



Evaluation of the cost and effectiveness of diverse recruitment methods for a genetic screening study

Hila Milo Rasouly, PhD¹, Julia Wynn, MS, CGC², Maddalena Marasa, MD¹, Rachel Reingold, BS¹, Debanjana Chatterjee, PhD¹, Sheena Kapoor, BS¹, Stacy Piva, BS¹, Byum Hee Kil, BS¹, Xueru Mu, MD¹, Maria Alvarez, MD¹, Jordan Nestor, MD¹, Karla Mehl, MD¹, Anya Revah-Politi, MS, CGC³, Natalie Lippa, MS, CGC³, Michelle E. Ernst, MS, CGC³, Louise Bier, MS, CGC³, Aileen Espinal, BS², Bianca Haser, BS², Anoushka Sinha, BS⁴, Ian Halim, BA⁴, David Fasel, BS⁵, Nicole Cuneo, BS¹, Jacqueline J. Thompson, BS⁴, Miguel Verbitsky, PhD¹, Elizabeth G. Cohn, Ph.D⁶, Jill Goldman, MS, CGC⁶, Karen Marder, MD, MPH⁶, Robert L. Klitzman, MD⁷, Manuela A. Orjuela, MD, MPH^{2,8}, Yat S. So, BS⁵, Alex Fedotov, PhD⁹, Katherine D. Crew, MD¹, Krzysztof Kiryluk, MD, MS¹, Paul S. Appelbaum, MD⁷, Chunhua Weng, PhD⁵, Karolynn Siegel, PhD¹⁰, Ali G. Gharavi, MD¹ and Wendy K. Chung, MD, PhD^{1,2}

Purpose: Recruitment of participants from diverse backgrounds is crucial to the generalizability of genetic research, but has proven challenging. We retrospectively evaluated recruitment methods used for a study on return of genetic results.

Methods: The costs of study design, development, and participant enrollment were calculated, and the characteristics of the participants enrolled through the seven recruitment methods were examined.

Results: A total of 1118 participants provided consent, a blood sample, and questionnaire data. The estimated cost across recruitment methods ranged from \$579 to \$1666 per participant and required a large recruitment team. Recruitment methods using flyers and staff networks were the most cost-efficient and resulted in the highest completion rate. Targeted sampling that emphasized the importance of Latino/a participation, utilization of translated

materials, and in-person recruitments contributed to enrolling a demographically diverse sample.

Conclusions: Although all methods were deployed in the same hospital or neighborhood and shared the same staff, each recruitment method was different in terms of cost and characteristics of the enrolled participants, suggesting the importance of carefully choosing the recruitment methods based on the desired composition of the final study sample. This analysis provides information about the effectiveness and cost of different methods to recruit adults for genetic research.

Genetics in Medicine (2019) 21:2371–2380; <https://doi.org/10.1038/s41436-019-0497-y>

Keywords: genomic research; personalized medicine; recruitment methods; socioeconomic diversity; cost-efficiency

INTRODUCTION

Precision medicine holds many promises, including the implementation of genomic medicine for all.¹ Like genetic risk factors that vary based on ancestry, concerns and expectations about genomic medicine may also differ based on individuals' characteristics such as ethnicity, education, religion, or culture.² Therefore, enrollment of cohorts reflecting the heterogeneity of the United States' population is crucial to ensure the wide applicability of research findings.

As enrollment inclusiveness is a problem across medical research, multiple recruitment methods have been tested, including targeted recruitment of minorities, recruitment in health-care facilities serving underrepresented communities, engagement of community leaders, and incentives.^{3–9} To overcome the historic disparities in participation in medical research, some strategies have also focused on educating prospective participants about the value of research, increasing communication and engagement with participants

¹Department of Medicine, Columbia University Medical Center, New York, NY, USA; ²Department of Pediatrics, Columbia University Medical Center, New York, NY, USA; ³Institute for Genomic Medicine, Columbia University Medical Center, New York, NY, USA; ⁴College of Physician & Surgeons, Columbia University, New York, NY, USA; ⁵Department of Biomedical Informatics, Columbia University Irving Medical Center, New York, NY, USA; ⁶Department of Neurology, Aging & Dementia, Columbia University Medical Center, New York, NY, USA; ⁷Department of Psychiatry, Columbia University Medical Center, New York, NY, USA; ⁸Department of Epidemiology, Columbia University Medical Center, New York, NY, USA; ⁹Irving Institute of Clinical and Translational Research, Columbia University Medical Center, New York, NY, USA; ¹⁰Department of Sociomedical Sciences, Columbia University Mailman School of Public Health, New York, NY, USA. Correspondence: Wendy K. Chung (wkc15@cumc.columbia.edu)

These authors contributed equally: Hila Milo Rasouly and Julia Wynn

Submitted 31 October 2018; accepted: 12 March 2019

Published online: 1 April 2019

throughout the course of the study, and employing ethnically diverse and stable research teams.^{3,7,10}

Enrolling minorities in genomic research has been reported to be especially challenging, in part due to concerns and mistrust regarding the use of DNA and genetic information.^{2,11,12} Banked samples from participants enrolled in prior studies are therefore often used as they provide a cost-efficient option.^{13,14} However, depending on the scope of the original study, including the ability to recontact participants or return research results, these cohorts may not help address questions important for the introduction of genomic medicine for all.¹⁵ Overall, genomic studies have reported variable success in enrolling diverse participants in terms of gender, age, race, ethnicity, health status, and education, with most reporting mainly on white, non-Latino individuals from high socioeconomic groups often with prior exposure to genetics or high health literacy.^{12,16–22} It is thus unknown whether the concerns and expectations of those participating in genetic studies differ from those who do not, and consequently whether the factors and considerations reported to influence use of genetic testing adequately reflect those that operate in the general population. If researchers cannot engage all stakeholders in genomic research, findings based on a homogeneous group of early adopters of genomic technologies may mislead future decisions regarding the implementation of genomic medicine.

The eMERGE Network is a national consortium focused on integrating genomics with the electronic health record (EHR).²³ For the Columbia eMERGE site, we developed a protocol to enroll adult participants for genetic screening of adult-onset conditions with available treatment or preventive interventions. Both positive and negative results were returned to all participants and placed in their EHR. We strove to enroll a sample with greater diversity in ethnicity, age, and socioeconomic status than cohorts previously enrolled in genetic research.^{12,16–22}

Currently, there is a scarcity of research literature on the cost and effectiveness of different methods for recruiting a diverse study sample for genetic research, so we began to address this gap by retrospectively examining data we had on the seven recruitment methods we employed at our single site within eMERGE. These data, although not prospectively collected as part of a study on recruitment strategies, contained information relevant to four outcomes: (1) consent rate; (2) completion rate for all components of the first phase of the study; (3) diversity among participants with respect to age, gender, race and ethnicity, education, and employment; and (4) estimated average cost per participant.

MATERIALS AND METHODS

Columbia study design

The Columbia University Medical Center is located in upper Manhattan, a socioeconomically disadvantaged area of New York City, and serves both the local community and individuals traveling to the hospital for specialized care. We developed seven recruitment methods to enroll 1475

participants for free preventive genetic screening; enrollment had to be completed over a 13-month period (Supplementary Methods; Table 1). As one important identified barrier to enrollment of minorities or individuals of low socioeconomic status is convenience,^{3,7} we developed a participant web portal accessible after invitation to allow self-guided genetic education and electronic study consent at any time. To accomplish one of the eMERGE Consortium's goals of having the ability to interpret genetic data across ancestry,²³ we included a targeted study arm restricted to individuals who self-identified as Latino/a and/or Ashkenazi Jewish, and a general arm for participants of any ancestry. We originally aimed to enroll 475 participants into the targeted arm and 1000 participants of any ancestry into the general arm. Both arms of the study were approved by Columbia University institutional review board.

Eligibility

Individuals 18 and older, able to speak and read English or Spanish, and to provide informed consent, were eligible to participate. All recruitment and patient educational materials, including the web portal, were available to potential participants in both English and Spanish and were modified to be culturally appropriate for Latino/as, under the guidance of the Community Engagement Core Resource of the Columbia University Irving Institute for Clinical and Translational Research.

Recruitment workforce

Participant recruitment and follow-up were conducted by eight clinical research coordinators and four research assistants. Both research coordinators and research assistants invited participants and explained the study, but they also had different roles and salaries: research coordinators obtained written consent and drew blood samples, while research assistants directed participants who verbally expressed interest in the study to the web portal or mailed them written educational materials and consent forms, and scheduled blood draws. The research coordinators and research assistants were from diverse backgrounds including Latina, Jewish, Asian, and European ancestry, and some were Spanish-speaking (Supplementary Methods).

Recruitment methods

The seven methods differed in sampling type, sample arm, invitation method, deployment period, and number of attempted contacts per potential participants (Table 1). In the general arm, blood collection occurred at the consent visit and a questionnaire was mailed or emailed, according to participants' preference (Figure S1). In the targeted arm, following verbal interest, participants were mailed a paper consent and questionnaire or could choose to complete the consent and questionnaire via the web portal. Following consent, blood collection was scheduled at multiple locations and times, including early, late, and weekend hours.

Table 1 Description of the seven recruitment methods

Method	EHR	Flyers: targeted	Community events	Clinic	Biobank	Staff network	Flyers: general
Deployment period	6/2016–4/2017	7/2016–7/2017	4 events: 8/2016, 12/2016, 4/2017, 5/2017	11/2016–7/2017	11/2016–7/2017	1/2017–7/2017	1/2017–7/2017
Sampling type	Targeted	Targeted	Targeted	General	Targeted	General	General
Invitation type	Active	Passive	Active invitation of a group	Active	Active	Active	Passive
Sample frame	CUMC patients	CUMC community (patients, employees, and students)	Washington Heights and Yeshiva University communities attending genetic education events	CUMC patients of nephrology, hepatology, and oncology subspecialty outpatient clinics	Participants in a Washington Heights (Latino community) breast cancer biobank	CUMC community (family and friends of study staff)	CUMC community (patients, employees, and students)
Electronic consent option	Yes	Yes	Yes	No	Yes	No	No
Last step to complete enrollment (CBQ)	Blood draw	Blood draw	Blood draw	Questionnaire	Blood draw	Questionnaire	Questionnaire
Maximum number of attempted contacts	Up to 9 (letter, 2 phone calls, 6 reminders)	Up to 7 (flyer, 6 reminders)	Up to 7 (event, 6 reminders)	Up to 4 (invitation, 3 reminders)	Up to 9 (letter, 2 phone calls, 6 reminders)	Unknown	Up to 4 (flyer, 3 reminders)

For “Sampling Type”, “Targeted” refers to the eligibility restricted to individuals self-identified as Ashkenazi Jewish and/or Latino/a; and “General” indicates that there were no race or ethnicity eligibility criteria; “Maximum number of attempted contacts” refers to all attempted contacts to encourage potential participants and consented participants to complete blood and questionnaire (not counting potential viewing of the flyers across all recruitment types; Supplementary Methods).
 CBQ consented, provided blood, and completed the questionnaire, CUMC Columbia University Medical Center, EHR electronic health records.

The methods were deployed in a stepwise manner (Figure S2). They were characterized as either active or passive invitations on the part of the researcher. In the former case, potential participants were invited to the study in-person or by phone, while in the latter case they were made aware of the study through flyers. Active invitations and targeted recruitment were employed first. Two additional methods were subsequently deployed when the number enrolled by the first five methods indicated that, based on the study team capacity, the desired total sample size might not be attained within the study time frame.

A full description of the methods is provided in the Supplementary Methods and Figure S1. Each is succinctly described here in order of deployment.

1. **Targeted active invitation of individuals identified through the EHR:** A letter in English and Spanish was sent to patients identified through an EHR query of patients living in zip codes with a high proportion of Latino/a and/or Ashkenazi Jewish individuals (Appendix 1). Research assistants followed up with phone calls to invite them to participate and to obtain verbal agreement to receive the study materials.
2. **Targeted passive invitation through flyers:** Two different formats of flyers were posted at Columbia University Medical Center, in both English and Spanish: one describing eligibility for people of Latino/a and/or Ashkenazi Jewish ancestry and one specifically encouraging people of Latino/a ancestry to enroll (Figure S3). Interested individuals contacted the research assistants to learn about enrollment procedures.
3. **Targeted active invitation of attendees at genetic educational events in the community:** Community organizers hosted events at which the study was described and attendees were invited to participate. Interested attendees shared their contact information and were later contacted by a research assistant.
4. **General active invitation of patients in medical clinics:** All patients seen in participating outpatient clinics were screened for eligibility. Patients approved by their provider were invited in-person immediately after their clinic appointment.
5. **Targeted active invitation of prior research participants from the breast cancer biobank (biobank):** Biobank participants were sent an invitation letter in English and Spanish (Appendix 1). Research assistants followed up with phone calls to invite them to the study and obtain verbal agreement to receive the study materials.
6. **General active invitation of staff network:** Study staff invited their friends, family members, and colleagues to the study. Interested individuals were given information on how to contact the study team to schedule an enrollment visit.
7. **General passive invitation through flyers:** Flyers in English were posted at around Columbia University Medical Center (Figure S3). Interested individuals

contacted the research coordinators to learn about enrollment procedures.

Consent rate and decliners

Only a subset of the recruitment methods (EHR, clinics, biobank, and community events) enabled the calculation of consent and decline rates. The consent rates for the clinics, EHR, and biobank methods were calculated as the proportion of participants who consented to the study of those invited and eligible. For the community events, the total number of attendees was counted as invited and eligible participants, even though probably not all were eligible. For all methods, if a potential participant declined the study, they were not approached again, though they may have seen the flyers afterwards. Reasons for declining were solicited for clinics, EHR, and biobank with a multiple-choice question (Supplementary Methods).

Enrollment completion

Even though all participants who provided a blood sample received free genetic screening, for the purpose of this report, participants were considered to have completed enrollment only after they gave written consent, provided a blood sample, and answered a questionnaire (CBQ participants; Supplementary Methods). A \$25 gift card incentive was offered for questionnaire completion. In the general arm, up to three follow-up attempts were made to remind participants to complete the questionnaire (Table 1, Figure S1). In the targeted arm, methods included up to three follow-up attempts after verbal interest in participating was expressed to remind potential participants to complete consent and questionnaire and up to three follow-up attempts after the written consent to schedule the blood draw. CBQ completion rates were calculated as the proportion of individuals who completed enrollment (CBQ participants) of those who consented to the study (Supplementary Methods).

Cost analysis

A full description of the cost analysis can be found in the Supplementary Methods. In short, study start-up costs (fixed costs), recruitment cost per CBQ participant (marginal cost) and supply costs per CBQ participant were retrospectively estimated. The fixed cost was based on an estimate of the time each of the contributing researchers spent developing the study materials and their salaries, as well as outsourced work costs. The marginal cost was based on the estimated time spent on recruitment by each method, and the estimated percent efforts of each staff member (full-time equivalent, FTE). The supply costs were calculated for each method separately. Extrapolated recruitment numbers for each method per one FTE were estimated.

Participant demographics

Demographic information, including age, gender, race, ethnicity, address, employment status, and education level,

was collected at the time of informed consent or through the study questionnaire (Appendix 2, Supplementary Methods). Ashkenazi Jewish ancestry information was collected only for the targeted arm. Socioeconomic background was evaluated using three variables: education level, employment status, and geo-income, as income was not collected from the participants.²⁴

Evaluation of the recruitment methods

Using the full CBQ sample characteristics as a reference, we evaluated how each method contributed to the CBQ sample. The participants’ characteristics and the cost per CBQ participant were summarized for each method.

Statistical analysis

Study data were collected and managed using REDCap electronic data capture tools hosted at Columbia University.²⁵ Results are presented as *Ns* and frequencies. Chi square and Fisher’s exact test analyses were used to explore the differences between participants who did and did not complete CBQ. Analysis was completed in SAS²⁶ and R.²⁷

RESULTS

Study and recruitment costs

The development of the study materials required an estimated 3600 hours of work by 24 people, including faculty and staff, for a fixed cost of approximately \$408,665 (Supplementary Results, Table S1). Of the total time devoted to study

development, 30% was dedicated to web portal development, 18% to questionnaire development, 17% to study design, 17% to materials translation, and the remaining 18% was used for eight additional tasks. More training time was invested in the research coordinators than in the research assistants (120 hours versus 40 hours, respectively; Table S1a).

The length of time the recruitment methods were deployed varied from 4 to 12 months (Figure S2). Recruitment efforts involved 16 people, including research coordinators, research assistants, research scientists, a bioinformatician, and a genetic counselor, which totaled 3.8 FTE employees, corresponding to an estimated marginal cost of \$430,663 (Table S2). Due to the different protocols and their stepwise deployment, the number of FTEs invested in the study varied, with over a third of the study staff deployed in clinics (1.4 FTEs; Table S3). The extrapolated estimated number of CBQ participants per one FTE ranged between 91 participants/FTE for the EHR method to over 630 participants/FTE for the general flyer (Fig. 1a).

Depending on the method, the estimated cost per CBQ participant ranged from \$579 to \$1666 (Fig. 1b). For most methods, at least 50% of the cost originated from the fixed cost (\$366/participant). Despite the \$25 gift cards, the supply costs accounted for only a minority of the expenses (\$31–\$49 per CBQ participant, Table S4). The high FTE investment in the clinics enabled the recruitment of a high number of participants. The estimated cost per consented participant who provided a blood sample regardless of questionnaire

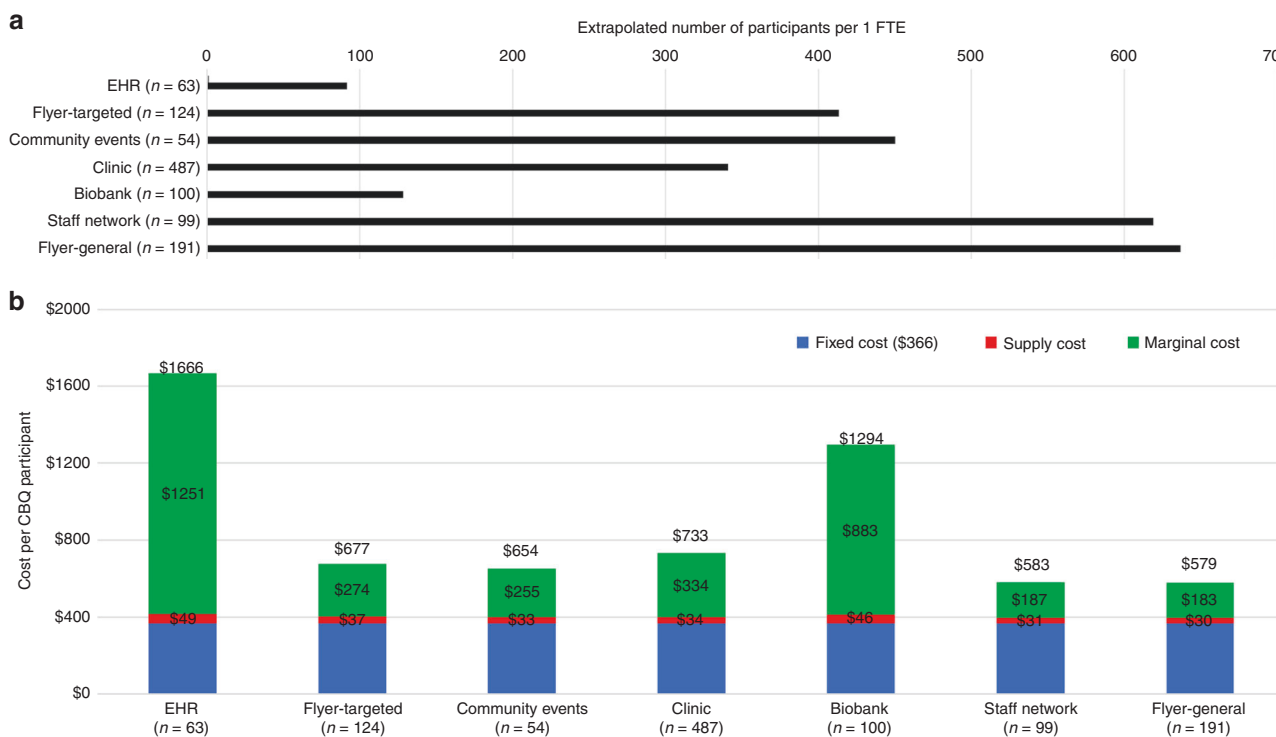


Fig. 1 Variable costs across seven recruitment methods. (a) Extrapolated number of participants who completed enrollment per one FTE for each recruitment method (assuming the same average number of CBQ participants over 12 months). (b) Average cost including fixed start-up cost, supply cost, and marginal cost, per CBQ participant for each recruitment method. (n: number of CBQ participant per recruitment method). CBQ consent, blood, and questionnaire complete, EHR lectronic health records, FTE full-time equivalent.

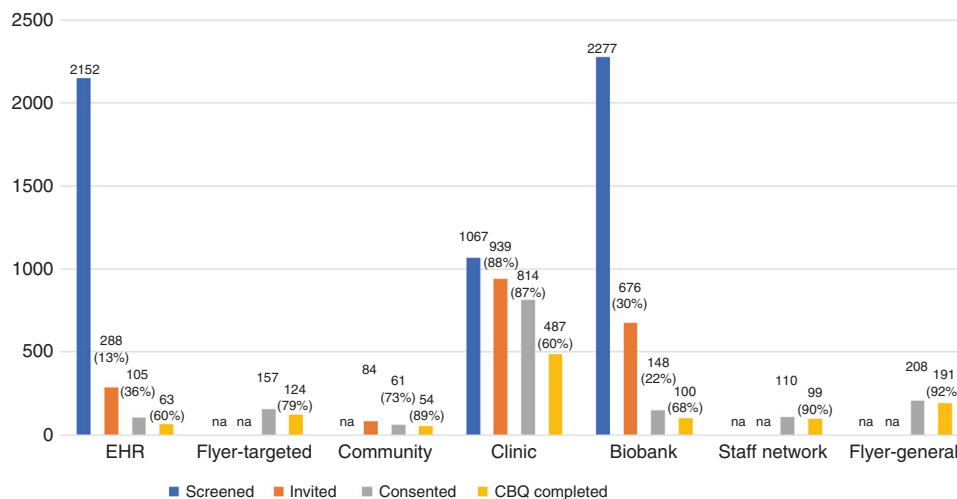


Fig. 2 Number of participants screened, invited, consented, who provided a blood sample and completed a questionnaire (CBQ) by each recruitment method. Only eligible individuals were invited, and for the EHR and biobank only individuals who answered the phone call were counted as invited. *CBQ completed* individuals who provided written consent, blood, and completed the questionnaire, *EHR* electronic health records, *na* not available for certain methods.

completion was similar for the clinics, staff network, and general flyers (\$205, \$173, and \$173, respectively, Table S4).

Variable consent, decline, and completion rates

Across all methods, a total of 1603 participants consented to the study, of whom 1118 were CBQ participants, 355 provided blood only, and 130 completed the questionnaire only (Table S5). We observed a wide range of consent rates (22–87%), with the highest rate among those invited in-person in the clinics (Fig. 2, Table S6). Furthermore, when examining all 4429 individuals who were sent invitation letters (EHR, Biobank), regardless of whether they were reached by phone, fewer than 6% consented (Supplementary Results). Twenty-five consented participants who were enrolled through the clinics, targeted flyers, general flyers, and staff network (17, 4, 3, and 1, respectively) were also sent an invitation letter through the EHR or the biobank methods, but it is unknown whether they received and read it.

Of the 517/909 decliners who provided a reason, a common reason in all three methods was “lack of interest” ($n = 183$) and for the EHR and biobank it was “time constraints,” which was not a reason offered to the decliners in the clinics ($n = 143$; Figure S4). Only 37 decliners cited concerns about privacy or discrimination, and 21 additional decliners indicated they did not want to learn their genetic results.

Enrollment rates varied over time and across methods (Figure S2). For some of the methods still being used at the time we reached our enrollment goal, we had not yet observed any decline in the number of participants being enrolled. We observed a peak of recruitment approximately a month after sending the invitation letters or holding the community events. The clinics’ enrollment required a period of familiarizing the health-care team with the study during which recruitment was slow, but thereafter the recruitment was stable until the end of the study period. There were two peaks

for the staff network enrollment, the first one soon after the method was deployed and the second close to the end of the study. Enrollment through the flyers began immediately after posting and appeared to be associated with number of flyers posted.

The CBQ completion rates varied across recruitment methods (60–92%, Fig. 2, Table S5). Because the order of completion differed between the general and the targeted arms (Figure S1), the CBQ completion rates were analyzed separately for each arm. In the general arm, of the 1132 participants who provided a blood sample at time of consent, only 777 also completed the questionnaire. The proportions of consented participants who did and did not complete the questionnaire across methods were significantly different (p value = $1.90E^{-23}$). Among the 471 in the targeted arm who completed the questionnaire at time of consent, only 341 provided a blood sample. Similarly, the proportions of consented participants who did and did not provide a blood sample across methods were significantly different (p value = $7.19E^{-05}$). Though the sample size goal was almost reached (1473 participants provided a blood sample), the CBQ sample consisted of only 1118 participants (Table 2).

Different demographic distribution

The age range of the CBQ sample was 18–94 years-old, with 10% of the participants younger than 25 years old, 21% older than 65 years old, and the remainder evenly distributed in 10-year intervals between 25 and 65 years old (Table 2). Though none of the methods were designed to recruit a specific age group, a higher proportion of individuals younger than 44 years were enrolled through the community, staff network, targeted and general flyers than through the other methods (Table S6, Figure S5).

A high proportion of women, who represented two-thirds of the CBQ sample (Table 2), were recruited through the

Table 2 Demographics of the 1118 participants who provided written consent and blood samples and completed the questionnaire (CBQ participants): number and percentage of the total study sample

		N	%
Total		1118	
Gender	Male	404	36.1%
Age	18–24	109	9.7%
	25–44	410	36.7%
	45–64	363	32.5%
	65+	235	21.0%
	Race and ethnicity	White, non-Latino/a	513
	Black/African American, non-Latino/a	70	6.3%
	Asian, non-Latino/a	118	10.6%
	American Indian or Alaska Native, non-Latino/a	1	0.1%
	More than one race, non-Latino/a	12	1.1%
	Other, non-Latino/a	8	0.7%
	White, Latino/a	118	10.6%
	Black/African American, Latino/a	26	2.3%
	Asian, Latino/a	4	0.4%
	American Indian or Alaska Native, Latino/a	5	0.4%
	More than one race, Latino/a	64	5.7%
	Other, Latino/a	179	16.0%
Not born in the United States		442	40%
Education	Less than high school	85	7.6%
	High school or GED	110	9.8%
	College some or degree	550	49.2%
	Graduate some or degree	370	33.1%
	NR	3	0.3%
Employment	Employed, full or part time	670	59.9%
	Unemployed	45	4.0%
	Disabled, temporary or permanent	88	7.9%
	Retired	174	15.6%
	Student/other	134	12.0%
	NR	7	0.6%
Geo-income	<\$33,948	176	15.7%
	\$33,948–100,000	632	56.5%
	\$100,000+	284	25.4%
	NA	26	2.3%
Questionnaire format	English online	881	78.8%
	English on paper	99	8.9%
	Spanish online	19	1.7%
	Spanish on paper	119	10.6%
Recruitment modality	EHR	63	5.6%
	Flyers: targeted	124	11.1%
	Community events	54	4.8%
	Clinic	487	43.6%

Table 2 continued

	N	%
Biobank	100	8.9%
Staff network	99	8.9%
Flyers: general	191	17.1%

EHR electronic health records, GED general equivalency degree, NA not applicable, NR not reported.

flyers (targeted and general), staff network, EHR, and, of course, the biobank (which almost exclusively invited women; Table S6, Figure S5). Only the clinics and the community events recruited male and female participants almost equally.

The CBQ sample was diverse in terms of ancestry. Fewer than half self-identified as white, non-Latino/a and 35% self-identified as Latino/a (Table 2). Highlighting the utility of the Spanish translation and the ability to complete the study offline, 30% of the Latino/a participants completed study materials in Spanish and on paper. Consistent with the targeted enrollment, 44% participants of this arm self-identified as Ashkenazi Jewish, including nine participants who self-identified as both Latino/a and Ashkenazi Jewish. Similarly, the targeted flyer and the biobank enrolled the highest proportions of participants who self-identified as Latino/a (76% and 69%, respectively; Table S6, Figure S6).

Different socioeconomic backgrounds

In terms of socioeconomic background, 18% of the CBQ sample had a high school education or less, 8% reported a disability preventing them from working, and 16% lived in a neighborhood with a low average annual household income (Table 2). The targeted flyers and the biobank methods resulted in the highest proportion of participants living in low income areas (26% and 31%, respectively; Table S6, Figure S6). The EHR and the staff network methods recruited the highest proportion of participants with graduate degrees (54% and 56%, respectively).

Effectiveness and cost evaluation of the seven recruitment methods

No two methods had the same combination of participants’ characteristics, CBQ completion rate, and recruitment cost (Table 3, Figure S5 and Figure S6). CBQ completion rates were higher and recruitment costs were lower for methods resulting in a sample with comparatively higher education levels and employment rates, and lower ethnic diversity (Table S6). More expensive methods enrolled a sample with comparatively lower education levels and employment rates, and greater ethnic diversity.

DISCUSSION

Evaluation of seven recruitment methods for a genomic screening study revealed that each method differently contributed to the enrollment of a sample more demographically diverse than samples of previous studies offering

Table 3 CBQ sample: recruitment effectiveness, participants’ characteristics, and cost per CBQ participant for all methods combined

Invitation	Type	Full CBQ Sample	EHR		Flyer-Targeted	Community Events	Clinic	Biobank	Staff Network	Flyer-General
			Active	Passive	Active invitation of a group	Active	Active	Active invitation of network	Passive	
Recruitment effectiveness	N	1118	63	124	54	487	100	99	191	
	Consent Rate	56%	36%	na	73%	88%	22%	na	na	
	CBQ Rate	70%	60%	79%	89%	60%	68%	90%	92%	
Participants characteristics	Age	< 45 y old	47%	-	+	+	-	-	+	+
	Gender	> 65 y old	21%	+	-	-	+	+	-	-
		male	36%	-	-	+	+	-	-	-
	Race and Ethnicity	Latino/a	35%	-	+	-	-	+	-	-
		not White non-Latino/a	46%	-	+	-	-	-	+	+
	Education	High school or less	18%	-	-	-	+	+	-	-
Employment	not employed	40%	+	-	+	+	+	-	-	
Cost (marginal and start-up)	Geo-Income < \$33,948	16%	-	+	+	+	+	-	-	
		\$787	\$1,669	\$676	\$656	\$733	\$1,295	\$582	\$578	

CBQ consented, provided blood, and completed the questionnaire, EHR electronic health records, NA not available. Dark grey (+): higher consent or CBQ rate, more diverse cohort, and lower cost per CBQ participant for the recruitment method compared with the full CBQ sample. White (-): lower consent or CBQ rate, less diverse cohort, and higher cost per CBQ participant for the recruitment method compared with the full CBQ sample.

genetic testing.^{12,16–22} While the retrospective nature of this assessment and the variability across the methods was a limitation and prevented a systematic comparative effectiveness analysis, we think this report of our experience with different recruitment methods for a genetic study may aid in the design of future studies.

Multiple barriers have been discussed as limiting recruitment of diverse samples, and several approaches have been developed to address them.^{11,13,28–30} The approaches implemented in this study included employment of a diverse and bilingual, stable study staff, online and offline bilingual study materials, diversified recruitment methods, and incentives. The clinics, which had a high consent rate of a relatively diverse sample, were convenient for participants and provided in-person recruitment by a demographically diverse staff, all factors shown to increase enrollment and help foster trust among prospective participants.^{2,4,7,28,31} Utilizing additional approaches reported to be important in recruitment and retention of minority groups^{5–7,9}—targeted recruitment emphasizing the importance of Latino/a participation, bilingual study staff and study materials allowing participation in Spanish—likely contributed to the successful enrollment of a high proportion of Latino/as. Finally, the option for offline completion of the study, which was used by the majority of participants who completed the study in Spanish, may have also contributed to our sample’s ethnic diversity. The study design did not allow us to isolate the importance of each of these approaches. It is also unknown whether those approaches would provide effective enrollment of other underrepresented populations. Future studies should evaluate the cost and effectiveness of different recruitment strategies with other populations and of additional recruitment strategies, such as the use of social media.

The reported costs are only estimates as they were retrospectively calculated. Nevertheless, consistent with prior reports, the methods differed in cost, largely related to the number of attempted contacts when the active invitation was not in-person.^{4,13} In contrast to previous reports suggesting that biobanked samples are cost-effective,¹³ the need in this case to recontact participants and obtain an additional blood draw and consent resulted in this approach being one of the most expensive.

The high start-up cost supports reusing protocols and materials. Compilation of genomic medicine education, recruitment materials, and enrollment protocols into a central toolkit available to all researchers could leverage prior experience, minimize future study start-up costs, leading to their continuous improvement and ultimately introduction to clinical care.

Since previous reports have described concerns about privacy and discrimination as being potential barriers to participation in genetic research,^{25,32,33} we were apprehensive about how placement of genetic results in the EHR might impact recruitment. Interestingly, few potential participants declined when invited in-person, possibly because they had a trusting relationship with the medical team, and few of them reported privacy or discrimination concerns. While this may be a function of the method (i.e., face-to-face recruitment), it may also reflect a shift in attitudes or the differences between this sample and those previously reported.

The EHR approach, though targeted to specific zip codes and individuals of specific ancestry, was the method most similar to recruitment of the general population. Its low consent rate may be attributable to limited awareness of or interest in genetic and medical research.^{25,34,35} Public education may be helpful across all recruitment strategies and may facilitate the implementation of genomic medicine.

Future studies should systematically assess awareness of or interest in genetics and reasons for declining study participation and how enrollment rates differ by demographics characteristics.

This analysis has several limitations, including the utilization of data from a study that was not designed to assess recruitment cost and efficiency. Self-selection bias is a common problem in studies, especially those that rely on passive recruitment methods, as we did in the majority of our strategies. Because we could not know who was made aware of the opportunity to participate (e.g., saw a flyer) or their characteristics, we are unable to assess the nature of the self-selection bias this introduced. For instance, it was not possible to calculate what proportion of those who viewed the flyer went on to enroll. Other limitations precluded drawing firm conclusions about the relative effectiveness of different recruitment methods. For example, some methods were deployed longer; others had more follow-up attempts. Future prospective, randomized studies implementing different recruitment methods on defined populations, with a uniform workforce, identical timelines and eligibility would provide more rigorous assessment of recruitment methods.

Because the available data on the effectiveness of different recruitment strategies for genomic research studies and their associated costs is sparse, we conducted a retrospective analysis of our recruitment data. The study limitations and biases may limit the generalizability of our findings to other clinical settings, geographic regions of the country, or studies with other aims. However, much of our experience is likely transferable to recruitment for other genomic studies.^{36,37} We hope our observations will provide information that stimulates thinking about the importance of diverse recruitment options, helps to improve efficiency and effectiveness, and ultimately contributes to recruitment of samples that better reflect the heterogeneity of the US population¹².

SUPPLEMENTARY INFORMATION

The online version of this article (<https://doi.org/10.1038/s41436-019-0497-y>) contains supplementary material, which is available to authorized users.

ACKNOWLEDGEMENTS

Research was funded by U01HG008680 (principal investigators C. W., George Hripcsak, and A.G.G.). We would like to thank the participants of this research study. We also thank Evan Joiner, Emily Webster, Andria Reyes, Alejandra Aguirre, Nayrobi Ribera, Katrina Celis, Rachel Yarmolinsky, and Ilana Chilton for their guidance in the development and translation of the educational materials. We thank Ian Halim and Anoushka Sinha and the staff of the nephrology clinics for assistance with recruiting participants.

AUTHOR CONTRIBUTIONS

CW, AG, WKC, HMR, JW, MM conceived the study. JW, HMR, MM, EGC, AF, JM, KM, RLK, RA, PAM, MAO, YSS, AF, PSA, AG, CW, KK, WKC designed the study and developed study

materials. JW, HMR, MM, RR, DC, SK, BHK, XM, MA, JN, KM, AE, BH, AS, IH, DF, NL, ME, ARP enrolled participants and collected data. JW, HMR, MM, JTT analyzed data and interpreted the results. JW, HMR, MM wrote the manuscript. All authors contributed and discussed the results and critically reviewed the manuscript

DISCLOSURE

The authors declare no conflicts of interest.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Health Sciences Policy, Roundtable on Genomics and Precision Health. Implementing and evaluating genomic screening programs in health care systems: proceedings of a workshop. Washington, DC: National Academies Press; 2018. <http://www.ncbi.nlm.nih.gov/books/NBK500393/>. Accessed 24 December 2018.
2. Hindorff LA, Bonham VL, Brody LC, et al. Prioritizing diversity in human genomics research. *Nat Rev Genet.* 2018;19:175–185.
3. Hoyo C, Reid ML, Godley PA, Parrish T, Smith L, Gammon M. Barriers and strategies for sustained participation of African-American men in cohort studies. *Ethn Dis.* 2003;13:470–476.
4. Huynh L, Johns B, Liu S-H, Vedula SS, Li T, Puhan MA. Cost-effectiveness of health research study participant recruitment strategies: a systematic review. *Clin Trials.* 2014;11:576–583.
5. Lopez-Owens M, Starkey K, Gil C, Armenta K, Maupomé G The VidaSana Study: Recruitment Strategies for Longitudinal Assessment of Egocentric Hispanic Immigrant Networks. *Int J Environ Res Public Health.* 2018;15:2878.
6. Nicholson LM, Schwirian PM, Groner JA. Recruitment and retention strategies in clinical studies with low-income and minority populations: progress from 2004–14. *Contemp Clin Trials.* 2015;45(Pt A):34–40.
7. George S, Duran N, Norris K. A systematic review of barriers and facilitators to minority research participation among African Americans, Latinos, Asian Americans, and Pacific Islanders. *Am J Public Health.* 2014;104:e16–31.
8. UyBico SJ, Pavel S, Gross CP. Recruiting vulnerable populations into research: a systematic review of recruitment interventions. *J Gen Intern Med.* 2007;22:852–863.
9. Brown SD, Lee K, Schoffman DE, King AC, Crawley LM, Kiernan M. Minority recruitment into clinical trials: experimental findings and practical implications. *Contemp Clin Trials.* 2012;33:620–623.
10. Troost JP, Hawkins J, Jenkins DR, et al. Consent for genetic biobanking in a diverse multisite CKD cohort. *Kidney Int Rep.* 2018;3:1267–1275.
11. Amendola LM, Berg JS, Horowitz CR, et al. The Clinical Sequencing Evidence-Generating Research Consortium: integrating genomic sequencing in diverse and medically underserved populations. *Am J Hum Genet.* 2018;103:319–327.
12. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature.* 2016;538:161–164.
13. Salowe R, O'Keefe L, Merriam S, et al. Cost and yield considerations when expanding recruitment for genetic studies: the primary open-angle African American glaucoma genetics study. *BMC Med Res Methodol.* 2017;17:101.
14. Carey DJ, Fetterolf SN, Davis FD, et al. The Geisinger MyCode Community Health Initiative: an electronic health record-linked biobank for precision medicine research. *Genet Med.* 2016;18:906–913.
15. Wolf SM, Crock BN, Van Ness B, et al. Managing incidental findings and research results in genomic research involving biobanks and archived data sets. *Genet Med.* 2012;14:361–384.
16. Dewey FE, Murray MF, Overton JD, et al. Distribution and clinical impact of functional variants in 50,726 whole-exome sequences from the DiscovEHR study. *Science.* 2016;354. <https://doi.org/10.1126/science.aaf6814>. Accessed 28 March 2019.

17. Amendola LM, Dorschner MO, Robertson PD, et al. Actionable exomic incidental findings in 6503 participants: challenges of variant classification. *Genome Res.* 2015;25:305–315.
18. Henderson G, Garrett J, Bussey-Jones J, Moloney ME, Blumenthal C, Corbie-Smith G. Great expectations: views of genetic research participants regarding current and future genetic studies. *Genet Med.* 2008;10:193–200.
19. Amiri L, Cassidy-Bushrow AE, Dakki H, et al. Patient characteristics and participation in a genetic study: a type 2 diabetes cohort. *J Invest Med.* 2014;62:26–32.
20. Espeland MA, Dotson K, Jaramillo SA, et al. Consent for genetics studies among clinical trial participants: findings from Action for Health in Diabetes (Look AHEAD). *Clin Trials.* 2006;3:443–456.
21. Lewis KL, Hooker GW, Connors PD, et al. Participant use and communication of findings from exome sequencing: a mixed-methods study. *Genet Med.* 2016;18:577–583.
22. Sanderson SC, Linderman MD, Suckiel SA, et al. Psychological and behavioural impact of returning personal results from whole-genome sequencing: the HealthSeq project. *Eur J Hum Genet.* 2017;25:280–292.
23. National Human Genome Research Institute. eMERGE Network. 2018. <https://emerge.mc.vanderbilt.edu/>. Accessed 10 October 2018.
24. Brokamp C, Wolfe C, Lingren T, Harley J, Ryan P Decentralized and reproducible geocoding and characterization of community and environmental exposures for multisite studies. *J Am Med Inform Assoc.* 2017;25:309–314.
25. Hann KEJ, Freeman M, Fraser L, et al. Awareness, knowledge, perceptions, and attitudes towards genetic testing for cancer risk among ethnic minority groups: a systematic review. *BMC Public Health.* 2017;17:503.
26. SAS Institute Inc. SAS 9.4 [computer program]. Cary, NC: SAS Institute Inc.; 2014.
27. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018.
28. Johnson VA, Powell-Young YM, Torres ER, Spruill IJ. A systematic review of strategies that increase the recruitment and retention of African American adults in genetic and genomic studies. *ABNF J.* 2011;22:84–88.
29. Skinner CS, Schildkraut JM, Calingaert B, et al. Factors associated with African Americans' enrollment in a national cancer genetics registry. *Community Genet.* 2008;11:224–233.
30. Levy D, Splansky GL, Strand NK, et al. Consent for genetic research in the Framingham Heart Study. *Am J Med Genet A.* 2010;152a:1250–1256.
31. De La Rosa M, Babino R, Rosario A, Martinez NV, Aijaz L. Challenges and strategies in recruiting, interviewing, and retaining recent Latino immigrants in substance abuse and HIV epidemiologic studies. *Am J Addict.* 2012;21:11–22.
32. Lowstuter KJ, Sand S, Blazer KR, et al. Influence of genetic discrimination perceptions and knowledge on cancer genetics referral practice among clinicians. *Genet Med.* 2008;10:691–698.
33. Amendola LM, Robinson JO, Hart R, et al. Why patients decline genomic sequencing studies: experiences from the CSER Consortium. *J Genet Couns.* 2018;27:1220–1227.
34. Hamilton JG, Shuk E, Arniella G, et al. Genetic testing awareness and attitudes among Latinos: exploring shared perceptions and gender-based differences. *Public Health Genomics.* 2016;19:34–46.
35. Cobb EM, Singer DC, Davis MM. Public interest in medical research participation: differences by volunteer status and study type. *Clin Transl Sci.* 2014;7:145–149.
36. Malterud K. Qualitative research: standards, challenges, and guidelines. *Lancet.* 2001;358:483–488.
37. Hellström T. Transferability and naturalistic generalization: new generalizability concepts for social science or old wine in new bottles? *Qual Quant.* 2008;42:321–337.