Unmasking Fraud: A Multimethod Pilot Study on the Problem of Fraudulent Participants in Healthcare Research

Vithusa Kumarasamy, Nicole Goodfellow, Era Mae Ferron, Amy L. Wright

Submitted to: JMIR Formative Research
on: August 02, 2023

Table of Contents

Unmasking Fraud: A Multimethod Pilot Study on the Problem of Fraudulent Participants in Healthcare Research

Vithusa Kumarasamy[1] BScN, MN; Nicole Goodfellow[1] BScN, MN; Era Mae Ferron[1] BNSc, MN, PhD; Amy L. Wright[1] MScN, PhD, NP-Pediatrics, NCC-BC, CNeoN(C)

[1]Lawrence S. Bloomberg Faculty of Nursing University of Toronto Toronto CA

Corresponding Author:
Amy L. Wright MScN, PhD, NP-Pediatrics, NCC-BC, CNeoN(C)
Lawrence S. Bloomberg Faculty of Nursing
University of Toronto
155 College Street
Toronto
CA

Abstract

Background: The shift towards online recruitment methods, accelerated by the COVID-19 pandemic, has brought to the forefront the growing concern of encountering fraudulent participants in healthcare research. The increasing prevalence of this issue poses a serious threat to the reliability and integrity of research data and subsequent findings.

Objective: This study aimed to explore the experiences of healthcare researchers (HCRs) who have encountered fraudulent participants while using online recruitment methods and platforms. The primary objective was to gain insights into how researchers detect and mitigate fraudulent behavior in their work and provide prevention recommendations.

Methods: A multimethod sequential design was employed for this pilot study, comprising a quantitative arm involving an online survey, followed by a qualitative arm featuring semi-structured interviews. The qualitative description approach framed the qualitative arm of the study. Sample size determination was based on information power and data redundancy, with content analysis used to analyze open-ended survey questions and interview data.

Results: A total of 37 HCRs participated with 13 of them engaging in qualitative interviews. Online platforms such as Facebook, email, Twitter, and online newsletters were the most used methods for recruitment. 84% of fraudulent participation occurred in studies that mentioned incentives in their recruitment communications, with 71% of HCRs offering physical or electronic gift cards as incentives. Researchers identified several indicators of suspicious behavior, including email surges, discrepancies in contact or personal information, geographical inconsistencies, and suspicious responses to survey questions. HCRs emphasized the need for a comprehensive screening protocol that extends beyond eligibility checks and is seamlessly integrated into the study protocol, grant applications, and Research Ethics Board (REB) submissions.

Conclusions: This study sheds light on the intricate and pervasive problem of fraudulent participation in healthcare research using online recruitment methods. The findings underscore the importance of vigilance and proactivity among HCRs in identifying, preventing, and addressing fraudulent behavior. To effectively tackle this challenge, researchers are encouraged to develop a comprehensive prevention strategy and establish a community of practice, facilitating real-time access to solutions, support, and the promotion of ethical research practices. This collaborative approach will enable researchers to effectively address the issue of fraudulent participation, ensuring the conduct of high-quality and ethically sound research in the digital age.

(JMIR Preprints 02/08/2023:51530)
DOI: https://doi.org/10.2196/preprints.51530

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Unmasking Fraud: A Multimethod Pilot Study on the Problem of Fraudulent Participants in Healthcare Research

Abstract

Background:
The shift towards online recruitment methods, accelerated by the COVID-19 pandemic, has brought to the forefront the growing concern of encountering fraudulent participants in healthcare research. The increasing prevalence of this issue poses a serious threat to the reliability and integrity of research data and subsequent findings.

Objective:
This study aimed to explore the experiences of healthcare researchers (HCRs) who have encountered fraudulent participants while using online recruitment methods and platforms. The primary objective was to gain insights into how researchers detect and mitigate fraudulent behavior in their work and provide prevention recommendations.

Methods:
A multimethod sequential design was employed for this pilot study, comprising a quantitative arm involving an online survey, followed by a qualitative arm featuring semi-structured interviews. The qualitative description approach framed the qualitative arm of the study. Sample size for the quantitative and qualitative arms were based on pragmatic considerations that in part stemmed from experiencing fraudulent participants in a concurrent study. Content analysis was used to analyze open-ended survey questions and interview data.

Results:
A total of 37 HCRs participated with 13 of them engaging in qualitative interviews. Online platforms such as Facebook, email, Twitter, and online newsletters were the most used methods for recruitment. 84% of fraudulent participation occurred in studies that mentioned incentives in their recruitment communications, with 71% of HCRs offering physical or electronic gift cards as incentives. Researchers identified several indicators of suspicious behavior, including email surges, discrepancies in contact or personal information, geographical inconsistencies, and suspicious responses to survey questions. HCRs emphasized the need for a comprehensive screening protocol that extends beyond eligibility checks and is seamlessly integrated into the study protocol, grant applications, and Research Ethics Board (REB) submissions.

Conclusions:
This study sheds light on the intricate and pervasive problem of fraudulent participation in healthcare research using online recruitment methods. The findings underscore the importance of vigilance and proactivity among HCRs in identifying, preventing, and addressing fraudulent behavior. To effectively tackle this challenge, researchers are encouraged to develop a comprehensive prevention strategy and establish a community of practice, facilitating real-time access to solutions, support, and the promotion of ethical research practices. This collaborative approach will enable researchers to effectively address the issue of fraudulent participation, ensuring the conduct of high-quality and ethically sound research in the digital age.

Keywords:
fraudulent participants; threats to data integrity; online recruitment; multimethod study; healthcare research, bots, social media

Introduction

Recruiting participants for healthcare research through online methods and platforms is becoming increasingly prevalent, particularly in recent years.[1–3] In response to the COVID-19 lockdown in 2020, many researchers had to shift their recruitment strategies from traditional approaches to online methods like email, listservs, websites, and social media platforms like Facebook, Instagram, X (formally known as Twitter), and crowdsourcing websites including Amazon mTurk and Prolific. [1,4–11] While this sudden shift was unexpected, researchers are recognizing the numerous benefits associated with online recruitment strategies. Online recruitment methods offer the potential to reach a broader and more diverse participant pool including vulnerable populations.[1,3,12–20] They also enable targeted sampling,[4,21] offer convenience for participants,[3,4,14,15,18–20,22] enhance participant anonymity,[1,12,14,23] facilitate faster recruitment,[18,19,24] and lower recruitment costs.[2–4,14,15,18–20,22] Despite its numerous benefits, online recruitment brings with it several challenges. One challenge is the increased risk of fraudulent participants, defined as ineligible persons or computer bots designed to pose as real people to participate in research studies, threatening data quality.[3,11,17] For instance, Pozzar et al[2] investigated the extent of fraudulent participants in healthcare research using online recruitment methods, specifically social media, and found that almost their entire sample of 271 survey respondents were either fraudulent (94.5%) or suspicious (5.5%). In a qualitative study of 19 research experts, Teitcher et al[25] found that participant misrepresentation was a common problem in online research, with various forms of misrepresentation present, such as duplicate responses, fraudulent demographic information, and dishonesty about eligibility criteria.

Fraudulent participants generally fall into two main categories: real humans who participate in a disingenuous manner and computer bots designed to impersonate human participants.[18] Both human and computer bots attempt to participate in research studies in which they are not qualified or attempt to participate multiple times in the same study.[17,26] The presence of fraudulent participants in research studies can lead to various detrimental outcomes. These include the increased financial burden of identifying and addressing the impacts of fraudulent participants, as well as the obligation to compensate such individuals, including bots, that generate multiple invalid responses, in accordance with the study protocol.[27] Additionally, the presence of fraudulent participants can compromise the validity and reliability of research findings, potentially resulting in misguided recommendations that may have harmful consequences.[1,3,17]

The presence of fraudulent participants imposes significant stress on research teams,[16] such as requiring additional resources to manage and mitigate their impacts. This unforeseen allocation of resources could limit the availability of research staff for other tasks, potentially causing delays in study recruitment and data analysis.[16] Furthermore, fraudulent participants in research studies can lead to the misallocation of funding and human resources through the recruitment and compensation of ineligible individuals.[2,16]

Numerous ethical dilemmas that have emerged from this issue include the misuse of research funds to compensate fraudulent participants[1,11,28,29] and implementing invasive participant verification strategies to validate subjects' identities.[25] Underlying ethical principles such as 'respect for persons' also comes into question as research teams attempt to manage fraudulent participant encounters. For instance, while researchers are required to outline all study methods to individuals interested in participating, researchers are faced with the ethical dilemma of whether they should disclose their fraudulent participant detection and prevention strategies (e.g., tracking Internet Protocol (IP) addresses) to potential participants.[25] Although respecting individuals' rights to be informed about all study components is necessary, researchers also want to avoid deterring anyone

from participating in their studies and having fraudulent participants bypass their security measures.[25] The need to adequately balance respect for persons and respect for privacy with data integrity and researcher transparency, calls for further attention to ensure that ethical principles are upheld throughout the research process.[3,13,25]

Although the issue of fraudulent participants in healthcare research is not a new issue, as healthcare researchers are increasingly using online strategies to recruit participants, they are becoming more aware of the inherent issues in the approach. Further research to investigate and develop strategies to combat these issues is necessary. One such strategy, the REAL framework, presented by Lawlor and colleagues'[11] provides a structured approach for researchers to assist in prevention and identification of fraudulent participants within their samples. Researchers are first asked to "REFLECT" on the inherent vulnerabilities of the recruitment plan and consider built in survey design elements to avoid fraud. Next, researchers must give thought to the usual patterns in the data that they would "EXPECT" to see and what would present as unusual. The "ANALYZE" stage guides researchers to assess whether actual data patterns align with their expectations. Finally, the framework reinforces the importance of transparency through its "LABEL" stage, encouraging researchers to establish criteria for labeling and excluding fraudulent responses, addressing the issue of underreporting fraudulent participants. While not providing an exact count of fraudulent participation, the framework assists researchers in detecting and addressing the inclusion of fraudulent participants, contributing to the mitigation of this challenge. Considering the criticality of upholding research integrity, further investigation is required to identify and address fraudulent participants during the recruitment stages of healthcare studies where online methods are employed.

Aims

The aim of this study was to describe the issue of fraudulent participants in healthcare studies where online methods and platforms are used to recruit participants to provide healthcare researchers with prevention strategies. We aimed to answer the following research questions: (1) What are the experiences of healthcare researchers when encountering fraudulent participants through online recruitment strategies; (2) How do healthcare researchers identify fraudulent participants?; (3) How do healthcare researchers deter fraudulent participants?; and (4) How do healthcare researchers ensure the integrity of the data they collect when fraudulent participants are suspected?

Methods

Design

Using a multimethod sequential approach, we conducted a pilot study to explore the experiences of healthcare researchers with encountering fraudulent participants when using online recruitment strategies. The study consisted of sequential quantitative and qualitative phases involving an online survey which informed semi-structured interviews that used a qualitative descriptive design.

Participants and Recruitment

We recruited healthcare researchers from healthcare-related Faculties, schools, or departments at colleges or universities in Canada who had suspected at least one fraudulent participant in a study conducted in the past five years using online recruitment methods. For this study, a healthcare researcher (HCR) was defined as a member of a research team, including principal investigators, co-investigators, research managers, research coordinators, research officers, and research assistants, or some iteration of these roles or job titles. We used purposive and snowball sampling to recruit participants by distributing a digital recruitment flyer via targeted email outreach to individual researchers, and mass email distribution by Faculty/department/organization administrators to their researcher employees on our behalf. Potential participants were directed to complete an online

survey. Our sample size was based on pragmatic considerations including budget constraints and finding timely solutions for fraudulent participants in our concurrent study, strategies for pilot studies supported in the literature.[30] As such, we aimed to recruit 30 participants for the quantitative phase of the study and 10 participants for the qualitative arm. Participants who indicated their interest in participating in a follow-up interview and provided detailed and varied responses to open-ended questions were contacted for an interview. All participants provided informed consent before participation and were compensated with a $5 gift card in Canadian dollars (CAD) for completing the survey and an additional $15 CAD gift card for participating in an interview.

To safeguard against attracting fraudulent participants, several precautionary measures were implemented. First, a CAPTCHA was included at the beginning of the survey to deter automated submissions. Additionally, participants were contacted directly through their professional email addresses obtained from their respective employer's website. Recruitment was conducted through direct requests made to college or university Faculties, departments, or organization administrators, who in turn sent bulk emails to their staff members to invite participation. To enhance the authenticity of participants, those who consented to a virtual interview were interviewed with their cameras turned on, ensuring a form of verification. Data Collection

In the quantitative arm of the study, we used an online survey in the Research Electronic Data Capture (REDcap)[31] software platform to collect data using closed- and open-ended questions to identify the circumstances of participants' encounters with fraudulent participants, such as the number of studies in which a fraudulent participant was discovered, and the online platforms and strategies used to recruit participants. REDCap[31] is a secure web application that supports online data collection for research studies.[32,33] In the subsequent qualitative arm, participants were interviewed virtually on Zoom using a semi-structured interview guide to better understand healthcare researchers' experiences with encountering fraudulent participants and the approaches used to prevent fraudulent participants in future studies. The interview guide was informed by the results of the survey data and questions were designed to gain a further understanding of answers provided in the survey. All interviews were recorded with the permission of the participant and lasted approximately 30 minutes.

Analyses

Triangulation was performed during data analysis and interpretation. We analyzed the quantitative survey data by employing descriptive statistics, including means, frequencies, and percentages. Content analysis was then conducted on the interview and open-ended survey data. Three researchers read the transcripts multiple times to familiarize themselves with the data and decided on the analysis of manifest content, the visible, obvious components of the text,[34] following the methods described by Elo and Kyngäs[35] and Vaismoradi et al.[36] Open coding was then performed, and three researchers met several times to categorize and compare codes both with each other and with the entire data set,[35,36] and a categorical summary was devised.[34–36]

Ethics Approval

All procedures were reviewed and approved by the Health Sciences research ethics board of the University of Toronto (44014).

Results

The study results are presented in an integrated manner, where findings from both quantitative and qualitative components are combined to facilitate comparison and expansion. Similar results are grouped together, enhancing the comprehensive analysis of the data. A total of 43 individuals accessed and initiated the survey, of which 6 individuals were excluded (2 did not complete the consent form, 2 did not complete any part of the survey, and 2 did not meet eligibility criteria). This

resulted in a final sample of 37 participants for the quantitative arm of the study. Among these, 13 participated in an interview with authors VK or NG. The entire sample had a mean age of 35.8 years (SD=9.9), held various academic rankings and research positions including postdoctoral fellows and graduate students, with an average role duration of 4.2 years (SD=3.8), and the majority (35.1%) of participants were employed at a nursing Faculty (refer to Table 1 for a summary of the sample characteristics). Please note that Twitter underwent rebranding and became known as X corporation after conducting interviews for this study.[7] Participants refer to Twitter rather than X corporation in the following quotations.

Table 1. Characteristics of healthcare researchers.

Characteristic		Survey Participants (n=37)	Interview Participants (n=13)
Age in years, mean (SD)		35.8 (9.9)	39.4 (10.4)
Age range in years, n (%)			
	20-29	12 (32.4)	3 (23.1)
	30-39	13 (35.1)	4 (30.8)
	40-49	6 (16.2)	3 (23.1)
	50-59	6 (16.2)	3 (23.1)
Research role, n (%)			
	Pre-tenure Professor [a]	10 (27.0)	6 (46.2)
	Tenured Professor[b]	4 (10.8)	2 (15.4)
	Research Assistant[c]	10 (27.0)	3 (23.1)
	Research Manager[d]	7 (18.9)	2 (15.4)
	Lead Researcher[e]	3 (8.1)	0 (0.0)
	Graduate Student[f]	2 (5.4)	0 (0.0)
	Postdoc	1 (2.7)	0 (0.0)
Years in the role, mean (SD)		4.2 (3.8)	5.0 (4.7)
Years in the role, n			
	Min	<1	<1
	Max	17	10
Research Department, n (%)			
	Nursing	13 (35.1)	6 (46.2)
	Medicine	7 (18.9)	2 (15.4)
	Health Sciences	5 (13.5)	0 (0.0)
	Public Health	4 (10.8)	2 (15.4)
	Social Work	3 (8.1)	2 (15,4)
	Pharmacy	2 (5.4)	1 (7.7)
	Psychology	1 (2.7)	0 (0.0)
	Research Institute	1 (2.7)	0 (0.0)
	Social Science	1 (2.7)	0 (0.0)

[a] The pre-tenure professor role includes assistant professors.

[b] The tenured professor role includes professors and associate professors.

[c] The research assistant role includes the research coordinator. If a participant indicated that their role was both research assistant and PhD student, they were counted as a research assistant.

[d] The research manager role includes project officer, program manager, lab manager, and assistant director, research officer, and research administrator.

[e] The lead researcher role includes scientist, co-investigator, and staff scientist.
[f] The graduate student role includes PhD graduate students.

Approaches to Recruitment and Data Collection in Studies with Suspected Fraudulent Participants

Findings revealed that 45.9% of HCRs reported suspicions of fraudulent participants in qualitative studies only, 24.3% in quantitative studies only, and a further 29.7% in either mixed-methods, qualitative and/or quantitative studies (Table 2). Most HCRs suspected fraudulent participants in a single study. Facebook, email, Twitter, and online newsletters were the most used online methods/platforms. In recruitment communications, incentives were mentioned in approximately 84% of cases, and data collection methods were provided in about 70% of communications. Online surveys were the most widely employed data collection method and email was the predominant method for expressing interest in participation.

Table 2. Research approaches in studies with fraudulent participants.

	Survey Participants (n=37)		Interview Participants (n=13)	
	n	%	n	%
Number of studies with suspected fraudulent participants				
One	24	64.9	7	53.8
Two	6	16.2	2	15.4
Three	6	16.2	3	23.1
Four	0	0.0	0	0.0
Five	1	2.7	1	7.7
Online method or platform used to recruit participants				
Facebook	26	70.3	9	69.2
Email	24	64.9	8	61.5
Twitter	24	64.9	7	53.8
Online Newsletter	21	56.8	8	61.5
Instagram	15	40.5	4	30.8
LinkedIn	9	24.3	3	23.1
Paid Ads (e.g., Google or Facebook ads, etc.)	9	24.3	3	23.1
Other	15	40.5	6	46.2
Information included in recruitment communications				
Mention of incentive	31	83.8	12	92.3
Data collection method	26	70.3	11	84.6
All inclusion criteria	24	64.9	10	76.9
Some inclusion criteria	14	37.8	2	23.1
Link to an online survey	14	37.8	5	38.5
No incentive provided	6	16.2	2	15.4
Other	2	5.4	0	0.0
None of the inclusion criteria	1	2.7	0	0.0
Data collection methods				

Online survey	24	64.9	9	69.2
Virtual interview	22	59.5	9	69.2
Phone interview	11	29.7	5	38.5
Virtual focus group	8	21.6	3	23.1
In person interview	3	8.1	1	7.7
Paper and pencil survey	3	8.1	2	15.4
In person focus group	1	2.7	1	7.7
Methods used by participants to express interest in participating				
Email	30	81.1	12	92.3
Link to online survey	19	51.4	7	53.8
Phone	13	35.1	6	46.2
Social media message	6	16.2	4	30.8
Other	3	8.1	0	0.0
Research Methodology used when fraudulent participants were discovered				
Qualitative	17	45.9	5	13.5
Quantitative	9	24.3	3	23.1
Both Qualitative and Quantitative	11	29.7	5	13.5

In terms of incentives provided to study participants, most HCRs (71%) offered physical or electronic gift cards, while less than 1% offered the chance to win a prize such as Apple AirPods. Gift card values ranged from $5 CAD to $100 CAD, with draw prizes having higher values (e.g., AirPods valued at $250 CAD). Some HCRs employed a staggered compensation approach, offering gift cards for survey completion and higher-value compensation for participation in interviews or focus groups. One HCR allowed participants to choose between a gift card or an e-transfer for compensation.

Fraudulent Participation Tactics

HCRs speculated about the tactics employed by fraudulent participants to participate in research studies. One HCR speculated that fraudulent participation may be linked to the use of Artificial Intelligence (AI) technology that scanned the Internet for recruitment advertisements including compensation indicators, such as the dollar sign ($). Another HCR recounted an incidence she described as “scary”, when the research team received 17 fraudulent emails shortly after their partner organization sent a recruitment email to their own listserv. This participant reasoned that the listserv had somehow been accessed and infiltrated by bots.

Individuals attempted to participate fraudulently by employing various tactics, such as falsely claiming to reside within the study catchment area and asserting eligibility criteria related to race, specific healthcare conditions, or certain professional designations. For example, one HCR shared their observations about participants who pretended to be both a youth and a parent, representing two distinct target populations within the same study.

HCRs suggested that fraudulent participants may have searched for information online prior to their interview to convincingly respond to interview questions, such as familiarizing themselves with the responsibilities of an eligible individual, like a prenatal nurse. During some interviews, unusual pauses and typing sounds were observed, suggesting to one HCR that fraudulent participants may have been searching for answers online during the interview. An HCR described their experience: *"There was another thing to where we noticed that sometimes there would be pauses and typing*

sounds on the computer. So, it's almost like the person is Googling things or, you know, uncertain about certain things. It's like you shouldn't need to Google the city you're in."

Suspicions and Discovery of Fraudulent Participants

HCRs encountered various unusual behaviors and circumstances that raised suspicions of fraudulent participation across different study methodologies. Notably, suspicions were most common during online surveys (67.6%), virtual video interviews (62.2%), phone interviews (32.4%), and virtual focus groups (24.3%). In email communications, HCRs observed discrepancies between participant names and email addresses, unusual timing of emails (e.g., early morning) or receiving a batch of emails at the exact same time or within a short timeframe. Some emails lacked the typical contextual information, like how participants learned about the study or providing their eligibility, while other emails were excessively lengthy, nonsensical, and included the use of profanities and explicit sexual content in survey responses.

HCRs also suspected or discovered fraudulent participants when participants used different names within a string of email interactions (such as using the name, Nancy, in the first email, and the name, Sarah, in a subsequent email) and stereotypical Western names like John Smith, or the names of famous people (e.g., Britney Spears). Multiple signatures of different names on the same consent form and an unusually high representation of marginalized groups raised doubts about participant authenticity. HCRs noted other peculiar participant behaviors such as responses that would seem to suggest that the participant had a complete lack of understanding or familiarity with the study topic or stating that one's city of residence was "Ontario", which is a province, not a city, in Canada. Background laughter and statements like "oh, forget it" and logged off further reinforced the impression that these individuals were not fully committed to the research, behavior that, according to HCRs, is atypical from an individual from the target population.

HCRs grew skeptical of participants' authenticity when participants would not disclose their mailing address, despite researchers' explanations that a mailing address was required to receive the study compensation. HCRs also raised concerns about the identity of participants in instances where they refused to appear on camera during interviews. This behavior raised suspicions when participants did not possess the expected vocal characteristics associated with their claimed demographic, such as one participant claiming to be a child but having a voice that sounded like that of an adult. One HCR who encountered fraudulent participants who provided survey responses that didn't make sense, were unrelated to the question, or completed questions in an unusually quick time, was prepared for the possibility of fraudulent participants by using a recruitment website that allowed for the tracking of the amount of time a respondent takes to complete survey questions or activities.

Several participants spoke about detecting fraudulent participants due to surges in emails or completed surveys after the use of Facebook[6] and X[7] in particular, and at times, LinkedIn,[37] to publish a recruitment ad. These surges caused some HCRs to stop using these social media platforms to recruit. One participant was aware of the possible issues with social media platforms therefore avoided their use altogether: *The worst we saw was not LinkedIn, although LinkedIn also attracts [fraudulent participants], but we noticed spikes every time we would put it on Facebook or Twitter. In fact, the cannabis study, I stopped advertising on Twitter and Facebook because we were getting nothing that was eligible, and we were just bombarded with ineligible responses.*

Surges in fraudulent participants were not limited to Facebook and Twitter. Paid-for services such as Honeybee Health[38] (a digital clinical trial recruitment platform), Qualtrics[39] (an online survey platform), and paid-for ads in Facebook did not prevent the infiltration of fraudulent participants. HCRs also spoke about puzzling circumstances such as emails from several respondents within a short timeframe and these respondents attempting to schedule their interview quickly, which, according to one participant, is unlikely considering that prenatal nurses are very busy and typically require some time to schedule an interview. One HCR encountered inconsistencies related to participants' birthdates in a longitudinal study. As part of the study, participants were asked to

provide their birthdate in multiple surveys over time. To the HCR's surprise, some participants provided different birthdates across these surveys.

Two HCRs shared their notable experiences with participants who raised questions regarding the details of compensation. Specifically, participants inquired about whether the compensation would be in the form of a gift card and whether it would be sent via mail or electronically, which according to the HCRs, was unusual. Some participants expressed a strong insistence on receiving an electronic gift card from specific companies, such as Amazon, even when the HCRs were offering gift cards from different companies like Tim Horton's or Starbucks, or when no gift card was being provided as compensation at all. Lastly, participants provided responses to open-ended questions that ultimately had HCRs question the validity of the responses. *"And then of course, reading their qualitative responses because some of the questions…were open-ended, and the responses also just did not sound like responses we might have gotten from a real parent."*

Treatment of Fraudulent Participants and Their Data

Once research teams suspected or identified fraudulent participants, they implemented various actions to address the problem. One common action was consulting and discussing participant concerns with other research team members, including decision making around the appropriate handling of data collected from suspected fraudulent participants and additional verification processes. For instance, to assess participant eligibility, some HCRs conducted follow-up phone calls with suspected participants to ask additional screening questions with hopes that research teams could get a better sense of participants' eligibility. Some researchers requested proof of identity, such as photo identification or professional license number, to verify participants' identity and adherence to inclusion criteria. One HCR described that they used a screening protocol to classify responses from suspected fraudulent participants: "*We coded all responses as 'complete' or 'bad' (ineligible); we also had a 'PARTIAL' status for those who did not complete the questionnaire before the study was closed"*.

While most HCRs informed suspected fraudulent participants of their ineligibility, some HCRs chose to cease all communication, including not responding to emails, scheduling interviews, or sending survey links. Additionally, all HCRs excluded data from fraudulent participants in their data analysis, often storing data in a separate file. Two HCRs expressed interest in conducting a secondary analysis or separate study using the collected data from fraudulent participants with the intention of deepening their understanding of this emerging issue and facilitating open discussions among other HCRs regarding their experiences with fraudulent participant encounters.

Strategies Used to Prevent Fraudulent Enrollment, Identify Fraudulent Participants, and Verify their Identity

HCRs explored various strategies to prevent the fraudulent participation of individuals in their studies and safeguard the integrity of their findings. While some HCRs implemented specific measures, others reflected on lessons learned and discussed strategies they would employ in subsequent studies. See Table 3 for a detailed list of these strategies.

Table 3. Indicators leading to suspicion of fraudulent participants in healthcare research.

Key Areas	Fraud Indicators	Supporting Quotes
Identity Irregularities	• Unusual names, generic names in line with stereotypical Western names, or names of famous people (e.g., John John, Britney Spears). • Names do not match the names in their email address. • Different names used throughout	Consent Forms *And I think of the 200, 270 respondents a hundred came back with the signed consent form, but within that as well, we had multiple signatures, the same signatures, different emails on*

	the email thread. • Different names used in email interactions. • Different birthdates provided across longitudinal surveys, e.g., change in birth date and year. • Multiple signatures on consent forms.	*that same, so different people were sending me the same consent form but the different email address*". – *Research Administrator*
Unusual Participant Behaviors	• Attempts to negotiate the type of incentive and/or the method of incentive delivery. • Fails to provide mailing address despite it being required for incentive delivery. • Attempts to schedule the interview very quickly, contrary to typical behavior of busy healthcare professionals. • Demonstrates poor knowledge about the study topic. • Exhibits a lack of interest in the interview. • Does not turn on camera during interviews.	Negotiations *"They also asked if I could instead, we promised them a Tim Horton's and Starbucks gift card and they said that those are not, we don't want, I don't want those, we, I get an Amazon gift card instead". – Research Administrator* Schedule Interviews *"And my research coordinator noticed that they were responding very, very quickly to schedule the interview, which was quite strange because typically healthcare professionals are really busy and it's like, it's a long time to get responses from them. And then when they actually have time to schedule the interview, it's like a week or a few weeks in advance". – Associate Professor* Lack of interest *"But just in speaking with our research coordinator about the interactions that she's had with people, it kind of like- the way that they're acting, sort of saying funny things, joking around, laughing in the background. Sometimes she can hear more than one person there and then they'll just be like, oh, forget it and log off. It really has this kind of Halloween vibe". – Assistant Professor* Camera's Off

		"Then when we get to the interview with people, they'll leave their cameras off during interviews. They will say that they're child with cancer, but they have obviously an adult voice, but their camera is off." – Assistant Professor
Location Discrepancies	• Location of residence does not match participant's postal code, IP address, and/or phone numbers.	Location Data *"And then when you match up the location and the postal code, sometimes there's a mismatch in terms of they say that they're in Toronto, but then the post code starts is the V, which is in out of Vancouver". – Research Coordinator* *"So, people would say, oh, I'm in Ontario, but their IP address was in Texas or their IP address was in Nigeria or something like that. So that was another method that we were able to use". – Research Officer*
Suspicious Email Communication	• Provision of temporary email addresses. • Inconsistencies between participant's name and name used in the email address. • Surge of emails in a very short timeframe. • Clustering of emails around the same time or in close succession. • High volume of emails with a similar address structure, such as first name, last name, followed by numbers. • Receipt of emails at unusual times of the day, such as during the middle of the night. • Presence of near-identical wording across emails from different email addresses. • Unusual brevity and directness in emails, lacking contextual information, such as where the participant saw the advertisement or how they meet the study criteria, and minimal salutations or	Unusual brevity *"…the content of the email that doesn't really seem to jive, lots of typos, this type of thing. The email being very quick and not, doesn't really seem to address the point of the study". – Assistant Professor* Email surges *"And then we realized there would be a huge influx in emails at certain hours of the day, and it was typically 2:00 to 3:00 AM and all the email addresses were structured the exact same way. So, it would've been a generic first name, last name, so Brian James or things like that. And then a series of numbers @gmail.com and everything was the exact same." – Research Officer*

	greetings. • Lengthy and convoluted emails, or emails containing gibberish content (e.g., Chinese characters with no logical meaning). • Use of different names by the participant throughout the email correspondence.	Different names *"And then we started to* realize *sometimes they would forget to change the name or in one email they would say, my name's Kristen. And then in the follow up they would say Nancy, for example. So, there were just these little inconsistencies". – Research Officer*
Suspicious Responses	• Use of terminology not commonly used by legitimate participants, such as outdated or inappropriate language. • General and nonspecific responses that lack the expected level of detail from individuals within the target population. • Responses that do not align with the typical patterns of communication or knowledge expected from individuals in the sample population. • Consent forms with similar labels or multiple signatures with different names. • Responses that are unrelated to the question being asked. • Inappropriate responses, including the use of profanity or sexual references. • Multiple survey questions left unanswered with no contact information provided. • Signs of disinterest or lack of engagement during the study interview.	Uncommon terminology *"'Tribe' is a very outdated term to use for describing Indigenous communities". – Research Assistant* Unanswered survey questions *"We have people clicking on the link to the survey and it seems as if they're stopping at the yes and no questions. So, we have two yes and no questions in the beginning and they would just answer yes and then stop after that. It's just nothing. It's just blank. So, there's no email, there's no address, there's no way of reaching out to them." – Research Administrator*
Unlikely Circumstances	• Disproportionate number of participants self-identifying as belonging to rare or marginalized groups.	Rare or Marginalized Groups *"So to have that many people who were identifying as black and Indigenous, both, and also all in this age group of 30 to 39, and also parents...So to get kind of young parents that all had these identifiers, something just kind of did not sit right". – Assistant Professor*

One key approach to preventing the participation of fraudulent participants involved incorporating additional security features, such as CAPTCHA, into survey platforms to deter fraudulent participation. CAPTCHA stands for Completely Automated Public Turing test to tell Computers and Humans Apart.[40] It is a widely used method that helps verify the authenticity of participants by requiring users to complete a task or answer a challenge that is easy for humans to perform but difficult for automated computer programs (bots) to solve. HCRs also engaged in investigation of the authenticity of participants, whether prompted by their suspicion or as a preemptive strategy, by employing other survey platform features to track IP addresses, geolocation, latitude and longitude, and participants' postal codes, when they discovered that geographic markers/indicators did not match the participants' stated location of residence. As one Research Coordinator stated: *"And then when you match up the location and the postal code, sometimes there's a mismatch in terms of they say that they're in Toronto, but then the postal code starts is the V, which is in out of Vancouver."*
Other strategies were used to discourage fraudulent participation, including selectively using social media platforms or groups instead of advertising broadly and publicly, avoiding specific symbols or words like the dollar sign that could be detected by AI systems, and refraining from explicitly mentioning incentives in recruitment advertisements. As stated by an Assistant Professor from a Canadian University, *"…we removed the survey link from study advertisements to prevent any participant from filling it out without eligibility."*
When emails were used to communicate with potential participants, HCRs analyzed emails for specific patterns indicative of fraudulency such as duplication (multiple survey entries from the same email address and/or duplication of text within emails even if addresses were different) and their content to ensure the language matched researchers' expectations. Once fraudulent participation was suspected, some HCRs introduced additional requirements for participants to provide personal information or verify their identity, such as full names, photo identification (ID), mailing addresses, or professional/institutional email addresses, as an additional deterrent against fraudulent participation.
HCRs also highlighted the need for comprehensive screening protocols that go beyond eligibility screening and are integrated into the study protocol, grants, and Research Ethics Board (REB) applications. They underscored the importance of attentiveness to participants' verbal and non-verbal cues to assess their sincerity, genuineness, and ultimate authenticity. Concurrent analysis of data was also recommended to identify the potential for fraudulent participation to mitigate further issues with recruitment and data collection and safeguard data integrity.
When fraudulent participants were successful at passing eligibility criteria, some HCRs still requested participants to verify their identity. When participants refused to turn on their camera, this was often seen as a red flag. *"[W]e did at the beginning, and we did it with some people that pushed us a little bit 'cause there were a few that pushed. And as soon as we said that they had to turn on their camera and show photo ID, they never followed through"*.

Motivation for Participating Fraudulently

HCRs discussed the potential motives behind individuals choosing to partake in research despite not meeting the specified criteria for inclusion. Several participants emphasized the role of incentives as a significant driving force. One researcher expressed astonishment at the discovery of individuals purposefully participating in a fraudulent manner, fully aware that the receipt of the $100 gift card was not guaranteed. It was surprising to the researcher that these individuals would willingly invest their time in engaging in online interviews and completing surveys, all for the mere possibility of receiving the gift card. Another participant shared her experience of a participant who expressed intentions to encourage every member of her immediate family to participate in the project, even though her mother did not meet the inclusion criteria. The researcher went on to contemplate the benefits and challenges associated with offering financial incentives to participants. On the one hand, providing compensation demonstrates to participants that researchers value their time and effort.

However, it may inadvertently encourage individuals to misrepresent themselves to obtain compensation. *“I think the incentives are always a problem and a blessing. You do want to compensate people for time, but at the same time, you don’t know people’s motivations for why they do certain things and what their needs are. And especially during these times of inflation. And for somebody, [a] $30 gift card to a grocery store could make a difference in one week depending on the type of nutrition they’re going to consume.”*

The use of online platforms to recruit participants offers a level of anonymity and convenience that may not be readily achievable in traditional in-person recruitment and research settings. This aspect was highlighted by a healthcare researcher who explained how online participation allows individuals to engage in interviews without the need to turn on their camera. Additionally, participants can easily self-screen by simply clicking “a button”, rather than being screened by a researcher who might uncover their ineligibility to participate.

Participants speculated about the potential occurrence of fraudulent participation driven by personal amusement. One healthcare researcher shared about their research team's encounters with individuals displaying peculiar behavior during interviews, such as hanging up in the middle of the interview, joking around, making funny remarks, and background laughter. The researcher likened these situations to Halloween or prankster activities, such as *"knocking on people's doors and running away or toilet papering someone's house."*

Ethical and Practical Challenges of Fraudulent Participants

Identifying fraudulent participants presented significant resource challenges, especially in studies with high participant interest. An HCR highlighted the extensive resources needed to identify and exclude fraudulent participants from their study: *"We used LinkedIn to recruit, and within the first week of posting this ad, making it live, I got about 200 emails from potential participants, and I couldn't distinguish between imposters, between bot robots, between fraudulent participants and actual participants. So, we had to go back to the drawing board to decide how to deal with this situation because obviously we couldn't- our sample size was 45, we had 200 people responding to the ad, so we had to narrow it down somehow."* This issue not only caused delays in study timelines but also placed a strain on study budgets. Research teams had to allocate additional resources to effectively manage this challenge, posing a significant setback for studies with limited funding.

The presence of fraudulent participants posed a considerable source of stress for research teams, with several HCRs expressing limited knowledge of best practices on how to effectively address this issue. This was particularly pronounced in studies involving participant interviews. One participant noted the predominant focus of existing literature on preventive strategies applicable to quantitative research, suggesting a potential hesitation or lack of awareness of qualitative researchers to openly discuss these issues. The lack of transparency surrounding this problem contributed to the limited guidance available for research teams facing similar challenges.

Researchers often faced an ethical dilemma regarding the compensation of individuals identified as fraudulent participants. Some researchers chose not to provide an honorarium to these participants since they provided invalid data, particularly when survey responses indicated a high presence of bots. However, certain researchers decided to give an honorarium to fraudulent participants to preempt potential legal issues. Justifying this decision, one HCR explained that their REB protocol required providing an honorarium to all participants, including those identified as fraudulent. Nevertheless, the issue of compensating ineligible individuals was a subject of debate among HCRs, with frustrations expressed regarding diverting funds from individuals with genuine lived experiences and/or in need of financial support. As one participant highlighted, *"…it’s taking away money from people with lived experience who really need this money and research funding is limited. It's not like I have unlimited funds."*

HCRs raised another ethical concern about the potential invasiveness of verifying eligibility. While

requesting proof of identification could help minimize the risk of enrolling fraudulent participants, researchers also struggled to justify requesting proof of identity in their REB protocols, as obtaining participants' personal information did not directly contribute to the study beyond verification purposes. HCRs also viewed a stringent screening process as potentially insensitive, difficult to access, and discouraging for prospective participants. An HCR recruiting individuals with disabilities expressed concerns about the potential harm involved stating, *"But it feels a little bit like I'm asking people to confess that they have a disability, and they have to prove it to me that they're disabled. And that is such a hurtful and potentially harmful thing, right?"* Furthermore, HCRs encountered challenges in distinguishing between fraudulent participants and eligible ones, leading to feelings of doubt and guilt. The subjective nature of relying on a researchers' intuition further complicated the identification process, making it difficult to determine which participants should be included or excluded from a study. The limited interaction between HCRs and participants recruited from online platforms added another layer of complexity in understanding participants' true intentions.

Discussion

Study findings highlight the pervasiveness of fraudulent participation[12,16] across healthcare studies involving online methods, regardless of the research methodology, recruitment methods, social media platforms, incentives, and data collection techniques used, underscoring the complexity and ubiquitous nature of this problem. While online methods and platforms, including email, project websites, and social media platforms like Facebook, Instagram, X, and LinkedIn, offer several benefits, such as convenience,[3,4,14,16,20] cost-savings,[2–4,14,15,18–20,22] and the opportunity to reach diverse populations[3,13–20] they also expose HCRs to risks related to fraudulent participants.[2,24,25] These risks encompass the greater difficulty in determining eligibility,[3,4,14] collection of false data, which can significantly alter results and invalidate findings,[3,12,15,17,27] leading to inappropriate and harmful applications,[17] as well as the waste of time, funding, and human resources.[16,27] Furthermore, our study highlights the growing challenge of AI and bots in healthcare research involving online recruitment methods, and HCRs' relatively limited understanding of the various technologies and their capabilities. For instance, circumstances reported by HCR participants that were attributed to AI, may, in fact, have been carried out by web crawling bots (i.e., spiderbots) scanning the Internet for recruitment ads.[41–43] Despite the significant challenges posed by fraudulent participation, the use of online methods and platforms for recruitment offers substantial benefits[1,2,4,12–19,21,24] that typically outweigh the associated risks. Therefore, researchers using online strategies to recruit must proactively develop a comprehensive protocol to prevent and detect fraudulent behavior (e.g., Lawlor et al's[11] REAL framework approach to addressing survey fraud).[2,11] Such a protocol, detailed in the recommendations section below and summarized in Table 4, will help ensure the integrity and validity of the research by effectively addressing the challenges posed by fraudulent participants. It is important to note that these recommendations are based on the experiences of a limited number of HCRs.

Table 4. Recommendations for preventing and addressing fraudulent participation in healthcare research involving online recruitment methods.

Key Areas	Recommended Measures
Targeted Recruitment	• Implement targeted recruitment strategies, such as posting study advertisements in closed or topic-specific groups on social media platforms. • Establish partnerships with relevant organizations that have access to eligible research participants and recruit directly from those groups. • Disclose only essential eligibility criteria in study

	advertisements, providing enough information to attract individuals from the target population. • Do not include the survey link in study advertisements. Instead, require potential participants to contact the research team for further information and screening. • Avoid using research or incentive-related keywords in recruitment advertisements. • Replace symbols in email addresses with their spelled-out equivalents, such as replacing the at sign (@) with the word spelled out to mitigate the risk of automated systems harvesting email addresses for fraudulent purposes. (e.g., researchlab at utoronto dot ca).
Identification and Verification	• Collect comprehensive participant information, including full name, phone number, mailing address, photo identification, and/or professional email addresses. • Use direct communication methods, such as phone or video, for participant screening rather than email-only screening. • Observe participant behavior during the screening process, including factors such as response time and non-verbal cues. • Pay attention to participants' confidence levels and consistency in their responses. • Implement a requirement for participants to briefly appear on camera during virtual interviews for identification purposes.
Email Communication	• Use email verification services to identify and scrutinize temporary email addresses. • Monitor and track duplicate email addresses (same email address used by multiple interested participants). • Examine the format and structure of email addresses, paying attention to generic names followed by a seemingly random combination of letters and/or numbers. • Evaluate emails for coherence and correct syntax, ensuring that the email content aligns with the typical language and syntax patterns used by the sample population.
Incentives	• Minimize disclosure of financial incentives (type and value) in recruitment materials. • Avoid the "$" sign in recruitment materials. • Refrain from including research or incentive-related keywords or hashtags in recruitment advertisements and on social media.
Survey Platform Tools/Features	• Use a verification service like CAPTCHA or TransUnion's TLOxp[3] to verify human participation, identity, and deter automated bots. • Include honey pot questions (i.e., questions that only

	bots can see) allowing for the identification of potential fraudulent activity when a bot responds to these questions. • Track IP addresses, geolocation, or latitude and longitude data to detect suspicious or inconsistent participant locations.
Question Formatting	• Include check questions/attention questions (e.g., at the end of a question, instruct participants to choose option C in a multiple-choice type question). • Include open-ended questions to assess participants' knowledge and gauge their familiarity with the research topic. • Pose specific questions that only legitimate participants would be able to answer accurately.
Post-Survey Checks	• Analyze when surveys are completed, paying attention to those completed during unusual hours. • Note time spent completing surveys, to compare to average completion time and identify abnormally fast or slow responses. • Scrutinize a surge of surveys completed at approximately the same time.
Analysis of Survey Responses	• Monitor incoming data, such as IP addresses and completion locations, to identify unusual patterns. • Conduct concurrent data collection and analysis when feasible to identify fraudulent participants early and take mitigation measures. • Select and review survey responses to detect any inconsistencies or suspicious behavior at several points during data collection

These recommendations provide HCRs with strategies that may be beneficial in preventing or identifying fraudulent participants when using online methods and platforms for recruitment. However, not all recommended measures may be suitable for all studies and should be considered within the context of the study goals and how each may impact their expected results. Furthermore, HCRs should also be cognizant that although strategies may deter some fraudulent participants or bots, they are not full-proof as individual scammers and technology are rapidly changing and advancing (e.g., closed Facebook groups may contain fraudulent members, participant voices in phone calls or video calls make be “deep fakes” (AI-generated faces and voices)[44], avoiding symbols (e.g., $ and @) may only deter simple bots, tracking IP addresses may be limited due to the increased use of virtual private networks (VPNs)).

Our study sheds light on the various peculiar and irregular behaviors that indicate the presence of fraudulent participants in research. HCRs, however, cannot rely on any single indicator of fraudulent behavior to determine the existence of fraudulent participation. Instead, they must be aware of and consider several behaviors and circumstances simultaneously, which together, are a good indication of the presence of fraudulent participants. These behaviors include the use of unusual or very generic Western names, discrepancies between names and email addresses, attempts to negotiate incentive type or delivery, perceived lack of interest in the interview, use of temporary email addresses, surges in emails at unusual times, and inappropriate or outdated terminology, among others. HCRs should also be aware that the absence of incentives or their perceived low value (e.g., $5 Amazon gift card) from recruitment materials may not deter individuals from engaging in inauthentic behavior to obtain

compensation.[45] This is particularly true if a fraudulent participant works for or is part of a sophisticated operation that employs AI to systematically search the Internet for research studies, aiming to accumulate incentives across multiple studies.[45] Our study underscored this phenomenon, as one HCR speculated on the involvement of such operations. Additionally, in a 2019 blog post, the founder of a market research company highlighted the discovery of a website specifically designed to train individuals in fraudulently completing large volumes of online surveys.[46] It is important to note that while an operation using AI may indeed be more sophisticated, this example may not capture the full spectrum of fraudulent activity. Web crawlers (spiders or spiderbots) can search for information like studies without advanced AI capabilities.[41–43] Furthermore, individuals may also employ manual methods to scour the web in pursuit of accumulating incentives.

HCRs experience ethical conflicts for being too stringent and invasive in participant screening processes, financially compensating fraudulent participants, and potentially excluding legitimate participants from studies. It is recommended that all ethical considerations be outlined and addressed in the study's REB application and consent form[47] to ensure transparency regarding the steps taken to minimize this issue. In REB applications, researchers need to indicate the amount and type of personal information obtained from participants as part of the screening process, how teams will screen for fraudulent participants, and whether an incentive will be provided to those identified as fraudulent. Additionally, research teams should consider stating in their study letter that any participant who is discovered to be providing false or misleading information about their identity may forfeit the incentive, despite their time and/or effort contributed to the study. By considering the diverse tactics used by individuals attempting to participate fraudulently in research and recognizing the significance of incentives regardless of their perceived value, researchers can address the ethical dilemmas uncovered in this study.

Recommendations

Comprehensive Prevention Strategy

As demonstrated by Lawlor and colleagues,[11] to effectively prevent and address fraudulent participation in research studies, it is crucial to implement a comprehensive strategy that encompasses prevention, identification, and response measures (Table 4). This strategy should involve the utilization of various security features available on survey platforms to mitigate the risk of fraudulent enrollment. Incorporating tools such as CAPTCHA[3,40] and other security features acts as an initial defense against computer bots attempting to gain access to research studies.[2,4,12,27] It is important to note that while these security measures are valuable, they may not provide absolute protection against fraudulent participation as technology continues to evolve.[1,27]

To minimize the attraction of fraudulent participants, researchers can employ a targeted online recruitment approach in addition to security measures.[27,48] This approach involves avoiding the explicit publication of incentives,[1,17,25,48] not using some symbols such as the dollar sign, and publishing recruitments ads in closed social media groups rather than public ones.

Furthermore, when using online recruitment methods, healthcare researchers should operate under the assumption that some fraudulent participants may successfully bypass security features. To effectively identify such participants, it is recommended to use multiple strategies in the survey design. This includes incorporating honey pot questions, which are hidden questions detectable only by bots,[27,49,50] attention check questions that individuals are required to answer as a way to distinguish between high and low quality data,[19,20,27] and employing open-ended questions to assess the genuineness of responses.[2,49,51] Researchers can also leverage additional online survey features, such as tracking IP addresses, which can be cross-referenced with other data points, such as postal or zip codes, to verify participant eligibility and detect duplication.[18,25,52] By monitoring survey entries for time stamps, researchers can identify clustered entries and assess the time taken to

complete surveys compared to the average response time.[2,24,45]

In addition to survey design considerations, it is crucial to pay attention to key email markers that may indicate fraudulent participation. These markers include duplicate addresses, suspicious email formats (e.g., name1234@gmail.com) and temporary emails.[12,16,27] It is worth noting that screening for fraudulent participants requires considerable time and resources. To streamline the identification process, healthcare researchers may opt to categorize survey entries into three distinct categories: authentic, suspicious, and fraudulent entries.[52] This categorization approach can be particularly useful when dealing with large volumes of surveys, allowing researchers to focus their attention on reviewing entries that raise suspicion.

Communities of Practice in Academia

In addition to practical benefits, a community of practice focused on research strategies and ethical considerations when employing online recruitment and data collection methods, provides a supportive environment where researchers can openly share their experiences and learn from one another. A community of practice is a group of people who share a common concern, a set of problems, or a passion for a particular topic.[53] Through continuous interactions, the group engages in activities that facilitate the deepening of their knowledge and expertise in that area.[53] This collaborative approach not only supports acknowledging the existence and pervasive nature of fraudulent participation but also fosters a culture of collaborative learning, sharing, and continuous improvement. By engaging in ongoing discussions and interactions with colleagues in similar situations, HCRs can refine their prevention strategies, adapt to emerging trends, and contribute to the integrity and validity of healthcare research in the digital age.

By establishing a community of practice, researchers can bridge the gap between research and practice, allowing for timely access to practical knowledge and insights, by sharing experiences, exchanging information, and co-creating effective prevention strategies to stay ahead of the ever-evolving tactics employed by fraudulent participants.[53–55] Furthermore, a community of practice serves as a mechanism to collaboratively address the urgent need for conducting research with high-quality data that is free from the influence of fraudulent participation. This proactive approach not only ensures the credibility and reliability of healthcare research but also fosters a culture of continuous learning and improvement within the academic community.

Limitations

The present study has several limitations. First, the data collected for this research is limited to Canada only. While this provides valuable insights into fraudulent participation in the Canadian context, it may not fully capture the variations and complexities of fraudulent behaviors in different countries and cultural settings. Second, it is important to note that these recommendations stem from a pilot study involving a limited number of HCRs. Therefore, while these suggestions can be valuable for researchers conducting similar studies, the generalizability of the recommendations is limited to general population recruitment. Furthermore, we did not explicitly collect data on the HCRs level of expertise with online recruitment strategies nor data describing the samples that the HCRs were recruiting, such as whether they were recruiting health professionals, patients, or the general population. This is an area that should be explored in future research with a more extensive design to explore these aspects in greater depth to support the potential approaches and adaptations during the recruitment process. Similarly, specific data was not collected on whether participants were describing second-hand accounts reported to them by research staff or if they were involved first-hand in the day-to-day detection of fraudulent participants. Future research with a larger sample size may delve deeper into these nuances. Third, due to the inherently narrow scope of our pilot study, there is a need for studies of a larger scope to other countries, scientific fields, and larger sample sizes to gain a more comprehensive picture of fraudulent behaviors and the strategies to prevent, deter, and identity fraudulent participants. In addition, collaborating with Information Technology (IT) and security specialists would help build effective mitigation strategies. And last, this study does not include the direct input from fraudulent participants themselves. While our study

findings provide the reader with insights into potential motivations for fraudulent participation based on the experiences of researchers, not having direct access to the perspectives of these participants may limit the depth of understanding regarding their underlying motivations and tactics. Future research that incorporates the voices of fraudulent participants could provide valuable insights and enhance the development of more targeted prevention and mitigation strategies in studies involving online recruitment methods and platforms.

Conclusion

In conclusion, this study sheds light on the complex and pervasive problem of fraudulent participation in healthcare research when online recruitment methods are employed. The findings emphasize the need for healthcare researchers to be vigilant and proactive in identifying, preventing, and responding to fraudulent behavior. To address this challenge effectively, HCRs must go beyond relying on intuition and subjective methods that may introduce bias. Instead, they should develop and implement several prevention, verification, and mitigation strategies at once and not rely on post-hoc measures to verify participant data. Additionally, researchers should stay informed of the ever-changing landscape of the Internet, technology, and methods used by fraudulent participants to bypass safeguards. By understanding that the driving force for deception may be the prospect of gaining a financial incentive regardless of its value, and recognizing the diverse tactics employed by fraudulent participants to gain compensation, researchers can be prepared to encounter and effectively manage this problem by developing and implementing robust prevention and management strategies. HCRs should be encouraged to document the evidence of fraud within their studies, providing sufficient information for other researchers to become aware of the potential dangers when using online research methods and to establish their credibility and protect the integrity of their research. To address the challenge of fraudulent participation in a timely manner, researchers can establish a community of practice to access timely solutions and support, enabling them to address the problem more effectively and to conduct quality and ethically sound research. By taking collective action with other researchers and staying informed, researchers can safeguard the integrity and validity of healthcare research in the digital age.

Acknowledgements

We thank the Nursing Research Interest Group (NRIG) for funding this study through the NRIG Novice Research Grant Award. We also acknowledge the contributions of research assistants Ossaid Ahmed and Carlie Lukasik.

Conflicts of Interest

None declared.

Abbreviations

AI – artificial intelligence
CAD – Canadian dollars
CAPTCHA – completely automated public Turing test to tell computers and humans apart
HCR – healthcare researcher
IP – internet protocol
IT – Information Technology
REB – research ethics board
REDcap – research electronic data capture
SD – standard deviation
VPN – virtual private network

References

1. Griffin M, Martino RJ, LoSchiavo C, et al. Ensuring survey research data integrity in the era of internet bots. *Qual Quant.* 2022;56(4):2841-2852. doi:10.1007/s11135-021-01252-1
2. Pozzar R, Hammer MJ, Underhill-Blazey M, et al. Threats of bots and other bad actors to data quality following research participant recruitment through social media: Cross-sectional questionnaire. *J Med Internet Res*. 2020;22(10). doi:10.2196/23021
3. Glazer J V., MacDonnell K, Frederick C, Ingersoll K, Ritterband LM. Liar! Liar! Identifying eligibility fraud by applicants in digital health research. *Internet Interv*. 2021;25. doi:10.1016/j.invent.2021.100401
4. Bybee S, Cloyes K, Baucom B, Supiano K, Mooney K, Ellington L. Bots and nots: safeguarding online survey research with underrepresented and diverse populations. Psychol Sex. 2022;13(4):901-911. doi:10.1080/19419899.2021.1936617
5. Instagram. Instagram from Meta. Published 2023. Accessed July 30, 2023. https://www.instagram.com/
6. Meta. Connect with friends and the world around you on Facebook. Accessed July 30, 2023. https://facebook.com/Meta
7. XCorp. Happening now. Join Twitter today. Accessed July 30, 2023. https://twitter.com/?lang=eng-ca
8. Newman A, Bavik YL, Mount M, Shao B. Data collection via online platforms: Challenges and recommendations for future research. *Appl Psychol*. 2021;70(3):1380-1402. doi:10.1111/apps.12302
9. Moss AJ, Rosenzweig C, Jaffe SN, Gautam R, Robinson J, Litman L. *Bots or Inattentive Humans? 1 Bots or Inattentive Humans? Identifying Sources of Low-Quality Data in Online Platforms*.
10. Dennis SA, Goodson BM, Pearson CA. Online worker fraud and evolving threats to the integrity of MTurk data: A discussion of virtual private servers and the limitations of IP-based screening procedures. *Behav Rese in Accounting*. 2020;32(1):119-134. doi:10.2308/bria-18-044
11. 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). doi:10.1177/20597991211050467
12. 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). doi:10.2196/12344
13. Bush J, Blackwell CW. Social Media as a Recruitment Strategy with Transgender-Identified Individuals: Using an Ethical Lens to Direct Methodology. *Journal of Transcultural Nursing*. 2022;33(5):603-614. doi:10.1177/10436596221101928
14. Campbell CK, Ndukwe S, Dubé K, Sauceda JA, Saberi P. Overcoming challenges of online research: measures to ensure enrollment of eligible participants. *J Acquir Immune Defic Syndr (1988)*. 2022;91(2):232-236. www.jaids.com
15. Godinho A, Schell C, Cunningham JA. Out damn bot, out: Recruiting real people into substance use studies on the internet. *Subst Abus*. 2020;41(1):3-5. doi:10.1080/08897077.2019.1691131
16. Heffner JL, Watson NL, Dahne J, et al. Recognizing and Preventing Participant Deception in Online Nicotine and Tobacco Research Studies: Suggested Tactics and a Call to Action. *Nicotine and Tobacco Research*. 2021;23(10):1810-1812. doi:10.1093/ntr/ntab077
17. Hohn KL, Braswell AA, DeVita JM. Preventing and Protecting Against Internet Research Fraud in Anonymous Web-Based Research: Protocol for the Development and Implementation of an Anonymous Web-Based Data Integrity Plan. *JMIR Res Protoc*. 2022;11(9). doi:10.2196/38550
18. Levi R, Ridberg R, Akers M, Seligman H. Survey Fraud and the Integrity of Web-Based Survey Research. *American Journal of Health Promotion*. 2022;36(1):18-20. doi:10.1177/08901171211037531
19. Newman A, Bavik YL, Mount M, Shao B. Data Collection via Online Platforms: Challenges and

Recommendations for Future Research. *Applied Psychology*. 2021;70(3):1380-1402. doi:10.1111/apps.12302
20. Salinas MR. Are Your Participants Real? Dealing with Fraud in Recruiting Older Adults Online. *West J Nurs Res*. 2023;45(1):93-99. doi:10.1177/01939459221098468
21. Reagan L, Nowlin SY, Birdsall SB, et al. Integrative Review of Recruitment of Research Participants Through Facebook. *Nurs Res*. 2019;68(6):423-432. doi:10.1097/NNR.0000000000000385
22. Ellington M, Connelly J, Clayton P, et al. A systematic review of the use of social media for recruitment of participants in nutrition, obesity, and physical activity related studies. *Community and Public Health Nutrition*. 2021;5(Supplement 2):120. https://academic.oup.com/cdn/article/5/Supplement_2/120/6293507
23. Burnette CB, Luzier JL, Bennett BL, et al. Concerns and recommendations for using Amazon MTurk for eating disorder research. *International Journal of Eating Disorders*. 2022;55(2):263-272. doi:10.1002/eat.23614
24. Guest JL, Adam E, Lucas IL, et al. Methods for authenticating participants in fully web-based mobile app trials from the ireach project: Cross-sectional study. *JMIR Mhealth Uhealth*. 2021;9(8). doi:10.2196/28232
25. Teitcher JEF, Bockting WO, Bauermeister JA, Hoefer CJ, Miner MH, Klitzman RL. Detecting, preventing, and responding to “fraudsters” in internet research: Ethics and tradeoffs. *Journal of Law, Medicine and Ethics*. 2015;43(1):116-133. doi:10.1111/jlme.12200
26. Hardesty JJ, Crespi E, Nian Q, et al. The Vaping and Patterns of e-Cigarette Use Research Study: Protocol for a Web-Based Cohort Study. *JMIR Res Protoc*. 2023;12. doi:10.2196/38732
27. Storozuk A, Ashley M, Delage V, Maloney EA. Got Bots? Practical Recommendations to Protect Online Survey Data from Bot Attacks. *Quant Method Psychol*. 2020;16(5):472-481. doi:10.20982/tqmp.16.5.p472
28. Teitcher JEF, Bockting WO, Bauermeister JA, Hoefer CJ, Miner MH, Klitzman RL. Detecting, preventing, and responding to “fraudsters” in internet research: Ethics and tradeoffs. *Journal of Law, Medicine and Ethics*. 2015;43(1):116-133. doi:10.1111/jlme.12200
29. Wang J, Calderon G, Erin R. Hager 3, et al. *Identifying and Preventing Fraudulent Responses in Online Public Health Surveys: Lessons Learned during the COVID-19 Pandemic*.; 2022.
30. Vasileiou K, Barnett J, Thorpe S, Young T. Characterising and justifying sample size sufficiency in interview-based studies: Systematic analysis of qualitative health research over a 15-year period. *BMC Med Res Methodol*. 2018;18(1). doi:10.1186/s12874-018-0594-7
31. REDCap. REDCap: Research Electronic Data Capture. Accessed July 30, 2023. https://www.project-redcap.org/
32. 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. doi:10.1016/j.jbi.2008.08.010
33. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform*. 2019;95. doi:10.1016/j.jbi.2019.103208
34. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse Educ Today*. 2004;24(2):105-112. doi:10.1016/j.nedt.2003.10.001
35. Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs*. 2008;62(1):107-115. doi:10.1111/j.1365-2648.2007.04569.x
36. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nurs Health Sci*. 2013;15(3):398-405. doi:10.1111/nhs.12048
37. LinkedIn. Welcome to your professional community. Accessed July 30, 2023. https://ca.linkedin.com/?original_referer=https%3A%2F%2Fwww.bing.com%2F
38. Honeybee Health. Helping patients find alternative health treatments. Accessed July 30, 2023.

https://www.honeybeetrials.com/
39. Qualtrics XM. Make every interaction an experience that matters. Accessed July 30, 2023. http://www.qualtrics.com/
40. Carnegie Mellon University. CAPTCHA: Telling humans and computers apart automaticaly. Accessed July 30, 2023. http://www.captcha.net/
41. Piehlmaier D. *Bot Detection in Online Studies and Experiments*. SAGE Publications, Ltd.; 2022. doi:10.4135/9781529601312
42. Goodrich B, Fenton M, Penn J, Bovay J, Mountain T. Battling bots: Experiences and strategies to mitigate fraudulent responses in online surveys. *Appl Econ Perspect Policy*. 2023;45(2):762-784. doi:10.1002/aepp.13353
43. What is a web crawler? | How web spiders work. Cloudflare. Published 2023. Accessed October 18, 2023. https://www.cloudflare.com/en-ca/learning/bots/what-is-a-web-crawler/
44. Gow G. The Scary Truth Behind The FBI Warning: Deepfake Fraud Is Here And It's Serious—We Are Not Prepared For An Attack. *Forbes*. Published online May 2, 2021. Accessed October 4, 2023. https://www.forbes.com/sites/glenngow/2021/05/02/the-scary-truth-behind-the-fbi-warning-deepfake-fraud-is-here-and-its-serious-we-are-not-prepared/?sh=142f68253179
45. Brainard J, Houghton MJ, Mumford MS, O'brien SJ. The Wasps are Clever: Keeping Out and Finding Bot Answers in Internet Surveys Used for Health Research. *preprint*. Published online 2022. doi:10.20944/preprints202203.0243.v2
46. Pasternak O. Market research fraud: Distributed survey farms exposed. Published 2019. Accessed July 30, 2023. https://persona.ly/blog/2019/01/market-research-fraud-distributed-survey-farms-exposed/
47. Roehl JM, Harland DJ. Imposter Participants: Overcoming Methodological Challenges Related to Balancing Participant Privacy with Data Quality When Using Online Recruitment and Data Collection. *Qualitative Report*. 2022;27(11):2469-2485. doi:10.46743/2160-3715/2022.5475
48. Mournet AM, Kleiman EM. Internet-Based Mental Health Survey Research: Navigating Internet Bots on Reddit. *Cyberpsychol Behav Soc Netw*. 2023;26(2):73-79. doi:10.1089/cyber.2022.0173
49. Moss AJ, Rosenzweig C, Jaffe SN, Gautam R, Robinson J, Litman L. Bots or inattentive humans? Identifying sources of low-quality data in online platforms. *preprint*. Published online 2021. Accessed July 30, 2023. https://psyarxiv.com/wr8ds/
50. Parks AM, Duffecy J, McCabe JE, et al. Lessons Learned Recruiting and Retaining Pregnant and Postpartum Individuals in Digital Trials: Viewpoint. *JMIR Pediatr Parent*. 2022;5(2). doi:10.2196/35320
51. Wang J, Calderon G, Hager E, et al. Identifying and preventing fraudulent responses in online public health surveys: Lessons learned during the COVID-19 pandemic. *medRxiv*. Published online 2022:1-17. Accessed July 30, 2023. https://www.medrxiv.org/content/10.1101/2022.12.12.22283381v1.full-text
52. Mitchell JW, Chavanduka TMD, Sullivan S, Stephenson R. Recommendations from a Descriptive Evaluation to Improve Screening Procedures for Web-Based Studies with Couples: Cross-Sectional Study. *JMIR Public Health Surveill*. 2020;6(2). doi:10.2196/15079
53. Kensington-Miller B. Surviving the first year: new academics flourishing in a multidisciplinary community of practice with peer mentoring. *Professional Development in Education*. 2018;44(5):678-689. doi:10.1080/19415257.2017.1387867
54. Cantin CM, Brune S, Killam L, Glass T, Walker R, Vanderlee E. Establishing a Community of Practice for Doctoral Studies Amidst the COVID-19 Pandemic. *Quality Advancement in Nursing Education - Avancées en formation infirmière*. 2022;8(2). doi:10.17483/2368-6669.1320
55. McLoughlin C, Patel KD, O'Callaghan T, Reeves S. The use of virtual communities of practice to improve interprofessional collaboration and education: findings from an integrated review. *J Interprof Care*. 2018;32(2):136-142. doi:10.1080/13561820.2017.1377692