Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Reached but Not Recruited: Sociodemographic Correlates ofPrior Health and Social Care Research Participation in a Community-Recruited Sample in North West London: A Cross- Sectional Study

Sami Altalib
Imperial College London
Mohammed Adam
Imperial College London
Manisha Karki
Imperial College London
Aos Alaa
Imperial College London
Haitham Idriss
Imperial College London
Marcelle Costa Marinho
Barts Health NHS Trust
Tomas Harvey
Barts Health NHS Trust
Prital Patel
Barts Health NHS Trust
Jemima Hepburn-Scott
Barts Health NHS Trust
Carlos Romero Muñoz
Barts Health NHS Trust
Billy Hopkins
Barts Health NHS Trust
Gary Bellamy
Barts Health NHS Trust
Arti Sharma
Barts Health NHS Trust
Sandra O'Sullivan
Barts Health NHS Trust
Waseem Jerjes
Imperial College London
Austen El-Osta
a.el-osta@imperial.ac.uk

Imperial College London

Research Article

Keywords: Research participation, Health equity, Under-served groups, Ethnicity, Educational attainment, Patient and public involvement, Community-based recruitment, Cross-sectional study, United Kingdom

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11116648/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Persistent inequities in health and social care research participation undermine the representativeness and legitimacy of evidence. Participation is shaped by structured opportunity, particularly direct invitation from healthcare providers, but population-level estimates may obscure patterns within communities underserved by conventional recruitment.

Methods

We conducted a cross-sectional analysis of a purposively recruited community-based sample (N = 1,306; analytical samples 951–1,040), drawn predominantly from North West London and targeted toward groups historically underrepresented in research. Logistic regression examined predictors of active participation and of ever declining participation; perceptions of organisational communication were analysed using ordinal logistic regression. Missing data were addressed by multiple imputation by chained equations (50 datasets), with recruitment stream as an auxiliary variable.

Results

Only 31.7% of respondents had ever been invited to a research study by a healthcare provider and 32.8% had participated. Educational attainment showed the clearest gradient: relative to a university degree or above, odds of participation were lower for those educated to secondary level or below (OR 0.57, 95% CI 0.39–0.82) and for those with post-secondary or vocational education (OR 0.52, 95% CI 0.34–0.79). Other ethnicity was associated with lower odds of participation (OR 0.43, 95% CI 0.21–0.87) and disability with higher odds (OR 1.79, 95% CI 1.21–2.64). Prior work in health, care or research more than doubled the odds of participation in an exploratory extended model. One in five (21.0%) had ever declined, with higher odds among disabled respondents (OR 1.79, 95% CI 1.17–2.75) and lower odds in the Other ethnic category (OR 0.45, 95% CI 0.20–0.99). Ratings of organisational communication ran counter to participation, being more favourable among Asian and Other ethnicity respondents, those educated to secondary level or below, and men. Time constraints, lengthy processes and limited awareness were the most frequently reported barriers.

Conclusions

Within this community-recruited sample, participation remained associated with educational attainment, disability and the heterogeneous Other ethnic category, though the absence of a comparison recruitment strategy means the effect of community-based recruitment itself cannot be determined. The reported barriers support evaluating more systematic invitation pathways and measures to reduce procedural burden.

Trial registration

Not applicable.

Background

Equitable participation in health and social care research is widely regarded as a cornerstone of scientific legitimacy, ethical integrity, and external validity [1–3]. Evidence generated from populations that do not reflect the diversity of those ultimately affected by policy and clinical decisions risks entrenching inequities in care, misallocating resources, and weakening public trust in research institutions [4, 5]. Despite sustained policy emphasis on inclusion, including mandates from funders, regulators, and ethics bodies, participation in research remains socially patterned. Ethnic minority groups, individuals with lower educational attainment, migrants, people with limited English proficiency, and socioeconomically disadvantaged communities are consistently underrepresented across clinical trials, public health studies, and service evaluations [2, 6, 7]. These participation gaps are rarely attributable to a lack of interest alone; rather, they reflect structural, procedural, and communicative barriers embedded within health systems.

Existing evidence indicates that participation is shaped predominantly by structured opportunity rather than intrinsic willingness. Direct invitation from a healthcare professional is among the most powerful procedural determinants of engagement, and sociodemographic disparities commonly persist after adjustment for key covariates [1, 7–9].

The concept of “harder to reach” requires careful interrogation. It risks implying inherent disengagement or cultural resistance, when in many cases the challenge lies not in communities themselves but in the reach of research infrastructures [1, 8, 10]. Communities facing linguistic barriers, migration-related precarity, digital exclusion, disability, insecure housing, or mistrust rooted in historical inequities are often peripheral to mainstream recruitment channels. Traditional models of research enrolment rely heavily on institutional interfaces including general practice lists, hospital clinics, digital registries, or online platforms which may systematically undercapture individuals outside routine healthcare contact or those who experience structural marginalisation [11–13]. Thus, the descriptor “harder to reach” may more accurately reflect the limited reach of research systems than the dispositions of individuals.

Empirical evidence suggests that barriers to participation operate across multiple domains. Informational barriers include limited awareness of ongoing research, poor communication of study aims, and insufficient feedback of results [7, 14, 15]. Logistical barriers encompass time constraints, inflexible study procedures, transport limitations, and competing caregiving responsibilities. Structural barriers involve language discordance, digital exclusion, immigration status concerns, and mistrust of institutions perceived as distant or unresponsive. Importantly, these domains intersect: for example, linguistic exclusion can amplify informational asymmetry, while precarious employment can intensify the opportunity cost of participation [1, 16]. Moreover, underrepresentation is not uniform across all research modalities. Survey-based engagement, community consultations, and qualitative studies may reach different populations than interventional trials or registry-based recruitment.

Community-embedded recruitment strategies have been proposed as a means of addressing these inequities [1, 2, 14, 17]. Partnerships with voluntary and community sector organisations, faith-based institutions, local hubs, and culturally aligned networks may facilitate access to groups that traditional institutional channels overlook. Such approaches shift the locus of engagement from hospital or university settings into trusted community spaces. Yet evidence evaluating the demographic composition, participation patterns, and attitudinal profiles of cohorts recruited through these routes remains limited. Without systematic evaluation, it is difficult to determine whether community-engaged recruitment meaningfully alters the composition of research participants or simply reproduces existing gradients.

The present study addresses this gap by examining a purposively recruited, community-based cohort drawn predominantly from North West London which is an area characterised by substantial ethnic, linguistic, and socioeconomic diversity. This analysis focuses specifically on populations targeted to enhance representation of groups historically underrepresented in research. Rather than assuming disengagement, we examine patterns of awareness, invitation, active participation, refusal, perceived organisational communication, and self-reported barriers

within this cohort. By applying multivariable modelling alongside descriptive and bivariate analyses, we aim to identify independent predictors of participation and perceptions, while accounting for intersecting sociodemographic characteristics.

Understanding participation within harder-to-reach populations has implications beyond methodological optimisation. Research inclusion shapes whose experiences inform clinical guidelines, public health interventions, digital health tools, and service redesign. Underrepresentation can lead to evidence that is less generalisable, culturally incongruent, or insensitive to context-specific barriers. Moreover, perceptions of how research findings are communicated back to communities influence trust, reciprocity, and future engagement. If research is perceived as extractive (i.e. drawing data without visible community benefit), cycles of disengagement may be reinforced. Conversely, transparent communication and demonstrable local relevance may strengthen participation over time.

Importantly, this study does not conceptualise inclusion as a binary outcome (e.g., reached versus unreached) but as a continuum influenced by invitation pathways, institutional interfaces, and sociostructural positioning. By situating participation within a structural framework, we avoid attributing disparities solely to individual-level motivation. Instead, we examine how opportunity, access, and communication intersect to shape engagement among populations frequently positioned at the margins of research systems.

The primary aim of this study was to estimate adjusted associations between sociodemographic characteristics and self-reported prior participation in health and social care research within a purposively recruited community sample. Secondary aims were to examine factors associated with ever having declined participation, describe unadjusted sociodemographic differences in local research awareness, assess variation in ratings of organisational communication, and characterise reported barriers and facilitators. The analysis examined whether participation gradients described in national studies were also present within a community-recruited sample; it was not designed to estimate the effect of community recruitment relative to institutional recruitment. Together, these analyses aim to identify modifiable system-level levers to strengthen equitable research inclusion among harder-to-reach communities.

Methods

Study Design and Setting

A cross-sectional survey was conducted as part of the RESEARCH HEROES, a programme of work examining public awareness, experiences, facilitators and barriers regarding participation in health and social care research. The study recruited through two arms using the same instrument: community, primary care and homecare streams in North West London, reported here, and a national digital arm reported separately [18]. The study was a collaborative effort involving Imperial College London, the NIHR Clinical Research Network North West London (subsequently the NIHR North London Research Delivery Network), Al-Manaar Mosque, and the Nourish Hub.

Participant Recruitment

Participants were recruited through 14 data-collection sites and networks selected to broaden community reach; because recruitment was purposive and self-selected, the resulting sample should not be regarded as representative of the wider community. A purposive sampling approach was employed, specifically targeting underrepresented and culturally diverse groups to capture the barriers they face regarding research participation. The primary recruitment hubs included the Hillingdon Confederation, the North West London Integrated Care System Research Engagement Network Development (NWL ICS REND), various Homecare services, and the Integrated Care Board Team. To capture

seldom-heard community voices and enhance cultural diversity, recruitment extended to community-based organisations and events, including the Nourish Hub, the Museum of Brands, the Research Champions Network, and local community events. The survey was distributed both online and through direct community engagement. No a priori sample size calculation was performed; the analytical samples comprise all respondents who provided outcome data during the recruitment period. The contribution of each recruitment stream to the analytical samples is reported in **Additional file 1: Table S1**. Of the 1,308 individuals who accessed the survey, 1,306 (99.8%) provided consent, with two declining (0.2%). Of those who consented, 965 (73.9%) completed the survey in full. A further 236 (18.1%) consented but answered no substantive items, and 256 (19.6%) provided no data on any of the three outcome variables. As the survey was open and distributed through multiple channels, the number of individuals invited is unknown and a conventional response rate cannot be calculated; analytical samples therefore ranged from 951 to 1,040 rather than the full consenting cohort.

Because all analysed respondents had chosen to complete the survey, the study estimates correlates of previous research experience among survey participants and cannot directly identify the characteristics of people who were not reached or did not respond.

Data Collection Instrument

The survey instrument was designed to capture participant demographics, awareness of local research opportunities, motivations and barriers to research participation, and perceptions of how well organisations communicate research findings. The survey was made available in 15 languages to accommodate the linguistic diversity of the target population. Translations were produced using the Qualtrics translation service and reviewed by bilingual members of the study team before deployment. No formal cross-cultural validation of the instrument was undertaken. The distribution of survey completion languages is reported in Table 1. The instrument was developed de novo for the RESEARCH HEROES study and is not an adaptation of any previously published or validated questionnaire; no existing instrument combined awareness, invitation exposure, motivations, barriers and communication perceptions within a single public-facing survey. It was co-produced with researchers, community champions, voluntary and community sector organisations and research delivery partners, and pilot tested with 38 members of the public before deployment, with refinements made to terminology, question flow and length. The same instrument was administered in a companion analysis of the study's national digital arm [18]. To permit attribution of responses to their source, the instrument was deployed as a separate Qualtrics project for each recruitment stream. Item content, wording and response options were identical across projects; they differed only in the geographical item, which most community and primary care projects asked as a single North West London borough question in place of the region and conditional borough items. The full version is provided as Additional file 2. Participants were informed that completion should take no more than 10 minutes, and no items were mandatory: respondents could skip any question. The questionnaire was hosted on Qualtrics as a single block of items presented in a fixed order without randomisation. Display logic was applied so that questions about participation experience and motivation were shown only to respondents reporting previous participation, questions about reasons for declining only to those reporting a previous refusal, and the London borough item only to respondents located in a London region. Respondents were able to move backwards through the questionnaire to review and amend earlier answers before submission. Respondents answering that they did not consent were routed directly to the end of the survey. No incentive was offered for participation. No completeness or consistency checks were applied, since no items were mandatory. The questionnaire concluded with a separate, optional invitation for respondents to provide contact details if willing to be approached about a linked qualitative interview study.

Descriptive statistics

Quantitative survey data were extracted and analysed using Stata version 18. Descriptive statistics were used to summarise categorical variables, including participant demographics, awareness levels, motivations, and barriers, using frequencies and valid percentages. For questions permitting multiple responses, percentages were calculated using as the denominator the number of respondents who selected at least one option for that item, since respondents who discontinued before reaching a question cannot be distinguished from those who declined every option. Denominators are reported per item in **Additional file 1: Table S2**.

Data Preparation and Variable Recoding

Prior to analysis, extensive data cleaning and variable recoding were performed. Original string variables were converted to numeric formats with appropriate value labels. To ensure sufficient cell counts for statistical power, specific variable categories were merged based on distributional properties and theoretical relevance. The ethnicity item offered 22 response options, comprising 20 substantive categories, a free-text option and "Prefer not to say"; the substantive categories were consolidated into the five high-level 2021 Census categories for England and Wales (White, Asian, Black, Mixed, and Other) [19]. Free-text entries recorded under Other were reviewed against a prespecified rule: an entry was reassigned to a high-level census grouping only where it named a nationality, region or ancestry mapping to exactly one grouping without interpretation. Entries that were ambiguous, spanned more than one grouping, or named an Arab or Middle Eastern identity (which the Census places within Other) were retained in Other. Six of the 17 free-text entries were reassigned, reducing the Other category from 85 to 79 respondents. Educational attainment was collapsed into a three-tier ordinal variable (Up to Secondary, Post-Secondary/Vocational, and University Degree or above). The three outcome variables were derived from single survey items. Active participation was measured by "Regardless of how you learned about it, have you ever actively participated in any health and social care research studies?" (Yes/No). Refusal was measured by "Have you ever declined or refused to participate in a research study?" (Yes/No). Perceptions of communication were measured by "In your opinion, how well do health and social care organisations in your community communicate research findings to the public?", reverse-coded from the order presented in the instrument so that higher values denote more favourable ratings, from 1 (Very poorly) to 5 (Very well). None of these three items offered a "not sure" option. Local research awareness was measured by "Are you aware of health and social care research that is happening in your area?" (Yes / To some extent / No / Not sure); "Not sure" responses were combined with "No" to create a distinct non-awareness category, yielding a three-level ordinal variable (Unaware/Not sure, To some extent, Yes). "Prefer not to say" responses were treated as missing data across all variables unless otherwise specified.

Bivariate Analysis

Bivariate associations between demographic characteristics (age, gender, ethnicity, and English language skills) and awareness of local research opportunities were assessed using chi-square tests among respondents with observed awareness and data for the relevant characteristic, with Cramér's V reported as the effect-size measure.

Missing Data

Extent of Missingness

Missing data were present across both outcome and predictor variables, ranging from 20.4% for active participation to 27.2% for ratings of organisational communication among the outcomes, and from 25.0% for age to 52.1% for

previous work history among the predictors (**Additional file 1: Table S3**).

Missing Data Mechanism

For each model, the analytical sample was first restricted to respondents with observed outcome data. Little's global test [20] rejected the missing completely at random (MCAR) hypothesis across all three models (all $p < .001$). As this test assumes multivariate normality and the analysis variables are categorical, it was interpreted alongside pairwise diagnostic models rather than as definitive evidence. Subsequent diagnostic logistic regressions showed that missingness in specific predictors was predicted by observed outcome data, a pattern consistent with a Missing at Random (MAR) mechanism; MAR cannot be verified empirically and a Missing Not at Random mechanism cannot be excluded. Because complete-case analysis could introduce bias and reduce precision when missingness depends on observed variables, multiple imputation by chained equations (MICE) was used for all three models under an assumed missing-at-random mechanism.

Multiple Imputation Procedure

Fifty imputed datasets were generated for each model following the recommendation of White, Royston and Wood [21] that the number of imputations should be at least equal to the percentage of incomplete cases. Variables were imputed using methods appropriate to their measurement level: ordered logistic regression for ordinal variables; multinomial logistic regression for nominal categorical variables; and binary logistic regression for binary variables. The outcome variable was included as a predictor in all imputation models to preserve the identified MAR relationships. Because missingness was strongly clustered by recruitment stream, recruitment stream was additionally included in every imputation model as an auxiliary variable, so that the missing-at-random assumption is conditioned on the strongest observed predictor of missingness. All imputation procedures were conducted using the mi impute chained command in Stata with a fixed random seed (12345) to ensure reproducibility, and results were pooled across imputed datasets using Rubin's rules [22]. The augment option was applied to address perfect prediction arising from sparse cells in the categorical imputation models.

Model-specific missingness ranges and the diagnostic regressions testing whether each outcome predicted missingness in its predictors are reported in **Additional file 1: Table S4**.

Multivariable Regression Models

Three multivariable regression models were fitted to estimate adjusted associations between sociodemographic characteristics and the three research-engagement outcomes. Binary logistic regression was used to examine factors associated with active participation in research (Model 1) and with ever having declined or refused to participate (Model 2). Perceptions of organisational communication of research findings were analysed using an ordinal logistic (proportional odds) regression model (Model 3). The primary model for each outcome included five predictors: age group, gender, ethnicity, highest level of education, and disability status. Previous work in health, care, or research was excluded from the primary models because 52.1% of values were missing and missingness was strongly patterned by recruitment stream (chi-square(13) = 216.2, $p < .001$), with two streams contributing no analysable cases. It is a plausible consequence of research engagement as well as a determinant of it, whereas the remaining five predictors are structural characteristics fixed prior to any research contact. It is therefore reported for each outcome in an exploratory extended model, which should be interpreted with corresponding caution given the extent and patterning of the missingness. Reference categories were consistent across all three models: age 25–34 years, White ethnicity, female gender, university degree or above, no disability, and no prior health or research work history. Results are reported as odds ratios (OR) with 95% confidence intervals (CI), with statistical significance set at $p < 0.05$. For Model 3, the Brant test cannot formally evaluate the proportional odds assumption within the multiple imputation framework. The autofit procedure in gologit2 was therefore applied to complete cases and separately to five completed datasets.

The selection proved unstable: the terms released from the parallel-lines constraint differed across datasets, and the term released most frequently was the gender category comprising only 11 respondents, which is not a credible basis for a specification. A data-driven partial proportional odds specification was therefore not identifiable in these data, and an ordinal logistic (proportional odds) model was retained. No adjustment was made for multiple comparisons. Associations involving small or heterogeneous subgroups, in particular the 11 respondents identifying as non-binary or another gender and the 79 respondents in the Other ethnic category, should be regarded as exploratory. Recruitment stream was not included in the primary analysis models, because it is a feature of the sampling design rather than a characteristic of respondents, and adjusting for it removes variation in whom each stream reached, which is part of what this analysis describes. In a sensitivity analysis, each primary model was refitted with recruitment stream as a fixed effect, with streams contributing fewer than 30 analysable cases grouped into a single category.

Reporting and Ethical Considerations

This study is reported in accordance with the STROBE statement for cross-sectional studies and the CHERRIES checklist for internet-based surveys; completed checklists are provided as Additional files 3 and 4 [23, 24]. The survey was open rather than closed: it was distributed through multiple community and health service channels without a defined sampling frame, participants were self-selected, and no unique identifiers or IP-based deduplication were applied. The number of individuals who saw an invitation is therefore unknown and a conventional response rate cannot be calculated. Recruitment ran from 30 October 2023 to 11 June 2025. Participation in the survey was voluntary. Survey responses were collected without direct identifiers; at the end of the questionnaire respondents were separately invited to provide contact details if willing to be approached about a linked qualitative interview study, and these were stored separately, were not linked to the analytic dataset and are not used in the present analysis. Because full postcode and free-text responses were collected, the dataset was treated as potentially re-identifiable rather than fully anonymous. Responses were held on Imperial College London's secure Qualtrics account, with access restricted to named members of the study team. Informed consent was obtained electronically at the beginning of the survey, where participants were required to confirm their agreement to take part in the study before proceeding to the questions. The study protocol was reviewed and approved by the Imperial College Research Ethics Committee (ICREC reference 6658968) and was registered on the NIHR Central Portfolio Management System (CPMS ID 60665).

Use of large language models

Large language models were used during the preparation of this manuscript: Claude (Anthropic) supported the drafting and revision of analytical code and assisted with editing and reviewing manuscript drafts. ChatGPT (OpenAI) was used by some authors to review drafts. All analyses were specified, executed and verified by the authors, who take full responsibility for the content and interpretation of this work.

Results

Participant Characteristics

Of the 1,306 consenting participants, 965 (73.9%) completed the survey in full. Participant characteristics are shown in Table 1. The vast majority of participants completed the survey in English (94.4%), with Arabic being the second most common language (3.1%), followed by French (0.7%) and Spanish (0.5%). Other languages, including Bengali, Gujarati, Punjabi, Portuguese, Romanian, Somali, and Tamil, were each used by no more than three participants (≤ 0.2%). Of the 979 participants who reported their age, the most represented age group was 25–34 years (22.3%), followed by 35–44 years (19.5%) and 65 years and over (17.1%). Among those who disclosed their gender identity, females constituted the majority (68.2%). Regarding ethnicity, the sample was diverse. White ethnic groups collectively represented 31.0% of respondents, with the majority identifying as White British/English/Welsh/Scottish/ Northern Irish (23.2%). Asian or

Asian British participants constituted 28.5% of the sample, with Indian background being the largest subgroup (17.4%). Black, African, Caribbean or Black British participants accounted for 22.9%, with Black African being the most common subgroup (11.9%). Mixed or multiple ethnic groups represented 5.1%, and participants identifying as Arab or other ethnic groups comprised 8.7%. Among the 953 participants who answered the question on English language proficiency, the majority reported being fluent in English (66.7%), while 12.5% described their skills as advanced, 15.2% as intermediate and 5.6% as beginner.

Regarding employment status, participants could select multiple options. Employed full-time was the most common status (30.2%), followed by retired (11.0%) and employed part-time (9.3%). Unemployed participants accounted for 8.0%, while smaller proportions were self-employed (3.3%), students (2.7%), or volunteers (2.7%). Over half held a university degree or above (54.8%), making this the most common level of education in the sample. Secondary or high school education was reported by 18.0%, followed by apprenticeship or vocational training (9.5%) and A-levels or equivalent (8.2%). Concerning disability status, the majority of respondents did not consider themselves to have a disability as defined by the Equality Act 2010 (77.0%). Finally, among 625 participants who responded to the question on health, care, or research work experience, slightly more than half had never worked in these sectors (52.2%), while 47.8% reported current or past experience in health, care, or research.

Table 1
Participants' characteristics (N = 1306)

Variable	Frequency (Percentage)
User Language for completing the survey	
Arabic (□□□□□□□)	41 (3.1%)
Bengali (□□□□□)	1 (0.1%)
English (English)	1233 (94.4%)
Spanish-Spain (Español)	7 (0.5%)
French (Français)	9 (0.7%)
Gujarati (□□□□□□□)	3 (0.2%)
Punjabi-India (□□□□□□)	3 (0.2%)
Portuguese-Brazil (Português)	2 (0.2%)
Romanian (Română)	3 (0.2%)
Somali (Af-Soomaali)	2 (0.2%)
Tamil (□□□□□)	2 (0.2%)
Please select from the age range below	
N-Miss	327
17–24	78 (8.0%)
25–34	218 (22.3%)
35–44	191 (19.5%)
45–54	163 (16.6%)
55–64	162 (16.5%)
65 & over	167 (17.1%)
Which of the following best describes your gender identity?	
N-Miss	327
Female	668 (68.2%)
Male	283 (28.9%)
Non-binary	5 (0.5%)
Other	6 (0.6%)
Prefer not to say	17 (1.8%)
What is your ethnic group? Choose one option that best describes your ethnic group or background. These fields are taken from the 2021 Census.	
N-Miss	333

Variable	Frequency (Percentage)
User Language for completing the survey	
White – English/Welsh/ Scottish/ Northern Irish/British	226 (23.2%)
White - Irish	18 (1.9%)
White - Roma	1 (0.1%)
White - Any other White background	56 (5.8%)
Asian/Asian British - Indian	169 (17.4%)
Asian/Asian British - Bangladeshi	19 (2.0%)
Asian/Asian British - Chinese	11 (1.1%)
Asian/Asian British - Pakistani	44 (4.5%)
Asian/Asian British - Any other Asian background	34 (3.5%)
Black/African/Caribbean/Black British - African	116 (11.9%)
Black/African/Caribbean/Black British - Caribbean	75 (7.7%)
Black/African/Caribbean/Black British – Any other Black/African/Caribbean background	32 (3.3%)
Mixed/Multiple Ethnic Groups - White & Black African	21 (2.2%)
Mixed/Multiple Ethnic Groups - White & Black Caribbean	8 (0.8%)
Mixed/Multiple Ethnic Groups- White & Asian	6 (0.6%)
Mixed/Multiple Ethnic Groups - Any other mixed/multiple ethnic backgrounds	15 (1.5%)
Other Ethnic Group - Arab	43 (4.4%)
Other Ethnic Group - Any other ethnic group	42 (4.3%)
Prefer not to say	37 (3.8%)
How would you describe your English language skills	
N-Miss	353
Beginner	53 (5.6%)
Intermediate	145 (15.2%)
Advanced	119 (12.5%)
Fluent	636 (66.7%)
In which region of England are you currently located?	
N-Miss	128
East Midlands	4 (0.3%)
East of England	12 (1.0%)
Greater Manchester	2 (0.2%)

Variable	Frequency (Percentage)
User Language for completing the survey	
Kent, Surrey and Sussex	10 (0.8%)
North East and North Cumbria	1 (0.1%)
North Thames (London (North), Chelmsford, Luton)	32 (2.7%)
North West Coast	2 (0.2%)
North West London (London (West), Harrow, Hillingdon)	1004 (85.2%)
South London	28 (2.4%)
South West Peninsula (Plymouth, Exeter, Truro)	9 (0.8%)
Thames Valley and South Midlands	1 (0.1%)
Wessex (Southampton, Portsmouth, Bournemouth)	1 (0.1%)
West Midlands	2 (0.2%)
West of England	25 (2.1%)
Other	45 (3.8%)
What is your current employment status? (You can select one or more)	
Employed full-time	394 (30.2%)
Employed part-time	121 (9.3%)
Self-employed	43 (3.3%)
Retired	144 (11.0%)
Seeking opportunities currently	20 (1.5%)
Unemployed	105 (8.0%)
Student	35 (2.7%)
Unable to work	27 (2.1%)
Volunteer	35 (2.7%)
Full-time parent	16 (1.2%)
Caregiver	21 (1.6%)
Experiencing homelessness	8 (0.6%)
Prefer not to say	13 (1.0%)
What is your highest level of education?	
N-Miss	367
Basic level	60 (6.4%)
Secondary/ high school	169 (18.0%)

Variable	Frequency (Percentage)
User Language for completing the survey	
A Levels or equivalent	77 (8.2%)
Apprenticeship/ training	89 (9.5%)
University degree or above	515 (54.8%)
Prefer not to say	29 (3.1%)
Disability is defined by the Equality Act 2010 as a physical or mental impairment which has a substantial & long-term (at least 12 months) effect on your ability to carry out normal day-to-day activities. Do you consider yourself to be a person with a disability?	
N-Miss	356
Yes	163 (17.2%)
No	732 (77.0%)
Prefer not to say	55 (5.8%)
Do you currently or have you ever worked in health, care or research?	
N-Miss	681
Yes	299 (47.8%)
No	326 (52.2%)
Percentages are valid percentages calculated on non-missing responses, except for employment status, which permitted multiple selections and is expressed as a percentage of all 1,306 respondents. Prefer not to say responses are shown here but were treated as missing in all regression models.	

Awareness and Perceptions of Research

Among the 695 respondents who answered this item, participants were most familiar with Clinical/Healthcare Research (63.0%), Mental Health Research (55.7%) and Public Health Research (51.4%). Awareness of Economic Research (31.7%) and Charitable Research (27.6%) was comparatively lower, and 13.1% stated they were not aware of any of the listed research types. Regarding health and social care research within their local area, 57.5% reported no awareness or were unsure (49.7% no and 7.8% not sure), consistent with the recoded non-awareness category; a further 21.6% were aware to some extent and 20.9% reported definite awareness. Despite this, an overwhelming majority (87.9%) believed that health and social care research could benefit themselves, their families and their local communities. Of the 1,045 respondents who selected priority areas for their local community, mental health and wellbeing (68.8%) and health conditions or disease (65.5%) were chosen most often, followed by education (43.9%), crime and safety (38.9%), women's health (35.4%) and social isolation and loneliness (35.2%); **Additional file 1: Table S2**.

Research Participation and Experiences

Among the 1,041 respondents answering this item, fewer than one third (31.7%) had ever been informed about or invited to participate in a health and social care research study by a healthcare provider, 59.2% had not and 9.1% were unsure. Of the 1,040 respondents answering the participation item, 341 (32.8%) reported having actively participated in research. Among those who had participated, the experience was largely positive, with 72.0% reporting being either extremely satisfied (36.2%) or somewhat satisfied (35.8%). A majority of participants (68.2%) felt they were given

sufficient information before agreeing to take part, while 23.7% felt informed: "to some extent". Among the 322 participants who gave a reason, the primary motivations were a desire to help others (60.6%) and interest in the research topic (53.7%). Access to new treatments and technologies motivated 29.5%, while 23.6% were encouraged by a professional, friend or family member. Financial compensation was a motivator for only 12.7%. Conversely, 21.0% of respondents reported having declined or refused to participate in a research study. Of the 214 respondents who had declined, 197 gave at least one reason. The most common were a lack of time (54.8%), privacy and confidentiality concerns (28.4%) and a lack of interest in the topic (27.4%). Fear of potential risks or side effects (26.4%) and a lack of information about the study (23.9%) were also notable deterrents (**Additional file 1: Table S2**).

Facilitators and Preferences for Engagement

Participants identified several factors that could encourage broader participation in research studies. Among the 961 respondents answering this item, the most frequently selected facilitators were access to better care, new treatments and devices (53.1%), being paid for participation (52.5%), making surveys shorter with simplified medical terminology (49.7%) and gaining a better understanding of their own health conditions (49.6%). Better access to research findings was also a strong motivator (46.4%). When asked about their preferred methods of participation, completing online surveys was the most popular choice by a wide margin (66.5% of 946 respondents). Attending workshops or meetings (43.8%), engaging in face-to-face interviews (40.1%), participating in online interviews (36.2%) and joining focus groups (35.3%) were also favoured, while smaller proportions expressed interest in participating in research design (27.2%) or clinical trials (27.1%). For receiving information about studies and findings, email was heavily preferred (74.2% of 952 respondents), followed by online portals or apps (24.5%), in-person communication (23.6%) and postal mail (22.7%); **Additional file 1: Table S2**.

Barriers and Organisational Communication

Participants cited several barriers that hinder their involvement in health and social care research. Among the 942 respondents answering this item, lengthy research processes requiring too much involvement (45.3%) and time constraints or personal commitments (40.1%) were the most prominent obstacles. A lack of general awareness about research was highlighted by 35.2%, and 35.0% noted a lack of relevant studies for their specific health conditions. Concerns regarding safety and side effects (30.5%) and mistrust in health and social care organisations (22.1%) were also reported. Public perception of how well local health and social care organisations communicate research findings is mixed, leaning negative. While 40.3% of respondents held a neutral view, 31.8% felt communication was either poor (20.7%) or very poor (11.1%). Only 27.9% felt organisations communicated well (17.6%) or very well (10.3%). Positively, completing the survey itself served as an educational tool; 74.8% of respondents indicated that the survey increased their awareness of health and care research and participation opportunities. The vast majority of respondents (94.9%) found the survey easy to understand and complete (**Additional file 1: Table S2**).

Association Between Demographics and Local Research Awareness

Bivariate analyses were conducted to examine the distribution of local research awareness across key demographic subgroups (**Additional file 1: Table S5**). A statistically significant association was observed between age and research awareness ($X^2(10) = 26.04$, $p = 0.004$), although the effect size was small (Cramér's V = 0.118). In this unadjusted analysis, definitive awareness was reported by 29.9% of respondents aged 17–24 years, whereas more than 64% of respondents aged 55 years or older were unaware or unsure; the association was small (Cramér's V = 0.118) and should not be interpreted as an independent age effect. In contrast, local research awareness was relatively evenly distributed across gender ($X^2(4) = 7.02$, $p = 0.135$), ethnicity ($X^2(8) = 12.26$, $p = 0.140$), and English language skill levels ($X^2(6) = 9.71$, $p = 0.137$), with no statistically significant differences observed between these subgroups.

Active Participation in Research

A binary logistic regression was conducted to examine sociodemographic and background factors associated with active participation in research (N = 1,040); Fig. 1 and **Additional file 1: Table S6**. The overall model was statistically significant (F(14, 45369.0) = 2.69, p < .001). The average relative variance increase (RVI = 0.131) and largest fraction of missing information (FMI = 0.263) indicated modest additional uncertainty associated with missing data. Four predictors were significantly associated with active participation. Respondents with a disability were more likely to actively participate than those without (OR = 1.79, 95% CI [1.21, 2.64], p=.003). Regarding education, respondents educated to post-secondary or vocational level had significantly lower odds of active participation compared to those with a university degree or above (OR = 0.52, 95% CI [0.34, 0.79], p=.002). Those educated to secondary level or below showed a comparable reduction in odds (OR = 0.57, 95% CI [0.39, 0.82], p=.003), so the education gradient was present at both levels. Respondents of "Other" ethnicity had significantly lower odds of participation than White respondents (OR = 0.43, 95% CI [0.21, 0.87], p=.019). No statistically significant associations were observed for age group, gender, or the remaining ethnicity categories (all p > .05). In an extended model additionally adjusting for prior work in health, care, or research, work history was strongly associated with participation (OR = 2.23, 95% CI [1.54, 3.23], p < .001), F(15, 27792.5) = 3.86, p < .001. Both education terms attenuated: secondary level or below was no longer significant (OR = 0.69, 95% CI [0.47, 1.02], p=.063) and post-secondary or vocational education remained so (OR = 0.58, 95% CI [0.37, 0.89], p=.013). Disability was essentially unchanged (OR = 1.89, 95% CI [1.27, 2.82], p=.002), as was Other ethnicity (OR = 0.48, 95% CI [0.23, 0.98], p=.045). Full estimates are reported in **Additional file 1: Table S7**.

Refusal to Participate

A binary logistic regression was performed to investigate sociodemographic and background predictors of ever refusing or declining an invitation to participate in research (N = 1,021) (Fig. 2 and **Additional file 1: Table S8**). The model reached statistical significance (F(14, 51490.2) = 2.08, p=.010), with an average RVI of 0.122 and largest FMI of 0.302, indicating modest variance attributable to missing data. Apart from the unstable estimate for non-binary or another gender, two predictors reached statistical significance. Estimates for the 11 respondents identifying as non-binary or another gender were unstable and are reported in **Additional file 1: Table S8** without interpretation. Having a disability was independently associated with greater likelihood of having refused participation compared to those without a disability (OR = 1.79, 95% CI [1.17, 2.75], p=.008). Respondents classified under the Other ethnic group were significantly less likely to have refused participation than White respondents (OR = 0.45, 95% CI [0.20, 0.99], p=.048). This estimate was mildly sensitive to the ethnicity recoding rule: with free-text entries left unrecoded it was OR = 0.40 (95% CI [0.18, 0.88], p=.023), so the reclassification of six respondents attenuates the estimate slightly without altering its direction or significance. Age, education, and all remaining ethnicity categories (Asian, Black, and Mixed) showed no statistically significant association with declining research participation (all p> .05). In an extended model additionally adjusting for prior work in health, care, or research, work history was not associated with having declined participation (OR = 1.34, 95% CI [0.89, 2.01], p=.161) and no other estimate changed materially.

Organisational Communication

An ordinal logistic regression model was estimated to examine sociodemographic and background predictors of how respondents rated organisational communication of research findings (N = 951) (Fig. 3 and **Additional file 1: Table S9**). The model was statistically significant (F(14, 386223.1) = 6.43, p < .001). The average RVI (0.034) and largest FMI (0.179) indicated low variance inflation due to missing data, with fractions of missing information for the ethnicity terms ranging from 0.03 to 0.07. Odds ratios above 1 indicate higher odds of a more favourable rating, and under the proportional odds assumption, a single odds ratio applies across all response thresholds. Respondents educated to

secondary level or below had significantly higher odds of rating organisational communication favourably than those with a university degree or above (OR = 1.63, 95% CI [1.20, 2.20], p=.002). Regarding ethnicity, both Asian (OR = 1.61, 95% CI [1.17, 2.22], p=.003) and Other ethnicity respondents (OR = 2.03, 95% CI [1.25, 3.29], p=.004) rated communication more favourably than White respondents. Men also rated communication more favourably than women (OR = 1.56, 95% CI [1.19, 2.04], p=.001). In contrast, respondents aged 55 to 64 (OR = 0.52, 95% CI [0.35, 0.78], p=.002) and 65 and over (OR = 0.61, 95% CI [0.41, 0.90], p=.013) had lower odds of a favourable rating than those aged 25 to 34, although the estimate for respondents aged 65 and over did not persist after adjustment for recruitment stream. Estimates for the 11 respondents identifying as non-binary or another gender were unstable and are reported in **Additional file 1: Table S9** without interpretation. Disability status (OR = 0.74, 95% CI [0.53, 1.02], p=.067), post-secondary or vocational education, and the remaining age and ethnicity categories showed no statistically significant associations (all p > .05). In a sensitivity analysis refitting each primary model with recruitment stream as a fixed effect, estimates for education, disability, gender and ethnicity in the participation and communication models were materially unchanged (**Additional file 1: Table S10**). Two estimates were attenuated: the association between Other ethnicity and refusal (OR 0.55, 95% CI [0.24, 1.25], p = .152) and the lower odds of a favourable communication rating among respondents aged 65 and over (OR 0.77, 95% CI [0.51, 1.16], p = .215). Both are therefore reported as exploratory.

Discussion

Principal Findings

This study examined patterns of awareness, participation, refusal, and perceptions of organisational communication within a community-embedded cohort enriched for populations that national studies consistently under-recruit. Three principal findings emerge. First, even within a sample deliberately designed to enhance inclusion of ethnically and socioeconomically diverse communities, participation in health and social care research remained socially patterned. Educational attainment was the clearest independent gradient, present at both levels below a university degree, and respondents in the Other ethnic category had lower odds of participation than White respondents. Prior work in health, care, or research more than doubled the odds of engagement in an exploratory extended model and absorbed much of the education signal when included. Second, perceptions of how research is communicated ran in the opposite direction to participation: they were most favourable among the groups that participated least, and least favourable among respondents aged 55 to 64. Third, the most frequently cited barriers to participation were procedural and informational, including lengthy research processes, time constraints, and lack of awareness, rather than explicit rejection of research as a social enterprise. This accords with reviews finding procedural and informational obstacles to be reported more frequently than mistrust, even in populations where mistrust is often assumed to predominate [1, 7, 15]. Taken together, these findings suggest that while community-embedded recruitment may broaden initial contact, it does not, in isolation, dismantle the deeper structural barriers that shape research engagement. Participation appears to remain mediated by institutional familiarity, informational access, and the effort a study demands of those it invites, consistent with evidence that structured opportunity rather than willingness drives participation gradients [1, 7].

Community-Based Recruitment and Structural Persistence

Community-engaged recruitment strategies are often advanced as a corrective to the institutional narrowness of traditional research enrolment pathways. By situating research outreach within trusted community settings, including voluntary sector organisations, faith-based hubs, and culturally aligned networks, such approaches aim to reduce mistrust, increase visibility, and improve accessibility. The present findings provide partial support for this rationale:

the cohort achieved substantial ethnic diversity and engaged individuals across multiple language backgrounds. However, participation gradients persisted.

The association between prior work in health, care, or research and active participation, observed in an exploratory extended model, is instructive despite the caution its missingness demands. Individuals with prior institutional exposure are likely to possess higher baseline research literacy, greater familiarity with ethical safeguards, and clearer understanding of potential benefits and risks. They may also occupy professional or social networks where research participation is normalised. That this predictor remained strong within a community-based cohort indicates that structural proximity to research institutions continues to shape engagement, even when recruitment occurs outside formal healthcare settings, although the extent of missingness on this item means the magnitude should not be over-read. This tempers reports that community engagement strategies substantially improve recruitment among under-served groups [17], suggesting that such strategies broaden contact more readily than they change who converts an invitation into participation [10]. In effect, community outreach may widen the top of the funnel, but conversion to active participation remains conditioned by pre-existing institutional capital.

Educational attainment showed a similar gradient. Respondents with post-secondary or vocational education and those educated to secondary level or below both had significantly lower odds of active participation than those holding a university degree or above, a gradient that was clearest once work history was excluded from the model. Education likely operates through multiple pathways: health literacy, confidence in navigating research documentation, comprehension of consent materials, and perceived self-efficacy in engaging with academic processes. The persistence of this gradient suggests that procedural and communicative complexity within research design may disproportionately disadvantage individuals without higher education credentials, even when they are successfully reached through community channels.

Respondents in the Other ethnic category were less likely to have participated and, in the primary model, less likely to have ever declined, although the latter estimate did not persist after adjustment for recruitment stream. Since declining presupposes being asked, this combination is hard to reconcile with reluctance. Invitation and participation were recorded as separate self-reported ‘ever’ measures, so the present analysis cannot establish invitation-to-participation conversion; the lower participation odds in the heterogeneous Other ethnic category should therefore be interpreted as an exploratory association rather than evidence of a post-invitation mechanism. This is consistent with umbrella review evidence identifying post-invitation factors, including trust, comprehension and study design, as principal determinants of enrolment among ethnic minority populations [15]. That only 31.7% of respondents had ever been invited to take part in research, within a cohort recruited specifically through community channels, suggests structural invitation pathways remain limited even in community-engaged contexts.

Perceptions of Organisational Communication

Perceptions of organisational communication did not track participation. Respondents in the Other ethnic category had significantly lower odds of active participation than White respondents, yet both Asian and Other ethnicity respondents rated communication more favourably. Respondents educated to secondary level or below did the same, despite participating less. The least favourable ratings came from respondents aged 55 to 64; the corresponding estimate for those aged 65 and over was attenuated after adjustment for recruitment stream.

This divergence is informative. At the group level, communication ratings did not parallel participation patterns; however, the study did not directly model the individual-level association between communication ratings and prior participation, so explanations based on exposure, expectations, or response style remain hypotheses. Respondents with the least contact with research infrastructure have the fewest reference points against which to judge it and may also be more prone to acquiescence in survey responses. Older respondents reported less favourable communication

ratings, but healthcare-contact frequency was not measured, so the mechanism underlying this age pattern remains uncertain. Insufficient feedback of results is among the most consistently reported informational barriers in the literature [14, 15], and older adults are also disproportionately affected by digitally mediated communication channels [11, 12].

Two implications follow. Favourable ratings among groups that participate least should not be read as evidence that communication is working, since they may simply reflect limited engagement with the system being rated. And the age gradient points to a concrete, addressable failure: respondents in late working age, who have substantial healthcare contact, report the poorest feedback on research findings. Improving dissemination to older adults through routine clinical channels is a more specific target than a general call for culturally responsive communication. Dedicated language access provision offers a comparable specificity for linguistically diverse populations [16].

Heterogeneity within the aggregated ethnic categories warrants caution here. Collapsing 21 original categories into five high-level groups improves statistical power but reduces specificity, and the residual Other category is particularly mixed (see Limitations).

Disability and Participation Dynamics

An important and somewhat counterintuitive finding was that respondents with a disability had higher odds of active participation and were also more likely to have declined participation. This dual association suggests a more complex engagement dynamic. Individuals with disabilities may be more frequently invited to participate in research due to clinical contact or condition-specific studies. Increased exposure to invitation may raise both participation and refusal probabilities. Alternatively, individuals with disabilities may have heightened motivation to contribute to research relevant to their conditions, while also exercising agency in declining studies perceived as burdensome or misaligned with their priorities.

These findings challenge simplistic narratives that position disability solely as a barrier to research inclusion. Instead, it highlights the importance of differentiating between access, agency, and burden. Participation among disabled populations may be facilitated when studies are accessible, relevant, and responsive to lived experience. Conversely, refusal may signal misalignment between research design and participant needs rather than disengagement per se. Disabled people are frequently grouped with other under-served populations in inclusion frameworks [2], but this combination of elevated participation alongside elevated refusal does not appear to have been described previously and warrants direct investigation.

Refusal Patterns and Agency

Approximately one in five respondents reported having declined participation in a research study. Disability status was independently associated with higher odds of refusal. Importantly, refusal was most commonly attributed to lack of time, privacy concerns, lack of interest in the topic, and insufficient information.

These reasons indicate that refusal is not necessarily an expression of mistrust or opposition to research. Rather, it reflects opportunity costs and informational thresholds. Time scarcity, competing caregiving responsibilities, and precarious employment may amplify the cost of participation. Privacy concerns and perceived insufficient information point to communicative clarity and trust calibration. In this context, refusal may represent rational decision-making within constrained circumstances.

Barriers and Procedural Burden

The most frequently cited barriers, lengthy processes, time constraints, and lack of awareness, align with the multivariable findings. Procedural burden appears central. Research studies often require extended consent forms, multiple visits, complex questionnaires, or technical language. While ethically justified safeguards are essential, cumulative complexity may deter participation among individuals with limited time or lower confidence navigating formal documentation.

Mistrust in health and social care organisations was cited by 22.1% of respondents and cultural or religious reasons by 8.4%, both lower than procedural barriers such as lengthy processes (45.3%) and time constraints (40.1%). Mistrust is therefore not negligible, but it was less commonly reported than design-related obstacles. These findings challenge narratives that attribute underrepresentation primarily to cultural resistance. Instead, the data suggest that structural and procedural design features may be more salient obstacles. Simplification of materials, flexible scheduling, compensation for time, and clearer articulation of relevance may therefore yield greater equity gains than outreach messaging alone.

Implications for Policy and Practice

The findings have several implications for research governance and health system design. First, community engagement should be conceptualised as a necessary but insufficient condition for equitable inclusion. Without parallel reform of invitation infrastructure — including routine prompts within primary care systems, transparent eligibility communication, and systematic follow-up — outreach risks becoming episodic rather than structural.

Second, procedural simplification and linguistic responsiveness are critical. Reducing survey length, minimising technical terminology, and offering translated materials were identified by participants as facilitators. Designing research with accessibility in mind from inception, rather than retrofitting inclusivity, may reduce education-related gradients.

Third, communication of research findings requires attention. Feedback loops, including community briefings, accessible summaries, and digital dissemination tailored to local contexts, may strengthen relational trust. Inclusion should extend beyond enrolment to encompass reciprocal knowledge exchange.

Strengths and Limitations

This study benefits from an ethnically diverse cohort and robust statistical modelling with multiple imputation to address missing data. The inclusion of refusal and communication outcomes extends beyond binary participation metrics, offering a multidimensional view of engagement. The cross-sectional design precludes causal inference. The consolidation of ethnic categories may mask intra-group heterogeneity. Self-reported measures of participation and awareness are subject to recall bias. Missing data, particularly for work history, required imputation, introducing uncertainty despite methodological safeguards. Finally, while community-based recruitment enhanced diversity, the sample remains geographically concentrated in North West London and may not generalise nationally.

The Other ethnicity category was associated with lower odds of participation and, in the primary model only, with lower odds of refusal, yet remains heterogeneous, comprising 43 respondents identifying as Arab alongside free-text and unspecified entries. Applying the prespecified recoding rule reassigned six respondents and reduced the category from 85 to 79. This attenuated the refusal estimate slightly, from OR 0.40 ($p=.023$) to OR 0.45 ($p=.048$), without altering its direction or significance, but the residual category remains mixed and this estimate should be interpreted with that in mind. Missingness was strongly clustered by recruitment stream, both for the participation outcome ($\chi^2(13) = 165.1$, $p < .001$) and for work history ($\chi^2(13) = 216.2$, $p < .001$). Recruitment stream was therefore included in the imputation model as an auxiliary variable, although the missing-at-random assumption cannot be verified

empirically and a missing-not-at-random mechanism cannot be excluded. Adjusting for recruitment stream attenuated two estimates, indicating that part of the apparent association between ethnicity and refusal, and between older age and communication ratings, reflected differences in whom each recruitment stream reached rather than characteristics of respondents themselves. The remaining estimates were unchanged, but this illustrates that in a purposively recruited sample the recruitment pathway is itself a source of variation. Estimates for the 11 respondents identifying as non-binary or another gender are unstable and should not be read as reliable effect sizes. The proportional odds assumption for Model 3 cannot be formally tested within the imputation framework. Autofit selections on complete cases and on five completed datasets were inconsistent, and the term released most often was the gender category comprising 11 respondents, so a proportional odds model was retained rather than a partial specification driven by a sparse cell. Finally, 256 respondents (19.6%) consented but provided no outcome data, and the survey was open and distributed through multiple channels, so self-selection towards individuals already favourably disposed to research is likely.

Future Directions

Future research should examine longitudinal engagement trajectories to assess whether community-based recruitment fosters sustained participation over time. Linking survey responses with invitation data from healthcare systems could clarify the relative contribution of exposure versus willingness. More granular qualitative work may illuminate why perceptions of organisational communication diverge from participation patterns, and in particular why respondents in late working age rate the feedback of research findings most poorly, and why this pattern is less clear among those aged 65 and over. Additionally, experimental evaluation of simplified consent processes, culturally tailored communication strategies, and compensation models could test whether procedural modifications attenuate educational and ethnic gradients.

Conclusion

The community-recruited sample included substantial ethnic diversity, but prior research participation remained socially patterned. The prominence of time constraints, lengthy processes, and limited awareness identifies plausible, modifiable targets for future intervention studies. More equitable inclusion is likely to require community engagement alongside durable access pathways and participant-centred study design.

Abbreviations

ARC
Applied Research Collaboration
CHERRIES
Checklist for Reporting Results of Internet E-Surveys
CI
confidence interval
CPMS
Central Portfolio Management System
FMI
fraction of missing information
ICREC
Imperial College Research Ethics Committee
MAR

missing at random
MCAR
missing completely at random
MICE
multiple imputation by chained equations
MNAR
missing not at random
NIHR
National Institute for Health and Care Research
OR
odds ratio
RVI
relative variance increase
STROBE
Strengthening the Reporting of Observational Studies in Epidemiology

Declarations

Ethics approval and consent to participate

The study protocol was reviewed and approved by the Imperial College Research Ethics Committee (ICREC reference 6658968) and registered on the NIHR Central Portfolio Management System (CPMS ID 60665). Participation was voluntary. Survey responses were collected without names or contact details. espondents were separately invited at the end of the questionnaire to provide contact details if willing to be approached about a linked qualitative interview study; these were stored separately from survey responses and are not used in the present analysis. Full postcode was collected for geographical analysis and was not used to identify individuals. Informed consent was obtained electronically at the start of the survey, where participants confirmed their agreement to take part before proceeding. All methods were carried out in accordance with the Declaration of Helsinki and the UK Policy Framework for Health and Social Care Research.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. Austen El-Osta is supported by the National Institute for Health and Care Research (NIHR) Applied Research Collaboration (ARC) Northwest London. The views expressed are those of the authors and not necessarily those of the NHS or the NIHR or the Department of Health and Social Care.

Author Contribution

AEO, SOS and AS conceived the study. AEO, HI, SOS and AS designed the study. MK, AA, MCM, TH, PP, JHS, CRM, BH and GB contributed to recruitment and data acquisition. SA led the data analysis with support from MA. SA, MA, HI, WJ and AEO interpreted the data. AEO drafted the manuscript with support from SA. All authors critically reviewed the manuscript and read and approved the final version. AEO is the guarantor.

Acknowledgement

The authors thank everyone who took part in the survey and the community and health service partners who supported its distribution. We also thank Skills for Care for their general support.

Data Availability

The datasets generated and analysed during this study are not publicly available because the combination of recruitment stream, survey language, ethnicity, disability status and free-text responses could permit re-identification. De-identified data and the Stata do-file used to generate all reported estimates are available from the corresponding author on reasonable request, subject to the terms of the study ethical approval.

References

1. Bonevski B, Randell M, Paul C, Chapman K, Twyman L, Bryant J, et al. Reaching the hard-to-reach: a systematic review of strategies for improving health and medical research with socially disadvantaged groups. BMC Med Res Methodol. 2014;14:42. 10.1186/1471-2288-14-42.
2. Treweek S, Banister K, Bower P, Cotton S, Devane D, Gardner HR, et al. Developing the INCLUDE Ethnicity Framework-a tool to help trialists design trials that better reflect the communities they serve. Trials. 2021;22(1):337. 10.1186/s13063-021-05276-8.
3. Idnay B, Fang Y, Stanley E, Ruotolo B, Chung WK, Marder K, et al. Promoting equity in clinical research: The role of social determinants of health. J Biomed Inf. 2024;156:104663. 10.1016/j.jbi.2024.104663.
4. Ocloo J, Matthews R. From tokenism to empowerment: progressing patient and public involvement in healthcare improvement. BMJ Qual Saf. 2016;25(8):626–32. 10.1136/bmjqs-2015-004839.
5. Warner E, Marron JM, Peppercorn JM, Abel GA, Hantel A. Shifting from Equality toward Equity: Addressing Disparities in Research Participation for Clinical Cancer Research. J Clin Ethics. 2024;35(1):8–22. 10.1086/728144.
6. Redwood S, Gill PS. Under-representation of minority ethnic groups in research–call for action. Br J Gen Pract. 2013;63(612):342–3. 10.3399/bjgp13X668456.
7. Hussain-Gambles M, Atkin K, Leese B. Why ethnic minority groups are under-represented in clinical trials: a review of the literature. Health Soc Care Community. 2004;12(5):382–8. 10.1111/j.1365-2524.2004.00507.x.
8. Hussain-Gambles M, Leese B, Atkin K, Brown J, Mason S, Tovey P. Involving South Asian patients in clinical trials. *Health Technol Assess*. 2004;8(42):iii, 1–109. 10.3310/hta8420
9. Van Der Ven L, Van Baar J, Wispelweij R, Jambroes M. How to involve ‘hard to reach’ population groups in research projects. Eur J Pub Health. 2022;32(Supplement3). 10.1093/eurpub/ckac129.476.

10. Bekteshi V, Sifat M, Kendzor DE. Reaching the unheard: overcoming challenges in health research with hard-to-reach populations. Int J Equity Health. 2024;23(1):61. 10.1186/s12939-024-02145-z.
11. Heponiemi T, Jormanainen V, Leemann L, Manderbacka K, Aalto AM, Hyppönen H. Digital Divide in Perceived Benefits of Online Health Care and Social Welfare Services: National Cross-Sectional Survey Study. J Med Internet Res. 2020;22(7):e17616. 10.2196/17616.
12. Turnbull J, MacLellan J, Churruca K, Ellis LA, Prichard J, Browne D, et al. A multimethod study of NHS 111 online. Health Soc Care Deliv Res. 2023;11(5):1–104. 10.3310/ytrr9821.
13. Moseson H, Kumar S, Juusola JL. Comparison of study samples recruited with virtual versus traditional recruitment methods. Contemp Clin Trials Commun. 2020;19:100590. 10.1016/j.conctc.2020.100590.
14. Dawes P, Arru P, Corry R, McDermott JH, Garlick J, Guest H, et al. Patient and public involvement in hearing research: opportunities, impact and reflections with case studies from the Manchester Centre for Audiology and Deafness. Int J Audiol. 2024;63(2):146–54. 10.1080/14992027.2022.2155881.
15. Pardhan S, Sehmbi T, Wijewickrama R, Onumajuru H, Piyasena MP. Barriers and facilitators for engaging underrepresented ethnic minority populations in healthcare research: an umbrella review. Int J Equity Health. 2025;24(1):70. 10.1186/s12939-025-02431-4.
16. Gabler BS, Barrios A, Kakishita S, Cufino D, Frost CJ. Increasing diversity in research through dedicated language access services. Contemp Clin Trials. 2021;106:106439. 10.1016/j.cct.2021.106439.
17. Vance KG, Pedelty J, Van Gorp BJ, Vance CGT, Johnson EM, Jiang F, et al. Community engagement strategies improve recruitment and enrollment in a pragmatic clinical trial. J Clin Transl Sci. 2025;9(1):e184. 10.1017/cts.2025.10103.
18. El-Osta A, Adam M, Altalib S, Karki M, Alaa A, Idriss H et al. Invitation architecture and inequitable access to health research: a large cross-sectional study of 43,005 UK adults. Research Square. 2026. Preprint. 10.21203/rs.3.rs-10757130/v1
19. Office for National Statistics. Ethnic group, national identity, language, and religion: Census 2021 in England and Wales. Newport: Office for National Statistics. 2022. https://www.ons.gov.uk/releases/ethnicgroupnationalidentitylanguageandreligioncensus2021inenglandandwales. [Accessed 26 Aug 2026].
20. Little RJA. A Test of Missing Completely at Random for Multivariate Data with Missing Values. J Am Stat Assoc. 1988;83(404):1198–202. 10.2307/2290157.
21. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011;30(4):377–99. 10.1002/sim.4067.
22. Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley; 1987. 10.1002/9780470316696.
23. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol. 2008;61(4):344–9. 10.1016/j.jclinepi.2007.11.008.
24. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res. 2004;6(3):e34. 10.2196/jmir.6.3.e34.

Figures

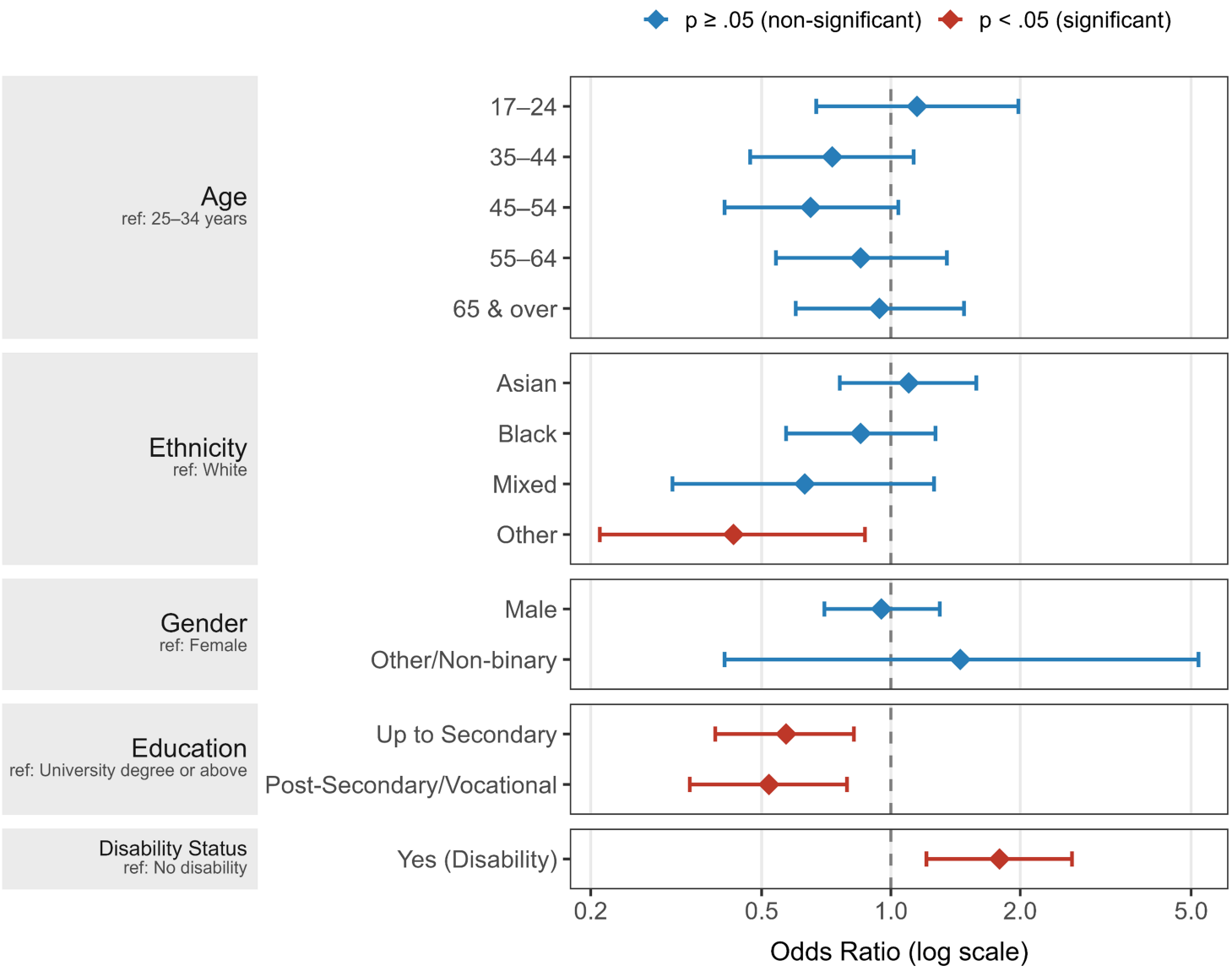


Figure 1

Sociodemographic and background predictors of active participation in health and social care research (N=1,040). Forest plot of odds ratios and 95% confidence intervals from a multivariable binary logistic regression, pooled across 50 imputed datasets using Rubin's rules. The model includes five predictors; prior work in health, care or research is reported separately in an exploratory extended model in Additional file 1: Table S7. Diamonds indicate point estimates and horizontal bars indicate 95% confidence intervals; the dashed vertical line marks an odds ratio of 1. The x-axis is on a logarithmic scale. Red=p < .05; blue=p ≥ .05. The estimate for Other/Non-binary gender is based on 11 respondents and is not interpreted in the text.

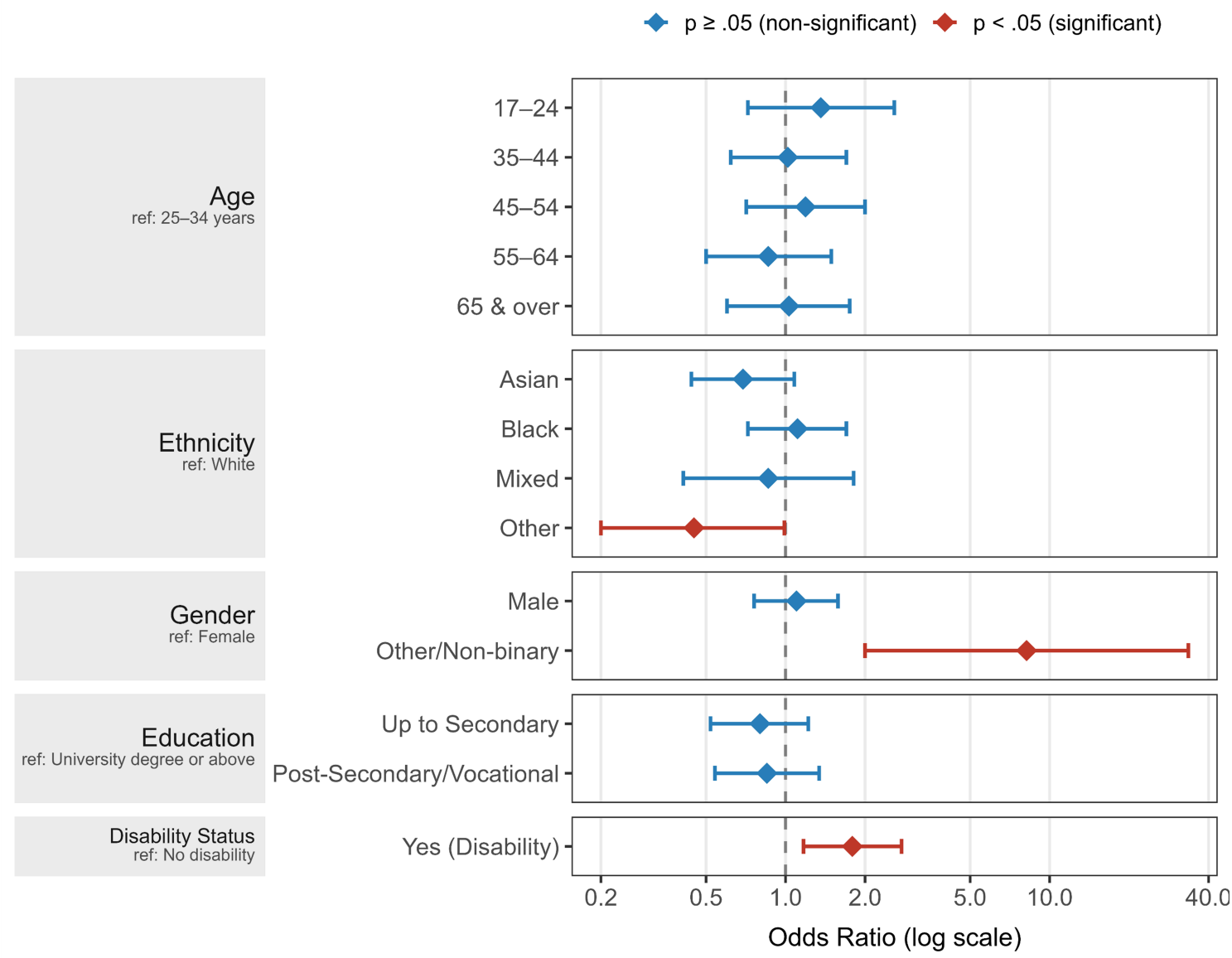


Figure 2

Sociodemographic and background predictors of ever declining participation in health and social care research (N=1,021). Forest plot of odds ratios and 95% confidence intervals from a multivariable binary logistic regression, pooled across 50 imputed datasets using Rubin's rules. Diamonds indicate point estimates and horizontal bars indicate 95% confidence intervals; the dashed vertical line marks an odds ratio of 1. The x-axis is on a logarithmic scale. Red=p < .05; blue=p ≥ .05. The estimate for Other/Non-binary gender is based on 11 respondents and is not interpreted in the text.

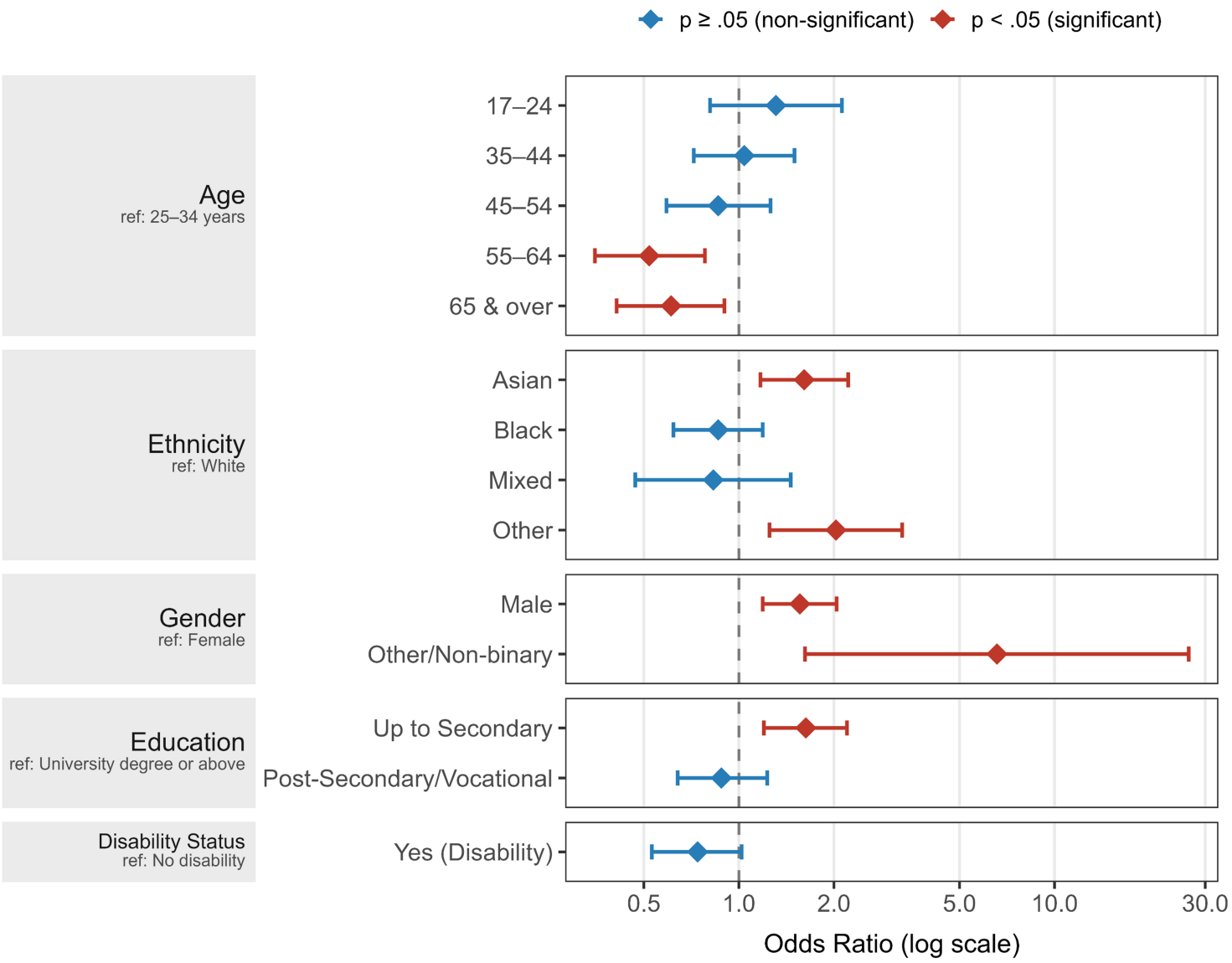


Figure 3

Sociodemographic and background predictors of ratings of organisational communication of research findings (N=951). Forest plot of odds ratios and 95% confidence intervals from an ordinal logistic regression, pooled across 50 imputed datasets using Rubin's rules. Under the proportional odds assumption, a single odds ratio applies across all four response thresholds; odds ratios above 1 indicate higher odds of a more favourable rating. Diamonds indicate point estimates and horizontal bars indicate 95% confidence intervals; the dashed vertical line marks an odds ratio of 1. The x-axis is on a logarithmic scale. Red=p < .05; blue=p ≥ .05. The estimate for Other/Non-binary gender is based on 11 respondents and is not interpreted in the text.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1Tables.docx
- Additionalfile2.RESEARCHHEROESsurveyinstrument.docx

- Additionalfile3STROBEchecklist.docx
- Additionalfile4CHERRIESchecklist.docx