

1 **Title:** Equality, diversity, and inclusion in oncology clinical trials: An audit of essential documents and
2 data collection against INCLUDE under-served groups in a UK academic trial setting.

3 **Authors:** Dhrusti Patel¹, Lucy Kilburn¹, Lisa Fox¹, Emma Hall¹, Judith Bliss¹, Rebecca Lewis¹

4 **Affiliation:**

5 ¹Clinical Trials and Statistics Unit, The Institute of Cancer Research, London, United Kingdom

6 **Corresponding author:** Rebecca Lewis, Rebecca.lewis@icr.ac.uk

7 **Authors' address:** Clinical Trials and Statistics Unit, The Institute of Cancer Research, London, United
8 Kingdom, SM2 5NG

9 **Abstract**

10 **Background:** Clinical trials should be as inclusive as possible to facilitate equitable access to research
11 and better reflect the population towards which any intervention is aimed. Informed by the UK's
12 National Institute for Health and Care Research (NIHR) Innovations in Clinical Trial Design and Delivery
13 for the Under-served (INCLUDE) guidance, we audited oncology trials conducted by the Clinical Trials
14 and Statistics Unit at The Institute of Cancer Research, London (ICR-CTSU) to identify whether essential
15 documents were overtly excluding any groups and whether sufficient data were collected to assess
16 diversity of trial participants from groups suggested by INCLUDE as under-served by research in the
17 UK.

18 **Methods:** Thirty cancer clinical trials managed by ICR-CTSU and approved between 2011-2021 were
19 audited. The first ethics approved version of each trial's protocol, patient information sheet, and
20 patient completed questionnaire, together with the first case report forms (CRFs) version were
21 reviewed. A range of items aligned with the INCLUDE under-served groups were assessed, including
22 age, sex and gender, socio-economic and health factors. The scope did not cover trial processes in
23 participating hospitals.

24 **Results:** Data relating to participants' age, ethnic group and health status were well collected and no
25 upper age limit was specified in any trials' eligibility criteria. 23/30 (77%) information sheets used at
26 least one gendered term to address patients. Most CRFs did not specify whether they were collecting
27 sex or gender and only included male or female categories. The median reading age for information
28 sheets was 15-16 years (IQR: 14-15 – 16-17). Socio-economic factors were not routinely collected and
29 not commonly mentioned in trial protocols.

30 **Conclusions:** No systemic issues were identified in protocols which would explicitly prevent any under-
31 served group from participating. Areas for improvement include reducing use of gendered words and
32 improving readability of patient information. The challenge of fully assessing adequate inclusion of

33 underserved populations remains, as socio-economic factors are not routinely collected because they
34 fall beyond the data generally required for protocol-specified trial endpoint assessments. This audit
35 has highlighted the need to agree and standardise demographic data collection to permit adequate
36 monitoring of the under-served groups identified by the NIHR.

37 **Word count:** 348

38 **Key words:** oncology clinical trials, equality, inclusivity, diversity, data collection, NIHR INCLUDE
39 framework.

40 **List of abbreviations**

41 CI: Confidence interval

42 CRFs: Case report forms

43 CTIMP: Clinical Trial of Investigational Medicinal Product

44 HRA: NHS Health Research Authority

45 ICR-CTSU: Clinical Trials and Statistics Unit at The Institute of Cancer Research

46 INCLUDE: Innovations in Clinical Trial Design and Delivery for the Under-served

47 MHRA: Medicines and Healthcare products Research Authority

48 NHS: National Health Service

49 NIH: National Institutes of Health

50 NIHR: National Institute for Health and Care Research

51 PIS: Patient information sheet

52 US: United States of America

53 UK: United Kingdom

54

55 **Background**

56 Clinical trials should be as inclusive as possible to allow robust assessment of the utility, efficacy, and
57 safety of an intervention in a sample representative of the population who may ultimately receive it.
58 This is critical in oncology trials, both to aid generalisability of results and from an equity and
59 accessibility perspective as they can provide access to alternative, though unproven, treatment
60 options when all other treatment has failed.

61 Whilst the US' National Institutes of Health (NIH) implemented guidelines in 1990 regarding inclusion
62 of women and minority groups in NIH supported research,(1) the focus of UK funders has been largely
63 on equality in terms of those applying for funding, rather than those participating. However the
64 importance of inclusivity was emphasised in the most recent strategy of the UK's NHS Health Research
65 Authority, the body responsible for ethics review of research involving the NHS.(2)

66 One of the largest UK funders of non-commercial research, the National Institute for Health and Care
67 Research (NIHR), recently released its first Equality, Diversity and Inclusion Strategy (2022-2027).(3)
68 This includes objectives around improved tracking and reporting of diversity of research participants,
69 focusing on protected characteristics according to the UK's Equality Act 2010 legislation (age,
70 disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race,
71 religion or belief, sex, sexual orientation). The NIHR also incorporate other groups of interest
72 consistent with those identified by their 'Innovations in Clinical Trial Design and Delivery for the Under-
73 served' (INCLUDE) project.(4) Whilst INCLUDE does not provide one single definition of a universally
74 under-served group in the UK, they note that key characteristics common to under-served groups are:
75 lower inclusion in research than expected from population estimates, high disease burden unmatched
76 by the volume of research, and differences in how some groups may engage with healthcare
77 interventions without research into these differences. Examples of potentially under-served groups
78 are essentially anyone outside the white male archetype traditionally considered a 'standard human'
79 in medical research, which historically aimed to limit heterogeneity amongst participants to try and

80 generate consistent results.(5) Suggested INCLUDE groups are intersectional and may comprise
81 people from less represented biological sexes, genders, ethnic groups, age extremes, people with
82 underlying health conditions and disabilities, people with less formal education or from less affluent
83 groups and those who are geographically isolated.

84 Despite the INCLUDE recommendations, developed via expert consensus, there remains a dearth of
85 quantitative data regarding UK trial inclusivity. In 2022 the NIHR reviewed data reported in 148 NIHR
86 funded randomised controlled trials conducted between 2007 and 2017, published 2019-2021.(6) Sex
87 and ethnicity were compared with population level data from the 2011 England and Wales decennial
88 census. Sex of participants matched proportions reported in the census (51% female, 49% male). Of
89 the 60% of trials reporting ethnicity, the proportion of non-white participants was found to be broadly
90 consistent with 2011 census data. The NIHR have recently released recommendations around data
91 collection to improve monitoring of representation within research, covering all protected
92 characteristics under the Equality Act, and others have recommended development of uniform
93 standards to capture aspects of diversity, including language, religious practices and sexual
94 orientation. (5)

95 In our area of focus, oncology, we are aware of only one quantitative study into diversity, conducted
96 at a single centre and published in 2010. This focused on ethnicity and compared admitted oncology
97 patients with oncology trial participants, using data from 2000-2005.(7) Ethnicity data were poorly
98 recorded, however an analysis adjusted for disease, age and gender found that patients from minority
99 ethnic groups had a lower chance of being in a research trial than white patients (Odds ratio (OR)=0.70
100 (95% confidence interval (CI): 0.53 to 0.94); p=0.01).

101 The large majority of published datasets demonstrating a lack of inclusion in trials use data from the
102 United States. Although these give indications of potential areas for improvement in the UK, it is likely
103 that patterns of participation differ due to the variations in historical, societal and demographic

104 context, together with the universal coverage of the UK's National Health Service - free at the point of
105 use in contrast to the US' private sector healthcare system.

106 Given the recommendations from INCLUDE, and findings in other research settings identifying
107 systematic exclusion of some populations,(8) we conducted an audit to determine whether our trial
108 protocols were explicitly excluding any under-served groups and our patient information provision
109 was inclusive, and to assess whether we would be able to identify under-representation from the data
110 we collect as standard.

111 **Methods**

112 At the methodologist-led academic Clinical Trials and Statistics Unit at The Institute of Cancer Research
113 (ICR-CTSU) we have designed, managed and analysed a substantial number of phase II and III oncology
114 clinical trials including over 30,000 adult participants worldwide. We conduct investigator-initiated
115 non-commercial trials funded by charities, government-funding schemes and pharmaceutical industry
116 partners. Phase II trials generally focus on repurposing existing treatments with known safety profiles,
117 we do not conduct trials for regulatory licensing purposes or on behalf of commercial sponsors.
118 Interventions comprise radiotherapy, chemotherapy, hormone therapy and immunotherapy and
119 disease sites include breast, prostate, bladder, head and neck, lung, and rare cancers.

120 All clinical trials managed by ICR-CTSU which gained UK regulatory approval between 2011 and 2022
121 and were sponsored by either The Institute of Cancer Research or our partner organisation The Royal
122 Marsden NHS Foundation Trust, were included in the audit. Trials with external sponsors were
123 excluded to ensure audited trials had used ICR-CTSU templates and processes.

124 Documents were reviewed for a broad range of items which could be mapped onto the INCLUDE
125 under-served groups (Table 1).(4)

126 ***Table 1: INCLUDE under-served groups mapped to items reviewed in the audit***

127 For each trial meeting the selection criteria the first UK ethics approved version of the protocol,
128 patient information sheet (PIS) and patient-completed questionnaire, together with the first version
129 of case report forms (CRFs), were reviewed.

130 Where trials had multiple patient information sheets, these were categorised as being either
131 comprehensive, summary, or substudy. The PIS using the most gendered terms was selected for
132 review, regardless of category. To investigate accessibility for those with educational disadvantage,
133 Flesch Kincaid grade level, the Flesch Reading Ease, word count, and average number of words per
134 sentence were determined using Office 365 Microsoft Word's inbuilt tools. Flesch-Kincaid grade level
135 refers to US school year, ranging from 0-18, with a grade of 9 equivalent to a reading age of 13-14.(9)
136 Flesch reading ease ranges from 0-100, with a higher score indicating easier readability.(10) Where
137 trials had more than one comprehensive PIS, a mean of the word counts and readability scores was
138 calculated.

139 Audit data were tabulated for all trials and analysed combined and separately: 1) according to the
140 time of trial approval (2011-2016 vs 2017-2022), to identify any trends or improvements in practice
141 and 2) by funding source, according to receipt of any funding from industry partners - industry funded
142 trials were the most likely to use potentially unlicensed agents or have eligibility requirements
143 imposed by external partners. (Figure 1)

144 ***Figure 1: Audit process and trial categorisation.***

145 **Statistical Methods**

146 Analyses were conducted using Stata version 17. A p-value of <0.01 was considered statistically
147 significant to make some account for multiple testing. Fisher's exact test was used to compare
148 frequencies of items in protocols, CRFs and questionnaires by time of approval and funding source.
149 For PIS review data, the median and quartiles were calculated for readability scores and word

150 counts. Wilcoxon rank sum test was used to compare PIS review data by time of approval and
 151 funding source.

152 **Results**

153 30 trials met the inclusion criteria for the audit conducted between March and May 2022 (Figure 1).

154 Trial characteristics are summarised in table 2.

155 **Table 2: Trial characteristics**

	Date of approval				Total
	2011-2016		2017-2022		
Type of funding	Non-commercial funding	Industry funding	Non-commercial funding	Industry funding	
	9	9	6	6	30
Phase					
II	4	8	0	5	17
II/III	1	0	2	0	3
III	4	1	4	2	11
Trial type					
CTIMP	3	9	1	6	19
Non-CTIMP	6	0	5	0	11
Cancer type					
Breast	1	3	2	3	9
Multiple: breast, lung & prostate	1	0	0	0	1
Gynaecological	0	1	0	1	2
Head and neck	1	0	1	0	2
Lung	0	0	1	0	1
Penile	1	0	0	0	1
Prostate	1	4	2	1	8
Urinary system	4	1	0	1	6

156

157 **Demographic factors**

158 90% (27/30) of protocols specified a lower age limit. 18/27 (67%) had a lower limit of 18, 8/27 (30%)
159 had a lower limit of 16. 1/27 (3%) protocol specified a lower age limit of 60 years or that participants
160 should be post-menopausal. Of the three trials that did not specify age, two were approved between
161 2017 and 2022 and had industry funding, the other was approved between 2011 and 2016, with
162 non-commercial funding. There was no evidence of a difference according to time of trial approval
163 or industry support. No protocol stated an upper age limit and all CRFs and patient-completed
164 questionnaires collected date of birth.

165 No trial's eligibility criteria specified ethnicity. Ethnic group was collected in 83% (24/29) of trials'
166 CRF booklets and was not captured in any patient-completed questionnaires.

167 Nine of the thirty trials reviewed were for cancers affecting people of male sex, seven included only
168 female sex patients and the remaining 14 trials were open to all sexes. All trials enrolling people who
169 could potentially become pregnant stated pregnancy was an exclusion for safety reasons. 14/30 (47%)
170 protocols used a gendered term such as 'women', 'men', 'male', or 'female' in their eligibility criteria,
171 none stated whether they referred to gender identity or biological sex. 50% (7/14) of trials using
172 gendered terms included patients of only one sex due to the type and location of tumour (eg prostate
173 cancer). The remaining seven protocols including gendered terms were for non-sex specific disease
174 areas and stated both men and women could be included. There was no difference in use of gendered
175 terms by the time of trial approval or type of funding.

176 83% (25/30) of patient information sheets used at least one gendered term (men, women, he, she).
177 23/30 (77%) trials used at least one gendered pronoun when referring to patients and 2/30 (7%) trials
178 used gendered terms for clinicians only. Five trials did not include any gendered terms in patient
179 information. 10/14 (71%) trials that were non-sex specific used a gendered term to address patients
180 (Table 3).

181 ***Table 3: Use of gendered terms for patients in patient information sheets***

	Male terms	Female terms	Male and female terms	No gendered terms	n (N=30)
Male sex specific trial	6	0	0	3 ^a	9
Female sex specific trial	0	7	0	0	7
Non-sex specific trial	0	1 ^b	9	4 ^c	14
Total	6	8	9	7	30

182 ^a Two trials approved >6 years ago, of which one was industry funded and the other non-industry funded. The other trial was
183 approved <6 years ago and industry funded.

184 ^b Trial approved <6 years ago and industry funded.

185 ^c Two trials approved <6 years ago and two approved >6 years ago. Three were industry funded and one was non-industry
186 funded.

187 15/29 (52%) CRFs collected whether the participant was male or female, of which 12/15 captured sex
188 and the remainder captured gender. Of the 12 trials capturing sex, one trial (2021 ethics approval)
189 collected "sex at birth". 12/14 (86%) trials that did not collect sex or gender were sex-specific due to
190 the anatomical site of the tumour, three included female patients and nine male. There was no
191 difference in collection of this information by the time of trial approval or by funding source.

192 Sexual orientation was not collected on any CRFs and neither sex/gender nor sexual orientation were
193 collected in any patient-completed questionnaires.

194 Average patient information sheet readability scores and word counts suggested that the median
195 information sheet was suitable for people with a reading age of 15-16, with a median Flesch reading
196 ease score of 55.6. Most trials' information sheets had between 5500-6000 words and
197 approximately 250 sentences (Figure 2). PIS for trials approved between 2011-2016 had a higher
198 Flesch reading ease score than that for trials approved between 2017-2022 (p=0.003). No
199 information was provided in a non-written format at time of first approvals and education level was
200 not collected in any trial's CRF but was collected in two questionnaires.

201 *Figure 2: Patient information sheet readability scores and word counts.*

202 **Social and economic factors**

203 Employment status was collected in one (3%) trial's CRF and three trials' (21%) questionnaires. One
204 questionnaire also captured marital, childcare, and carer status. Religion was not collected in any
205 trial's CRFs. No CRFs collected income level, residence details or military status. Geographic location
206 (postcode) was collected in 69% (20/29) of trial CRFs and 50% (7/14) of patient-completed
207 questionnaires. No trials specified a language requirement, however the first version of all trials'
208 patient information sheet and patient-completed questionnaire was only available in English. Main
209 language was not collected in any CRFs or patient questionnaires.

210 Ten percent (3/30) of trials had socio-geographic requirements, with an eligibility criterion stating:
211 *"Absence of any psychological, familial, sociological or geographical condition potentially hampering*
212 *compliance with the study protocol and follow-up schedule; those conditions should be discussed with*
213 *the patient before registration in the trial."* The time of trial approval or source of trial funding had no
214 influence on whether this was included and no explanation was provided in the protocols.

215 There was no difference in any socioeconomic factors being included in eligibility criteria or captured
216 by CRFs or patient-completed questionnaires either by the time of trial approval or the funding type.

217 **Health status**

218 All protocols' eligibility criteria included statements regarding non-permitted co-morbidities, which
219 were well collected in CRFs and patient-completed questionnaires (Figure 2). 87% (26/30) of protocols
220 permitted inclusion of patients with prior malignancies, subject to a disease-free duration varying
221 between 4-6 months to 5 years. 3 trials' protocols included a statement regarding psychological
222 conditions, as noted above. No trials mentioned learning disability, substance addiction or physical
223 requirements other than performance status. No trials provided information or patient questionnaires
224 in formats accessible to people with visual impairments.

225 Time of approval did not have any impact on whether trials specified comorbidity exclusions or
226 collected related details in CRFs or patient-completed questionnaires. Positive HIV status (Figure 3)
227 was the only co-morbidity in eligibility criteria found to be significantly different according to trial
228 funding ($p=0.002$), with all industry funded trials excluding people with HIV. A statistically significant
229 difference was found between trial funding and medical history collection in CRFs ($p<0.001$), with
230 industry funded trials collecting medical history more frequently than non-commercially funded trials.

231 **Figure 3: Health status eligibility criteria and data collection.**

232 *Abbreviations: Spec = specified. Not spec = not specified. ECOG = Eastern Cooperative Oncology Group;*
233 *WHO = World Health Organisation A) Pregnancy status eligibility. Pregnancy status N = 21, 9 trials*
234 *excluded as they had only male sex participants. E) Brain metastasis N=17, 13 trials excluded as they*
235 *were in early-stage non-metastatic cancer, so any metastases were excluded.*

236 **Discussion**

237 We found that our trial protocols did not have overt systematic exclusion for the majority of groups
238 identified in the INCLUDE guidance. Our patient information generally included gendered terms and
239 required a relatively high level of reading comprehension, with readability worsening for trials
240 approved after 2016. Data collected for the purposes of assessing the trials' endpoints, whether via
241 CRF or patient completed questionnaire, was insufficient to identify groups suggested as under-
242 served by research by INCLUDE. There was little difference in practice over time or by funding
243 source.

244 Cancer is largely a disease affecting older people - of the 375,400 people diagnosed with cancer each
245 year in the UK between 2016-2018, 194,500 were over 70 (11, 12). Evidence from the United States
246 suggests that older adults are under-represented in cancer trials.(13) Whilst it was encouraging that
247 none of our protocols specified an upper age limit, we recognise that some co-morbidity exclusions,
248 often required for safety when testing newer agents, may de-facto exclude some older people. Over
249 50% of those aged 65 and older have at least two chronic conditions, and this is projected to increase

250 in the future.(14) In addition, cancer incidence is strongly associated with a number of lifestyle choices,
251 such as smoking, that are likely to increase burden of co-morbidities,(15) so it is crucial that co-
252 morbidity exclusions are minimised when safe to do so. Patients with co-morbidities can be included
253 in oncology trials when clinically appropriate, for example in later phase trials or academic phase II
254 trials repurposing existing treatments with known safety profiles, to facilitate better representation
255 of the population that the intervention is aimed towards.

256 Lower age limits were dictated by the UK regulatory landscape with its differing requirements for
257 paediatric cancer trials. ICR-CTSU trials investigate cancers which are far more prevalent in older
258 populations and are highly unlikely to affect younger people, therefore most of our clinical trials had
259 18 (age of UK adulthood) as their lower age limit. Some had a lower limit of 16, this is likely as a result
260 of 2018 guidance from the UK's National Cancer Research Institute's Teenage, Young Adults and Germ
261 Cell Tumours Group which recommended that Cancer Research UK (CRUK), a major funder of oncology
262 trials in the UK, should request justification for lower age limits for studies they were supporting to
263 avoid inadvertent exclusion of adolescents.(16) Drug effects are similar in adolescents and adults and
264 by changing lower limits to 16 this would allow people to access new treatments earlier than
265 otherwise possible, due to the tendency for paediatric trials to be conducted several years after trials
266 in adults.(17) Despite this guidance, we did not see any association between time of trial approval or
267 source of funding and lower age limit. This was likely due to trials being funded by other non-
268 commercial funders, together with investigator consensus that changing the lower limit would be
269 irrelevant due to the lack of incidence in younger populations.

270 Ethnicity was not stipulated in any inclusion criteria and was generally well collected in CRFs. This is
271 encouraging as it should allow us to compare our trial participants' data with UK incidence statistics
272 to identify any signals of under-representation in current practice. However we recognise that UK
273 health records regarding ethnicity, from which the CRF data would likely be reported by hospital staff,

274 are not always accurate particularly for those outside the White British category,(18) so we also intend
275 to investigate direct collection from trial participants in future.

276 We identified a lack of clarity in our use of sex and gender nomenclature, both within protocols and
277 in data collection. Whilst some trials were necessarily restricted to one sex due to the nature of the
278 tumour, its location, or the intervention studied, around half were open to any sex. All trials enrolling
279 people who could potentially become pregnant stated pregnancy was an exclusion. Whilst this is a
280 group identified as under-served by research by INCLUDE, due to the nature of treatments studied in
281 our trials it would be very difficult to justify loosening this criterion due to the danger of foetal
282 exposure to cytotoxic agents, radiation or hormonal therapies.

283 All data collection regarding sex/gender used binary categories (male/female), which does not capture
284 the known range of gender identities within the UK (19) and also does not recognise existence of
285 intersex individuals, incidence of which is admittedly low, but too poorly measured in the UK to
286 provide robust figures.(20) In addition there was frequent use of gendered terms both in protocols
287 and patient information. Whilst we did not collect any data related to transgender identity, we
288 recognise that people identifying as different genders to the sex they were assigned at birth may still
289 be eligible for our single sex trials, patient information should ideally avoid use of gendered terms to
290 avoid alienating people. Prostate trials' eligibility criteria, for example, should include anyone with a
291 prostate, such as some intersex or non-binary people and trans-women, as the prostate is not
292 normally removed in gender affirming surgery.(21) A review of literature published between 1975 and
293 2017 identified only 10 published cases of transgender women with prostate cancer.(22) It is unlikely
294 that incidence in this group is so low, but the findings may suggest an issue with data collection and
295 reporting. As a result of our findings we have updated our templates and guidance to recommend the
296 removal of gendered terms in protocols and CRFs wherever not required, and to be clear about
297 whether they refer to sex or gender identity where they are used.

298 We did not collect any information related to participants' sexual orientation, however LGBTQ+
299 people who have had cancer have shown a preference for gender-neutral language to address
300 themselves and their partners.(23) We have therefore updated our patient information guidance to
301 recommend removal of all gendered terms to prevent any inadvertent discouragement of
302 participation of people from the LGBTQ+ community.

303 We assessed our patient information as a proxy for systematic exclusion of people with educational
304 disadvantage and those who were not fluent in English or had visual impairments. In the UK, all
305 medical research studies are expected to provide written information to potential participants,
306 reviewed by Research Ethics Committees during the approval process.(24) There are no requirements
307 to provide information in languages other than British English or in alternative formats, although
308 recommendations to consider the latter have recently been introduced.(25) Thus it is unsurprising
309 that we did not find any alternative formats or languages were available at first approval of our
310 studies, although some trials did provide short versions of the full patient information sheet. We are
311 currently exploring the best approach to providing information in alternative languages, about which
312 there is very little guidance for researchers in the UK, and introducing guidance around use of
313 validated non-English language patient-completed questionnaires. We are also beginning to introduce
314 audio-visual presentation of clinical trial information for selected trials as this may improve patient
315 understanding, although this remains a topic requiring further research.(26)

316 It was disappointing to observe a trend towards lower readability of patient information over time.
317 This may be associated with an increased level of complexity for more recent trials, although this was
318 not formally reviewed within this audit. Approximately 15% of UK adults had literacy levels of 9-11
319 years or younger in 2011, representing an estimated 5.1 million people.(27) Our patient information
320 sheets, with a median reading age of 15-16, are therefore likely to be too complex for a large
321 proportion of the UK population despite routine involvement of patient and public representatives in
322 their development. This is consistent with other research in the UK and Republic of Ireland finding the
323 median reading age for information sheets to be 16.1 years, with a median Flesch Reading Ease score

324 of 49.6.(28) We are seeking to improve readability by implementing more training in the use of plain
325 English and recommending use of Word readability statistics whilst preparing patient information. In
326 addition, whilst there used to be a suggested template provided by the NHS Health Research Authority
327 including mandatory sections which contributed to the overall length of information provided, current
328 advice recommends reducing length of information to be *“enough to make an informed choice about*
329 *taking part, and no more”*.(25) It is therefore likely that length of our patient information sheets could
330 be reduced in future, and we hope that readability will improve, both as a result of process
331 improvements we implemented as a result of this study and due to the changing landscape in ethics
332 review processes.

333 Cancer incidence is higher in socio-economically disadvantaged populations,(29) and according to
334 data from North America, patients from these backgrounds are less likely to take part in oncology
335 trials.(30) However there are few recent published data investigating the impact of socio-economic
336 status on research participation in the UK. Despite having universal healthcare coverage via the NHS,
337 this does not mean that trial participation is without cost, as it may involve additional visits
338 necessitating more time off work and leading to higher transport costs than associated with standard
339 of care.

340 Our relatively routine collection of participants’ postcodes will allow identification of those living in
341 remote areas, together with some indication of inclusivity via the Indices of Multiple Deprivation (IMD)
342 tools, available for each devolved UK nation.(31) However the IMD is not a reliable measure for
343 individuals’ level of deprivation, as within each area people’s circumstances will vary. Without
344 collecting social and economic factors directly from participants it is not possible to robustly assess
345 the majority of such groups identified within INCLUDE from our trial datasets.

346 Reassuringly, the majority of protocols did not include any statements regarding social and economic
347 factors, or overtly exclude patients in the majority of under-served health status groups proposed by
348 INCLUDE. However, three trial protocols did include statements regarding patients from certain socio-

349 geographical backgrounds, which could result in people being less likely to be invited, or to decide not
350 to join these trials based on clinicians' preconceptions regarding their ability to comply and resulting
351 challenging conversations. We were unable to identify why such a statement had been included as
352 there was no discernible pattern and have updated our protocol guidance to ensure these are not
353 included in future. Excluding populations from remote geographical locations may be due to the
354 locations of specialist cancer treatment centres, however, we are currently improving guidance on
355 obtaining support to allow patients to be enrolled at geographically-distant sites across the country
356 and to travel for the intervention. This approach is being taken in the TORPEDO trial(32) to avoid
357 excluding patients due to location but requires buy-in from non-commercial funders who have
358 historically been hesitant to fund participants' travel and/or accommodation costs due to restricted
359 budgets.

360 Health status was well collected in CRFs, and patient-completed questionnaires. The majority of co-
361 morbidities and medical history requirements listed in protocols appeared directly related to safety
362 requirements for the interventions under study, with an association between funding source and HIV
363 status in particular, likely to be due to industry funded trials using newer agents than those in non-
364 commercially funded trials. Participants' history of other cancers, with a disease-free duration prior
365 to trial entry was often specified, however the required disease-free duration ranged from a few
366 months to a few years. This did not appear to be associated with the stage of cancer being
367 investigated. Prior cancer disease-free duration could be shorter for trials in the metastatic setting,
368 where endpoint events such as cancer progression or death are unfortunately likely to be reached
369 within a short time after enrolment. In earlier stage disease, participants may need to be followed up
370 for many years and this could risk a recurrence of their prior cancer being conflated with a recurrence
371 of the cancer being studied within the trial. However, in our audit, trials in the metastatic setting did
372 not routinely stipulate a shorter cancer-free period than trials for patients with early-stage disease. A
373 more systematic approach for deciding the disease-free period for prior cancers in trial eligibility
374 criteria is a potential area for improvement in our protocols.

375 Inclusion of patients with a high burden of comorbid conditions with likely associated use of multiple
376 medications may mean that the effects of these cannot be reliably disentangled from symptoms
377 caused by the trial treatment or impact on survival outcomes. Concomitant medication may also
378 interfere with the intervention's mechanism, leading to exclusions for safety purposes. However
379 blanket exclusion of a high number of co-morbidities and concomitant medications could lead to
380 overestimating the safety of an intervention before it is rolled out to the wider population, so it is
381 critical to strike the right balance between protecting participants safety and ensuring inclusion of a
382 representative group of patients.

383 This audit had limitations as we could only consider factors over which we have influence, including
384 the design of our trials, development of research protocols and patient information and data capture
385 practices. We reviewed the first-approved version of trial documents - later versions of documents
386 may have been more inclusive, although we saw few discernible patterns when we looked at later
387 approved trials in comparison to earlier ones. The audit does not account for other barriers to
388 inclusion that are not possible to identify from trials' essential documents and data capture alone. We
389 have not systematically collected sufficient data to allow assessment of all under-served groups, as
390 this information is not needed to assess any of the reviewed trials' outcomes. We have historically
391 taken the approach of focused data collection for the purposes of ensuring protocol adherence, safety
392 and endpoint analysis, both to avoid collecting data that would not be used, which would not be
393 ethically justified, and to avoid over-burdening NHS hospitals with unnecessary data collection.

394 As our trial participants are not identified and recruited by ICR-CTSU, but by clinicians and their
395 research staff within participating NHS hospitals, we would require a different dataset to ensure
396 robust unbiased assessment of equitable inclusion, we are planning to access national electronic
397 healthcare datasets to explore this further.

398 Despite its limitations, this audit represents a starting point for our planned programme of work
399 investigating inclusivity in UK oncology trials and has identified several areas for improvement to our

400 current practice as described above. We plan to work with patient and public contributors of differing
401 backgrounds and life experiences to determine an acceptable level of enhanced demographic data
402 collection, informed by the NIHR's recent recommendations. We recognise we need to collect enough
403 information to better monitor representation in future trials whilst balancing proportionality of
404 potentially intrusive data collection and ensuring acceptability to participants.

405 There is a distinct lack of published quantitative data regarding inclusivity in UK cancer trials, across
406 all protected characteristics and under-served groups, which makes any impact of process
407 improvements difficult to discern. Obtaining and publishing these data is a key area of focus in our
408 future research plans.

409 **Conclusion**

410 Our trials' eligibility criteria were relatively inclusive. Data were routinely collected regarding co-
411 morbidities, age, ethnic group, and sex/gender. Other demographic and social and economic factors
412 were not frequently collected. Process improvements implemented as a result of this audit, such as
413 use of gender neutral terminology, recommendations around minimising co-morbidity exclusions
414 where possible, and considering collecting more demographic factors, may also be relevant to other
415 academic trial groups.

416 **Ethics approval and consent to participate**

417 Not applicable.

418 **Consent for publication**

419 Not applicable.

420 **Availability of data and material**

421 The datasets used and/or analysed during the current study are available from the corresponding
422 author on reasonable request.

423 **Competing interests**

424 None.

425 **Funding**

426 No project specific funding was obtained for this audit. The ICR-CTSU is supported by a Cancer
427 Research UK Clinical Trials Unit programme grant (C1491/A25351).

428 **Authors' contributions**

429 DP collected and analysed the data and drafted the publication. LK and RL contributed to data analysis
430 and interpretation. RL, LK, LF, EH and JB designed the audit and critically revised the draft publication.
431 All authors read and approved the final manuscript.

432 **Acknowledgements**

433 Grateful thanks to ICR-CTSU trial team members for providing documents for this audit and assisting
434 with queries. The ICR-CTSU is supported by a Cancer Research UK Clinical Trials Unit programme grant
435 (C1491/A25351).

436 **References:**

- 437 1. National Institutes of Health. NIH Guide For Grants and Contracts. In: U.S Department of
438 Health and Human Services, editor. 1990.
- 439 2. NHS Health Research Authority. We make it easy to do research that people can trust
440 Strategy 2022-25. NHS Health Research Authority,; 2023.
- 441 3. National Institute for Health and Care Research. Equality, Diversity and Inclusion Strategy
442 2022-2027. 2023 March 2023.
- 443 4. Witham MD, Anderson E, Carroll C, Dark PM, Down K, Hall AS, et al. Developing a roadmap
444 to improve trial delivery for under-served groups: results from a UK multi-stakeholder process.
445 *Trials*. 2020;21(1):694.
- 446 5. Sharma A, Palaniappan L. Improving diversity in medical research. *Nature Reviews Disease*
447 *Primers*. 2021;7(1):74.
- 448 6. National Institute for Health and Care Research. Randomised controlled trial participants:
449 Diversity data report. 2022 23/11/2022.
- 450 7. Sylvia G, Gareth A, Allyson MP. Recruitment of minority ethnic groups into clinical cancer
451 research trials to assess adherence to the principles of the Department of Health Research
452 Governance Framework: national sources of data and general issues arising from a study in one
453 hospital trust in England. *Journal of Medical Ethics*. 2010;36(6):358.
- 454 8. Ranganathan M, Bhopal R. Exclusion and Inclusion of Nonwhite Ethnic Minority Groups in 72
455 North American and European Cardiovascular Cohort Studies. *PLOS Medicine*. 2006;3(3):e44.
- 456 9. Weiss BD, Schwartzberg JG, Terry C. Davis TC, Parker RM, Sokol PE, Williams MV. Health
457 literacy and patient safety: Help patients understand. 2 ed2007.
- 458 10. Flesch R. A new readability yardstick. *J Appl Psychol*. 1948;32(3):221-33.

- 459 11. Cancer Research UK. Cancer Incidence Statistics 2023 [Available from:
460 <https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence>.
- 461 12. Cancer incidence by age [Available from: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence/age#heading-Zero>.
- 462
463 13. Parks RM, Holmes HM, Cheung KL. Current Challenges Faced by Cancer Clinical Trials in
464 Addressing the Problem of Under-Representation of Older Adults: A Narrative Review. *Oncol Ther*.
465 2021;9(1):55-67.
- 466 14. Kingston A, Robinson L, Booth H, Knapp M, Jagger C, for the Mp. Projections of multi-
467 morbidity in the older population in England to 2035: estimates from the Population Ageing and
468 Care Simulation (PACSim) model. *Age and Ageing*. 2018;47(3):374-80.
- 469 15. Cancer Research UK. Cancer risk statistics 2023 [Available from:
470 <https://www.cancerresearchuk.org/health-professional/cancer-statistics/risk>.
- 471 16. Laurence R. Eliminating age restriction in cancer clinical trials National Cancer Research
472 Institute 2018 [Available from: <https://www.ncri.org.uk/eliminating-age-restriction-in-cancer-clinical-trials/>.
- 473
474 17. Chuk MK, Mulugeta Y, Roth-Cline M, Mehrotra N, Reaman GH. Enrolling Adolescents in
475 Disease/Target-Appropriate Adult Oncology Clinical Trials of Investigational Agents. *Clin Cancer Res*.
476 2017;23(1):9-12.
- 477 18. Drummond R. How ethnicity recording differs across health data sources and the impact on
478 analysis: Office for National Statistics; 2023 [Available from:
479 <https://blog.ons.gov.uk/2023/01/16/how-ethnicity-recording-differs-across-health-data-sources-and-the-impact-on-analysis/>.
- 480
481 19. Office for National Statistics (ONS). Gender identity, England and Wales: Census 2021. 2023
482 06/01/2023.
- 483 20. Government Equalities Office. Variations in Sex Characteristics Technical Paper. In: Office GE,
484 editor.: UK Government; 2019.
- 485 21. UK PC. Trans women and prostate cancer [Available from:
486 [https://prostatecanceruk.org/prostate-information-and-support/risk-and-symptoms/trans-women-and-prostate-cancer#:~:text=The%20prostate%20is%20not%20removed,of%20a%20vagina%20\(vaginoplasty\)](https://prostatecanceruk.org/prostate-information-and-support/risk-and-symptoms/trans-women-and-prostate-cancer#:~:text=The%20prostate%20is%20not%20removed,of%20a%20vagina%20(vaginoplasty))].
- 487
488 22. Deebel NA, Morin JP, Autorino R, Vince R, Grob B, Hampton LJ. Prostate Cancer in
489 Transgender Women: Incidence, Etiopathogenesis, and Management Challenges. *Urology*.
490 2017;110:166-71.
- 491 23. Waters AR, Tennant K, Cloyes KG. Cultivating LGBTQ+ Competent Cancer Research:
492 Recommendations from LGBTQ+ Cancer Survivors, Care Partners, and Community Advocates. *Semin
493 Oncol Nurs*. 2021;37(6):151227.
- 494 24. NHS Health Research Authority. Participant Information Quality Standards: NHS Health
495 Research Authority,; 2023 [updated 16/08/2023. Available from: <https://www.hra.nhs.uk/planning-and-improving-research/research-planning/participant-information-quality-standards/>.
- 496
497 25. NHS Health Research Authority. Participant Information Design and Review Principles: NHS
498 Health Research Authority,; 2023 [updated 21/08/2023. Available from:
499 <https://www.hra.nhs.uk/planning-and-improving-research/research-planning/participant-information-design-and-review-principles/>.
- 500
501 26. Hutchison C, Cowan C, McMahon T, Paul J. A randomised controlled study of an audiovisual
502 patient information intervention on informed consent and recruitment to cancer clinical trials. *Br J
503 Cancer*. 2007;97(6):705-11.
- 504 27. Harding C, Romanou E, Williams J, Peters M, Winkley J, Burke P, et al. The 2011 Skills for Life
505 Survey: a Survey of Literacy, Numeracy, and ICT Levels in England. GOV.UK, Department of Business
506 laS; 2012.
- 507

- 508 28. O'Sullivan L, Sukumar P, Crowley R, McAuliffe E, Doran P. Readability and understandability
509 of clinical research patient information leaflets and consent forms in Ireland and the UK: a
510 retrospective quantitative analysis. *BMJ Open*. 2020;10(9):e037994.
- 511 29. Cancer Research UK. Cancer in the UK 2020: socio-economic deprivation. Cancer Research
512 UK; 2020.
- 513 30. Unger JM, Gralow JR, Albain KS, Ramsey SD, Hershman DL. Patient Income Level and Cancer
514 Clinical Trial Participation: A Prospective Survey Study. *JAMA Oncology*. 2016;2(1):137-9.
- 515 31. Ministry of Housing CLG. The English Indices of Deprivation 2019 (IoD2019) Statistical
516 Release. In: Ministry of Housing CLG, editor. 2019.
- 517 32. Thomson DJ, Cruickshank C, Baines H, Banner R, Beasley M, Betts G, et al. TORPEdO: A phase
518 III trial of intensity-modulated proton beam therapy versus intensity-modulated radiotherapy for
519 multi-toxicity reduction in oropharyngeal cancer. *Clin Transl Radiat Oncol*. 2023;38:147-54.

520

521