Article

Consent document translation expense hinders inclusive clinical trial enrolment

https://doi.org/10.1038/s41586-023-06382-0

Received: 14 December 2022

Accepted: 28 June 2023

Published online: 02 August 2023

Check for updates

Maria A. Velez[1], Beth A. Glenn[2,3,4,5], Maria Garcia-Jimenez[1,2,6], Amy L. Cummings[1,2], Aaron Lisberg[1,2], Andrea Nañez[7], Yazeed Radwan[1], Jackson P. Lind-Lebuffe[1], Paige M. Brodrick[1], Debory Y. Li[1], Maria J. Fernandez-Turizo[8], Arjan Gower[1], Maggie Lindenbaum[2], Manavi Hegde[2], Jenny Brook[9], Tristan Grogan[9], David Elashoff[2,9], Michael A. Teitell[2,10] & Edward B. Garon[1,2]✉

Patients from historically under-represented racial and ethnic groups are enrolled in cancer clinical trials at disproportionately low rates in the USA[1–3]. As these patients often have limited English proficiency[4–7], we hypothesized that one barrier to their inclusion is the cost to investigators of translating consent documents. To test this hypothesis, we evaluated more than 12,000 consent events at a large cancer centre and assessed whether patients requiring translated consent documents would sign consent documents less frequently in studies lacking industry sponsorship (for which the principal investigator pays the translation costs) than for industry-sponsored studies (for which the translation costs are covered by the sponsor). Here we show that the proportion of consent events for patients with limited English proficiency in studies not sponsored by industry was approximately half of that seen in industry-sponsored studies. We also show that among those signing consent documents, the proportion of consent documents translated into the patient's primary language in studies without industry sponsorship was approximately half of that seen in industry-sponsored studies. The results suggest that the cost of consent document translation in trials not sponsored by industry could be a potentially modifiable barrier to the inclusion of patients with limited English proficiency.

Cancer clinical trials are the primary means of developing diagnostic and therapeutic strategies, and trial participation is associated with improved patient outcomes[8,9]. Patients from traditionally under-represented racial and ethnic groups participate in clinical trials at disproportionately low rates[10–14], limiting the generalizability of results[2,15]. Although barriers to the inclusion of historically under-represented racial and ethnic groups have been extensively studied, there has been limited progress towards achieving equity[1,2,7,10,12,16,17]. While many important barriers are not easily addressed by individual clinical trial investigators[16,18], investigator-related barriers to equitable clinical trial enrolment have been less thoroughly studied[18,19].

The non-Hispanic white population in the USA has proportionally decreased, based in part on immigration from Asia and Latin America[20,21]. The percentage of residents speaking a language other than English at home rose from 11% in 1980 to 22% by 2018, with rates above 70% among individuals identifying as Hispanic or Asian[20,22,23]. Consequently, the relative importance of limited English proficiency, an established barrier to trial participation, has probably increased over time. Yet, the factors contributing to the under-representation of patients with limited English proficiency are understudied[7,9].

Ensuring that trial participants are appropriately informed regarding procedures and risks is a cornerstone of ethical research[24]. The Food and Drug administration (FDA) mandates that presented consent documents are in a language understandable to the patient[25–27]. The FDA recommends that institutional review boards (IRBs) ensure that translated consent documents are prepared by a qualified entity with a certification statement for each translation[25,28–30], a potentially costly and time-consuming process[29,31]. Recognizing the importance of timely participation, the FDA allows an alternative approach in which patients sign translated, non-study-specific documents to be promptly followed by appropriately translated study-specific consent documents (Supplementary Methods).

Whether delays, costs or other aspects of the consent document translation process discourages trial participation among patients with limited English proficiency is challenging to study. As limited data can be collected from patients who do not sign consent documents, it is difficult to establish how these patients differ from those who participate. As consent documents are often translated only after a prospective participant is identified[32], analyses assessing the impact of available translated consent documents are subject to the bias of reverse causation[33].

Approximately 70% of randomized oncology clinical trials are funded by industry[34,35], with most studies not sponsored by industry funded by

[1]Department of Medicine, Division of Hematology/Oncology, University of California, Los Angeles, Los Angeles, CA, USA. [2]Jonsson Comprehensive Cancer Center, University of California, Los Angeles, Los Angeles, CA, USA. [3]Department of Health Policy and Management, University of California, Los Angeles, Los Angeles, CA, USA. [4]UCLA Center for Cancer Prevention and Control Research, University of California, Los Angeles, Los Angeles, CA, USA. [5]UCLA Kaiser Permanente Center for Health Equity, University of Califonia, Los Angeles, Los Angeles, CA, USA. [6]Division of Hematology/Oncology, UCLA-Olive View Medical Center, Los Angeles, CA, USA. [7]Department of Obstetrics and Gynecology, University of California, Los Angeles, Los Angeles, CA, USA. [8]Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA. [9]Department of Medicine Statistics Core, University of California, Los Angeles, Los Angeles, CA, USA. [10]Department of Pathology and Laboratory Medicine, University of California, Los Angeles, Los Angeles, CA, USA. ✉e-mail: egaron@mednet.ucla.edu

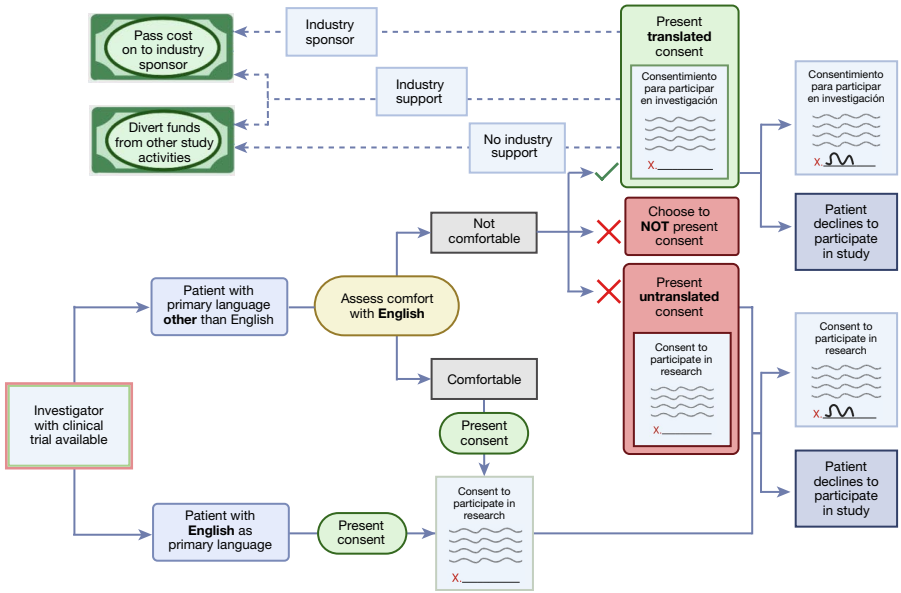


Fig. 1 | Consent process and cost allocation of consent document translation. An investigator meeting an eligible patient for a clinical trial should assess the patient's (or parent or guardian's) comfort with signing an English consent document. If the patient (or parent or guardian) is not comfortable signing consent documents in English, the investigator should translate the consent documents. Depending on the study funder, this cost can be either completely passed on to the industry sponsor, potentially covered by the industry sponsor or covered completely by the investigator.

a grant from either industry, philanthropic or governmental groups[36,37]. Industry can offer assistance for a study sponsored by an academic centre by providing study drug or device and/or additional financial support, although generally less funding than in industry-sponsored studies[38] (Fig. 1). In non-industry-sponsored studies, the principal investigator generally operates on a fixed, per-patient budget, whereas in industry-sponsored studies, the sponsor generally provides additional funds for consent document translation beyond the negotiated per-patient budget[39]. Although an investigator can request funds for consent document translation in a proposed grant, many grants have a budget cap, meaning that such a request would limit funds for other study activities[39]. Furthermore, funds intended for consent document translation could often be directed to other study activities if translation costs were below the budgeted amount[39].

Among several barriers to the participation of patients with limited English proficiency in clinical trials, we hypothesized that the additional costs incurred by investigators on studies not sponsored by industry could discourage investigators from offering trial participation to patients for whom consent document translation would be required[4,5,29]. Although prohibited by regulations[25], an investigator who lacks sufficient funds may not offer consent documents to patients with limited English proficiency (Fig. 1) or the investigator may utilize consent documents that are already available in a language in which the patient is not proficient (generally English).

To test our hypothesis, we assessed data from all consent events for studies conducted at the University of California Los Angeles (UCLA) Jonsson Comprehensive Cancer Center over a six-year period to determine patients' primary language, English proficiency and language of consent documents. We compared studies not sponsored by industry with those sponsored by industry to evaluate potential differences based on participant primary language and English proficiency.

Study population

Of 13,717 consent events between January 2013 and December 2018, 1,635 were excluded from further analysis based on lack of access to appropriate data (Fig. 2). Most of the remaining 12,082 consent events were for patients with English as their primary language (n = 11,340, 93.9%). Of the remaining 742 consent events, the patient met the definition for limited English proficiency in 481 consent events (64.8%).

Of 200 randomly selected consent events evaluated as a control to ensure English proficiency among patients with English as a primary language, 58 were for children. The need for an interpreter was found in only four consent events, all for paediatric patients with English as their primary language but limited English proficiency among their parents or guardians. Among 247 paediatric consent events for patients with English as their primary language, the need for an interpreter was documented in 17 (6.9%), and these patients were analysed as having a primary language other than English and limited English proficiency.

As some patients signed consent documents for multiple studies, the 12,082 consent events occurred in 9,213 patients, 63.4% of whom were non-Hispanic white. Although only 1.6% of non-Hispanic white patients had a primary language other than English, 18.3% of members of racial and ethnic groups other than non-Hispanic white signing consent documents had a primary language other than English, including nearly a quarter of Hispanic and Asian or Pacific Islander patients (Fig. 2 and Extended Data Table 1).

The most common primary languages other than English were Spanish (40.8%, n = 231) and Chinese (20.8%, n = 118) (Extended Data Table 2). The median number of words in the initial English consent document was 7,491.5 (range 598 to 20,382 words), with an estimated cost of US $1,498 per translation. Additional costs would be incurred to translate a consent document at the time of protocol amendments, an amount that would vary by trial.

Consent events based on study sponsor

Nearly half of consent events (n = 5,738) were for industry-sponsored studies (Extended Data Table 3). Of 758 studies for which patients signed consent documents, 261 (34.4%) had any available IRB-approved translated consent documents. Although most studies were sponsored by industry (n = 585), the median number of consent events per study was less compared with non-industry-sponsored studies (5.0 versus 8.0, P < 0.001). Yet, the proportion of consent events in studies that had translated consent documents available was higher

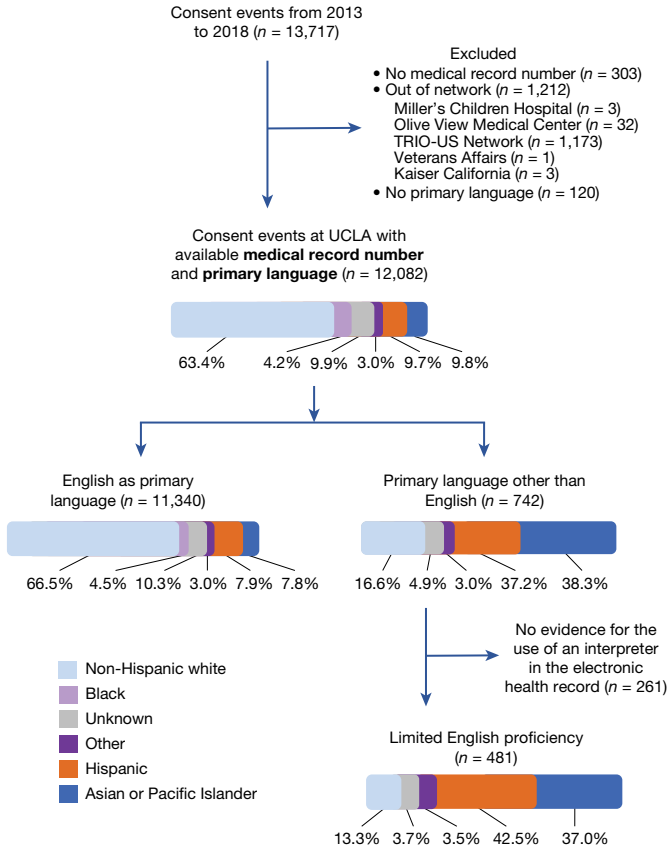


Fig. 2 | Consent events included in the study. Consent-event data for patients who signed consent documents for cancer centre studies from 2013 to 2018 were included in our analysis if they had a medical record number in our electronic health system as well as a documented primary language ($n = 12,082$). Patients were considered to have English as their primary language (English primary, $n = 11,340$) or to have a primary language other than English ($n = 742$). Patients with a primary language other than English were considered to have limited English proficiency if there was evidence for the use of an interpreter in the electronic health record. The racial and ethnic distribution of patients is depicted by colour with the representative colours described in the legend.

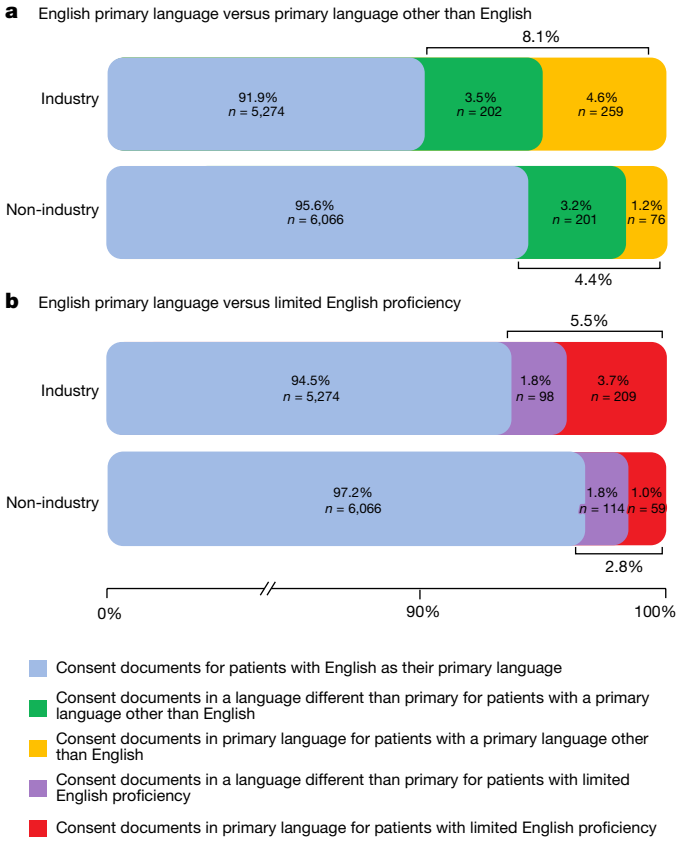


Fig. 3 | Comparison of the proportion of consent events based on primary language and English proficiency in industry-sponsored versus non-industry-sponsored studies. a, Blue indicates the proportion of consent events for patients with English as their primary language. The bracketed areas indicate the proportion of consent events for patients with a primary language other than English in industry-sponsored studies (top bar) versus non-industry-sponsored studies (bottom bar) (8.1% versus 4.4%, $P < 0.001$). Green indicates the proportion of consent events for patients with a primary language other than English signing consent documents in a language different than their primary in industry-sponsored studies (top bar) compared with non-industry-sponsored studies (bottom bar) (3.5% versus 3.2%, $P = 0.391$). Yellow indicates the proportion of consent events for patients with a primary language other than English signing consent documents in their primary language in industry-sponsored studies (top bar) compared with non-industry-sponsored studies (bottom bar) (4.6% versus 1.2%, $P < 0.001$). **b**, Blue indicates the proportion of consent events for patients with English as their primary language. The bracketed areas indicate the proportion of consent events for patients with limited English proficiency in industry sponsored studies (top bar) versus non-industry-sponsored studies (bottom bar) (5.5% versus 2.8%, $P < 0.001$). Purple indicates the proportion of consent events for patients with limited English proficiency signing consent documents in a language different than their primary in industry-sponsored studies (top bar) compared with non-industry-sponsored studies (bottom bar) (1.8% versus 1.8%, $P = 0.643$). Red indicates the proportion of consent events for patients limited English proficiency signing consent documents in their primary language in industry-sponsored studies (top bar) compared with non-industry-sponsored studies (bottom bar) (3.7% versus 1.0% $P < 0.001$). Logistic regression models with generalized estimating equations clustered by patient unique identifier compared the proportions above. The P values reported are two-tailed.

for industry-sponsored studies compared with those not sponsored by industry (51.4% versus 23.9%, $P < 0.001$).

Of 758 studies, 12 were paediatric-only studies, none of which were sponsored by industry. Of these 12 studies, 8 (66.7%) had translated consent documents at any point during the study. By contrast, among 718 adult-only studies, 580 (80.8%) were sponsored by industry and 241 (33.5%) had translated consent documents at any point during the study. The odds of a consent event for an industry-sponsored study having any available translated consent documents were greater than for a non-industry-sponsored study (odds ratio (OR) 3.4, 95% confidence interval (CI) 3.1 to 3.6, $P < 0.001$; data not shown).

Patients with a primary language other than English represented 8.1% of consent events in industry-sponsored studies versus 4.4% in studies not sponsored by industry ($P < 0.001$) (Fig. 3). Patients with limited English proficiency represented 5.5% of consent events in industry-sponsored studies versus 2.8% in studies not sponsored by industry ($P < 0.001$). Findings were similar when only interventional studies were analysed (Extended Data Fig. 1).

Consent documents at study opening

Only 18 studies had translated consent documents available at the time of study opening, 13 industry-sponsored and 5 non-industry-sponsored studies. Of these 18 studies, 12 had Spanish consent documents at study opening: 10 industry-sponsored and 2 non-industry-sponsored (Extended Data Table 4). Patients with Spanish as their primary language had higher odds of signing consent documents for studies that had Spanish consent documents at study opening than those without (OR 5.7, 95% CI 3.8 to 8.5, $P < 0.001$) (Extended Data Table 4).

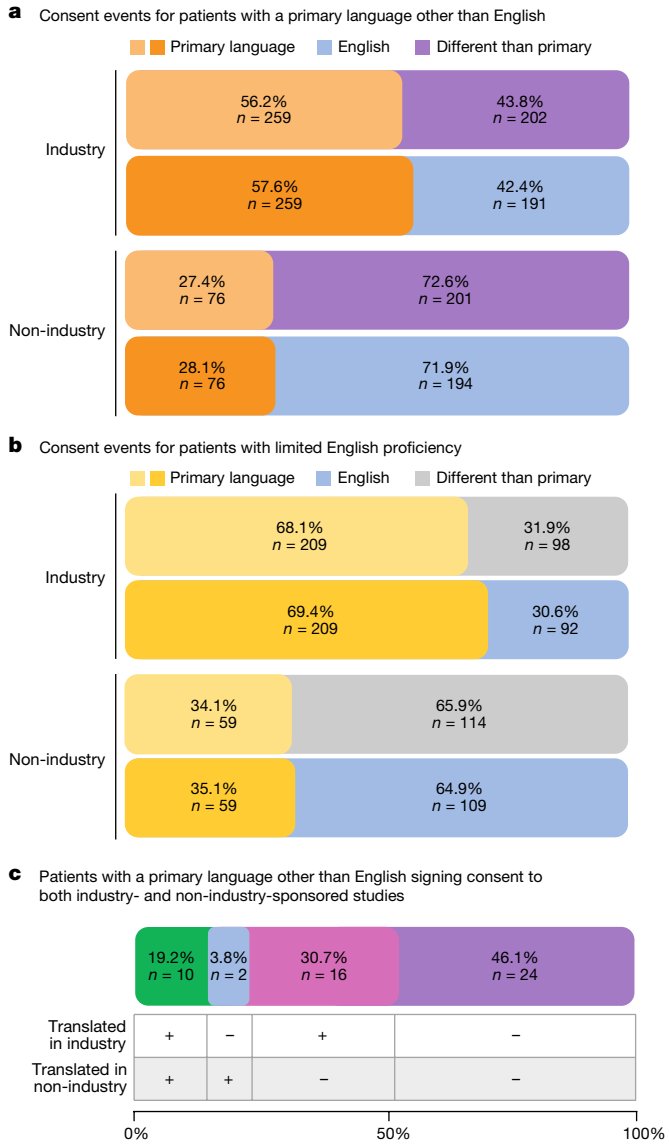


Fig. 4 | Comparison of the proportion of consent events by language. **a**, Orange bars indicate the proportion of consent events for which patients with a primary language other than English signed consent documents in their primary language in industry-sponsored versus non-industry-sponsored studies (light orange, 56.2% versus 27.4%, $P < 0.001$; dark orange, 57.6% versus 28.1%, $P < 0.001$). Purple bars indicate the proportion of consent events for which patients signed consent documents in a language different than primary in industry-sponsored versus non-industry-sponsored studies (43.8% versus 72.6%, $P < 0.001$). Blue indicates the proportion of consent events for which patients signed consent documents in English in industry-sponsored versus non-industry-sponsored studies (42.4% versus 71.9%, $P < 0.001$). **b**, Yellow bars indicate the proportion of consent events for which patients with limited English proficiency signed consent documents in their primary language in industry-sponsored versus non-industry-sponsored studies (light yellow, 68.1% versus 34.1%, $P < 0.001$; dark yellow, 69.4% versus 35.1%, $P < 0.001$). Grey bars indicate the proportion of consent events for which patients signed consent documents in a language different than primary in industry-sponsored versus non-industry-sponsored studies (31.9% versus 65.9%, $P < 0.001$). Blue bars indicate the proportion of consent events for which patients signed consent documents in English in industry-sponsored versus non-industry-sponsored studies (30.6% versus 64.9%, $P < 0.001$). Logistic regression models with generalized estimating equations clustered by patient unique identifier compared the proportions above. The P values reported are two-tailed. **c**, Among patients with a primary language other than English signing consent documents for both an industry-sponsored and a non-industry-sponsored study, 10 (green) signed in their primary language and 24 signed in a language different than primary for both (purple). Of the 18 patients who signed consent documents in discrepant languages, 16 (pink) signed in their primary language in the industry-sponsored study versus 2 (blue) in the non-industry-sponsored study (McNemar's test, $P = 0.002$).

Patients with a primary language other than English or Spanish did not have higher odds of signing consent documents for studies that had Spanish consent documents at study opening (OR 0.9, 95% CI 0.6 to 1.3).

Consent documents in primary language

Patients with a primary language other than English signed consent documents in a language different than the patient's primary language in 43.8% of consent events for industry-sponsored studies versus 72.6% in studies not sponsored by industry ($P < 0.001$). When analysing patients with limited English proficiency, rates were 31.9% versus 65.9%, respectively ($P < 0.001$) (Fig. 4). When evaluating only studies without any translated consent documents, the corresponding results were 42.4% versus 71.9% for patients with a primary language other than English ($P < 0.001$), 30.6% versus 64.9% ($P < 0.001$) in patients with limited English proficiency. This phenomenon of patients signing consent document in a language different than their primary language appears to be driven by lack of appropriately translated consent documents, as only 3% occurred when consent documents in the patient's primary language were available (data not shown). Patients with a primary language other than English had lower odds of signing consent documents in a language different than primary for studies with translated consent documents than those without them (OR 0.02, 95% CI 0.009 to 0.030; Extended Data Table 4).

Of 52 patients who signed consent documents for both industry-sponsored and non-industry-sponsored studies, 10 signed all in their primary language, 24 signed all in a language different than primary and 18 signed in their primary language for one study and a language different than primary for the other. Of these 18 patients, 16 signed consent documents in a language different than primary for the non-industry-sponsored study ($P = 0.002$; Fig. 4c).

Differences in the proportion of consent events by sponsor type were largely driven by a difference in consent events in the patient's primary language. The proportion of consent events for patients with a primary language other than English who signed consent documents in the patient's primary language was 4.6% versus 1.2% ($P < 0.001$) in industry-sponsored versus non-industry-sponsored studies, and 3.7% versus 1.0% ($P < 0.001$) for those with limited English proficiency (Fig. 3). However, the proportion of consent events for patients with a primary language other than English who signed consent documents in a language different than primary was similar between industry-sponsored and non-industry-sponsored studies (3.5% versus 3.2%, $P = 0.391$) and patients with limited English proficiency (1.8% versus 1.8%, $P = 0.643$). Patients with a primary language other than English had a higher proportion of consent events in which the patient signed consent documents in a language different than primary in studies not sponsored by industry across departments (Extended Data Table 5).

Consent odds based on language

A multivariable analysis evaluated whether associations were confounded by other factors. After adjusting for age at consent, gender, race, ethnicity, histology and study type (observational versus interventional), patients with a primary language other than English (OR 0.74, 95% CI 0.63 to 0.94, $P = 0.005$) and limited English proficiency (OR 0.74, 95% CI 0.58 to 0.95, $P = 0.021$) had lower odds of signing consent documents for non-industry sponsored studies than patients with

Table 1 | Multivariable analysis for odds ratio for the association between various factors and signing consent documents in a non-industry-sponsored study

	Multivariable analysis for patients with a primary language other than English signing consent documents			Multivariable analysis for patients with a primary language other than English signing consent documents in patient's primary language			Multivariable analysis for patients with limited English proficiency signing consent documents			Multivariable analysis for patients with limited English proficiency signing consent documents in patient's primary language		
Variable	**OR**	**95% CI**	***P* value**	**OR**	**95% CI**	***P* value**	**OR**	**95% CI**	***P* value**	**OR**	**95% CI**	***P* value**
Age												
Age at consent (per year)	0.97	0.97–0.98	<0.001	0.97	0.97–0.98	<0.001	0.97	0.97–0.98	<0.001	0.97	0.97–0.98	<0.001
Language												
English primary		Reference			Reference			Reference			Reference	
Primary other than English[a]	0.74	0.63–0.94	0.005	0.38	0.28–0.52	<0.001	–	–	–	–	–	–
Limited English proficiency[b]	–	–	–	–	–	–	0.74	0.58–0.95	0.021	0.35	0.25–0.50	<0.001
Race and ethnicity												
Non-Hispanic white		Reference			Reference			Reference			Reference	
Asian or Pacific Islander	0.64	0.54–0.75	<0.001	0.66	0.55–0.79	<0.001	0.65	0.55–0.77	<0.001	0.66	0.55–0.79	<0.001
Black	1.00	0.80–1.26	0.978	1.01	0.80–1.28	0.916	1.06	0.80–1.27	0.972	1.01	0.80–1.27	0.921
Hispanic	0.75	0.63–0.89	<0.001	0.76	0.63–0.90	0.002	0.73	0.62–0.88	<0.001	0.74	0.62–0.90	0.002
Other	1.15	0.87–1.54	0.324	1.16	0.87–1.56	0.321	1.15	0.86–1.53	0.345	1.2	0.87–1.55	0.320
Unknown	3.38	2.86–4.01	<0.001	3.39	2.86–4.02	<0.001	3.43	2.89–4.06	<0.001	3.41	2.87–4.05	<0.001
Study type												
Interventional		Reference			Reference			Reference			Reference	
Observational	36.2	28.3–46.4	<0.001	35.1	27.3–45.0	<0.001	35.7	27.7–37.3	<0.001	34.9	27.3–44.9	<0.001
Gender												
Male		Reference			Reference			Reference			Reference	
Female	0.38	0.35–0.42	<0.001	0.37	0.33–0.42	<0.001	0.38	0.35–0.42	<0.001	0.37	0.34–0.41	<0.001
Histology												
Single solid malignancy		Reference			Reference			Reference			Reference	
Healthy	1.78	1.35–2.35	<0.001	1.87	1.41–2.48	<0.001	1.86	1.37–2.52	<0.001	1.89	1.41–2.51	<0.001
Multiple histology	0.38	0.34–0.42	<0.001	0.38	0.34–0.42	<0.001	0.36	0.33–0.40	<0.001	0.38	0.35–0.43	<0.001
Single haem malignancy	0.06	0.04–0.08	<0.001	0.06	0.04–0.09	<0.001	0.06	0.04–0.09	<0.001	0.06	0.04–0.09	<0.001

[a]Patients with a primary language other than English compared with patients with English as their primary language. [b]Patients with limited English proficiency compared with patients with English as their primary language. ORs (with 95% CI) were estimated from a logistic regression model with generalized estimating equations clustered by patient identifier. 'Reference' indicates the reference category in the multivariable model; '–' indicates 'not applicable'. The *P* values reported are two-tailed.

English as their primary language. Younger age, women, and Asian or Pacific Islander and Hispanic (both compared with non-Hispanic white) patients also had lower odds of signing consent documents for non-industry-sponsored studies. The odds of signing consent documents for observational studies was higher in studies not sponsored by industry (Table 1).

The odds of signing consent documents in the patient's primary language for a non-industry-sponsored study were considerably lower for patients with a primary language other than English (OR 0.38, 95% CI 0.28 to 0.52, $P < 0.001$) and limited English proficiency (OR 0.35, 95% CI 0.25 to 0.50, $P < 0.001$) compared with patients with English as their primary language. The results were similar when evaluated by bivariable analysis (Extended Data Table 6). Patients with a primary language other than English, including those with limited English proficiency, had lower odds of signing consent documents for non-industry-sponsored than industry-sponsored studies across departments (Fig. 5). When looking at the distribution of patients with a primary language other than English, the proportion of those signing consent documents and signing consent documents in their primary language was decreased across non-industry-sponsored studies (Extended Data Fig. 2).

Additional potential confounders

The potentially confounding interactions between Medi-Cal insurance status and language interaction terms were evaluated, but they were not significant (Extended Data Table 7). Therefore, interaction terms were not included in the final model. While nesting consent events within studies led to *P* values that were somewhat higher for some analyses, the general trends seen were similar (Extended Data Table 8). Findings also remained consistent when studies that could have received some industry support for consent document translation were grouped with those that were sponsored by industry. The odds of signing consent documents for studies not sponsored or supported by industry were 0.61 (95% CI 0.52 to 0.72, $P < 0.001$) for patients with a primary language other than English and 0.64 (95% CI 0.53 to 0.79, $P < 0.001$) for patients with limited English proficiency compared with patients with English as their primary language (data not shown).

The safety net insurer Medi-Cal insured 48.8% of patients with a primary language other than English versus 6.9% among patients with English as their primary language ($P < 0.001$) (Extended Data Table 1). When Medi-Cal insurance status was added as a variable in

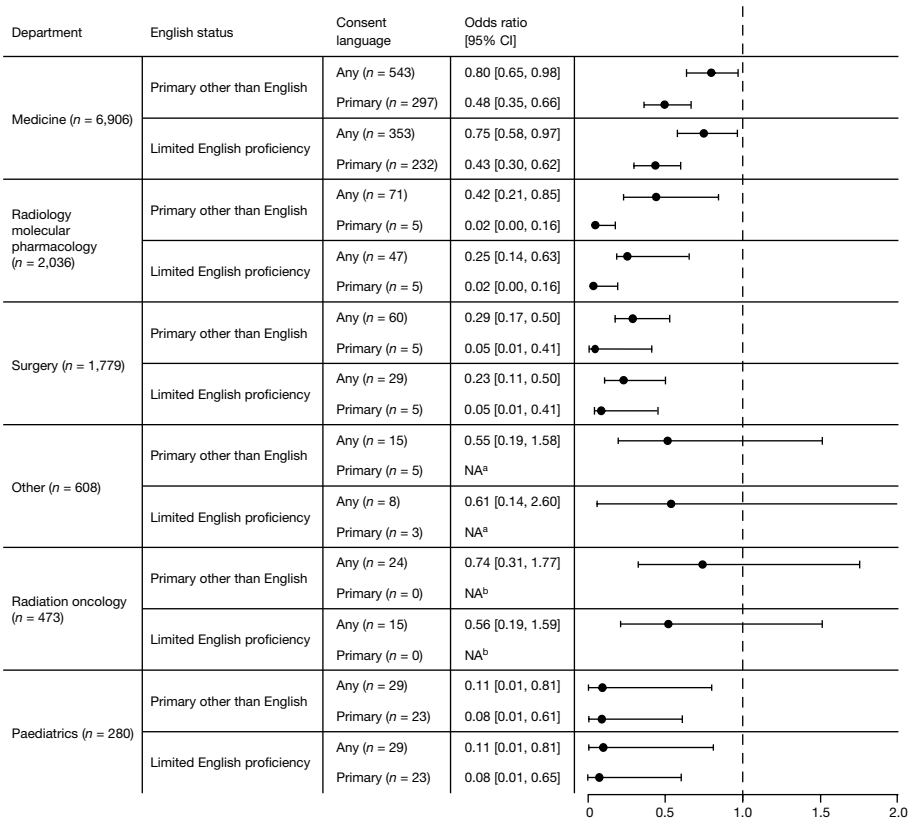


Fig. 5 | Odds ratios for patients with a primary language other than English and with limited English proficiency signing consent documents in non-industry-sponsored studies compared with patients with English as their primary language across the different departments. Odds ratios for patients with a primary language other than English (top row) and limited English proficiency (bottom row) of signing consent document in any language (top row) and in the patient's primary language (bottom row) were calculated using a logistic regression model with generalized estimating equations clustered by patient unique identifier. The dot denotes the odds ratio and the bars represent the 95% CI. [a]OR could not be calculated because there were no patients with a primary language other than English or limited English proficiency who signed consent documents in their primary language in industry-sponsored studies. [b]OR could not be calculated as no consent documents were translated into patient's primary language.

the multivariable model, results remained consistent (Extended Data Table 7).

Discussion

We found that the proportion of consent events for patients with a primary language other than English was lower in non-industry-sponsored versus industry-sponsored studies. For non-industry-sponsored studies, patients with a primary language other than English frequently signed consent documents in a language different than their primary language. Findings persisted when analyses were restricted to patients with limited English proficiency.

Standard economic theory argues that increasing the expense faced by an individual for an activity discourages the individual from engaging in that activity[40]. So, we tested the hypothesis that patients requiring translated consent documents would be less likely to sign consent documents for studies not sponsored by industry, studies for which the investigator would generally be responsible for the cost of consent document translation. While a retrospective study cannot prove causation, consistent associations across analyses support the hypothesis that patients requiring translated consent documents were selectively missing from studies not sponsored by industry.

It is unlikely that our observations were driven by differential enrolment by sponsor type, as the odds of having any translated consent documents available for non-industry-sponsored studies was substantially lower despite a greater median number of consent events per study when compared with industry-sponsored studies. It is also unlikely that our observations were driven by differences in the patient population by sponsor type, as when the same patient signed consent documents for both an industry-sponsored and non-industry-sponsored study, nearly all patients who signed consent documents in discrepant languages signed in a language different from their primary for the non-industry-sponsored study.

An approach that increases the participation of patients with a primary language other than English in non-industry-sponsored to the level seen in industry-sponsored studies would be expected to lead to a modest increase in the representation of patients from ethnic or racial groups other than non-Hispanic white. If either efficacy or toxicity substantially differed in these populations compared with non-Hispanic white patients, in aggregate, this increased representation could facilitate recognizing such a difference. Moreover, as patients with limited English proficiency may form a distinct subpopulation that is more likely to have poor social determinants of health within traditionally under-represented racial and ethnic groups, differential clinical outcomes observed in this subpopulation could be even more pronounced than in an unselected population from that racial or ethnic group[41,42].

Increased representation could be particularly important in paediatric studies as approximately 30% of the Hispanic population living in the USA are children[43]. Although patients with Spanish as their primary language were more likely to sign consent documents for studies with Spanish consent documents available at study opening, this result should be interpreted cautiously. First, the presence of translated consent documents at the time of study opening in a single study in less common languages, such as Thai, suggests that this analysis is subject

to the bias of reverse causation. Second, it is possible that a study anticipated to enrol a disproportionate number of Hispanic patients would be more likely to have Spanish consent documents available at study opening. The potential for translated consent documents at study opening facilitating increased inclusion in clinical trials should be an area for future investigation.

Although most industry-sponsored studies have a therapeutic intent, non-industry-sponsored studies often focus on biobanking, assessing screening and prevention strategies, and survivorship and quality of life issues[44], study types in which the inclusion of a diverse patient population is highly relevant. Although our analysis focused on cancer studies, investigators studying other diseases face similar pressures. Whether our findings extend beyond oncology studies should be investigated.

Our results raise concern about the quality of information conveyed to patients with limited English proficiency. The National Institutes of Health (NIH) Policy and Guidelines on the Inclusion of Women and Minorities clearly indicates that the cost of inclusion of participants with limited English proficiency in clinical research should not hinder their participation[45]. However, no additional resources are typically provided to investigators to cover the cost of consent document translation on studies not sponsored by industry, which are typically funded through federal grants or cooperative groups[29,46]. We are not aware of data so far that have explored whether the cost of consent document translation is commonly requested in NIH grant applications, but it would be helpful if those data could be made available.

The FDA does not specifically mandate who should perform consent document translation, and IRB requirements vary across institutions[47,48]. At some institutions, IRBs require that consent documents are translated by a professional translation service, whereas others rely on investigators to determine what constitutes an adequate translation[26]. Investigators at some institutions could have members of the research team fluent in another language translate consent documents, especially for minimal risk studies, at a lower cost than professional translation services. This could potentially decrease the barrier of cost of consent translation.

The strengths of the current dataset include a large number of consent events based on six years of heavily curated data, the high number of translated consent documents and the large number of patients signing consent documents for studies not sponsored by industry. In addition, inclusion of all consent events for which the appropriate data were available increases confidence in our results and reduces potential biases. The primary weakness of our analysis is its single-centre nature. Sensitivities regarding patient health information, study-related data and differences in regulatory structures make cross-centre studies difficult. The general consistency across departments suggests that these findings are widespread. However, data from additional cancer centres would enhance confidence in our findings. Although Southern California has particularly high racial and ethnic diversity[49], increasing non-Hispanic white populations are not limited to this region.

Significant findings for the Asian or Pacific Islander race and Hispanic ethnicity in multivariable analyses suggests that our models may not have optimally separated the effects of race and ethnicity from language. The effect of language in the multivariable analysis may have persisted for Asian or Pacific Islander and Hispanic patients based on perceived limited English proficiency. This will be an important topic for future research. Another limitation is the retrospective nature of our study and reliance on electronic health record data. For instance, Medi-Cal insurance status, a dynamic variable, was gathered retrospectively and may not accurately reflect insurance status at the time of the consent event. Furthermore, some data, such as language proficiency, may not be documented accurately in the electronic health record.

As all data included were from patients who signed consent documents for cancer centre studies, important barriers preventing patients from participating in any cancer centre study were not assessed. Barriers such as delays associated with consent document translation and lack of training for research staff on appropriate consent practices for patients with limited English proficiency may have had important roles. As such, additional impediments should be explored to inform possible future interventions.

Conclusion

Our findings suggest that an important barrier for patients with limited English proficiency to participate in cancer studies may be the cost that consent document translation presents to investigators, particularly in studies not sponsored by industry. This work identifies a potentially modifiable barrier to enrolling these patients on studies, which is of particular importance in an increasingly multicultural and multilingual population.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-023-06382-0.

1. Boulware, L. E. et al. Combating structural inequities—diversity, equity, and inclusion in clinical and translational research. *N. Engl. J. Med.* **386**, 201–203 (2022).
2. Oyer, R. A. et al. Increasing racial and ethnic diversity in cancer clinical trials: an American Society of Clinical Oncology and Association of Community Cancer Centers joint research statement. *J. Clin. Oncol.* **40**, 2163–2171 (2022).
3. Wendler, D. et al. Are racial and ethnic minorities less willing to participate in health research? *PLoS Med.* **3**, e19 (2006).
4. Frayne, S. M., Burns, R. B., Hardt, E. J., Rosen, A. K. & Moskowitz, M. A. The exclusion of non-English-speaking persons from research. *J. Gen. Intern. Med.* **11**, 39–43 (1996).
5. Glickman, S. W. et al. Perspective: The case for research justice: inclusion of patients with limited English proficiency in clinical research. *Acad. Med.* **86**, 389–393 (2011).
6. Muthukumar, A. V., Morrell, W. & Bierer, B. E. Evaluating the frequency of English language requirements in clinical trial eligibility criteria: a systematic analysis using ClinicalTrials.gov. *PLoS Med.* **18**, e1003758 (2021).
7. Roy, M. et al. Limited English proficiency and disparities in health care engagement among patients with breast cancer. *JCO Oncol. Pract.* **17**, e1837–e1845 (2021).
8. Fleurence, R. L. et al. Engaging patients and stakeholders in research proposal review: the patient-centered outcomes research institute. *Ann. Intern. Med.* **161**, 122–130 (2014).
9. Staples, J. N. et al. Language as a barrier to cancer clinical trial accrual: assessing consenting team knowledge and practices for cancer clinical trial consent among low English fluency patients. *Appl. Cancer Res.* **38**, 14 (2018).
10. Clark, L. T. et al. Increasing diversity in clinical trials: overcoming critical barriers. *Curr. Probl. Cardiol.* **44**, 148–172 (2019).
11. Corbie-Smith, G., Miller, W. C. & Ransohoff, D. F. Interpretations of 'appropriate' minority inclusion in clinical research. *Am. J. Med.* **116**, 249–252 (2004).
12. Davis, T. C., Arnold, C. L., Mills, G. & Miele, L. A qualitative study exploring barriers and facilitators of enrolling underrepresented populations in clinical trials and biobanking. *Front. Cell Dev. Biol.* **7**, 74 (2019).
13. Murthy, V. H., Krumholz, H. M. & Gross, C. P. Participation in cancer clinical trials: race-, sex-, and age-based disparities. *JAMA* **291**, 2720–2726 (2004).
14. Parada, H. Jr., Vu, A. H., Pinheiro, P. S. & Thompson, C. A. Comparing age at cancer diagnosis between Hispanics and non-Hispanic whites in the United States. *Cancer Epidemiol. Biomarkers Prev.* **30**, 1904–1912 (2021).
15. FDA takes important steps to increase racial and ethnic diversity in clinical trials. *US FDA* https://www.fda.gov/news-events/press-announcements/fda-takes-important-steps-increase-racial-and-ethnic-diversity-clinical-trials (2022).
16. Unger, J. M., Vaidya, R., Hershman, D. L., Minasian, L. M. & Fleury, M. E. Systematic review and meta-analysis of the magnitude of structural, clinical, and physician and patient barriers to cancer clinical trial participation. *J. Natl Cancer Inst.* **111**, 245–255 (2019).
17. Vuong, I. et al. Overcoming barriers: evidence-based strategies to increase enrollment of underrepresented populations in cancer therapeutic clinical trials—a narrative review. *J. Cancer Educ.* **35**, 841–849 (2020).
18. Ford, J. G. et al. Barriers to recruiting underrepresented populations to cancer clinical trials: a systematic review. *Cancer* **112**, 228–242 (2008).
19. Durant, R. W. et al. Perspectives on barriers and facilitators to minority recruitment for clinical trials among cancer center leaders, investigators, research staff, and referring clinicians: enhancing minority participation in clinical trials (EMPaCT). *Cancer* **120**, 1097–1105 (2014).
20. Selected characteristics of the foreign-born population by period of entry into the United States. *US Census Bureau* https://data.census.gov/cedsci/table?q=AMERICAN%20COMMUNITY%20SURVEY&tid=ACSST5Y2020.S0502 (2020).
21. Jones, N., Marks, R., Ramirez, R. & Rios-Vargas, M. 2020 census illuminates racial and ethnic composition of the country. *US Census Bureau* https://www.census.gov/library/stories/2021/08/improved-race-ethnicity-measures-reveal-united-states-population-much-more-multiracial.html (2021).

22. 2019: ACS 1-year estimates selected population profiles. *US Census Bureau* https://data.census.gov/table?q=Language+Spoken+at+Home&tid=ACSST1Y2019.S1601 (2019).
23. Zeigler, K. & Camarota, S. A. 67.3 million in the United States spoke a foreign language at home in 2018. *Center for Immigration Studies* https://cis.org/Report/673-Million-United-States-Spoke-Foreign-Language-Home-2018 (2019).
24. Department of Health, Education, and Welfare; National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research The Belmont Report. Ethical principles and guidelines for the protection of human subjects of research. *J. Am. Coll. Dent.* **81**, 4–13 (2014).
25. A guide to informed consent. *US FDA* https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guide-informed-consent (1998).
26. Klitzman, R. How US institutional review boards decide when researchers need to translate studies. *J. Med. Ethics* **40**, 193–197 (2014).
27. McMillan, G. IRB policies for obtaining informed consent from non-English-speaking people. *Ethics Hum. Res.* **42**, 21–29 (2020).
28. Association, A. T. What is a certified translation? *American Translators Association* https://www.atanet.org/client-assistance/what-is-a-certified-translation/ (2023).
29. Mistretta, S. Amending federal regulations to counteract language barriers in the informed consent process. *Voices Bioethics* https://doi.org/10.52214/vib.v8i.8815 (2022).
30. Resnik, D. B. & Jones, C. W. Research subjects with limited English proficiency: ethical and legal issues. *Account. Res.* **13**, 157–177 (2006).
31. Spiegel, M. L. et al. Non-small cell lung cancer clinical trials requiring biopsies with biomarker-specific results for enrollment provide unique challenges. *Cancer* **123**, 4800–4807 (2017).
32. Informed consent, draft guidance for IRBs, clinical investigators, and sponsors. *US FDA* https://www.fda.gov/regulatory-information/search-fda-guidance-documents/informed-consent#nonenglish (2014).
33. Abadie, A. in *Encyclopedia of Social Measurement* (ed. Kempf-Leonard, K.) 259–266 (Elsevier, 2005).
34. Fundytus, A. et al. Industry funding of oncology randomised controlled trials: implications for design, results and interpretation. *Clin. Oncol. (R. Coll. Radiol.)* **34**, 28–35 (2022).
35. Bodenheimer, T. Uneasy alliance — clinical investigators and the pharmaceutical industry. *New Engl. J. Med.* **342**, 1539–1544 (2000).
36. Califf, R. M. et al. Characteristics of clinical trials registered in ClinicalTrials.gov, 2007–2010. *JAMA* **307**, 1838–1847 (2012).
37. Ehrhardt, S., Appel, L. J. & Meinert, C. L. Trends in National Institutes of Health funding for clinical trials registered in ClinicalTrials.gov. *JAMA* **314**, 2566–2567 (2015).
38. Hakoum, M. B. et al. Characteristics of funding of clinical trials: cross-sectional survey and proposed guidance. *BMJ Open* **7**, e015997 (2017).
39. Nevens, H. et al. Budgeting of non-commercial clinical trials: development of a budget tool by a public funding agency. *Trials* **20**, 714 (2019).
40. Marshall, A. *Principles of Economics: An Introductory Volume* 8th edn (Macmillan, 1920).
41. Fischer, A., Conigliaro, J., Allicock, S. & Kim, E. J. Examination of social determinants of health among patients with limited English proficiency. *BMC Res. Notes* **14**, 299 (2021).
42. Proctor, K., Wilson-Frederick, S. M. & Haffer, S. C. The limited English proficient population: describing medicare, medicaid, and dual beneficiaries. *Health Equity* **2**, 82–89 (2018).
43. Quick facts. *US Census Bureau* https://www.census.gov/quickfacts/fact/table/US# (2022).
44. Schilsky, R. L. Publicly funded clinical trials and the future of cancer care. *The Oncologist* **18**, 232–238 (2013).
45. NIH policy and guidelines on the inclusion of women and minorities as subjects in clinical research. *NIH* https://grants.nih.gov/policy/inclusion/women-and-minorities/guidelines.htm (1993).
46. Schmidt, C. Cooperative groups say NCI trials funding inadequate; some turn to industry. *J. Natl Cancer Inst.* **99**, 830–837 (2007).
47. Txabarriaga, R. *IMIA Guide on Medical Translation*, https://www.imiaweb.org/uploads/pages/438.pdf (IMIA, 2009); .
48. Brelsford, K. M., Ruiz, E. & Beskow, L. Developing informed consent materials for non-English-speaking participants: an analysis of four professional firm translations from English to Spanish. *Clin. Trials* **15**, 557–566 (2018).
49. Staff, A. C. California: 2020 census. *US Census Bureau* https://www.census.gov/library/stories/state-by-state/california-population-change-between-census-decade.html#:~:text=Race%20and%20ethnicity%20(White%20alone,or%20More%20Races%2010.2%25) (2020).

Methods

Study population

After approval by the UCLA IRB, data were collected for all patients signing consent documents for studies conducted at the Jonsson Comprehensive Cancer Center from 1 January 2013 to 31 December 2018, and data on consent events and investigator-reported patient demographics were extracted from the clinical trials database, OnCore (OnCore Enterprise Research, Advarra; Supplementary Methods). Patient characteristics, including primary language, need for a translator, insurance provider and date of birth were obtained from the demographic section of the Epic (Epic Systems) electronic health record. Using each patient's medical record number, patient data were matched to consent event data retrieved from OnCore. Study data was collected and managed using the Research Electronic Data Capture (REDCap) system, and protected health information was manipulated by a third party through the UCLA Department of Biostatistics[50,51].

Language designations

Definitions for primary language can be found in Supplementary Methods. Patients were considered to have limited English proficiency if the demographic section of the electronic health record indicated the need for an interpreter or the medical record review indicated the need for an interpreter during any encounter within six months of the consent date. Chart review on 200 randomly selected consent events for patients with English as their primary language evaluated whether there was an identifiable group requiring an interpreter six months before or after the consent date. On the basis of this analysis, adult patients with English as a primary language were considered proficient in English, whereas English proficiency in paediatric patients was evaluated regardless of the patient's primary language. Paediatric patients with limited English proficiency included those for whom the electronic health record indicated that the patient, or parents or guardians, required an interpreter within six months of the consent date, as the parents or guardians sign the primary consent documents. When a paediatric patient had a primary language documented as English but limited English proficiency (based on the parents or guardians), the patient was considered to have a primary language other than English and limited English proficiency.

Consent language and sponsor assessment

For all patients with a primary language other than English, consent documents were reviewed to determine whether the patient signed consent documents in their primary language. When this information was not available, all IRB-approved translated consent documents were reviewed. We considered patients to have signed consent documents in their primary language if IRB-approved consent documents were available at the time of consent or within the subsequent 30 days (Supplementary Methods).

An additional analysis was restricted to consent events for which there were no translated consent documents at the time of consent or within the subsequent 30 days to identify patients who definitively signed English consent documents. Another analysis evaluated the odds of patients with Spanish as their primary language signing consent documents to studies that had Spanish consent documents at study opening.

Study type and sponsor assessment

The cancer centre labels studies as interventional when a clear pharmacologic, dietary, lifestyle intervention, procedural or diagnostic intervention was performed with other studies labelled as observational. We lacked access to complete budgetary data, but the study sponsor was documented. Studies considered to be industry sponsored had a biopharmaceutical company that evaluated a drug, device or procedure serve as the principal funding sponsor. All other studies were considered non-industry sponsored. An additional analysis was performed, dividing studies based on whether any funds for consent document translation could have been provided by an industry partner (that is, the study did not receive funding from industry beyond study drug or device) versus studies in which funds for consent document translation from industry could not be ruled out. Studies including only patients younger than 18 were considered to be paediatric only, whereas studies that included only patients 18 or older were considered to be adult only. Studies were also reviewed to assess whether they included a single solid or haematologic malignancy, multiple histologies or healthy patients.

Assessment of cost of consent document translation

For simplicity, we assumed that every study had the initial consent document translated at 20 cents per word, the median cost for translation paid by the cancer centre during the evaluated period (Supplementary Methods).

Statistical analyses

Patient characteristics were summarized using frequency (%) and compared using Pearson chi-squared tests (Supplementary Methods). The median number of consent events between studies sponsored and not sponsored by industry were compared using the Wilcoxon rank sum test.

Logistic regression models with generalized estimating equations clustered by patient unique identifier to adjust for repeated measures compared consent events for non-industry-sponsored versus industry-sponsored studies. As a sensitivity analysis, the same generalized estimating equation models were run specifying patients nested within each study as the repeated effect. Models were constructed in two consent-event groupings: all consent events and the subset in which patients signed consent documents in their primary language. The main explanatory variable was a language-grouping variable (English primary versus primary other than English or limited English proficiency). Additional covariates were prospectively identified: age at consent, a single category for race and ethnicity in which Hispanic patients were coded as such regardless of race (that is, Hispanic, Black, Asian or Pacific Islander, other (which included race or ethnicities in whose proportion in the evaluated population was less than 4.0%) or non-Hispanic white), female versus male, interventional versus non-interventional, and the study's included histologies (single haematologic malignancy, solid malignancy, multiple histologies or healthy patients). A variable evaluating whether each patient had Medi-Cal as their payor was subsequently added for specific analyses (patient with Medi-Cal as their payor, Yes or No). For each set of models, we first constructed bivariable and then multivariable models. Additional analyses estimated the effect of the language-grouping variable within subgroups based on the department conducting the study and interventional studies. Consent events missing the primary language were excluded from all analyses. Other methods for handling missing data are described in Supplementary Methods.

The McNemar's test compared the subset of patients who signed consent documents for both industry-sponsored and non-industry-sponsored studies to identify the probability of signing translated consent documents for a study based on whether or not the study had industry sponsorship (Supplementary Methods).

For all tests, a two-tailed P value < 0.05 was considered statistically significant. Data were analysed using SAS software, version 9.3 (SAS Institute) and JMP Pro 16.0 (SAS Institute).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All consent-event data associated with this study and data contained in Extended Data tables are freely available at https://doi.org/10.5281/zenodo.7992491.

50. Harris, P. A. et al. The REDCap consortium: building an international community of software platform partners. *J. Biomed. Inform.* **95**, 103208 (2019).
51. Harris, P. A. et al. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* **42**, 377–381 (2009).

Acknowledgements We thank J. Allen and R. Flores Stella for their assistance. E.B.G. was funded by NIH-NCI R01 CA276917, the Cancer Center Support Grant P30 CA016042, and NIH-NCATS UL1TR001881, CLIN-10784. A.L.C. was funded by a Specialty Training and Advanced Research (STAR) Award. A.L. was funded by NIH-NCI K08 CA245249-01A1 and a LUNGevity 2019 Career Development Award. E.B.G. is also funded by CLIN-10784. M.A.T. was funded by P30 CA016042- Cancer Center Support Grant, UCLA Jonsson Comprehensive Cancer Center. D.E. was funded by the National Center for Advancing Translational Science (NCATS) of the National Institutes of Health under the UCLA Clinical and Translational Science Institute grant number UL1TR001881. Figure 1 was created using BioRender.com. This work was funded by E.B.G.'s UCLA Jonsson Comprehensive Cancer Center Seed Grant.

Author contributions M.A.V., B.A.G. and E.B.G. contributed substantially to the conception and design of the work. M.A.V., A.N., Y.R., J.P.L.-L., P.M.B., D.Y.L., M.J.F.-T., M.L., M.H., J.B., T.G., D.E. and E.B.G. contributed to data acquisition, analysis and interpretation. M.A.V. and E.B.G. drafted the paper. B.A.G., A.L.C., M.G.-J., A.G., M.A.T., A.L. and E.B.G. contributed substantially to paper revision.

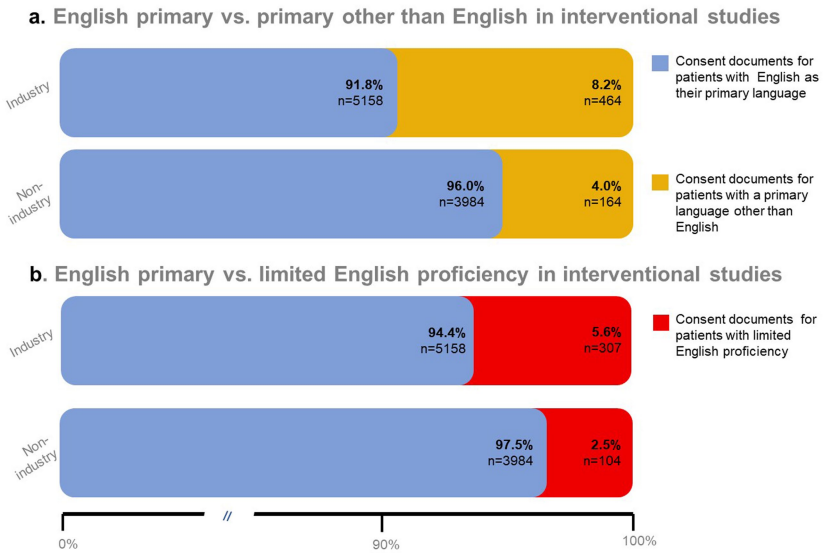
Competing interests E.B.G. has been a consultant and/or advisor for Abbvie; ABL-Bio; AstraZeneca, Boehringer-Ingelheim; Bristol-Myers Squibb; Dracen Pharmaceuticals; EMD Serono; Eisai; Eli Lilly; Gilead; GlaxoSmithKline; Ipsen; Merck; Natera; Novartis; Personalis; Regeneron; Sanofi; Shionogi; and Xilio. E.B.G. has received grant or research support from ABL-Bio; AstraZeneca; Bristol-Myers Squibb; Dynavax Technologies; Eli Lilly; EMD Serono; Genentech; Iovance Biotherapeutics; Merck; Mirati Therapeutics; Neon; and Novartis. A.L. has received commercial research grants from Daiichi Sankyo, Calithera Biosciences, AstraZeneca, Dracen Pharmaceuticals, WindMIL, eFFECTOR Therapeutics. A.E.L. has served as a consultant or on the advisory board for, AstraZeneca, Bristol-Myers Squibb, Leica Biosystems, Jazz Pharmaceuticals, Novocure, Pfizer, MorphoSys, Eli Lilly, Oncocyte, Novartis, Regeneron, Janssen oncology, Sanofi group of companies. M.A.V., B.A.G., M.G.-J., A.N., Y.R., J.P.L.-L., P.M.B., D.Y.L., M.J.F.-T., A.G., M.L., M.H., J.B. and T.G. declare no competing interests.

Additional information

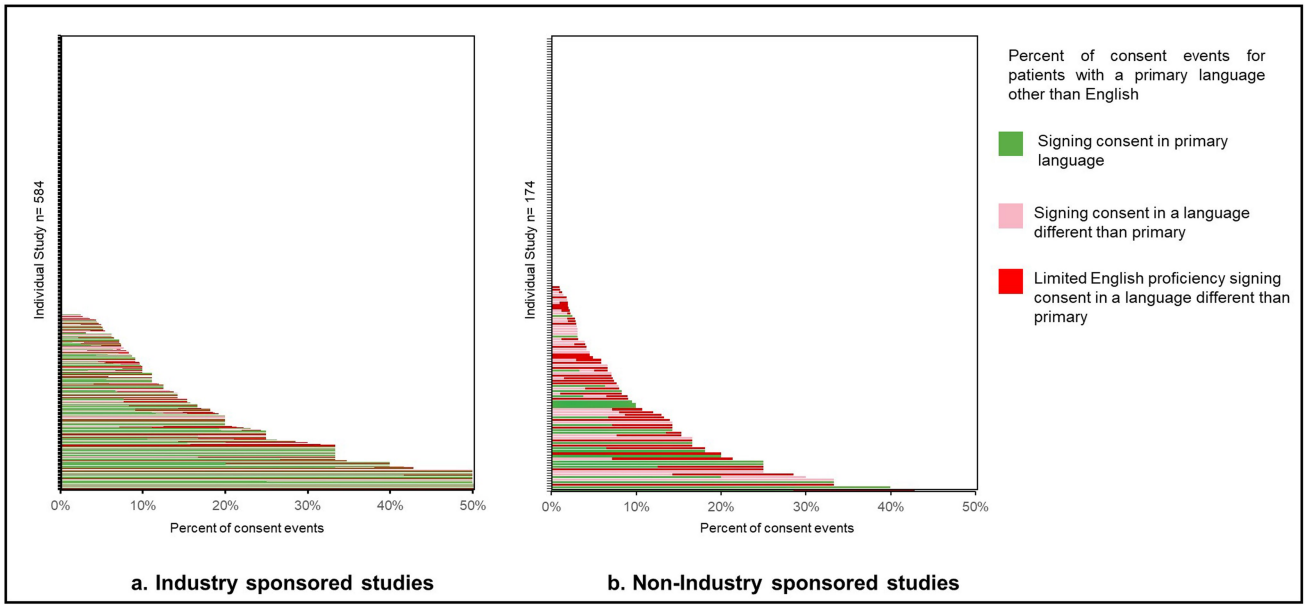
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-023-06382-0.

Correspondence and requests for materials should be addressed to Edward B. Garon.

Peer review information *Nature* thanks Jason Abaluck, Jan Eberth and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.



Extended Data Fig. 1 | Comparison of the proportion of consent events based on primary language and English proficiency in interventional industry versus non-industry sponsored studies. a. Blue indicates the proportion of consent events for patients with English as their primary language. Yellow indicates the proportion of consent events for patients with a primary language other than English in industry sponsored studies (top bar) versus non-industry sponsored studies (bottom bar) (8.2% versus 4.0%, $p < 0.001$). **b**. Blue indicates the proportion of consent events for patients with English as their primary language. Red indicates the proportion of consent events for patients with limited English proficiency in industry sponsored studies (top bar) versus non-industry sponsored studies (bottom bar) (5.6% versus 2.5%, $p < 0.001$). Logistic regression models with Generalized Estimating Equations clustered by patient unique identifier were used to test comparisons above, *P* values reported are two-tailed.



Extended Data Fig. 2 | Comparison of the percentage of patients signing consent documents per study between industry and non-industry sponsored studies. The X axis depicts the percentage of patients signing consent documents for each study who had a primary language other than English in industry (a) and non-industry sponsored studies (b). Each study is represented by a row in the Y axis. The rows are taller for non-industry sponsored studies compared to industry sponsored studies as there are fewer of them. Green denotes the percentage of patients who signed consent in their primary language, pink represents the percentage of patients signing consent documents in a language different than their primary language and red represents the percentage of patients with limited English proficiency signing consent document in a language different than primary.

Extended Data Table 1 | Characteristics of patients included in the analysis who signed consent documents to Cancer Center studies

Characteristic, n (%)	Total n=9213	English Primary n=8636 (93.7)	Primary Other than English n=577(6.3)	P Value[a]	Limited English Proficiency n=376 (4.0)	P Value[b]
Gender[c]						
Female	3513 (38.1)	3255 (37.7)	260 (45.1)	<0.001[d]	175 (46.5)	<0.001[e]
Male	5686 (61.7)	5368 (62.2)	316 (54.8)		201 (53.4)	
Race and Ethnicity[f]				<0.001[g]		<0.001[h]
Asian or Pacific Islander	901 (9.8)	680 (7.8)	221 (38.3)	<0.001	139 (37.0)	<0.001
Black	389 (4.2)	389 (4.5)	0	<0.001	0	<0.001
Other[i]	279 (3.0)	262 (3.0)	17 (3.0)	0.725	13 (3.5)	0.804
Hispanic	902 (9.7)	687 (7.9)	215 (37.2)	<0.001	160 (42.5)	<0.001
Non-Hispanic White	5840 (63.4)	5744 (66.5)	96 (16.6)	<0.001	50 (13.3)	<0.001
Medi-Cal as payor[j]						
Yes	879 (9.5)	597 (6.9)	282 (48.8)	<0.001[k]	206 (55.1)	<0.001[l]
No	8270 (89.8)	7983 (92.4)	287 (49.7)		162 (43.3)	

[a]Comparison between patients with primary language other than English and English as primary language. [b]Comparison between patients with limited English proficiency and English as primary language. [c]Unknown Total: 14 (0.2%). [d]Comparison of the proportion of female and male patients between patients with a primary language other than English and English as primary language. [e]Comparison of the proportion of female and male patients between patients limited English proficiency and patients with English as primary language.
[f]Unknown Total: 902 (9.7%). [g]Comparison of the proportion by racial and ethnic groups among patients with primary language other than English and English as primary language.
[h]Comparison of the proportion by racial and ethnic groups among patients with limited English proficiency and English as primary language. [i]Other: American Indian; 16 (0.2%), Multiracial; 28 (0.3%), Other; 235 (2.5%). [j]Unknown Total: 64 (0.7%). [k]Comparison of the proportion of patients with Medi-Cal as their payor between patients with a primary language other than English and patients with English as primary language. [l]Comparison of the proportion of patients with Medi-Cal as their payor between patients with limited English proficiency and patients with English as primary language. All statistical comparisons were performed using Pearson's chi-squared test, no adjustments were made for multiple comparisons. Abbreviations: n, number. Racial and ethnic groups representing less than 4.0% of the study population were included as "other".

Extended Data Table 2 | Primary languages spoken by patients with a primary language other than English

Language	Number of patients	Percent
Spanish	231	40.8%
Chinese[a]	118	20.8%
Korean	63	11.1%
Farsi, Persian	33	5.8%
Armenian	30	5.3%
Russian	20	3.5%
Vietnamese	19	3.4%
Japanese	12	2.1%
Arabic	8	1.4%
Other	6	1.1%
Tagalog	6	1.1%
Hindi	<5	<0.4%
Hungarian	<5	<0.4%
Lithuanian	<5	<0.4%
Thai	<5	<0.4%
Afar	<5	<0.4%
Burmese	<5	<0.4%
Cambodian	<5	<0.4%
Danish	<5	<0.4%
Ethiopian	<5	<0.4%
French	<5	<0.4%
Greek	<5	<0.4%
Hebrew	<5	<0.4%
Indonesian	<5	<0.4%
Italian	<5	<0.4%
Laotian	<5	<0.4%
Ukrainian	<5	<0.4%

[a]Chinese includes Mandarin, Cantonese, and Simplified Chinese.

Extended Data Table 3 | Characteristics and distribution of consent events in non-industry and industry sponsored studies

	Non-Industry Sponsored Study	Industry Sponsored Study	P Value
Studies, N (%)	173 (22.9)	585 (77.1)	
Studies with at least one translated consent document, N (% per sponsor type)	39 (23)	222 (38)	<0.001[a]
Consent events, N (%)	6344 (52.5)	5738 (47.5)	
Consent events for studies that translated at least one consent document, N (%, per sponsor type)	1513 (23.9)	2951 (51.4)	<0.001[a]
Median number of consent events per study (range)	8.0 (1-791)	5.0 (1-206)	<0.001[b]
Median number of consent events for studies that translated at least one consent document (range)	9.5 (1-510)	8.0 (1-206)	0.577[b]
Median number of consent events for studies that did not translate at least one consent (range)	8.0 (1-791)	4.0 (1-93)	<0.001[b]
Number of Interventional studies, N (%)	143 (19.8)	577(80.2)	
Median number of consent events for interventional studies (range)	8.0 (1-585)	5.0 (1-206)	<0.001 [b]

[a]Statistical comparisons were performed using Pearson's chi-squared. [b]Statistical comparisons were performed using the Wilcoxon Rank Sum Test. *P* values reported are two tailed. No adjustments were made for multiple comparisons. Abbreviations, OR; odds ratio, CI; confidence Interval.

Extended Data Table 4 | Analysis of studies with translated consent documents in Spanish and odds of signing consent based on consent document translation availability at study opening

	Non-Industry Sponsored Study	Industry Sponsored Study
Studies with translated consent documents in Spanish available at the start of study opening, N[a]	2	10
Consent events for studies with Spanish consent documents available at the start of study opening, N (%, per sponsor type)[b]	127 (2.0)	142 (2.4)
Consent events for patients with a primary language other than English[c]	10 (3.6)	31 (6.7)
Consent events for patients with limited English proficiency[d]	5 (2.8)	26 (8.4)
Consent events for patients with Spanish as their primary language[e]	9 (6.9)	23 (13.0)
Consent events for patients with Spanish as their primary language who had limited English proficiency[f]	5 (5.7)	20 (16.1)
At study opening	**OR**	**95% CI**
Odds for patients with Spanish as their primary language of signing of consent to studies with Spanish translation compared to studies without them.	5.7	3.8 – 8.5
Odds for patients with a primary language other than English or Spanish of signing consent to studies with a Spanish translation compared to studies without them.	0.9	0.6 – 1.3
At any time during the study		
Odds for patients with a primary language other than English of signing consent in a language different than primary to studies that had translated consent documents.	0.02	0.009 – 0.030
Odds for patients with limited English proficiency of signing consent in a language different than primary to studies that had translated consent documents.	0.02	0.009 – 0.037

[a]Other languages in which studies had available translated consent documents at study opening included: Chinese (n=1), Farsi (n=1), Hebrew (n=1), Japanese (n=1), Korean (N=1), Thai (n=1). [b]Denominator is consent events per sponsor type (non-industry sponsored studies n = 6344, industry sponsored studies n = 5738). [c]Denominator is consent events for patients with a primary language other than English within that sponsor type (non-industry sponsored studies n=278, industry sponsored studies n = 464). [d]Denominator is consent events for patients with limited English proficiency within that sponsor type (non-industry sponsored studies n=174, industry sponsored studies n=307). [e]Denominator is consent events for patients with Spanish as their primary language within that sponsor type (non-industry sponsored studies n=129, industry sponsored studies n=176). [f]Denominator is consent events for patients with Spanish and limited English proficiency within that sponsor type (non-industry sponsored studies n=88, industry sponsored studies n=124).

Odds ratios (with 95% CI) were estimated from a logistic regression model with Generalized Estimating Equations clustered by patient unique identifier. *P* values reported are two tailed. No adjustments were made for multiple comparisons. Abbreviations, OR; odds ratio, CI; confidence Interval.

Extended Data Table 5 | Proportion of consent events in which patients with a primary language other than English and limited English proficiency who signed consent documents in a language different than primary in industry sponsored and non-industry sponsored studies across Departments

Department	Proportion (%)	95% CI	P Value
Primary language other than English			
Medicine			<0.001
Industry Sponsored Study	39.5	34.8 – 44.4	
Non-Industry Sponsored Study	62.5	53.5 – 70.9	
Radiology and Molecular Pharmacology			0.003
Industry Sponsored Study	60.0	26.2 – 87.8	
Non-Industry Sponsored Study	98.4	91.2 – 99.9	
Pediatrics			NA[a]
Industry Sponsored Study	0.0	0.0 – 84.1	
Non-Industry Sponsored Study	19.2	6.5 – 39.5	
Radiation Oncology			NA[b]
Industry Sponsored Study	100.0	63.0 – 100	
Non-Industry Sponsored Study	100.0	79.4 – 100	
Surgical specialties			0.082
Industry Sponsored Study	82.6	6.1.2 – 95.5	
Non-Industry Sponsored Study	97.3	85.8 – 99.9	
Other			NA[c]
Industry Sponsored Study	100	54.0 – 100	
Non-Industry Sponsored Study	44.4	13.7 – 78.8	
Limited English Proficiency			
Medicine			<0.001
Industry Sponsored Study	27.4	22.2 – 33.0	
Non-Industry Sponsored Study	58.2	46.5 – 69.2	
Radiology and Molecular Pharmacology			0.013
Industry Sponsored Study	55.6	21.2 – 86.3	
Non-Industry Sponsored Study	97.4	86.1 – 99.9	
Pediatrics			NA[a]
Industry Sponsored Study	0.0	0.0 – 84.2	
Non-Industry Sponsored Study	19.2	6.6 – 39.4	
Radiation Oncology			NA[b]
Industry Sponsored Study	100.0	54.0 – 100	
Non-Industry Sponsored Study	100.0	66.3 – 100	
Surgical specialties			0.112
Industry Sponsored Study	69.2	38.5 – 90.9	
Non-Industry Sponsored Study	93.7	69.7 – 99.8	
Other			NA[c]
Industry Sponsored Study	100.0	29.2 – 100	
Non-Industry Sponsored Study	40.0	5.3 – 85.3	

[a]*P* value could not be generated because there were no patients who signed consent documents in a language different than primary in industry sponsored studies. [b]*P* value could not be generated because there were no consent documents translated. [c]*P* value could not be generated because there were no patients who signed consent documents in a language different than primary in industry sponsored studies. A logistic regression model with Generalized Estimating Equations clustered by patient unique identifier was used to compare the proportions above. *P* values reported are two tailed. No adjustments were made for multiple comparisons.

Extended Data Table 6 | Bivariable analysis odds ratio for the association between various factors and signing consent into a non-industry sponsored study

	Bivariable Analysis for signing consent documents			Bivariable Analysis for signing consent documents in patient's primary language		
Variable	**OR**	**95% CI**	**P Value**	**OR**	**95% CI**	**P Value**
Age						
Age at consent (per year)	0.99	0.98-0.99	<0.001	0.99	0.98-0.99	<0.001
Language						
English Primary		**Reference**			**Reference**	
Primary Other than English[a]	0.50	0.43-0.59	<0.001	0.25	0.19-0.32	<0.001
Limited English Proficiency[b]	0.47	0.38-0.57	<0.001	0.24	0.18-0.32	<0.001
Race and Ethnicity						
Non-Hispanic White		**Reference**			**Reference**	
Asian or Pacific Islander	0.61	0.53-0.69	<0.001	0.61	0.53-0.70	<0.001
Black	1.22	0.92-1.36	0.250	1.13	0.93-1.37	0.215
Hispanic	0.81	0.71-0.92	0.002	0.79	0.69-0.90	<0.001
Other	1.44	1.13-1.83	0.004	1.44	1.13-1.85	0.004
Unknown	3.34	2.89-3.94	<0.001	3.36	2.84-3.97	<0.001
Study Type						
Interventional		**Reference**			**Reference**	
Observational	32.4	25.1-41.8	<0.001	31.1	24.1-40.1	<0.001
Gender						
Male		**Reference**			**Reference**	
Female	0.37	0.34-0.40	<0.001	0.36	0.33-0.39	<0.001
Histology						
Single Solid Malignancy		**Reference**			**Reference**	
Healthy	1.62	1.23-2.14	<0.001	1.70	1.28-2.27	<0.001
Multiple Histology	0.46	0.42-0.50	<0.001	0.45	0.41-0.49	<0.001
Single Heme Malignancy	0.07	0.05-0.09	<0.001	0.07	0.05-0.09	<0.001

[a]Patients with a primary language other than English compared to patients with English as their primary language. [b]Patients with limited English proficiency compared to patients with English as their primary language. Odds ratios (with 95% CI) were estimated from a logistic regression model with Generalized Estimating Equations clustered by patient unique identifier. *P* values reported are two tailed. No adjustments were made for multiple comparisons. Abbreviations, OR; odds ratio, CI; confidence Interval.

Extended Data Table 7 | Multivariable analysis for odds ratio for the association between various factors including Medi-Cal status and signing consent into a non-industry sponsored study nested by individual patient

	Multivariable Analysis for patients with a primary language other than English signing consent documents			Multivariable Analysis for patients with a primary language other than English signing consent documents in patient's primary language			Multivariable Analysis for patients with limited English proficiency signing consent documents			Multivariable Analysis for patients with limited English proficiency signing consent documents in patient's primary language		
Variable	OR	95% CI	P Value	OR	95% CI	P Value	OR	95% CI	P Value	OR	95% CI	P Value
Age												
Age at consent (per year)	0.97	0.97-0.98	<0.001	0.98	0.97-0.98	<0.001	0.97	0.97-0.98	<0.001	0.98	0.97-0.98	<0.001
Language												
English Primary		Reference			Reference			Reference			Reference	
Primary Other than English[a]	0.78	0.63-0.98	0.033	0.39	0.29-0.54	<0.001	-	-	-	-	-	-
Limited English Proficiency[b]	-	-	-	-	-	-	0.80	0.62-1.00	0.105	0.38	0.27-0.54	<0.001
Race and Ethnicity												
Non-Hispanic White		Reference			Reference			Reference			Reference	
Asian or Pacific Islander	0.64	0.54-0.76	<0.001	0.66	0.56-0.79	<0.001	0.65	0.55-0.77	<0.001	0.66	0.55-0.79	<0.001
Black	1.03	0.81-1.29	0.833	1.03	0.81-1.30	0.787	1.03	0.82-1.30	0.805	1.03	0.81-1.30	0.776
Hispanic	0.75	0.63-0.89	0.001	0.76	0.63-0.91	0.003	0.74	0.62-0.88	<0.001	0.75	0.63-0.90	<0.002
Other	1.16	0.87-1.54	0.306	1.16	0.87-1.56	0.306	1.15	0.87-1.54	0.323	1.16	0.87-1.56	0.304
Unknown	3.36	2.83-3.98	<0.001	3.36	2.83-3.99	<0.001	3.40	2.86-4.04	<0.001	3.39	2.89-4.04	<0.001
Study Type												
Interventional		Reference			Reference			Reference			Reference	
Observational	36.3	28.4-46.5	<0.001	35.2	27.4-45.1	<0.001	35.5	27.7-45.4	<0.001	35.0	27.3-45.8	<0.001
Gender												
Male		Reference			Reference			Reference			Reference	
Female	0.38	0.34-0.42	<0.001	0.37	0.33-0.41	<0.001	0.37	0.34-0.42	<0.001	0.37	0.38-0.45	<0.001
Histology												
Single Solid Malignancy		Reference			Reference			Reference			Reference	
Healthy	1.73	1.30-2.3	<0.001	1.82	1.38-2.43	<0.001	1.77	1.34-2.30	<0.001	1.83	1.38-2.40	<0.001
Multiple Histology	0.38	0.35-0.43	<0.001	0.38	0.34-0.42	<0.001	0.39	0.35-0.43	<0.001	0.38	0.35-0.42	<0.001
Single Heme Malignancy	0.06	0.04-0.09	<0.001	0.06	0.04-0.09	<0.001	0.06	0.04-0.09	<0.001	0.06	0.04-0.09	<0.001
Medi-Cal[c]												
No		Reference			Reference			Reference			Reference	
Yes	0.81	0.68-0.97	0.021	0.81	0.67-0.98	0.034	0.78	0.66-0.95	0.015	0.83	0.66-0.97	0.025
Unknown	3.66	1.73-7.64	<0.001	3.73	1.97-8.12	<0.001	3.42	1.61-7.2	<0.001	3.49	1.55-7.72	<0.001

[a]Patients with a primary language other than English compared to patients with English as their primary language. [b]Patients with limited English proficiency compared to patients with English as their primary language. [c]The *P* value for the overall interaction between Medi-Cal and primary language other than English or English as primary language was p=0.163, and the overall *P* value for the interaction between Medi-Cal and limited English proficiency, English as primary language was *P*=0.275. Odds Ratios were estimated with a Generalized Estimating Equation Model clustered by patient. *P* values reported are two tailed. No adjustments were made for multiple comparisons. Abbreviations, OR; odds ratio, CI; confidence Interval.

Extended Data Table 8 | Multivariable analysis for odds ratio for the association between various factors and signing consent into a non-industry sponsored study nested by study

	Multivariable Analysis for patients with a primary language other than English signing consent documents			Multivariable Analysis for patients with a primary language other than English signing consent documents in patient's primary language			Multivariable Analysis for patients with limited English proficiency signing consent documents			Multivariable Analysis for patients with limited English proficiency signing consent documents in patient's primary language		
Variable	OR	95% CI	P Value	OR	95% CI	P Value	OR	95% CI	P Value	OR	95% CI	P Value
Age												
Age at consent (per year)	0.97	0.97-0.98	<0.001	0.98	0.97-0.98	<0.001	0.97	0.97-0.98	<0.001	0.98	0.97-0.98	<0.001
Language												
English Primary		**Reference**			**Reference**			**Reference**			**Reference**	
Primary Other than English[a]	0.79	0.65-0.96	0.019	0.40	0.30-0.54	<0.001	-	-	-	-	-	-
Limited English Proficiency[b]	-	-	-	-	-	-	0.78	0.61-0.99	0.045	0.38	0.27-0.53	<0.001
Race and Ethnicity												
Non-Hispanic White		**Reference**			**Reference**			**Reference**			**Reference**	
Asian or Pacific Islander	0.55	0.48-0.66	<0.001	0.58	0.49-0.69	<0.001	0.63	0.53-0.73	<0.001	0.58	0.49-0.69	<0.001
Black	0.95	0.76-1.28	0.608	0.94	0.76-1.17	0.605	1.05	0.84-1.30	0.660	0.94	0.76-1.18	0.600
Hispanic	0.67	0.58-0.78	<0.001	0.68	0.58-0.80	<0.001	0.75	0.64-0.88	<0.001	0.68	0.57-0.80	<0.001
Other	0.99	0.77-1.28	0.972	1.00	0.77-1.29	0.996	1.13	0.87-1.45	0.358	1.00	0.77-1.29	0.996
Unknown	1.12	0.86-1.41	0.315	1.12	0.89-1.40	0.318	3.31	2.78-3.94	<0.001	1.12	0.89-1.40	0.318
Study Type												
Interventional		**Reference**			**Reference**			**Reference**			**Reference**	
Observational	30.3	24.8-37.2	<0.001	29.2	23.8-35.9	<0.001	31.3	25.6-28.8	<0.001	29.1	23.7-35.8	<0.001
Gender												
Male		**Reference**			**Reference**			**Reference**			**Reference**	
Female	0.36	0.33-0.39	<0.001	0.36	0.32-0.39	<0.001	0.38	0.35-0.42	<0.001	0.35	0.32-0.39	<0.001
Histology												
Single Solid Malignancy		**Reference**			**Reference**			**Reference**			**Reference**	
Healthy	1.73	1.30-2.3	<0.001	1.83	1.37-2.44	<0.001	1.78	1.37-2.52	<0.001	1.84	1.41-2.45	<0.001
Multiple Histology	0.33	0.30-0.37	<0.001	0.33	0.30-0.36	<0.001	0.35	0.32-0.39	<0.001	0.33	0.30-0.37	<0.001
Single Heme Malignancy	0.06	0.04-0.09	<0.001	0.06	0.04-0.09	<0.001	0.06	0.04-0.09	<0.001	0.06	0.04-0.09	<0.001

[a]Patients with a primary language other than English compared to patients with English as their primary language. [b]Patients with limited English proficiency compared to patients with English as their primary language. Odds ratios (with 95% CI) were estimated from a logistic regression model with Generalized Estimating Equations clustered by study, *P* values reported are two-tailed. Abbreviations, OR; odds ratio, CI; confidence Interval.

nature portfolio

Corresponding author(s): Edward B Garon

Last updated by author(s): May 31, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Study data were collected and managed using REDCap electronic data capture tools hosted at UCLA.
Data analysis	SAS software, version 9.3 (SAS Institute) and JMP Pro 16.0 (SAS Institute Inc., Cary, NC, USA).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The data will be made publicly available on the public repository zenodo.org

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	The find- ings of our study apply to only both male and female gender. These were investigator-reported. Our study spans from 2013 to 2018. The electronic health record and the clinical trial repository were not linked during this period. As such, demographic data in this study is considered to be investigator rather than patient reported as research personnel manually inputted data in the clinical trial repository where data was collected for this analysis. Gender was included in both univariate as well as multivariable models are predictor variables. There were 3513 female patients and 5686 male patients included in our study there were 14 patients for which the gender was marked as "Unkown" in the electronic health record.
Reporting on race, ethnicity, or other socially relevant groupings	Race and ethnicity for patients who participated in studies during the study period were collected from the clinical trial repository Oncore. These were investigator reported as our study spans from 2013 to 2018 when the electronic health record were not linked. As such we only have data that was collected manually from clinical trial research personnel. There were consent for patients for which this variable was not collected and as such these were left as "Unknown".
Population characteristics	The evaluation assessed 12,082 consent events in 9213 patients. This represented six years of data from the Cancer Center that was analyzed for the Cancer Center support grant competitive renewal. All consent events in all Cancer Center Departments were included as long as there was evaluable data for primary language. Population characteristics as predictor variables: - Age at the time of consent (analyzed as a continuous variable) - Primary language: Patients whose primary language was not English, including patients whose primary language was "Other" in the electronic health record, were coded as "primary language other than English" (note: this group included pediatric patients whose primary language was English, but whose parent or guardian required an interpreter) Patients whose primary language was English were coded as "English as primary language" - English proficiency For all analyses, the comparator group for patients with limited English proficiency were patients coded as "English as primary language" as above Those falling into either group below were coded as having "Limited English proficiency": Adult patients for whom the electronic health record indicated that they needed interpreter or chart review indicated a need for an interpreter within 6 months of the consent date. Pediatric patients for whom the electronic health record indicated that the patient needed an interpreter or for whom the parents or guardians required an interpreter within 6 months of the consent date. - Race and ethnicity Hispanic ethnicity (coded as such regardless of race) Asian or Pacific Islander (as a single category) Black Non-Hispanic White Other (American Indian, Multiracial, Other) Unknown - Gender Female Male Unknown -Safety net insurance status (Medi-Cal) Yes No Unknown
Recruitment	This was a retrospective study that included patients who signed consent to studies at our institution from 2013 to 2018.
Ethics oversight	Protocol was submitted and accepted by the University of California, Los Angeles, IRB on March 2020 prior to any data collection

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size calculation was not performed as we prospectively sought to evaluate and include all consent events for patients who participated in clinical trials and for whom data was available. The time period elected for our study was the last NCI competing renewal

	period. Our study included over 12000 consent events and even though sample size calculation was not prospectively performed this represented enough power for our analysis. However because our study was retrospective in nature and we planned on including all consent events were data was available we did not limit the sample to a certain pre-specificied sample size.
Data exclusions	Of 13,717 consent events between January 2013 and December 2018, those excluded from further analysis included 303 for which no medical record number was available, 1,212 at affiliated sites for which electronic health record access was not available, and 120 for which the primary language could not be identified.
Replication	Replication was not applicable to our study was we only had one dataset in which the statistical analyses were run. Statistical tests were performed by the first author (Maria A Velez) and by Tristan Grogan separately on the same dataset to ensure that results were consistent.
Randomization	Randomization was not relevant to our study as this was a retrospective observational (cross-sectional) study.
Blinding	Blinding was not relevant to our study as this was a retrospective observational (cross-sectional) study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	N/A, this was not a clinical trial
Study protocol	submitted with previous versions of the manuscript
Data collection	Data was collected from the electronic health record and clinical trial repository.
Outcomes	The outcome was sponsor type: signing consent documents in an industry sponsored study versus a non-industry sponsored study