Management of fraudulent participants in online research: Practical recommendations from a randomised controlled feasibility trial

Molly R. Davies[1] (0000-0003-3483-9907), Dina Monssen[1] (0000-0003-0080-0799), Helen Sharpe[2] (0000-0001-6980-1699), Karina L. Allen[1,3] (0000-0003-2896-6459), Beki Simms[4], Kimberley A. Goldsmith[5] (0000-0002-0620-7868), Sarah Byford[6] (0000-0001-7084-1495), Vanessa Lawrence[6] (0000-0001-7852-2018), Ulrike Schmidt[1,3,*] (0000-0003-1335-1937)

1. Centre for Research in Eating and Weight Disorders (CREW), Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK

2. Department of Clinical and Health Psychology, School of Health in Social Science, University of Edinburgh, Edinburgh, UK

3. The Eating Disorders Service, Maudsley Hospital, South London and Maudsley NHS Foundation Trust, London, UK

4. Fruitful Studio, The Design Chapel, Cemetery Road, Southampton, Hampshire, UK

5. Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

6. King's Health Economics, Institute of Psychiatry, Psychology, & Neuroscience, King's College London, London, UK

*Corresponding author: Ulrike Schmidt
Email: Ulrike.Schmidt@kcl.ac.uk
Address: Centre for Research in Eating and Weight Disorders, Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King's College London, SE5 8AF, UK

Abstract

Objective: Fraudulent participation is an escalating concern for online clinical trials and research studies and can have a significant negative impact on findings. We aim to shed light on the risk and to provide practical recommendations for detecting and managing such instances.

Method: The FREED-Mobile (FREED-M) study is an online, randomised controlled feasibility trial to assess a digital early intervention for young people (aged 16-25) in England or Wales with eating problems. The trial involved baseline (week 0), post-intervention (week 4), and follow-up (week 12) surveys, alongside weekly modules provided over four weeks on the study website. Study completers were compensated with £20 shopping vouchers. Despite the complexity of the trial design, two instances of fraudulent sign-ups occurred in January and March 2023. To counter this, we developed a "fraudulent participants protocol" following internal investigations and discussions with collaborators.

Results: The implementation of prevention measures such as reCAPTCHA updates, IP address review, and changes in reimbursement effectively halted further fraudulent sign-ups. Our protocol facilitated the systematic identification and withdrawal of suspected or clear fraudsters and was demonstrably robust at distinguishing between fraudsters and genuine responders.

Discussion: All remote, online trials or studies are at risk of fraudulent participation. Drawing from our experience and existing literature, we offer practical recommendations for researchers considering online recruitment and data collection. Vigilance and the integration of deterrents,

and data quality checks into the study design from the outset are advised to safeguard research integrity.

1 Introduction

Online research studies and clinical trials are becoming more common. Remote, online recruitment and data collection enable studies to have a greater reach, increased accessibility for participants, larger sample sizes, and require less time and resources both for the participants and the research team (Evans & Mathur, 2018; Wright, 2006). For example, advertising on websites and social media is typically easier and less labour-intensive than recruitment via clinical services. In particular, online research provides an opportunity for researchers recruiting hard-to-reach or low-prevalence populations, such as eating disorders (Qian et al., 2022), to more efficiently identify and recruit participants (e.g., King, O'Rourke, & DeLongis, 2014; Williams, Clausen, Robertson, Peacock, & McPherson, 2012). To help motivate participants to take part and to remove barriers to access (e.g., limited time, work commitments, financial instability), financial incentives to reimburse participants for their time are often used in online research studies. Past research has shown that financial incentives can improve recruitment and increase sample diversity (Göritz, 2006; Manzo & Burke, 2012; More, Burd, More, & Phillips, 2022).

Alongside the increase in online research, there has been a parallel increase in studies reporting fraudulent participants who have responded to these surveys (e.g., Burnette et al., 2022; Glazer, MacDonnell, Frederick, Ingersoll, & Ritterband, 2021; Goodrich, Fenton, Penn, Bovay, & Mountain, 2023; Salinas, 2023; Storozuk, Ashley, Delage, & Maloney, 2020). Online

data collection often relies fully on participant self-reporting without supervision or assistance. Fraudulent (also known as imposter) participants refer to individuals or bots who take advantage of this anonymity by signing up to studies and deliberately giving false responses. Bots are software applications (i.e., robots) that have been programmed and automated to complete a specific task, e.g., to complete a survey. In this paper, we will use the term “fraudster” to encompass both individuals and bots attempting to defraud a research study.

Fraudsters generally sign up multiple times to online research studies in order to accumulate financial compensation. The negative impact of fraudulent responses on research findings can be substantial, as these responses can invalidate and falsify results of a study or trial, introduce bias, and interfere with attempts to recruit participants meeting specific criteria (Chandler, Sisso, & Shapiro, 2020; Dupuis, Meier, & Cuneo, 2019). Furthermore, in clinical trials fraudulent participants can affect randomisation and parity between groups if more fraudsters are randomised to one group than the other. Fraudulent participation can be an issue for qualitative, as well as quantitative studies (Flicker, 2004; Jones et al., 2021; Roehl & Harland, 2022). For example, Roehl and Harland (2022) discovered that a fraudster had completed multiple interviews by keeping their camera off and taking advantage of the online, virtual format. For qualitative research, fraudulent participants could potentially raise further major ethical issues, for instance if a fraudulent participant takes part in a focus group with other, authentic participants where personal experiences are shared.

Despite prior reports of fraudulent participants in quantitative and qualitative research studies, there is a lack of evidence on the risk that this poses to clinical trials and how it can be managed. In this report, we describe our process for detecting and managing fraudulent participants within a randomised controlled feasibility trial focused on early intervention for young people with eating disorders. Our aim is to raise awareness of this issue to other researchers, both within and beyond the field of mental health and eating disorders research, and to provide practical recommendations to aid future researchers in addressing this issue more effectively. To our knowledge, we are the first online clinical trial to report fraudulent participation.

2 The FREED-M study

2.1 Trial design

FREED-Mobile (FREED-M) is an online, early intervention for young people with eating disorders to improve help-seeking and motivation for change. The FREED-M *study* is a randomised controlled feasibility trial conducted entirely online with multiple time points and a target sample size of 176 participants. Participants were recruited by advertising through relevant charity websites, social media, schools, universities, and participating NHS services including general practitioners (GPs) and specialist eating disorders services. Interested participants signed up and took part via the study website (https://freedm.uk). They first completed an online consent form and screening survey to assess eligibility. Eligibility criteria

required participants to be between 16-25 years old, live in England or Wales, be experiencing eating problems, and never have sought or received treatment for an eating disorder. Eligible participants were taken directly to the baseline survey (week 0) and then randomly allocated via the King's Clinical Trials Unit (KCTU) online randomisation service to either the control or the intervention arm. After randomisation, participants received access to weekly online modules for four weeks, which included a brief questionnaire, and were notified by email once each module became available. At the end of the fourth module (week 4) and at 12-weeks post randomisation (week 12), participants completed a more detailed post-intervention and follow-up assessment, respectively. Participants who completed the week 12 follow-up survey were compensated with a £20 shopping voucher. Full details of the trial design are described by Gryzuk et al (under review).

2.2 Existing security measures and deterrents

There were several basic security measures integrated into the trial design and the website at the start of the study. Participants were required to create an account with an email address and password in order to take part. Although email addresses weren't explicitly validated, participants without a valid email would not receive notice when the modules became available. The website included Google reCAPTCHA v2 software at the consent stage of the sign-up to prevent bot attacks. Additionally, the study involved responses to surveys across multiple timepoints and included free-text responses, making bot completion less likely. Due to the

complexity and longitudinal design of the trial, we did not believe that the FREED-M study was at high risk of fraudulent sign-ups.

3 The influx of fraudulent sign-ups

Despite the existing security measures, the FREED-M study received multiple fraudulent sign-ups between 20-23 January 2023, and again between 5-6 March 2023. There were 76 baseline surveys completed in the January attack, and 22 in March. The sign-ups were quickly suspected to be fraudulent due to the large number of responses that came through within a short time period relative to the previous rate of recruitment (>20 per day vs ~7 per week). The influx of sign-ups occurred in specific intervals which suggested a level of coordination; new baseline surveys were completed every 3-20 minutes for 3-6 hour blocks, and then the sign-ups would pause for a few hours before resuming again. Some of these sign-up intervals occurred at unlikely times of day, such as mornings before 7:00AM or evenings after 9:00PM. In addition, these influxes of participant sign-ups did not correspond to an increase in recruitment efforts (e.g., a period of advertising) and therefore appeared to come out of nowhere. Several of the fraudsters also emailed the study team with identical email text asking about financial compensation. Taking all of these factors into consideration, our team suspected that FREED-M had been targeted by fraudsters.

4 Response to fraudulent accounts

In response to the January attack, we paused randomisation to assess the situation and develop a plan of action. We contacted the study's website developer at Fruitful Studio (BS) to discuss the FREED-M website security, spoke with collaborators and other researchers who had a similar experience, and reviewed the literature for advice on how to manage the situation. We conducted internal investigations to better understand factors which may have led to this attack and reviewed the baseline surveys of all participants who signed up during this period. Following these discussions and our investigations, we developed a "fraudulent participants protocol" which aimed to: 1) improve prevention of future attacks, 2) outline how to identify suspected fraudulent participants, and 3) detail the procedure for managing suspected fraudulent participants at the various stages of study participation. This protocol is detailed in the section below. We first describe the actions which were considered or implemented, and then the outcome.

Our "fraudulent participants protocol" was reviewed by the primary investigators on the project, the Trial Steering Committee, and the South London and Maudsley Research and Development team. It was approved by the Camden and Kings Cross Research Ethics Committee on 13th April 2023 (22/LO/0655; Substantial Amendment 1).

4.1 Prevention (pre-sign up)

<u>Software update</u>

In collaboration with the FREED-M website developer, we identified and discussed various options for increasing the website security. Following the January attacks, the developer first updated our Google reCAPTCHA software at the consent stage to the newest version (v3) in order to improve the prevention of bots.

Outcome

The January sign-ups stopped after the reCAPTCHA software was updated, which suggests that it was temporarily effective. However, the second attack in March indicates that the sign-ups either come from human fraudsters (i.e., humans who fake their responses or sign up multiple times) or that more sophisticated bots are able to bypass the reCAPTCHA update. Following the March attack, an additional reCAPTCHA was added at the screening stage.

Internet protocol (IP) address

We noticed that several of the fraudulent participants signed up using the same internet protocol (IP) address. We therefore considered updating the website to only allow sign-ups from unique IP addresses. Although this can easily be circumvented by using a virtual private network (VPN), it would create an additional barrier for an individual and more advanced coding for the bots and therefore could act as a deterrent. However, we noticed that some duplicate IP addresses were from students at a specific university who had received a text message from their university GP surgery. The study team were aware of the text message campaign and reviewed the students' responses, which verified that the responses were meaningful and

distinct. These students were likely using a university WiFi network, such as eduroam or a university VPN.

Decision

We ultimately decided not to implement a block on duplicate IPs, as this would have created an undue barrier for students in taking part, particularly those living in university halls. Instead, we chose to continue with a manual review of IP addresses. There are many third-party software providers on the market that may have been able to provide a solution by blocking the use of VPNs or verifying that the user's location was within England or Wales. This would not impact students using eduroam or other university WiFi and would be robust to multiple sign-ups from the same computer. However, implementing this software would require additional development time, annual subscription costs, and testing. As a result, we chose not to implement this for the feasibility trial but note it here as a potential option for future researchers.

Changing the reimbursement method

Prior to the identification of the fraudulent accounts, we had planned to reimburse participants by sending a voucher code via email. However, we wanted a further check of a participant's veracity prior to providing the financial reimbursement in case the fraudsters adapt their methods to slip past our detection and complete the study. At this stage, our participant information sheet and protocol did not specify the method of delivery. Other researchers who have encountered fraudulent or imposter participants (e.g., Roehl & Harland, 2022) have

recommended sending vouchers only via the post. Although this created additional work for the research team to post the vouchers and had minor cost implications (i.e., postage, envelopes), we decided to implement this change. Sending vouchers via the post required the participant to provide a valid English or Welsh address which enabled us to verify residence, ensured only one voucher was sent per address, and created a barrier for fraudulent participants in receiving financial compensation. The chosen reimbursement vouchers were also only valid for use in the UK.

No participants had completed the week 12 follow-up survey (the point at which vouchers were provided) prior to January 2023, therefore all participants who completed the study were asked to provide their postal address to receive the voucher. The email noted that the voucher would be sent in a discreet envelope without study details to ensure privacy. After the March attack, to deter fraudsters from taking part, all participant-facing study materials (i.e., the participant information sheet, study advertising materials, and website sign-up page) were updated to clarify the type of voucher and that it would be sent in the post.

Outcome

After the participant-facing study materials were updated in March, we did not receive any further identified fraudster sign-ups.

Summary

Since the reCAPTCHA updates and the change in reimbursement method were implemented simultaneously, we are unable to distinguish the individual impact of each action. Overall, these combined changes appear to have been effective at preventing future fraudster sign-ups.

4.2 Identification and withdrawal of fraudulent participants (post-baseline)

We developed a procedure to systematically identify and withdraw fraudulent participants. This procedure is outlined in Figure 1 and detailed in the text below.

Figure 1. Flowchart for identifying and withdrawing suspected fraudulent participants.

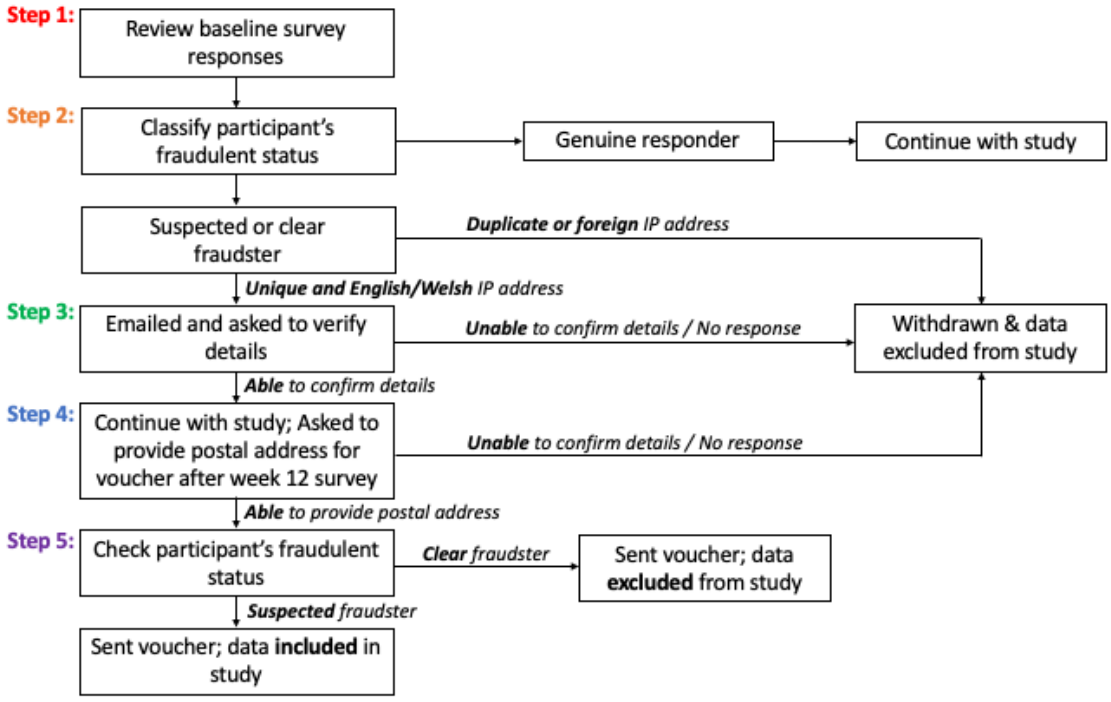

Figure 1 displays the steps undertaken by the research team for identifying and withdrawing suspected fraudsters or clear fraudsters. IP addresses were manually checked using a publicly available website

which displays the hosting server and device location. The duplicate IP addresses referenced above only refer to those that were not hosted by a university.

Step 1: Review baseline survey responses

When the first attack occurred in January, two members of the study team (MRD and DM) manually reviewed the survey responses for all participants who had signed up prior to 23rd January in order to identify which accounts were potentially fraudulent. Some participants had completed the baseline (week 0) and post-intervention (week 4) assessments, therefore we were able to compare responses at different time points to assess consistency. All of the suspected fraudulent responders from the January attack were identified as having expressed interest in the study through a large UK charity recruitment website. When expressing interest in the study through this website, individuals provided their name, email, date of birth, gender, and ethnicity and consented to this information being shared with the study team. They then received an email with the website link to sign-up. In addition to reviewing the survey responses, we were also able to compare these details with the information they signed up to FREED-M with to identify discrepancies.

Outcome

We created a list of characteristics or patterns in the personal details and survey data that were invalid, inconsistent, illogical, or followed specific, clear patterns which appeared within and across multiple accounts identified as likely to be fraudulent. We compiled this list into a

"fraudsters' profile" (Table 1). Although the website did not include software to detect IP address location, the research team used a publicly available website to check the IP address location and hosting server, therefore this information was still used to assess fraudulent status. Of note, the individual responses outlined in Table 1 were not enough to identify someone as a potential fraudster; it was the combination of multiple of these characteristics within a single account and similarities between accounts signing up in short succession that raised suspicion.

Table 1. Fraudulent participants' profile.

Data category	Variable(s)	Response characteristic or pattern
Personal details	Postcode	Invalid postcode (e.g., SE1, A220 1288, or 2138), or full, invalid address entered in postcode.
	IP address	IP address from another country, or the internet service provider was a VPN or hosting service.
	Email address	Email addresses followed a similar format, including the first name and last name, ending with numbers, and from the same email provider. For example, for someone who signed up under the name Jane Smith, the email address would typically look like one of the below: - janesmith256@emailprovider.com - smithjane34@emailprovider.com - jsmith897@emailprovider.com - smithj023@emailprovider.com - janes56@emailprovider.com
	Phone number*	Phone number started with +, 44, 7, or 021.
Demographics	Age	Age from birthdate did not match age reported on screener.

	Inconsistent date of birth and/or ethnicity	The date of birth and/or ethnicity provided to the FREED-M study and those provided to the charity when expressing interest did not match
Inconsistent or illogical survey responses	Weight	Current weight was higher than the highest weight. Current or highest weight varied widely at different time points (e.g., current and highest weight were 50kg at week 0 and 192kg at week 4).
	ED symptoms	The combination of reported ED symptoms were illogical or unlikely. For example, the participant reported feeling fat and an over-evaluation of weight and shape but no fear of gaining weight, or reported feeling depressed or guilty after overeating but not feeling upset about the overeating.
	Anxiety and depressive symptoms	Reported anxiety or depressive symptoms on the PHQ-4 but reported not feeling anxious or depressed on the quality of life questionnaire (ED-5D), or vice versa.
	Employment and service use	Illogical responses. For example, reported being 25 and employed, but reported seeing a Special Education Needs Coordinator (suggesting they were in education) in the last 3 months.
	Service use	Reported a large number of appointments on the service use questionnaire that seemed unlikely within a three-month time period. For example: - Multiple hospital appointments for their eating disorder, mental health, and non-mental health reasons - Treatment by ambulance crews but no contact with Accident and Emergency (A&E) - Over 10 encounters with A&E
	Medications	Implausible changes in medications between surveys. For example, only took medication for anxiety or depression at week 0, but at week 4 reported taking medication for anxiety/depression, ADHD, and Tics/Tourette's.
Response patterns within a single survey	Eating disorder diagnostic scale (EDDS)	Responses followed a pattern, for example: - All questions were answered "Yes" - All questions were answered "Yes", "No", "Yes", "No" etc

		- Most or all of the binge/compensatory behaviour frequencies were the same - All binge/compensatory behaviour frequencies were relatively high (e.g., vomit 7x, laxatives/diuretics 10x, fasted 8x, over-exercised 15x)
	Medications	Reported taking most (i.e., 4 out of 5) or all medication options. These options included medications for 1) depression or anxiety, 2) ADHD, 3) sleep disorders, 4) Tics/Tourette's, and 5) behaviour, irritability, aggression or psychosis.
Response patterns across participants signing up in a short timeframe	Name	Suspicious names, for example: - Similar names (e.g., same last name, or first name and last name switched around, such as Bob Eric and Eric Bob) - Names with strange or unlikely spellings (e.g., Groria) - Names based on famous people (e.g., Tom Cruise, Marilyn Monroe)
	Referral source	Reported hearing about the study from a friend or family member.
	Weight	Current and highest weight were the same value.
	BMI	BMI was high (>40).

Table 1 displays the characteristics or patterns of responses of accounts which showed clear signs of fraud. The exact profile (i.e., the combination of response characteristics or patterns) varied between fraudsters. A combination of characteristics were used to assess fraudster status; each characteristic individually was not enough to raise suspicion. The "data category" column provides additional context to the listed characteristics. Some characteristics were not suspicious in isolation, but were noted due to the discrepancies within a single response (i.e., response patterns within a single survey) or the number of accounts which signed up in a short period of time with similar response patterns (i.e., response patterns across participants signing up in a short timeframe).

*Note: although these phone numbers are not all invalid, we observed that all genuine responders input phone numbers beginning with “07” (the standard UK mobile phone format), whereas fraudsters often used these alternate formats.

Abbreviations: IP = internet protocol; VPN = virtual private network; kg = kilogramme; ED = eating disorder; PHQ-4 = 4-item patient health questionnaire; EQ-5D = EuroQol-5 Dimension questionnaire; A&E = accident and emergency; ADHD = attention deficit hyperactivity disorder; BMI = body mass index.

Step 2: Identifying fraudulent participants

Following this review, the two researchers (MRD and DM) independently classified accounts into three categories: clear fraudsters, suspected fraudsters, or genuine responders. The researchers met to confer regarding the likelihood of fraud and to resolve any discrepancies.

Clear fraudsters included those whose data followed many of the patterns from Table 1.

Suspected fraudsters were accounts whose data followed some of the patterns of responses highlighted in Table 1, but overall were consistent and logical. **Genuine responders** refer to participants whose responses followed very few of the patterns of responses in Table 1 and overall were non-contradictory, plausible, and consistent between the survey(s). Additionally, genuine responders often reported postcodes that were within catchment areas of organisations actively recruiting on behalf of the study, or reported learning about the study from a known, legitimate source (e.g., from an actively recruiting GP practice, eating disorder service, or university). These clear links to active recruitment efforts yielded further confidence that these genuine responders were trustworthy.

We debated creating an algorithm based on this profile to identify fraudulent accounts, but felt that was not feasible for several reasons. First, it was too time consuming to develop and properly test given the short duration (<1 year) of the FREED-M feasibility trial. Second, there was no absolute or foolproof method for detecting fraud. Some genuine responders included one or two of the response characteristics or patterns mentioned in Table 1 due to their severity of symptoms (e.g., “Yes” to all eating disorder symptoms) or to errors in data entry (e.g., differences between self-reported age and age calculated from date of birth). Finally, the clear or suspected fraudsters were often identified due to the overall impression of the response patterns (e.g., a combination of different response patterns listed in Table 1) and/or in the context of *other* accounts that had been created in the same time period (e.g., names being suspicious because they were similar to other accounts created on the same day). An algorithm would be unlikely to pick up on this.

Outcome

When the January attack occurred, the surveys of all participants who had signed up before 23 January 2023 were reviewed. Of the 76 participants who signed up during the 20-23 January attack, three were categorised as genuine responders and 73 as suspected or clear fraudsters. All 73 identified at this time had signed up through the charity recruitment platform. Of these, 32 (44%) had discrepancies on date of birth and 23 (32%) on ethnicity between the details reported to the charity versus FREED-M. Of the 39 participants who signed up and were assessed as

eligible (so were included in the study) before the 20th January, nine participants who had not raised suspicion at the time were classified as suspected fraudsters once we checked their responses against the fraudsters profile. Worryingly, it is unlikely that we would have identified these responses as fraudulent if the January attack had not occurred, or would only have noticed at the data analysis stage. In addition, of the 27 participants who had signed up and were assessed as ineligible (so were not included in the study) before the January attack, 10 participants were found to have personal details that matched the fraudsters profile. It is possible that the fraudster(s) had spent some time prior to the January incident to find a pattern of responses that met eligibility criteria and bypassed detection.

Following the January attack, all baseline survey responses were reviewed by at least one of the researchers before participants were randomised to a group and allowed to continue with the study. Suspicious responses were flagged and reviewed by both researchers.

No fraudulent responses were identified after the January attack and prior to the second attack in March 2023. All 22 participants who completed the baseline survey during the attack between the 5th and 6th of March were clear fraudsters. After the March attack, the researchers continued reviewing all new baseline surveys prior to randomisation. We identified one participant in April 2023 who was classified as a suspected fraudster, although their profile differed from the ones during the attacks (i.e., their personal details were potentially genuine, but their responses to the questions followed an obvious pattern). This participant was

contacted to confirm details and did not respond, and was therefore withdrawn from the study. Additionally, conducting these reviews helped us identify data quality errors. Participants with data quality errors were classified as suspected fraudsters and details were then checked with the participants. In April and May 2023 we identified two participants who entered an implausibly low weight value (i.e., <20kg) and two others who had discrepancies between self-reported age and date of birth.

As of May 2023, the two researchers have been in full agreement regarding the fraudulent or genuine categorisations for all participants who have been reviewed.

Step 3: Verifying personal details

Participants that were categorised as clear or suspected fraudsters who either had the same IP address as another participant (not associated with a university) or an IP address from another country were immediately withdrawn from the study and not given an opportunity to continue. Duplicate or foreign IP addresses were an objective indicator of fraud and therefore we felt that withdrawing these participants was ethically appropriate. For the fraudsters without duplicate or foreign IP addresses, we decided not to automatically withdraw these participants, even in cases of clear fraud in the survey responses, to avoid potential ethical issues as we did not state in our study protocol, information sheet, or consent form that participants' survey responses would be evaluated for truthfulness and used as a basis for withdrawal. Instead, these fraudsters were emailed and asked to verify their date of birth, postcode, and/or phone

number in order to continue with their participation. Those who were able to confirm could continue taking part. The participants with data entry errors but no other signs of fraud were additionally asked to correct these details.

Outcome

As of May 2023, only three (3%) of the 96 total suspected or clear fraudsters responded to confirm their details and were allowed to continue. Their progress through the study was monitored according to steps 4 and 5 below. One of the 96 suspected fraudsters replied with the wrong date of birth and was withdrawn. Of the four participants with data quality errors, one responded and continued with the study, while the other three did not and were withdrawn in line with the fraudulent participants protocol.

Step 4: Providing postal address

All participants who completed the final survey were asked to provide a postal address in order to receive their financial compensation. Suspected or clear fraudsters who were unable to provide a postal address in England or Wales or did not respond were withdrawn from the study.

Outcome

None of the three suspected fraudsters who had been allowed to continue with the study after Step 3 were able to provide a postal address. One fraudster responded to the email and stated

they were uncomfortable sharing their address details and asked for the voucher to be sent digitally. This participant's postcode (provided at sign-up) indicated that they lived in London, so they were offered the opportunity to pick up the voucher in person at university reception, so no interaction with study staff. The suspected fraudster then did not respond further. These participants were withdrawn.

In contrast, as of May 2023, there were 19 genuine responders who completed the final survey and all responded to provide their address. None of the genuine responders expressed concerns about this reimbursement method.

Step 5: Including or excluding response data

Flicker (2004) described three different approaches to managing data from suspected fraudulent or dishonest participants, depending on the researcher's interpretation of the event: the cynic, the sceptic, and the seeker. The cynic approach applies when the researcher believes the data is entirely fabricated, and therefore suggests the data should be excluded. The sceptic approach applies when the researcher thinks that some, but not all, data may be misrepresented or falsified, and therefore data should be included but treated with scepticism (e.g., compared to responses from other participants to check for large discrepancies). Alternatively, the seeker approach assumes that participants have a story to tell and their data may be important in the context of the research regardless of the truthfulness, and therefore

they should be included. Although these approaches were originally proposed for qualitative data, we believe they apply in this situation as well.

Based on this framework, we decided on two different approaches for managing data depending on our confidence of its veracity. We chose to follow the “cynic” approach for all clear fraudsters, regardless of whether they were able to confirm their personal details and even if they had been able to provide a valid postal address, based on our strong belief that these data were entirely fabricated (Flicker, 2004; Roehl & Harland, 2022). Suspected fraudsters who were withdrawn at any stage of the study (i.e., in steps 2-4 listed above) were similarly excluded from analyses, as we interpreted the inability to confirm details as confirmation of falsehood. In contrast, we planned to follow the sceptic approach for suspected fraudsters if they were able to provide a postal address. In these cases, we planned to compare their data and outcomes to those of genuine responders and consider conducting sensitivity analyses in which data from suspected fraudsters were excluded to assess the impact on our findings.

Outcome

As mentioned above, as of May 2023 none of the suspected or clear fraudsters were able to provide a postal address and were withdrawn, therefore all of their data will be excluded from analyses.

Summary

Our new procedure for identifying and managing fraudulent participants appears effective. The participants who were classified as suspected or clear fraudsters were unable to confirm their details or provide a postal address, in direct contrast with those classified as genuine responders. This suggests that these classifications are robust. Furthermore, this procedure enabled the research team to improve data quality by identifying and providing an opportunity to correct errors in data entry. However, manually reviewing surveys and contacting suspected fraudsters are time consuming for the research team and could create an additional barrier to participation (i.e., with true participants having to verify details). Eating disorders are often characterised by secrecy and shame and therefore these checks could accidentally dissuade some genuine responders from taking part.

5 Practical recommendations

All remote, online research studies are at risk of fraudulent participation. Based on our experience and previous literature, we have compiled a list of practical recommendations for future researchers considering online recruitment and data collection. These recommendations are summarised in Table 2.

Table 2. Practical recommendations for preventing or identifying fraudulent responses.

Aspect of study	Recommendation
Study design	Integrate fraudster prevention and detection methods from the outset

	Utilise Patient and Public Involvement (PPI) to help balance participants' needs
Financial compensation	Send vouchers in the post
	Compensate using vouchers only valid in the host country
	Consider how compensation is mentioned in public advertisements
Survey design	Include traps for bots hidden within the survey design, such as repeated questions to assess for consistency in response, "honeypot" questions (questions which are hidden from human participants, e.g., using JavaScript, but not from bots), open-ended or free-text questions, attention check questions (e.g., instructing participants to select a specific answer option to check they're paying attention to the survey), and illogical options on multiple choice questions
Data quality checks	Build in early and ongoing checks of the incoming data, incorporating manual review by a member of the research team
	Examine participant demographics
Website security measures	Inclusion of reCAPTCHA software
	Consider blocking VPNs
	Consider blocking foreign IP addresses

Abbreviations: PPI = Patient and Public Involvement; VPN = virtual private network

Study design

Most importantly, future researchers should integrate fraudulent or bot prevention and detection methods into the study design from the outset. Researchers often prioritise participants' convenience, anonymity, and privacy, aiming to make studies as accessible as possible and limit the collection of personal data to a minimum. Taking FREED-M as an example, we had

initially planned to send study vouchers via email to avoid asking participants for address details unnecessarily (and also make it easier for the study team). Ultimately, this left us vulnerable to fraud. Researchers should consider how best to balance participants' right to privacy versus data quality (Roehl & Harland, 2022; Teitcher et al., 2015). Utilising Patient and Public Involvement (PPI) could help researchers design the study in a way that maintains this balance and is acceptable to the population under study. Research has shown that PPI can significantly increase recruitment and retention rates (Crocker et al., 2018; Domecq et al., 2014; Pizzo, Doyle, Matthews, & Barlow, 2015; Price et al., 2018). For example, researchers could conduct focus groups to ask service users about whether certain security measures for protecting against fraudsters, such as asking for additional personal details or sending vouchers in the post, would be acceptable or a barrier to participation.

Financial compensation

Although financial incentives can help with recruitment, they can also increase the likelihood of fraud (Bowen, Daniel, Williams, & Baird, 2008; Rogers, Trubey, & Oard, 2022). If the study design includes a financial reimbursement for participation, consider methods for compensating people that would make it more difficult or impossible for fraudsters or bots. In this study, we changed our protocol to send vouchers exclusively by post. However, this may not be possible in countries with limited or unreliable postal systems and may create barriers to participation that are inappropriate for the population under study (e.g., those with insecure housing). Alternate options may include compensating with a voucher that is only valid in the country

where the study is taking place or conducting identity checks prior to sending online vouchers (e.g., over video calling). We would also recommend excluding mention of financial compensation from online, public advertisements (e.g., on social media) or making the compensation method clear to deter potential fraudsters.

Survey design

Specific survey design approaches could also be used to help identify fraudulent responders and bots. Including repeated or similar questions within or across surveys enables researchers to check for inconsistencies. For FREED-M, we avoided repeating unnecessary questions (e.g., height) to keep the surveys as short as possible. However, some of our measures within each survey assessed similar constructs (e.g., anxiety and depression) and we had repeated measures across time points that were less likely to vary drastically over time (e.g., weight). These were a fortunate accident rather than specifically designed to catch fraudsters, but proved highly beneficial.

There are many other suggestions which can be found online, such as: incorporating ‘honeypot’ questions which are hidden from human participants (e.g., using JavaScript) but not from bots, including open-ended questions where nonsensical responses would be more obvious, adding questions with instructions on which response to select (attention check questions), or having an illogical option on multiple choice questions (GitLab, 2023; Perkel, 2020; Storozuk et al., 2020). Additionally, we recommend recording IP addresses to be able to detect location and

VPN use, and also record the time it takes to complete surveys to inspect participants who respond more quickly than others. Consider limiting sign-ups to require unique email addresses to add further barriers.

Data quality checks

Data quality checks should be conducted regularly from the outset of the study to more quickly identify suspicious sign-ups and patterns. The survey design approaches mentioned above would make these checks more straightforward and make it more feasible to develop an algorithm to help screen responses for signs of fraud. However, we would still recommend oversight from a member of the study team, as fraudsters are capable of adapting their methods or response patterns to avoid detection (Storozuk et al., 2020). Beyond checking the individual responses, examining the overall demographics of participants could also help to identify a potential issue. In FREED-M, we had a much higher proportion of males and greater ethnic diversity than is typically observed in eating disorders research. Once the fraudulent participants were withdrawn, the demographics more closely matched what we had expected. For randomised clinical trials, it is particularly important to implement these checks prior to randomisation to avoid disrupting the randomisation protocol. Randomisation protocols often include minimisation factors, which are characteristics that are incorporated into the randomisation process to ensure they are balanced across the groups. In best practice, withdrawn participants should not be replaced on the list (i.e., randomisation allocations should

not be re-used). Although extra participants could be recruited to counteract the number of fraudulent randomisations, as we did for FREED-M, this may still introduce bias to the analyses.

If data checks find evidence of suspected fraudsters, then it is also important to carefully consider in advance how this would be managed and communicated. Care needs to be taken that potentially genuine responders do not feel that the authenticity of their experiences or perspective are being challenged. PPI would be helpful in designing this process and tailoring messages for requesting further information or confirmation of details.

Website security measures

For studies which utilise a bespoke website, we recommend implementing security measures specifically against bots. Ensure that the latest version of reCAPTCHA software is in use on public facing forms. Consider blocking sign-ups from other countries, which is more highly recommended than blocking sign-ups from duplicate IP addresses. However, be aware of VPNs which can easily circumvent IP address restrictions. There may be more advanced security software available that blocks VPN use.

Summary

The recommendations listed above are not exhaustive. Similar lists of recommendations have been published elsewhere which we found useful when developing this protocol (see Roehl & Harland, 2022; Storozuk et al., 2020; Teitcher et al., 2015). Technological advances have led to

more sophisticated bots and other programmes to enable fraud, and these advances are likely to continue. This means more sophisticated methods for detecting fraudulent responses will likely be required. Procedures for identifying and managing potential fraud may need to be developed iteratively. Other researchers have proposed detailed protocols for planning online research studies which address potential fraud, such as the REAL framework (Lawlor et al., 2021). The REAL framework provides guiding questions to aid researchers when thinking about how to prevent, detect, and manage fraud. Overall, early and ongoing monitoring of data and sign-ups will be crucial to enable researchers to identify issues quickly and adjust procedures as needed. With FREED-M, we were fortunate that our trial was relatively small and thus enabled the researchers to perform individual checks of the data and develop an ad hoc protocol for managing fraud. Larger online studies without comparable resources may wish to consider conducting a smaller scale pilot to develop, test, and refine methods for checking and verifying data quality before proceeding to large-scale data collection.

6 Conclusion

Fraudulent participants and bots are a real issue and can potentially impact any study or trial (either quantitative or qualitative), particularly those conducted remotely online with financial incentives. Complex study designs and multiple timepoints are not sufficient to deter or prevent these attacks. We did not believe FREED-M was at high risk of being targeted and thus were caught by surprise. We caution other researchers not to make the same mistake. Future researchers must be aware of this risk when using online data collection methods and should

implement deterrents and data quality checks in the study design from the outset.

Acknowledgements

This work was supported by the National Institute of Health Research (NIHR) under its Research for Patient Benefit (RfPB) Scheme (grant number NIHR201175). US, HS and KA receive support via the Medical Research Council/Arts and Humanities Research Council/Economic and Social Research Council Adolescence, Mental Health and the Developing Mind initiative as part of the EDIFY programme (grant number MR/W002418/1). US receives salary support from the NIHR Biomedical Research Centre (BRC) at the South London and Maudsley (SLaM) NHS Foundation Trust and King's College London (KCL). The views expressed herein are those of the authors and not necessarily those of the NHS, NIHR or Department of Health and Social Care.

Data availability statement

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

Bowen, A. M., Daniel, C. M., Williams, M. L., & Baird, G. L. (2008). Identifying multiple submissions in Internet research: preserving data integrity. *AIDS and Behavior*, *12*(6), 964–973. https://doi.org/10.1007/s10461-007-9352-2

Burnette, C. B., Luzier, J. L., Bennett, B. L., Weisenmuller, C. M., Kerr, P., Martin, S., Keener, J., & Calderwood, L. (2022). Concerns and recommendations for using Amazon MTurk for eating disorder research. *The International Journal of Eating Disorders*, *55*(2), 263–272. https://doi.org/10.1002/eat.23614

Chandler, J., Sisso, I., & Shapiro, D. (2020). Participant carelessness and fraud: Consequences for clinical research and potential solutions. *Journal of Abnormal Psychology*, *129*(1), 49–55. https://doi.org/10.1037/abn0000479

Crocker, J. C., Ricci-Cabello, I., Parker, A., Hirst, J. A., Chant, A., Petit-Zeman, S., Evans, D., & Rees, S. (2018). Impact of patient and public involvement on enrolment and retention in clinical trials: systematic review and meta-analysis. *BMJ (Clinical Research Ed.)*, *363*, k4738. https://doi.org/10.1136/bmj.k4738

Domecq, J. P., Prutsky, G., Elraiyah, T., Wang, Z., Nabhan, M., Shippee, N., Brito, J. P., Boehmer, K., Hasan, R., Firwana, B., Erwin, P., Eton, D., Sloan, J., Montori, V., Asi, N., Dabrh, A. M. A., & Murad, M. H. (2014). Patient engagement in research: a systematic review. *BMC Health Services Research*, *14*, 89. https://doi.org/10.1186/1472-6963-14-89

Dupuis, M., Meier, E., & Cuneo, F. (2019). Detecting computer-generated random responding in questionnaire-based data: A comparison of seven indices. *Behavior Research Methods*, *51*(5), 2228–2237. https://doi.org/10.3758/s13428-018-1103-y

Evans, J. R., & Mathur, A. (2018). The value of online surveys: a look back and a look ahead. *Internet Research*, *28*(4), 854–887. https://doi.org/10.1108/IntR-03-2018-0089

Flicker, S. (2004). “Ask Me No Secrets, I’ll Tell You No Lies:” What Happens When a Respondent’s Story Makes No Sense. *The Qualitative Report*, *9*(3), 528–537.

https://doi.org/10.46743/2160-3715/2004.1922

GitLab. (2023). *Qualtrics tips and tricks*. https://about.gitlab.com/handbook/product/ux/qualtrics/#prevent-spam-responses-before-sharing-the-survey

Glazer, J. V., MacDonnell, K., Frederick, C., Ingersoll, K., & Ritterband, L. M. (2021). Liar! Liar! Identifying eligibility fraud by applicants in digital health research. *Internet Interventions : The Application of Information Technology in Mental and Behavioural Health*, *25*, 100401. https://doi.org/10.1016/j.invent.2021.100401

Goodrich, B., Fenton, M., Penn, J., Bovay, J., & Mountain, T. (2023). Battling bots: Experiences and strategies to mitigate fraudulent responses in online surveys. *Applied Economic Perspectives and Policy*. https://doi.org/10.1002/aepp.13353

Göritz, A. (2006). Incentives in Web Studies: Methodological Issues and a Review. *International Journal of Internet Science*, *1*(1), 58–70.

Jones, A., Caes, L., Rugg, T., Noel, M., Bateman, S., & Jordan, A. (2021). Challenging issues of integrity and identity of participants in non-synchronous online qualitative methods. *Methods in Psychology*, *5*, 100072. https://doi.org/10.1016/j.metip.2021.100072

King, D. B., O'Rourke, N., & DeLongis, A. (2014). Social media recruitment and online data collection: A beginner's guide and best practices for accessing low-prevalence and hard-to-reach populations. *Canadian Psychology/Psychologie Canadienne*, *55*(4), 240–249. https://doi.org/10.1037/a0038087

Lawlor, J., Thomas, C., Guhin, A. T., Kenyon, K., Lerner, M. D., UCAS Consortium, & Drahota, A. (2021). Suspicious and fraudulent online survey participation: Introducing the REAL framework. *Methodological Innovations*, *14*(3), 205979912110504. https://doi.org/10.1177/20597991211050467

Manzo, A. N., & Burke, J. M. (2012). Increasing Response Rate in Web-Based/Internet Surveys. In L. Gideon (Ed.), *Handbook of survey methodology for the social sciences* (pp. 327–343).

Springer New York. https://doi.org/10.1007/978-1-4614-3876-2_19

More, K. R., Burd, K. A., More, C., & Phillips, L. A. (2022). Paying participants: The impact of compensation on data quality. *Testing, Psychometrics, Methodology in Applied Psychology*, *29*(4), 403–417.

Perkel, J. M. (2020). Mischief-making bots attacked my scientific survey. *Nature*, *579*(7799), 461. https://doi.org/10.1038/d41586-020-00768-0

Pizzo, E., Doyle, C., Matthews, R., & Barlow, J. (2015). Patient and public involvement: how much do we spend and what are the benefits? *Health Expectations*, *18*(6), 1918–1926. https://doi.org/10.1111/hex.12204

Price, A., Albarqouni, L., Kirkpatrick, J., Clarke, M., Liew, S. M., Roberts, N., & Burls, A. (2018). Patient and public involvement in the design of clinical trials: An overview of systematic reviews. *Journal of Evaluation in Clinical Practice*, *24*(1), 240–253. https://doi.org/10.1111/jep.12805

Qian, J., Wu, Y., Liu, F., Zhu, Y., Jin, H., Zhang, H., Wan, Y., Li, C., & Yu, D. (2022). An update on the prevalence of eating disorders in the general population: a systematic review and meta-analysis. *Eating and Weight Disorders*, *27*(2), 415–428. https://doi.org/10.1007/s40519-021-01162-z

Roehl, J., & Harland, D. (2022). Imposter Participants: Overcoming Methodological Challenges Related to Balancing Participant Privacy with Data Quality When Using Online Recruitment and Data Collection. *The Qualitative Report*. https://doi.org/10.46743/2160-3715/2022.5475

Rogers, K. M., Trubey, E., & Oard, D. W. (2022). A user study in a pandemic: some lessons learned. *Proceedings of the Association for Information Science and Technology*, *59*(1), 785–787. https://doi.org/10.1002/pra2.726

Salinas, M. R. (2023). Are Your Participants Real? Dealing with Fraud in Recruiting Older Adults Online. *Western Journal of Nursing Research*, *45*(1), 93–99. https://doi.org/10.1177/01939459221098468

Storozuk, A., Ashley, M., Delage, V., & Maloney, E. A. (2020). Got Bots? Practical Recommendations to Protect Online Survey Data from Bot Attacks. *The Quantitative Methods for Psychology*, *16*(5), 472–481. https://doi.org/10.20982/tqmp.16.5.p472

Teitcher, J. E. F., Bockting, W. O., Bauermeister, J. A., Hoefer, C. J., Miner, M. H., & Klitzman, R. L. (2015). Detecting, preventing, and responding to "fraudsters" in internet research: Ethics and tradeoffs. *The Journal of Law, Medicine & Ethics : A Journal of the American Society of Law, Medicine & Ethics*, *43*(1), 116–133. https://doi.org/10.1111/jlme.12200

Williams, S., Clausen, M. G., Robertson, A., Peacock, S., & McPherson, K. (2012). Methodological reflections on the use of asynchronous online focus groups in health research. *International Journal of Qualitative Methods*, *11*(4), 368–383. https://doi.org/10.1177/160940691201100405

Wright, K. B. (2006). Researching Internet-Based Populations: Advantages and Disadvantages of Online Survey Research, Online Questionnaire Authoring Software Packages, and Web Survey Services. *Journal of Computer-Mediated Communication*, *10*(3), 00–00. https://doi.org/10.1111/j.1083-6101.2005.tb00259.x