Combating fraudulent participant enrollment when using electronic recruitment

Objectives: To describe an unanticipated response to electronic survey recruitment and strategies we implemented to identify and exclude ineligible participants.

Methods: For an NIH-funded trial testing an online continence promotion program, we sought to recruit 500 participants for an exclusively electronic study. Eligible participants had to be assigned female at birth, 50 years or older, able to read and understand English, and have access to an internet-connected device with which to access the continence promotion. Participants were recruited through a study website promoted via online and in-person outreach by community organizations and social media advertisements. The study website provided information about the study and contained a link to an electronic screening survey where eligible participants provided electronic consent and then were invited to complete a baseline questionnaire and create a study account. Once participants created a study account, they were randomized and received their first participation stipend ($25).

In response to a sudden influx of participants overnight and subsequent inquiries about stipends, we suspected and then confirmed fraudulent activity and implemented a participant verification process. Participants who could not confirm their identity were removed from the trial.

Results:
Figure 1 presents the recruitment timeline and participants entering the screener daily. Recruitment launched 11/30/23 with 2 weeks of scheduled Facebook advertisements and emails sent to 1,361 community partner organizations. We closed enrollment on 12/22/23 following a surge of participants that entered the study between 12/14 and 12/22, having met our enrollment target.

As we processed the surge of new participants, several flags prompted us to question whether they were truly unique, eligible participants (Table 1). We worked with IRB and REDCap officials to implement a data integrity recovery plan. Almost 70% of enrolled participants did not pass an identity verification survey and were subsequently excluded. We re-opened recruitment with 3 strategies: 1) reCAPTCHA, 2) minimum survey completion time threshold, 3) identity confirmation survey with additional personal screening if discrepant. Even with these strategies, fraudulent participants continued to enter the screening link long after Facebook advertisements had stopped, and we ultimately created a new study link and sent the invitation only to university employees to complete enrollment.

We ultimately had 1,393 fraudulent participants (373 inconsistent information; 433 survey time; 587 failed personal screen) attempt to enroll, over 400 of whom received a participant stipend. Most (986, 71%) used Gmail or Outlook (368, 26%) emails.

Conclusions: Research recruitment via social media is particularly vulnerable to exploitation by

those who seek to earn compensation through participation regardless of eligibility. When recruiting for electronic studies, it is imperative to implement identity and eligibility confirmation checks to protect human subjects and preserve data integrity.

Figure 1. Number of participants that clicked into screening questionnaire across time
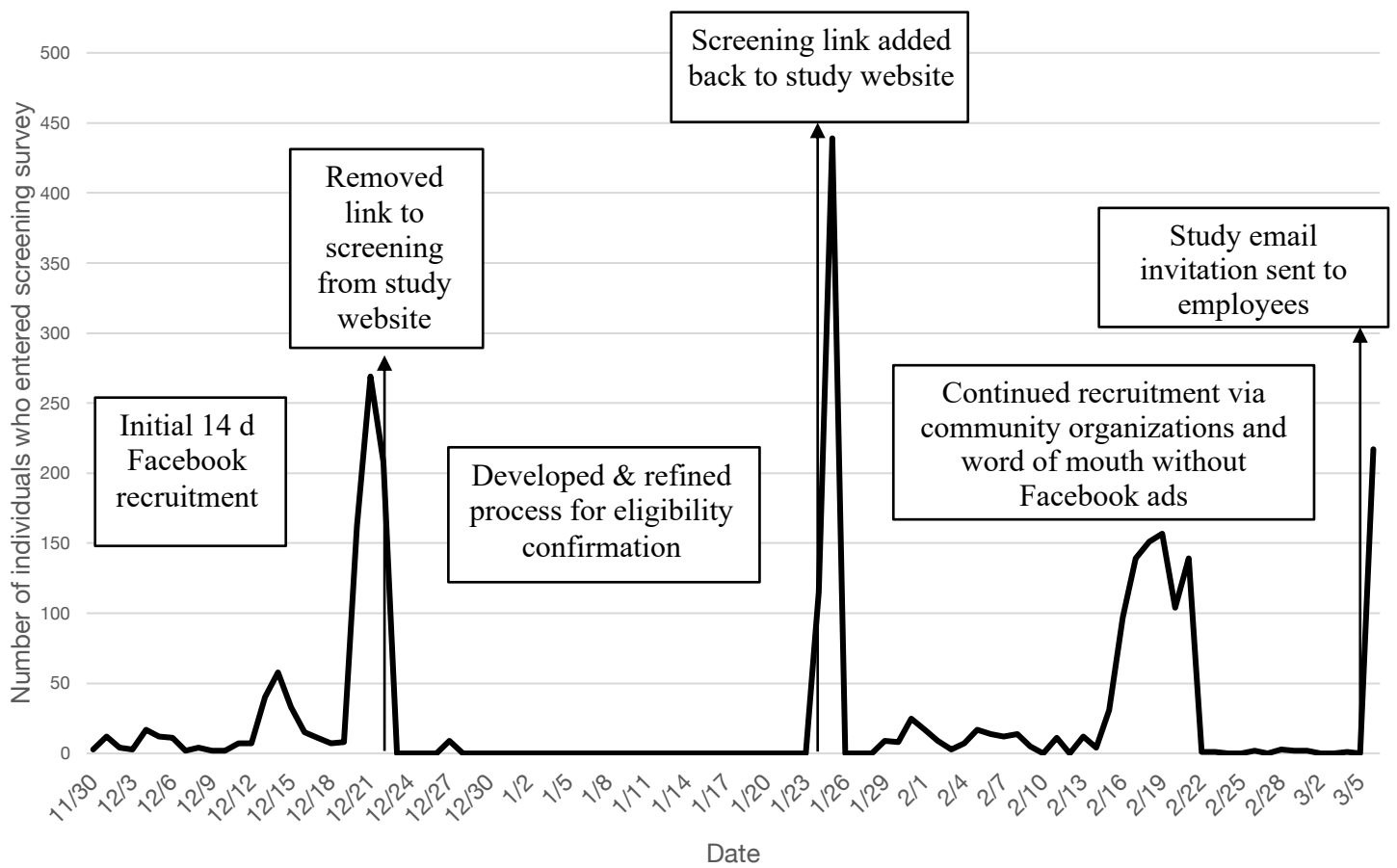
Table 1. Flags indicating potential fraudulent / ineligible participants

Category	Sub-category	Description and Examples
Emails	Name associated with email address	✗Name associated with email address differs from first and last name on informed consent.
	Subject line	✗Using all capitalized letters in the subject. ✗ "Interested" ✗ "Study" ✗ "Promised compensation" ✗ "About the compensation" ✗ "Requesting for the compensation"
	Body of email	Legitimate ✓Example: "Morning, I am a [xx y/o] women with a prolapsed bladder, as of [month and year]. Surgery improved it a little, but it is still very much "at the door" when I am standing. At [age] I had a hysterectomy after discovering cancer of the uterus, and I also had an episiotomy while giving birth to my only child, a son. Given my history, I would be interested in helping myself strengthen muscles in that area to help my bladder. If you believe I would make a good candidate for your study, please let me know. Sincerely, [first and last name]" ✓Example: [Last paragraph of an email received from participant] "Also, as a side note ... I got a message after I filled out the initial survey saying that I would receive the $25 gift card, but have not received it. Please know that receipt of the card has absolutely nothing to do with my participation, but did just want to let you know that if I was understanding correctly, it sounds like it should have reached me. I'm thinking it may have just slipped through the cracks, and I'm not worried about even receiving it. That's not why I'm participating, but just wanted you to be aware in case others also experienced it." Fraudulent ✗Example: "Hi there, Greetings, I came across your Women on the Go Study Website today and I saw your online program that helps women build knowledge and skills to improve bladder and bowel health. I'm deeply interested in this study program. I will be glad if permitted to participate actively and share my opinion with you. Thank you so much for your consideration" ✗Example: "Good day, I emailed you because of the promised compensation for completing this study, After completing this study you promised to provide and we will receive a compensation but till date I haven't seen anything, receiving and email that I have been removed from the study but haven't been compensate So all this thing aren't Good I haven't received my promised compensation and you removed me from the study because of am the last or what please explain And I completed this study long ago that is last year December which is unfair Take it to be like you how will you feel about this? Thank you for your response."
	Identical emails received within a short time	✗Verbatim emails received within a short period of time from multiple email addresses in response to a study email. ▪ Example: "I've created a previous account" ⇒ This email was received from five different email addresses, some of which were not sent the study email ⇒ All five emails were received by the study team on the same date between 5:10AM and 5:16AM ⇒ All five emails were verbatim and did not have a period at the end of the sentence
Survey responses	Processing speed (response time)	<table><tr><td></td><td>Real participants</td><td>Fraudulent participants</td></tr><tr><td>Average time (minutes)</td><td>16.57</td><td>5.50</td></tr><tr><td>Median time (minutes)</td><td>12.14</td><td>5.60</td></tr><tr><td>Inter-quartile range (minutes)</td><td>7.13 – 18.03</td><td>4.31 – 6.61</td></tr><tr><td>Range (minutes)</td><td>2.82 – 281.62</td><td>2.77 – 63,735</td></tr></table>
	Phone number	Utilizing online phone lookup can determine if the phone number is: ▪ Valid (✓yes/ ✗no) ▪ Line type (✗phone numbers listed as VoIP) ▪ Fraud score (✓0 – not likely fraudulent → ✗100 – likely fraudulent) ▪ Risky (✗yes/✓no)
	Mailing address	✗Themes our study team observed through mailing address responses: ▪ Wrote they lived in metropolitan locations, specifically Los Angeles, Miami, Brooklyn, and New York ▪ Provided two different addresses in Address line 1 and Address line 2 ▪ Only provided the street name, not a street number ✗Structure of mailing address provided: ▪ [##]W [##]th St or [##]th [##]W St ▪ Blvk [####] ▪ 10th street, 11th street, 12th street, etc ▪ 123 Main Street
	Common response errors	✗Individuals who were identified to be fraudulent often shared the same spelling errors: "Navada" instead of "Nevada," "Carlifornia" instead of "California," "Lowa" instead of "Iowa," "North California" instead of "North Carolina," and commonly mistook cities in the United States as states. 73 fraudulent participants wrote the state they lived in was "United State" or "USA."

✓: Likely a real participant; ✗: Likely a fraudulent participant.