteria based on their responses.

### ***Secondary Fraud Prevention and Initial Fraud Detection***

The research team detected fraudulent activity approximately 2 weeks after the survey recruitment launch when a researcher noted duplicate entries (ie, responses with the same identifiers and responses with identical answers) while reviewing the survey data to arrange respondent remuneration. The team continued to implement planned primary fraud detection strategies and additionally began to conduct basic manual screening to check respondents who met the eligibility criteria if they were flagged by fraud detection features in Qualtrics (ie, ID Duplicate, Relevant ID Duplicate Score, Relevant ID Fraud Score, and ReCAPTCHA score).

### **Case Study 4: SmartMom Randomized Controlled Trial**

#### ***Study Design***

Launched in November 2023, the SmartMom randomized controlled trial (ClinicalTrials.gov NCT05793944, UBC REB H22-00603) aimed to evaluate the effectiveness of the SmartMom prenatal education texting program [24]. The main goals were to determine if healthy pregnant people receiving the SmartMom program have improvements in (1) knowledge of healthy pregnancy, labor, and birth, (2) standardized measures of mental health, (3) adoption of positive health behaviors in pregnancy, and (4) maternal, fetal, and newborn health outcomes. Inclusion criteria included (1) the ability to read and understand English at a Grade 8 level; (2) living in Canada but outside of British Columbia and Nova Scotia (SmartMom was already offered in these provinces); (3) 15 weeks pregnant or less; (4) singleton pregnancy; (5) aged 15 years or older; (6) without prepregnancy conditions requiring medication including hypertension, cardiac disease, diabetes, mental health disorder, or neurological disorder; (7) having a cell phone with cellular and internet access; and (8) not having been previously enrolled in SmartMom. Participants randomized to the Treatment arm received the SmartMom texting program, which includes evidence-based text messages 3 times per week through their pregnancy. Control participants received general interest messages once per week with study updates and pregnancy topics that were not aimed at promoting behavior change. Participants were asked to provide their personal health number to link to third-party health authority databases, and complete 3 web-based surveys: after enrolment, in late pregnancy, and after birth. Surveys took about 15-20 minutes to complete in total. There is one question per screen, with multiple screens to complete the questionnaire. Participants could edit answers to previously answered questions. Completion rates for initial surveys were 90% and for final surveys 72%. Surveys are available in [Multimedia Appendix 2](#). All surveys were conducted and data stored using REDCap, hosted by BC Children's Hospital Research Institute, in a manner compliant with Canadian privacy regulations. PHNs were encrypted and sent through dedicated secure file transfer servers to provincial or national data registries for linking with personal health data. Personal health data will be sent to the researchers using an encrypted file without the PHN or any other identifying information.

### ***Recruitment***

To maximize geographic and demographic representation, the study was advertised via paid ads on Instagram, Facebook, and Reddit, targeting women aged 18-44 years in Canada. Ads mentioned that gift cards would be provided to participants and included a link to the study website. Individuals interested in the study could complete a web-based screening form via a generic link to a REDCap survey consisting of eligibility questions. If responses indicated the individual met eligibility criteria, they were automatically directed to a second survey to provide their phone number, province, due date, and whether they were older than 18 years. Study staff reviewed the screening information and then contacted the individuals by text initially and then via a phone call. Study staff described participation in the study, answered questions, and then provided a link to the digital e-consent form via text. Once the e-consent form was complete, we randomized the participant to a study arm. Participants received survey links by text for treatment or control arms. Recruitment for the trial is ongoing.

### ***Ethical Considerations***

This study received ethical approval (H22-00603) from BC Children's and Women's REB for both original data collection and subsequent fraud detection analysis. Participants signed a consent form prior to randomization. Respondents were offered a gift certificate for each of 3 surveys completed (CAD \$1, CAD \$20, and CAD \$25; equal to US \$0.74, US \$14.82, and US \$18.52, respectively) as financial compensation, as detailed in the consent form but not advertised in recruitment materials. All surveys were conducted and data stored using REDCap, hosted by BC Children's Hospital Research Institute, aligned with Canadian privacy regulations.

### ***Primary Fraud Prevention Strategies***

Web-based screening and survey forms included CAPTCHA to minimize activity by bots. Study staff reviewed information provided on the screening form, and if information was inconsistent or nonsensical, it was reviewed with the individual for accuracy. If the individual did not, then provide information that was consistent (ie, due date with gestational age), study staff deemed them ineligible. Survey responses included the participant's phone number, allowing the removal of duplicates. Cookies were not used, nor did we check the IP addresses of respondents' computers.

### ***Secondary Fraud Prevention and Initial Fraud Detection***

The team first detected fraudulent activity following a large increase in web-based screening activity over one weekend, with the number of screens exceeding the number of ad clicks, a high number of screens indicating ineligibility, and numerous screening forms filled out just minutes apart. In addition, the researchers found multiple screens with the same contact phone number, an increase in screens with inconsistent information (ie, due date and gestational age not matching), and an increase in responses to an introductory text asking for communication via text without a phone call. Similar to case study 2, the team learned that the study information had recently been shared via web-based platforms (eg, Facebook groups) related to completing surveys for remuneration, and not restricted to the

population relevant for study eligibility (ie, pregnant people). The researchers identified the need for a rigorous review of enrolled participants and for updated screening prior to enrolment.

In response, the team updated the initial web-based screening, with REB approval, to include phone number and due date in addition to the second screening form, to help detect multiple screening attempts by the same person. A detailed phone script included questions about how the pregnancy was going, whether the pregnancy had been confirmed, and what the due date was based on. If individuals provided inconsistent, nonsensical, vague, or minimal information, then the study staff requested documentation of their due date or gestational age, for example, a copy of their ultrasound report.

## Results

### Overview

These 4 cases of web-based surveys from Canada and Australia all used primary prevention strategies to prevent fraud, secondary prevention to detect suspicious activity for fraud, and once fraudulent responses were identified, developed tertiary prevention methods to identify and remove fraudulent data. All 4 studies implemented multiple primary and secondary fraud prevention techniques into the study design, such as the use of CAPTCHA questions and incorporation of a time stamp. While all 4 studies monitored incoming survey responses for fraud using secondary prevention techniques, as highlighted above, an increase in the number of incoming survey responses was the first detected sign of fraud in 2 of the 4 cases. Nonsensical or noncoherent responses and detection of duplicate responses alerted researchers to possible fraud in 3 of the 4 surveys. The researchers adapted several of the primary and secondary prevention strategies into tertiary strategies, as highlighted in Table S1 in [Multimedia Appendix 1](#). In the following paragraphs and Table S1 in [Multimedia Appendix 1](#) we describe the tertiary prevention strategies of each case study in more detail.

### Case Study 1: Tertiary Fraud Prevention and Identification and Removal of Fraudulent Responses

The 2019 Canadian Abortion Provider Survey: Following a literature review on fraud in web-based surveys, this team mapped applicable criteria as listed in Table S1 in [Multimedia Appendix 1](#) to their survey in collaboration with the applicable IRB and coinvestigators. The team created and tested an algorithm aiming to protect valid responses and detect fraud, then allocated negative fraud points to protective criteria (eg, sensical free-text responses) components and positive fraud points to components considered suspicious (eg, nonsensical answer combinations such as being a maternal-fetal medicine subspecialist but registered with the College of Family Physicians of Canada), with a higher number of points in either direction the stronger the assumed predictive value. The team focused on the mandatory questions that all respondents had completed. Additionally, IRB permission was obtained to temporarily link the email addresses provided when requesting remuneration to the respondent's answers. The algorithm was piloted, designating respondents with known or institution

emails as negative controls, and those with highly suspicious pattern-based emails as positive controls to assess each fraud criterion. Sensitivity and specificity analyses were conducted, respectively, on each criterion component and are presented in Table S3 in [Multimedia Appendix 1](#). Free text answers had the highest sensitivity (92.7%) and specificity (100%) while nonsensical answer combinations had lower sensitivity (54.5% combined with a high specificity (99.1)). Where possible the algorithm criteria were programmed in R Statistical Software (R Foundation for Statistical Computing) and at least 2 independent researchers manually assigned fraud points to the remaining criteria. Disagreements were discussed among multiple team members until a consensus was reached. The pilot data was reviewed and the team reached consensus on the criteria to include in the final algorithm.

The final fraud detection algorithm (Figure S1 in [Multimedia Appendix 1](#)) included the assessment of responses or response combinations for being nonsensical or nonprobable, and of those protected by sensical free text (eg, providing explanations to response options where "other" answer option was selected). It included the survey submission date before or after we detected fraud and remuneration criteria, in which not asking for remuneration was protective. No single criterion was used to exclude or include respondents. Depending on the combination of scoring results from the different criteria, the team included (low fraud score; n=342), excluded (high fraud score; n=302), or marked a participant for further review (midrange fraud score; n=271). These cut points were determined using the results from the sensitivity/specificity analysis described above. Then, responses marked for further review were manually assessed for additional free-text protective criteria responses or further nonsensical/nonprobable criteria responses. This survey had 1050 respondents, and after removal of participants due to data cleaning (n=135; duplicates, etc) and fraud detection (415/915, 45.3%), 500 respondents were included for analysis.

### Case Study 2: Tertiary Fraud Prevention and Identification and Removal of Fraudulent Responses

The Knowledge, Attitudes and Practices of Australian Primary Care Clinicians in the Provision of Long-Acting Reversible Contraception and Medication Abortion: Within the first 2 weeks after survey launch, the general practitioner and practice nurse surveys recorded over 300 responses over 4 days, in which respondent names did not match email addresses and were completed in batches from the same email account. Following advice from the research team and Monash University's information technology and REDCap staff, the team added a text-based CAPTCHA and a "honeypot" question not visible to human respondents to detect bots using the REDCap @HIDDEN-SURVEY function for improved primary fraud prevention. Changes were also made to the respondent explanatory statement, indicating full name verification would be requested via email prior to reimbursement. Following a reply email from the respondent, the team began to verify each respondent's name alongside the practice postcode given in the survey with their details on the AHPRA website. Active recruitment resumed following ethics committee approval of these changes. The team also developed an internal protocol for respondent verification and reimbursement. Across 3 surveys,



5598 (100%) respondents started the survey. From these, 45.5% (n=2547) could not be verified or were discontinued. Other reasons for excluded responses included duplicates or if the respondent did not complete the mandatory questions or did not work in general practice. A total of 27.2% (n=1521) responses were included in the final analyses.

### Case Study 3: Tertiary Fraud Prevention and Identification and Removal of Fraudulent Responses

**The EXTEND-PREFER Project:** Throughout the recruitment phase, the research team continued to implement the initiated primary and secondary fraud prevention measures and additionally conducted rigorous manual screening. This included a researcher checking all new responses daily for duplicate entries and if entries were flagged by fraud detection features in Qualtrics, including “ID Duplicate,” “Relevant ID Duplicate Score,” “Relevant ID Fraud Score,” and “ReCAPTCHA score”). Team members decided to conduct a more rigorous manual screening than initially planned due to a sudden influx (over a 48 hour period) of more than 800 likely fraudulent responses roughly one month into the recruitment phase. During this screening process, the researchers identified and excluded any completed responses deemed likely to be fraudulent or a duplicate entry, based on personal information provided (eg, respondent’s name, email address, and postcodes), computer information available (eg, IP address and timestamp), and those with unrealistically short response times (eg, <8 minutes).

For the rural and remote surveys, 3614 (100%) respondents started the initial survey. From these, 8.3% (n=300) respondents did not meet the eligibility criteria or could not be verified and were therefore discontinued. For the CALD surveys, 2348 (100%) respondents started the initial survey. From these, 16.4% (n=385) of respondents did not meet eligibility criteria or could not be verified and were therefore discontinued. Responses were excluded if there were multiple entries with a duplicated name, mobile number, or email address (retaining first response, and excluding additional entries), if it took the participant less than 800 s (approximately 13 mins) to complete the survey or if there were multiple responses with the same IP address (retaining first response, and excluding additional entries).

### Case Study 4: Tertiary Fraud Prevention and Identification and Removal of Fraudulent Responses

**SmartMom Randomized Controlled Trial:** To identify fraud, the researchers manually reviewed enrolled participants and categorized likelihood of fraud based on suspicious activity including (1) more than one screening attempt; (2) failure to provide personal health number; (3) inconsistency of information provided across surveys and screening; (4) same response provided for all questions in multiple choice standard scale surveys; (5) notably high interest in gift cards, manifested through asking several questions; and (6) low engagement with the texting program (ie, not texting keywords to receive optional content or notify of baby’s birth). Multiple screening attempts, high interest in remuneration and inconsistent or nonsensical responses on screening were assessed by the team to be most consistent with fraudulent activity. Repeat questions in surveys over time were less indicative, as the team noted high inconsistency across their study. It appeared that poor recall,

high error rates in responses, and lack of attention to detail were common in this perinatal study population. Based on the high fraud risk criteria above, each participant was categorized as low, medium, or high likelihood of being fraudulent. Data for participants with high fraud likelihood was excluded from analysis, and the team will undertake sensitivity analysis with and without those with medium fraud likelihood.

### Synthesis of Findings Across Case Studies

Across the 4 case studies, the researchers implemented primary fraud prevention strategies, but all surveys experienced fraud. We found that when primary and secondary fraud prevention techniques were more robust and included personal identifiers, researchers were able to save time and use less complicated tertiary fraud prevention techniques. Many fraud prevention and detection strategies, such as IP address monitoring and screening questions relevant to the inclusion criteria, can be used across prevention levels (primary, secondary, and tertiary) and have been deliberately included in the study design phase. Protecting participant identity by foregoing the collection of personal identifiers came at the cost of weakening fraud prevention strategies. Strategies to recruit an unidentifiable population, such as the use of a generic survey link and recruiting on social media, likely increased the risk of fraud by sharing the study with a wider audience than those who were eligible. Closely monitoring response patterns, while being a time-sensitive and time-consuming technique, proved critical in all 4 case studies in order to promptly identify the onset of suspicious and ultimately fraudulent respondent patterns. All 4 case studies offered remuneration for completing their surveys. Once fraud was detected, inclusion of checking questions that verify previous responses, mandatory free text questions, minimum time specification or timestamps, and inclusion of personal identifier fields were among the most commonly used and helpful tertiary prevention methods. Ultimately, 45.3% (n=415), 45.5% (n=2547), and 16.4% (n=385) of respondents were removed in case studies 1, 2, and 3, respectively. Case study 4 had not completed data analysis and, therefore, is unable to report fraud detection rates.

## Discussion

### Principal Findings

Our findings highlight that prevention and correct identification of fraud in web-based surveys are essential to maintain data integrity. Despite using primary prevention strategies to prevent fraud, all 4 research projects experienced substantial fraud, and needed to remove up to 45% of survey responses. Each study used secondary prevention methods that identified participants likely to be fraudulent. In consultation with members of our respective coinvestigator teams, ethics committees/IRBs, and, in some cases, biostatisticians, each study team developed tertiary prevention strategies to remove fraudulent respondents from the dataset and was able to move forward with recruitment and analysis.

### How to Prevent Fraud in Future Studies?

There is limited literature on survey fraud mitigation, with a particular gap for any papers that discuss detection and



prevention methods. A further gap is evident for the paucity (or lack) of articles that provide recommendations on how to reduce the impact of experiencing fraud on the validity of survey results across studies with varying degrees of personal identifiers and web-based recruitment strategies. Our manuscript builds on previously described fraud removal approaches and showcases how these can be tailored to surveys with different recruitment methods and data collection, including degrees of personal identifiers. Our cases highlight the importance of carefully weighing the pros and cons of study design, survey instruments, and recruitment decisions, while balancing the individual study needs versus the risk of experiencing fraud. A proactive, multilayered approach to fraud prevention throughout the study design, recruitment, and data analysis is crucial.

While case studies 1 and 2 recruited clinicians, and case studies 3 and 4 recruited end users of health care, all encountered fraud, indicating the pervasiveness of fraudulent behavior. All studies recruited using a generic survey link to widely recruit a population of at least partially unknown individuals. This is known to be a higher risk for fraud compared with sending individual survey invitation links [25]. Recruitment on social media further compounded the fraud risk in studies 2-4, although using only professional networks did not sufficiently protect study 1 from fraudulent respondents either [26]. Given that each study offered remuneration for participation, monetary incentives appear to have been a motivator for fraudsters. This aligns with prior research suggesting that financial compensation can drive fraudulent behavior in survey-based studies [9,27,28]. Avoiding remuneration, not advertising remuneration, or choosing options such as a lottery has been found to be protective [29]. Fortunately, there are evolving mechanisms to identify or remove bots, including the use of CAPTCHA or Honeypot questions [30].

One significant risk for fraudulent participation and a serious obstacle to identifying fraudulent responses is exemplified by the first case study, in which the absence of personal identifiers during data collection made the identification of fraudulent respondents more complex and time-consuming. While this approach was designed with the IRB to protect participant anonymity when researching a sensitive study population, it potentially increased the risk of fraud and further hindered the research team's ability to quickly and effectively remove fraudulent responses, and required the implementation of tertiary strategies. This contrasts with case studies 2 through 4, which were able to rely on well-documented methods to detect fraud, such as personal identifiers and IP addresses that may have been protective and required a less rigorous tertiary prevention protocol [6,28,31]. Contacting respondents to verify their information and eligibility criteria (as in case studies 2 and 4), has been used as an effective method to detect fraud [3].

These 4 case studies highlight the conflict between safeguarding participant privacy, especially when surveying vulnerable populations, and ensuring the accuracy and reliability of collected data. While the first case study, which did not use identifiers, effectively used tertiary prevention strategies, it increased the time commitment and complexity of detecting fraud. Similar to our first case study, Lawlor et al [28] described a framework to identify suspected web-based survey fraud, in

which they emphasize the importance of using multiple criteria combined in an algorithm to carefully determine the threshold to exclude a respondent. Among the multiple criteria in the tertiary fraud detection and removal algorithm of the first case, the following were especially effective: multiple screening attempts; implausible or inconsistent responses for clinical information (though can also be a sign of survey fatigue); refusal to provide requested identifier; several questions about remuneration; and lack of interest or questions about the purpose and nature of the study. Some of these effective criteria, such as open-ended questions or embedding several questions that can be linked to expose nonsensical answer combinations, are simple to implement and should be considered at the survey development stage. Unusual email addresses with patterns that switched between numbers and letters and did not appear to include initials or a name have been found to be associated with fraudsters in other studies, and are consistent with several of our case studies [3,10,32,33].

### Strengths and Limitations

We presented 4 case-studies from 2 countries, which surveyed clinician as well as patient respondent populations, enhancing the relevance of the findings across different reproductive health topic areas, populations, and geographical contexts. We included large-scale national web-based surveys, providing valuable insights into the scalability and effectiveness of fraud prevention strategies. All case studies are recent and applicable to contemporary issues in fraud prevention. Additionally, the variation in study design between case studies enabled us to conduct a rich exploration of primary, secondary, and tertiary fraud prevention strategies, offering a nuanced description of how these approaches can be applied at different stages of study design and how they are interconnected.

The case studies are based on a convenience sample of studies conducted by teams that included one or more of our author group, and all focus on reproductive health, which limits the generalizability of the findings to other areas of public health. Furthermore, the tertiary prevention strategies presented are specifically tailored to each case study context and may require adaptation to be effective in other contexts. This limitation is mitigated to some extent by our presentation of broad components, offering a framework that can be adapted in different contexts. While all studies removed fraudulent responses, as they were unable to definitively determine which participants were fraudsters, it is possible they excluded some valid respondents.

### Conclusions

Web-based surveys are at substantial risk for fraudulent responses, and prevention strategies are essential to enable researchers and IRBs to protect data integrity. Ideally, strategies should be incorporated into the survey and recruitment during the study design phase. Developing strong primary fraud prevention strategies, paired with having a predefined approach to monitoring for fraud (secondary prevention), can reduce the reliance on tertiary methods to remove fraudulent responses and ensure data integrity. The criteria we have presented in the fraud detection algorithm can be adapted to other web-based surveys, survey topics, and populations.

---

## Acknowledgments

Partner organizations include the Australian Commission on Safety and Quality in Health Care, Australian Government Department of Health, Australian Primary Health Care Nurses Association, Australasian Sexual Health and HIV Nurses Association, Australian Women's Health Nurse Association, Bayer Australia and New Zealand, Children by Choice, Family Planning New South Wales, Family Planning Tasmania, Family Planning Welfare Association of NT, Jean Hailes, Marie Stopes Australia, Organon, Pharmaceutical Society of Australia, Royal Australian College of General Practitioners, Sexual Health Victoria, Sexual Health Quarters, SHINE SA, and The Royal Australian and New Zealand College of Obstetricians and Gynecologists.

The individual case study funders had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

---

## Funding

The 2019 Canadian Abortion Provider Survey was supported by the Canadian Institutes of Health Research (PJT-162201). WVN was supported by a Canadian Institute for Health Research and Public Health Agency of Canada Chair in Applied Public Health Research (2014-2024, CPP-329455-107837) and is currently supported as a Tier 1 Canada Research Chair in Family Planning Innovation (2024-2032, CRC-2023-00135). In-kind support was provided by the Women's Health Research Institute of the British Columbia Women's Hospital, the Society of Obstetricians and Gynaecologists of Canada, and the Canadian Nurses Association.

The Australian Contraception and Abortion Primary Care Practitioner Support Network knowledge, attitudes, and practices surveys were conducted with support from the National Health and Medical Research Council (grant ID 1191793) and partner organization monetary or material support. This EXTEND-PREFER project was funded by the Commonwealth of Australia Department of Health, through the National Women's Health Strategy and National Men's Health Strategy Health Promotion Grant (grant ID 4-ECPEF7J).

The SmartMom randomized controlled trial was funded by the Canadian Institutes for Health Research (application number 470340).

---

## Data Availability

The datasets generated or analyzed in these studies have varying availability based on ethical and funding approvals. For the 2019 Canadian Abortion Provider Survey (case study 1) and the Australian Contraception and Abortion Primary Care Practitioner Support Network knowledge, attitudes, and practices (case study 2), the datasets are not publicly available due to a lack of approval from the respective research ethics boards. Similarly, for EXTEND-PREFER (case study 3), the datasets cannot be shared publicly without approval from the funder or the research ethics board. In contrast, the datasets from The SmartMom randomized controlled trial (case study 4) are available from the corresponding author upon reasonable request.

---

## Authors' Contributions

We guarantee that all authors have substantially contributed to this manuscript. All authors were coinvestigators in at least one of the 4 case studies and contributed to the analysis presented in this manuscript. ME, RR, CMS, and SJ drafted the first manuscript, and all authors contributed to revisions and accepted the final manuscript.

---

## Conflicts of Interest

None declared.

---

## Multimedia Appendix 1

Additional tables.

[\[PDF File \(Adobe PDF File\), 543 KB-Multimedia Appendix 1\]](#)

---

## Multimedia Appendix 2

The surveys for case studies 1-4.

[\[PDF File \(Adobe PDF File\), 1333 KB-Multimedia Appendix 2\]](#)

---

## References

1. Eysenbach G, Wyatt J. Using the internet for surveys and health research. *J Med Internet Res*. 2002;4(2):E13. [\[FREE Full text\]](#) [doi: [10.2196/jmir.4.2.e13](https://doi.org/10.2196/jmir.4.2.e13)] [Medline: [12554560](https://pubmed.ncbi.nlm.nih.gov/12554560/)]
2. van Gelder MMHJ, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? *Am J Epidemiol*. 2010;172(11):1292-1298. [doi: [10.1093/aje/kwq291](https://doi.org/10.1093/aje/kwq291)] [Medline: [20880962](https://pubmed.ncbi.nlm.nih.gov/20880962/)]

3. Ballard AM, Cardwell T, Young AM. Fraud detection protocol for web-based research among men who have sex with men: development and descriptive evaluation. *JMIR Public Health Surveill*. 2019;5(1):e12344. [FREE Full text] [doi: [10.2196/12344](https://doi.org/10.2196/12344)] [Medline: [30714944](https://pubmed.ncbi.nlm.nih.gov/30714944/)]
4. Fitzgerald D, Hockey R, Jones M, Mishra G, Waller M, Dobson A. Use of online or paper surveys by Australian women: longitudinal study of users, devices, and cohort retention. *J Med Internet Res*. 2019;21(3):e10672. [FREE Full text] [doi: [10.2196/10672](https://doi.org/10.2196/10672)] [Medline: [30869647](https://pubmed.ncbi.nlm.nih.gov/30869647/)]
5. Regmi PR, Waithaka E, Paudyal A, Simkhada P, van Teijlingen E. Guide to the design and application of online questionnaire surveys. *Nepal J Epidemiol*. 2016;6(4):640-644. [FREE Full text] [doi: [10.3126/nje.v6i4.17258](https://doi.org/10.3126/nje.v6i4.17258)] [Medline: [28804676](https://pubmed.ncbi.nlm.nih.gov/28804676/)]
6. Teitcher JEF, Bocking WO, Bauermeister JA, Hofer CJ, Miner MH, Klitzman RL. Detecting, preventing, and responding to "fraudsters" in internet research: ethics and tradeoffs. *J Law Med Ethics*. 2015;43(1):116-133. [FREE Full text] [doi: [10.1111/jlme.12200](https://doi.org/10.1111/jlme.12200)] [Medline: [25846043](https://pubmed.ncbi.nlm.nih.gov/25846043/)]
7. Fielding N, Lee R, Blank G. *The SAGE Handbook of Online Research Methods*. The University of British Columbia. 2008. URL: [http://GW2JH3XR2C.search.serialssolutions.com/?sid=sersol&SS\\_jc=TC0000784208&title=The%20SAGE%20handbook%20of%20online%20research%20methods852](http://GW2JH3XR2C.search.serialssolutions.com/?sid=sersol&SS_jc=TC0000784208&title=The%20SAGE%20handbook%20of%20online%20research%20methods852) [accessed 2025-10-22]
8. Singh S, Sagar R. A critical look at online survey or questionnaire-based research studies during COVID-19. *Asian J Psychiatr*. 2021;65:102850. [FREE Full text] [doi: [10.1016/j.ajp.2021.102850](https://doi.org/10.1016/j.ajp.2021.102850)] [Medline: [34534919](https://pubmed.ncbi.nlm.nih.gov/34534919/)]
9. Quach S, Pereira JA, Russell ML, Wormsbecker AE, Ramsay H, Crowe L, et al. The good, bad, and ugly of online recruitment of parents for health-related focus groups: lessons learned. *J Med Internet Res*. 2013;15(11):e250. [FREE Full text] [doi: [10.2196/jmir.2829](https://doi.org/10.2196/jmir.2829)] [Medline: [24231040](https://pubmed.ncbi.nlm.nih.gov/24231040/)]
10. Bauermeister JA, Pingel E, Zimmerman M, Couper M, Carballo-Diéguez A, Strecher VJ. Data quality in web-based HIV/AIDS research: handling invalid and suspicious data. *Field Methods*. 2012;24(3):272-291. [FREE Full text] [doi: [10.1177/1525822X12443097](https://doi.org/10.1177/1525822X12443097)] [Medline: [23180978](https://pubmed.ncbi.nlm.nih.gov/23180978/)]
11. Levi R, Ridberg R, Akers M, Seligman H. Survey fraud and the integrity of web-based survey research. *Am J Health Promot*. 2022;36(1):18-20. [doi: [10.1177/08901171211037531](https://doi.org/10.1177/08901171211037531)] [Medline: [34372667](https://pubmed.ncbi.nlm.nih.gov/34372667/)]
12. Pratt-Chapman M, Moses J, Arem H. Strategies for the identification and prevention of survey fraud: data analysis of a web-based survey. *JMIR Cancer*. 2021;7(3):e30730. [FREE Full text] [doi: [10.2196/30730](https://doi.org/10.2196/30730)] [Medline: [34269685](https://pubmed.ncbi.nlm.nih.gov/34269685/)]
13. Nur AA, Leibbrand C, Curran SR, Votruba-Drzal E, Gibson-Davis C. Managing and minimizing online survey questionnaire fraud: lessons from the triple C project. *Int J Soc Res Methodol*. 2024;27(5):613-619. [doi: [10.1080/13645579.2023.2229651](https://doi.org/10.1080/13645579.2023.2229651)] [Medline: [39494158](https://pubmed.ncbi.nlm.nih.gov/39494158/)]
14. Crowe S, Cresswell K, Robertson A, Huby G, Avery A, Sheikh A. The case study approach. *BMC Med Res Methodol*. 2011;11:100. [FREE Full text] [doi: [10.1186/1471-2288-11-100](https://doi.org/10.1186/1471-2288-11-100)] [Medline: [21707982](https://pubmed.ncbi.nlm.nih.gov/21707982/)]
15. Implementing the population health approach. Government of Canada. 2025. URL: <https://www.canada.ca/en/public-health/services/health-promotion/population-health/implementing-population-health-approach.html> [accessed 2025-10-10]
16. Renner R, Wagner M, Dunn S, Guilbert E, Munro S, Grewal S, et al. Development and testing for a national survey: the Canadian abortion provider survey (CAPS). *J Obstet Gynaecol Can*. 2020;42(5):690. [FREE Full text] [doi: [10.1016/j.jogc.2020.02.095](https://doi.org/10.1016/j.jogc.2020.02.095)]
17. BCCHRI REDCap datasystem. BC Children's Hospital Research. URL: <https://rc.bcchr.ca/> [accessed 2023-06-30]
18. Canadian Abortion Providers Support-Communauté de pratique canadienne sur l'avortement. URL: <https://caps-cpca.ubc.ca/> [accessed 2025-11-15]
19. Kramer J, Rubin A, Coster W, Helmuth E, Hermos J, Rosenbloom D, et al. Strategies to address participant misrepresentation for eligibility in web-based research. *Int J Methods Psychiatr Res*. 2014;23(1):120-129. [FREE Full text] [doi: [10.1002/mp.1415](https://doi.org/10.1002/mp.1415)] [Medline: [24431134](https://pubmed.ncbi.nlm.nih.gov/24431134/)]
20. Mazza D, James S, Black K, Taft A, Bateson D, McGeechan K, et al. Increasing the availability of long-acting reversible contraception and medical abortion in primary care: the Australian Contraception and Abortion Primary Care Practitioner Support Network (AusCAPPS) cohort study protocol. *BMJ Open*. 2022;12(12):e065583. [FREE Full text] [doi: [10.1136/bmjopen-2022-065583](https://doi.org/10.1136/bmjopen-2022-065583)] [Medline: [36521891](https://pubmed.ncbi.nlm.nih.gov/36521891/)]
21. Harris P, Taylor R, Minor B, Elliott V, Fernandez M, O'Neal L, et al. REDCap Consortium. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform*. 2019;95:103208. [FREE Full text] [doi: [10.1016/j.jbi.2019.103208](https://doi.org/10.1016/j.jbi.2019.103208)] [Medline: [31078660](https://pubmed.ncbi.nlm.nih.gov/31078660/)]
22. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377-381. [FREE Full text] [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
23. Mazza D, Botfield JR, Zeng J, Morando-Stokoe C, Arefadib N. Evaluating the effectiveness of a tailored online educational video on the contraceptive knowledge and decision making of young women from culturally and linguistically diverse backgrounds: findings from the EXTEND-PREFER study. *BMJ Sex Reprod Health*. 2025;51(1):18-26. [doi: [10.1136/bmjsexrh-2024-202236](https://doi.org/10.1136/bmjsexrh-2024-202236)] [Medline: [38871444](https://pubmed.ncbi.nlm.nih.gov/38871444/)]

24. Janssen P, Lecke S, Renner R, Zhang W, Vedam S, Norman WV, et al. Teaching by texting to promote positive health behaviours in pregnancy: a protocol for a randomised controlled trial of SmartMom. *BMJ Open*. 2024;14(1):e081730. [FREE Full text] [doi: [10.1136/bmjopen-2023-081730](https://doi.org/10.1136/bmjopen-2023-081730)] [Medline: [38238058](https://pubmed.ncbi.nlm.nih.gov/38238058/)]
25. Campbell CK, Ndokwe S, Dubé K, Saucedo JA, Saberi P. Overcoming challenges of online research: measures to ensure enrollment of eligible participants. *J Acquir Immune Defic Syndr*. 2022;91(2):232-236. [FREE Full text] [doi: [10.1097/QAI.0000000000003035](https://doi.org/10.1097/QAI.0000000000003035)] [Medline: [36094490](https://pubmed.ncbi.nlm.nih.gov/36094490/)]
26. Mizerek E, Wolf L, Moon MD. Identifying and mitigating fraud when using social media for research recruitment. *J Emerg Nurs*. 2023;49(4):530-533. [doi: [10.1016/j.jen.2023.04.002](https://doi.org/10.1016/j.jen.2023.04.002)] [Medline: [37393079](https://pubmed.ncbi.nlm.nih.gov/37393079/)]
27. Reips U. Standards for internet-based experimenting. *Exp Psychol*. 2002;49(4):243-256. [doi: [10.1026//1618-3169.49.4.243](https://doi.org/10.1026//1618-3169.49.4.243)] [Medline: [12455331](https://pubmed.ncbi.nlm.nih.gov/12455331/)]
28. Lawlor J, Thomas C, Guhin AT, Kenyon K, Lerner MD, Drahota A. Suspicious and fraudulent online survey participation: introducing the REAL framework. *Method Innov*. 2021;14(3):205979912110504. [doi: [10.1177/20597991211050467](https://doi.org/10.1177/20597991211050467)]
29. Sobolewski J, Rothschild A, Freeman A. The impact of incentives on data collection for online surveys: social media recruitment study. *JMIR Form Res*. 2024;8:e50240. [FREE Full text] [doi: [10.2196/50240](https://doi.org/10.2196/50240)] [Medline: [38963924](https://pubmed.ncbi.nlm.nih.gov/38963924/)]
30. Parker JN, Rager TL, Burns J, Mmeme O. Data verification and respondent validity for a web-based sexual health survey: tutorial. *JMIR Form Res*. 2024;8:e56788. [FREE Full text] [doi: [10.2196/56788](https://doi.org/10.2196/56788)] [Medline: [39651967](https://pubmed.ncbi.nlm.nih.gov/39651967/)]
31. Waggoner P, Kennedy R, Clifford S. Detecting fraud in online surveys by tracing, scoring, and visualizing IP addresses. *JOSS*. 2019;4(37):1285. [FREE Full text] [doi: [10.21105/joss.01285](https://doi.org/10.21105/joss.01285)]
32. Heffner J, Watson N, Dahne J, Croghan I, Kelly MM, McClure JB, et al. Recognizing and preventing participant deception in online nicotine and tobacco research studies: suggested tactics and a call to action. *Nicotine Tob Res*. 2021;23(10):1810-1812. [FREE Full text] [doi: [10.1093/ntr/ntab077](https://doi.org/10.1093/ntr/ntab077)] [Medline: [33876244](https://pubmed.ncbi.nlm.nih.gov/33876244/)]
33. Bowen AM, Daniel CM, Williams ML, Baird GL. Identifying multiple submissions in internet research: preserving data integrity. *AIDS Behav*. 2008;12(6):964-973. [FREE Full text] [doi: [10.1007/s10461-007-9352-2](https://doi.org/10.1007/s10461-007-9352-2)] [Medline: [18240015](https://pubmed.ncbi.nlm.nih.gov/18240015/)]

## Abbreviations

**CALD:** culturally and linguistically diverse  
**CHERRIES:** Checklist for Reporting Results of Internet E-Surveys  
**IRB:** institutional review board  
**LARC:** long-acting reversible contraception  
**REB:** Research Ethics Board  
**REDCap:** Research Electronic Data Capture

*Edited by A Stone, A Mavragani; submitted 06.Jun.2025; peer-reviewed by X Jing, J Hardesty; comments to author 31.Jul.2025; accepted 16.Oct.2025; published 01.Dec.2025*

### *Please cite as:*

Ennis M, Renner R-M, Morando-Stokoe C, James S, Janssen PA, Leckie S, Dunn S, Mazza D, Norman WV  
*Exploring Methods to Mitigate Fraud in Web-Based Surveys: Multicase Study Analysis*  
*J Med Internet Res* 2025;27:e78671  
URL: <https://www.jmir.org/2025/1/e78671>  
doi: [10.2196/78671](https://doi.org/10.2196/78671)  
PMID:

©Madeleine Ennis, Regina-Maria Renner, Claudia Morando-Stokoe, Sharon James, Patricia A Janssen, Sara Leckie, Sheila Dunn, Danielle Mazza, Wendy V Norman. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org/>), 01.Dec.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.