

# UCLA

## UCLA Previously Published Works

### Title

Novel recruitment approaches and operational results for a statewide population Cohort for cancer research: The Healthy Oregon Project.

### Permalink

<https://escholarship.org/uc/item/3bw4c9xb>

### Journal

Journal of Clinical and Translational Science, 8(1)

### Authors

Zhang, Zhenzhen

Shafer, Autumn

Johnson-Camacho, Katie

et al.

### Publication Date

2024

### DOI

10.1017/cts.2024.9

Peer reviewed

## Translational Research, Design and Analysis Research Article

**Cite this article:** Zhang Z, Shafer A, Johnson-Camacho K, Adey A, Anur P, Brown KA, Conrad C, Crist R, Farris PE, Harrington CA, Marriott LK, Mitchell A, O'Roak B, Serrato V, Richards CS, Spellman PT, and Shannon J. Novel recruitment approaches and operational results for a statewide population Cohort for cancer research: The Healthy Oregon Project. *Journal of Clinical and Translational Science* 8: e32, 1–9. doi: [10.1017/cts.2024.9](https://doi.org/10.1017/cts.2024.9)

Received: 18 September 2023

Revised: 13 December 2023

Accepted: 8 January 2024

### Keywords:

Cohort; population; HOP; genetics; app

### Corresponding author:




P. Spellman, PhD;

Email: [pspellman@mednet.ucla.edu](mailto:pspellman@mednet.ucla.edu)

© The Author(s), 2024. Published by Cambridge University Press on behalf of Association for Clinical and Translational Science. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.



# Novel recruitment approaches and operational results for a statewide population Cohort for cancer research: The Healthy Oregon Project

Zhenzhen Zhang<sup>1,4</sup> , Autumn Shafer<sup>5</sup>, Katie Johnson-Camacho<sup>2,4,6</sup>, Andrew Adey<sup>1,4,6,7,8</sup>, Pavana Anur<sup>6</sup>, Kim A. Brown<sup>4,9</sup>, Casey Conrad<sup>4,6</sup>, Rachel Crist<sup>4,6</sup>, Paige E. Farris<sup>4,10</sup>, Christina A. Harrington<sup>7,11</sup>, Lisa K. Marriott<sup>10</sup>, Asia Mitchell<sup>7</sup>, Brian O'Roak<sup>7</sup>, Vanessa Serrato<sup>4</sup>, C. Sue Richards<sup>7,12</sup>, Paul T. Spellman<sup>2,3,4,6</sup>  and Jackilen Shannon<sup>1,4,6</sup> 

<sup>1</sup>Division of Oncological Sciences, Knight Cancer Institute, Oregon Health & Science University, Portland, OR, USA;

<sup>2</sup>Department of Medicine, University of California Los Angeles, Los Angeles, CA, USA; <sup>3</sup>Department of Human

Genetics, University of California Los Angeles, Los Angeles, CA, USA; <sup>4</sup>Knight Cancer Institute, Oregon Health &

Science University, Portland, OR, USA; <sup>5</sup>School of Journalism and Communication, University of Oregon, Eugene,

OR, USA; <sup>6</sup>Cancer Early Detection Advanced Research Center, Oregon Health and Science University, Portland, OR,

USA; <sup>7</sup>Department of Molecular and Medical Genetics, Oregon Health and Science University, Portland, OR, USA;

<sup>8</sup>Knight Cardiovascular Institute, Oregon Health and Science University, Portland, OR, USA; <sup>9</sup>Oregon Clinical and

Translational Research Institute, Oregon Health & Science University, Portland, OR, USA; <sup>10</sup>OHSU-PSU School of

Public Health, Oregon Health & Science University, Portland, OR, USA; <sup>11</sup>Gene Profiling Shared Resource, Oregon

Health & Science University, Portland, OR, USA and <sup>12</sup>Knight Diagnostic Laboratories, Oregon Health & Science

University, Portland, OR, USA

### Abstract

**Background:** Cancer health research relies on large-scale cohorts to derive generalizable results for different populations. While traditional epidemiological cohorts often use costly random sampling or self-motivated, preselected groups, a shift toward health system-based cohorts has emerged. However, such cohorts depend on participants remaining within a single system. Recent consumer engagement models using smartphone-based communication, driving projects, and social media have begun to upend these paradigms. **Methods:** We initiated the Healthy Oregon Project (HOP) to support basic and clinical cancer research. HOP study employs a novel, cost-effective remote recruitment approach to effectively establish a large-scale cohort for population-based studies. The recruitment leverages the unique email account, the HOP website, and social media platforms to direct smartphone users to the study app, which facilitates saliva sample collection and survey administration. Monthly newsletters further facilitate engagement and outreach to broader communities. **Results:** By the end of 2022, the HOP has enrolled approximately 35,000 participants aged 18–100 years (median = 44.2 years), comprising more than 1% of the Oregon adult population. Among those who have app access, ~87% provided consent to genetic screening. The HOP monthly email newsletters have an average open rate of 38%. Efforts continue to be made to improve survey response rates. **Conclusion:** This study underscores the efficacy of remote recruitment approaches in establishing large-scale cohorts for population-based cancer studies. The implementation of the study facilitates the collection of extensive survey and biological data into a repository that can be broadly shared and supports collaborative clinical and translational research.

### Introduction

Historically, the development of large research cohorts has been limited by the enormous costs of recruitment and retention [1,2]. Additionally, limited bandwidth for interactions and potential loss of participant interest have resulted in reduced follow-up rates, as participants can become overburdened [3,4]. To address these challenges, the prominent epidemiological cohorts of the late 20<sup>th</sup> century, such as the Framingham Heart Study [5], Physicians' Health Study [6–9], Nurses' Health Study [10], and more recent cohort studies, have relied on highly selected populations to enhance compliance and engagement. Of concern, these populations may exhibit behaviors and exposures that are substantively different from the overall population.

In recent years, there has been a transition towards health system-based research and clinical trial cohorts. Nevertheless, these systems usually require participants to be in some sense captive, making it challenging to follow them when they move out of the network, which leads to potential biases in participant retention. Furthermore, these systems are often decoupled from the academic research networks. Fortunately, recent advancement in consumer engagement

models utilizing smartphone-based communication, driving projects, and social media have begun to upend these paradigms [11–13].

We sought to develop a research cohort by leveraging smartphones for participant engagement coupled with a specific participant engagement strategy based on inherited cancer DNA screening. We report a highly cost-effective strategy that enables the rapid building of a research cohort, adding 1,500 new participants per month, and reaching more than 1% of the adult population of Oregon in less than two years. Unlike traditional recruitment methods for cohort studies that relied on mass invitation mailings or digital dialing, which are expensive and time-consuming with low response and follow-up rates, our innovative approach overcomes recruitment limitations, making it an ideal fit for translational science.

The research findings from the cohort we are building have the potential to yield valuable research resources for clinical and translational investigations involving diverse patient populations. The goal is to improve participant outcomes by optimizing treatment approaches and enabling accurate outcome predictions, thus providing direct translational insights. Ultimately, our project aims to advance the field of clinical and translational research, fostering a deeper understanding of various health conditions and driving progress toward personalized and effective interventions.

### **Innovative population-based cohort-building approaches are needed history of large Cohorts**

Large-scale population cohorts established in the second half of the 20<sup>th</sup> century have facilitated a broad understanding of the major risk factors for diseases, including cancer. In the USA, where no national healthcare system exists, these studies utilized central control and relied on preexisting groups to identify their participants (e.g. Framingham Heart Study [5], Nurses' Health Study [10], and California Teachers Study [14,15]). This approach clearly facilitates communication, engagement, trust, and ultimately compliance with study goals; however, the use of preexisting groups introduces biases that do not reflect the overall population.

### **History of electronic medical record (EMR) data repositories**

The movement to electronic medical records (EMRs) for billing and treatment histories has allowed the creation of systematically collected data as part of medical care and has facilitated the broad usage of medical records in research. However, EMR-based research systems are not designed for research, and the limitations of engaging with patients have been well-documented [16]. Only the largest health systems can effectively utilize these systems in building research cohorts (e.g. Kaiser Permanente [17], Geisinger Health [18], and the Veterans Administration [19]). Again, these systems are limited by the biases of the individuals covered in these plans as well as the treatment schedules available in these systems.

### **New approaches**

More recently, projects from the National Institute of Health (NIH)'s All of Us Program [11], and two major companies, Ancestry and 23andMe [20], have engaged directly with participants by offering something of value in exchange for participant enrollment. Epidemiological research is becoming a

reciprocal process in which participants benefit more directly from studies. These approaches utilize newer technologies, particularly smartphone apps, to decrease barriers to engagement and communication. These approaches actively encourage direct participation in research by community members. We sought to capitalize on these trends to engage with the population of the state of Oregon, enabling broad research engagement at a low cost by utilizing a specific driving project, genetic health screening.

## **Materials and methods**

### **Study overview**

The Healthy Oregon Project (HOP) study was approved by the Oregon Health & Science University (OHSU) Institutional Review Board (IRB #STUDY00018473). Data from HOP surveys and DNA samples are securely stored by OHSU in a research repository accessible to scientists approved by the HOP and OHSU's research ethics board who are working on cancer early detection research and effective treatments for inherited diseases.

### **Recruitment strategy**

The driving recruitment model for this study was to market the study across the state to drive individuals to download the study app which allowed for eligibility assessment, consent, and surveys. The inclusion criteria encompass all adults in Oregon, which includes cancer survivors. Participants consenting to genetic health screening either visited a location in the community for a collection kit returned on-site, or utilized a mail-in option, where the participant requested a kit within the app to be sent to their Oregon address.

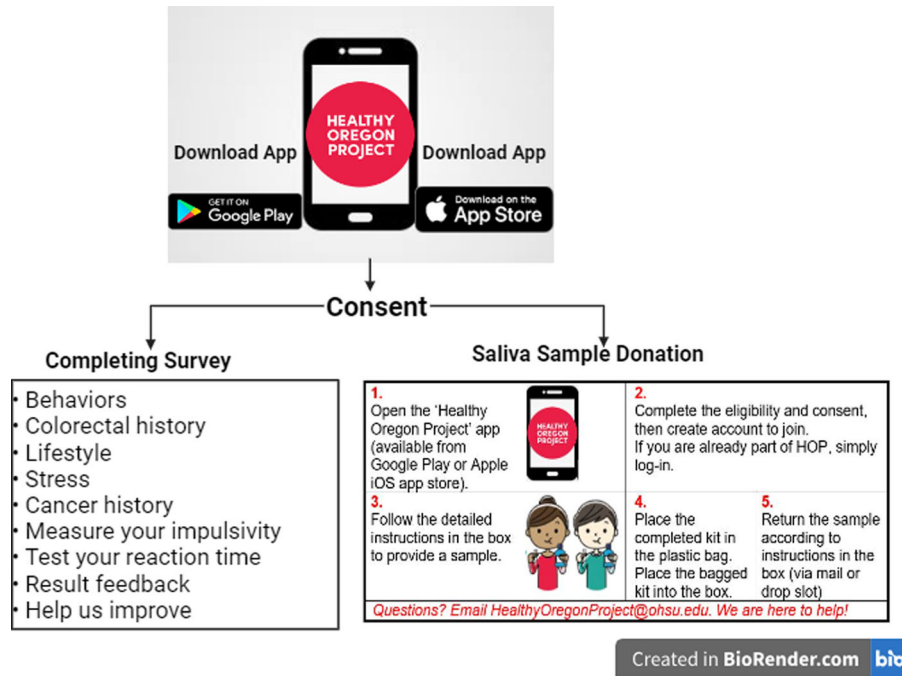
### **Internet/media**

The recruitment and participant support for the study were facilitated through the unique email account [healthyoregonproject@ohsu.edu](mailto:healthyoregonproject@ohsu.edu) and the webpage <https://healthyoregonproject.com/>. We utilized a study landing page on the HOP website that directed smartphone users to the appropriate app store for their devices. HOP uses a variety of social media platforms, including Facebook and Twitter, to engage with participants and direct them to the appropriate app store.

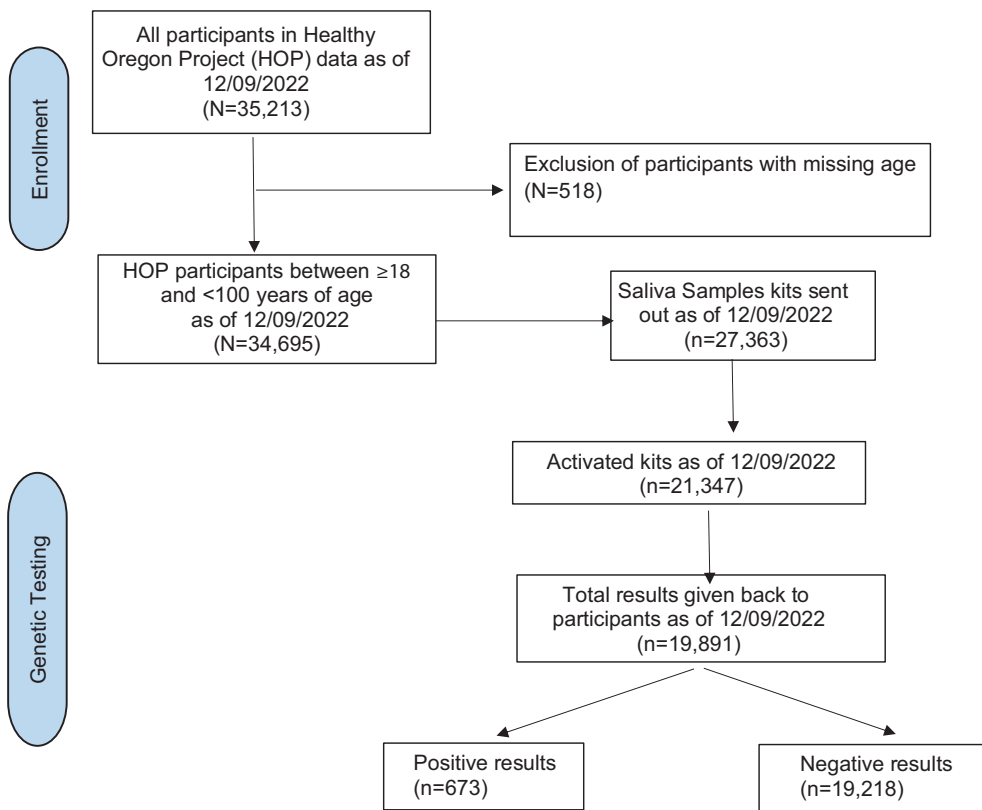
The HOP project also delivers monthly newsletters by email to participants to share with friends, family, or community groups to help spread the word about the HOP. Participants were encouraged to share the study broadly by sharing a post on their social media, sending the HOP website (HealthyOregonProject.com) to their go-to group chat, or telling their families about the results they received.

### **Healthy Oregon Project (HOP) app**

The HOP app can be downloaded on participants' smartphones. After consent, the participants can engage with the app through two components of the study: (a) one is a series of surveys and (b) the other is saliva sample collection. (a) *HOP Surveys*: The primary data collection administered through the HOP app includes questions on (1) behaviors: examining the relationship between chronic disease risk and behaviors like smoking cigarettes and drinking alcohol; (2) cancer history: helping researchers to understand past cancer diagnosis and known family history; (3) colorectal history: personal and family history related to colorectal cancer and disease; (4) lifestyle: everyday lifestyle to help researchers understand the participants and the health



**Figure 1.** Overview of the Healthy Oregon Project study process.



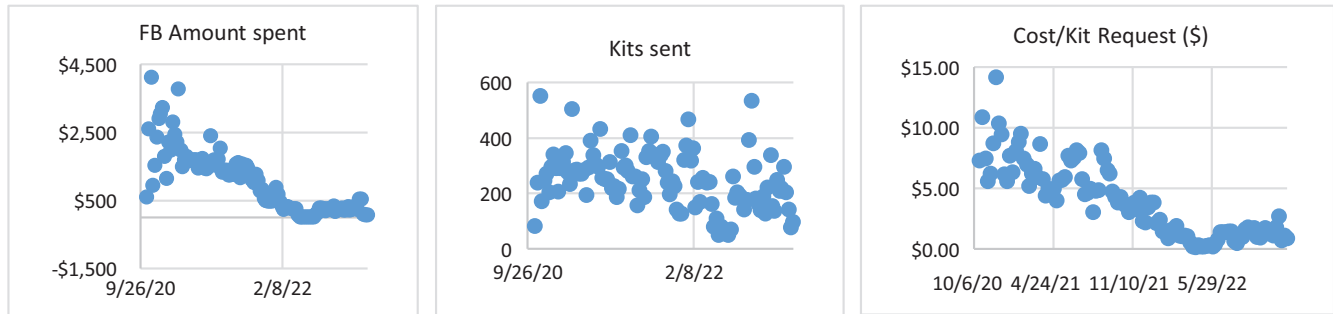
**Figure 2.** Study population flowchart.

impacts; (5) stress: experiences with stress over the last month before the survey; and (6) COVID-19: understanding how the pandemic has affected cancer risk, prevention, and screening behaviors. Later, (7) a measure of impulsivity and (8) a test of reaction time were added to provide tips for healthy sleep,

decision-making, and lifestyle choices. With the ongoing study, more questions will be supplemented through the HOP app, and participants will receive notifications through email. We aimed to obtain detailed data regarding risk factors impacting early cancer detection to enhance the breadth of the data with additional

**Table 1.** Table of genes screened for deleterious variants

|               |               |              |              |               |              |             |
|---------------|---------------|--------------|--------------|---------------|--------------|-------------|
| <i>APC</i>    | <i>BRCA2</i>  | <i>CHEK2</i> | <i>MSH6</i>  | <i>POLD1</i>  | <i>RET</i>   | <i>TSC2</i> |
| <i>ATM</i>    | <i>BRIP1</i>  | <i>MEN1</i>  | <i>MUTYH</i> | <i>PTEN</i>   | <i>SMAD4</i> | <i>LDLR</i> |
| <i>BAP1</i>   | <i>CDH1</i>   | <i>MITF</i>  | <i>NBN</i>   | <i>RAD51C</i> | <i>STK11</i> |             |
| <i>BMPR1A</i> | <i>CDK4</i>   | <i>MLH1</i>  | <i>PALB2</i> | <i>RAD51D</i> | <i>TP53</i>  |             |
| <i>BRCA1</i>  | <i>CDKN2A</i> | <i>MSH2</i>  | <i>PMS2</i>  | <i>RB1</i>    | <i>TSC1</i>  |             |

**Figure 3.** Cost per kit request was used as a surrogate of overall effectiveness. Weekly spend (a) was adjusted regularly to maintain a running average of 300 kit requests per week (b). Spending per week ranged from a high of \$4,123 (week of 11/4/2020) to a low of \$14 (week of 5/16/22).

targeted support. Our ultimate goal was to identify strategies for early detection of cancer. (b) *Biological specimen collections for inherited genetic health screening*: Following initial consent, the participants may click “Enroll: Genetic Screening” in the app’s Activities tab to complete additional consent for genetic testing. Following consent, participants can request saliva sample collection kits through the “Activity: Request a DNA Kit” action on the HOP app. The study implemented a mail option, where participants requested a DNA self-collection kit to be delivered to an Oregon address and then used the provided return mailing envelope to send to the screening lab. The participants were instructed that within 24 h of collecting the samples, they would drop-off their sealed envelopes through home mailboxes, at local post offices, or in a U.S. Post Office blue collection box. Participants were informed that results from DNA health screening would be returned at no cost within six months of receipt.

### Genes screened and reporting

DNA samples were sequenced at OHSU Knight Diagnostic Laboratories, a Clinical Laboratory Improvement Amendments lab at OHSU. Processes for genetic screening were covered in O’Brien *et al.* [21] According to the IRB-approved protocol, only pathogenic and likely pathogenic variants for which National Comprehensive Cancer Network (NCCN) guidelines exist are returned to the participants. All other participants received a negative report explaining that no variants that affected their care were detected. All participants with a positive test received genetic counseling at no cost from the OHSU genetic counseling staff.

### Results

The HOP is an IRB-approved research protocol (IRB #STUDY00018473) that addresses the challenge of engaging with the catchment area of the Knight Cancer Institute (KCI), an NCI Comprehensive Cancer Center that serves the state of Oregon. The

KCI leadership contracted with a Health Insurance Portability and Accountability Act-compliant smartphone app vendor to deploy a consenting and enrollment system available on iOS and Android devices through their respective app stores (Fig. 1). Consents were developed in close collaboration with IRB staff and included general research consent as well as additional consent for secondary studies, including a driving project to identify community members with inherited diseases who have established interventions that significantly alter patient health.

Primary genetic health screening was performed to detect genetic variants that increased the risk of cancer, particularly hereditary breast and ovarian cancer syndrome, Lynch syndrome, and less frequent syndromes such as Cowden’s syndrome and Li Fraumeni Syndrome. Prior to the launch of genetic health screening, we performed a community readiness assessment including focus groups [22]. As genetic health screening is offered to those who do not meet the NCCN genetic testing guidelines, only variants that are deemed pathogenic or likely pathogenic are identified by the testing lab and returned to the participants (see Methods). The full list of genes screened is shown in Table 1 and additional information on genetic health screening has recently been published [21]. HOP was designed to utilize direct participant engagement strategies, including direct mail, earned media, direct engagement, and social media. Enrollment in HOP occurred in two phases: a pre-COVID phase of in-person events and vending machine kit dispensing, and a post-COVID phase where mail-based recruitment became the standard. The in-person phase employed both small-scale human interactions (e.g., health fairs), engagement with large employers, and the use of kit dispensing machines situated at specific locations. Figure 2 provides an overview of the effectiveness of recruitment and total numbers. In total, 4,810 participants were recruited between 12/01/2018 and 10/16/2020 using in-person approaches. The shift to mail-based enrollment on Oct 17, 2020, resulted in the successful recruitment of a total of 34,695 participants (as of December 09, 2022), with valid age

**Table 2.** Improvement in cost-effectiveness associated with outreach

| Date     | Cost per kit request   |                       | Decrease in cost (%) |
|----------|------------------------|-----------------------|----------------------|
|          | Week before newsletter | Week after newsletter |                      |
| 10/26/20 | 10.87                  | 7.48                  | 0.31                 |
| 11/30/20 | 14.15                  | 10.36                 | 0.27                 |
| 12/21/20 | 6.17                   | 5.61                  | 0.09                 |
| 1/26/21  | 8.85                   | 9.51                  | -0.07                |
| 2/22/21  | 5.21                   | 6.22                  | -0.20                |
| 3/22/21  | 8.64                   | 5.78                  | 0.33                 |
| 4/26/21  | 4.97                   | 4.01                  | 0.19                 |
| 5/24/21  | 5.91                   | 7.68                  | -0.30                |
| 6/28/21  | 7.93                   | 5.78                  | 0.27                 |
| 7/26/21  | 4.98                   | 3.05                  | 0.39                 |
| 8/30/21  | 7.46                   | 6.48                  | 0.13                 |
| 9/27/21  | 4.29                   | 3.81                  | 0.11                 |
| 10/25/21 | 3.53                   | 3.03                  | 0.14                 |
| 11/29/21 | 4.25                   | 2.32                  | 0.45                 |
| 1/31/22  | 1.46                   | 0.90                  | 0.38                 |
| 2/28/22  | 1.89                   | 1.13                  | 0.41                 |
| 3/28/22  | 1.07                   | 0.52                  | 0.52                 |
| 6/27/22  | 1.41                   | 1.39                  | 0.02                 |
| 7/25/22  | 0.66                   | 0.48                  | 0.27                 |
| 8/29/22  | 1.81                   | 1.53                  | 0.16                 |
| 9/28/22  | 0.96                   | 1.48                  | -0.54                |
| 10/25/22 | 1.34                   | 1.09                  | 0.18                 |
| 11/28/22 | 1.11                   | 0.86                  | 0.23                 |
| Median   |                        |                       | 0.19                 |

information. Individuals who downloaded the app but did not provide consent were not considered study participants.

After switching to mail-based recruitment, we were able to directly measure the efficacy of social media, earned media, and digital communication with participants on engagement. One key parameter of such an engagement is the cost per enrolled participant. We found that after a relatively short training period of approximately eight months, social media systems moved from being very effective (approximately \$10/kit request) to extremely effective (approximately \$1.5/kit request), as shown in Figure 3. The saliva kit costs, including labor, around \$5 per participant. When two-way postage and a 70% return rate are taken into consideration, the total cost is approximately \$10/participant. The cost of genetic testing per participant was around \$75. The genetic counseling was provided at a cost of \$500 per positive result. Given the positivity rate of approximately 5%, the average cost per participant for genetic counseling was around \$25.

Due to laboratory constraints that limited throughput, we dramatically curtailed advertising in April and May of 2022; the average weekly spend was \$22, compared to \$283 in February and March, and \$224 in June and July. Enrollment decreased to 80 participants/week during the lowered spending versus 238 and 251

**Table 3.** Completion rates of study surveys by survey items

| Survey items <sup>1</sup>                      | N (Percent)    |
|--|----------------|
| Total N accessing the survey app               | 35,973         |
| HOP general consent                            | 31,147 (86.6%) |
| Behaviors                                      | 25,266 (70.2%) |
| Cancer history                                 | 22,621 (62.9%) |
| Colorectal history                             | 19,746 (54.9%) |
| Lifestyle                                      | 20,099 (55.9%) |
| Perceived Stress                               | 18,358 (52.1%) |
| COVID and cancer <sup>2</sup>                  | 1900 (5.3%)    |
| Impulsivity <sup>3</sup>                       | 11,665 (33.1%) |
| Test of reaction time (vigilance) <sup>3</sup> | 9945 (28.2%)   |
| Results feedback                               | 9200 (25.6%)   |

<sup>1</sup>The survey completion rates results were mainly based on data from the start of the study 12/01/2018 to 01/20/2023. The perceived stress, impulsivity and vigilance results were based on data from the study start date to 12/09/2022 (total N = 35,213). All surveys are available to all participants via the app, including those added at different dates. However, participants are only required to complete the initial demographics and consents surveys.

<sup>2</sup>The COVID and cancer survey was later added into the HOP app as part of an NIH supplement grant, which was mostly active during Aug. 2020–Jan. 2021.

<sup>3</sup>Impulsivity and vigilance surveys were later added into the HOP app on 08/03/2021.

per week in the two months prior and following, respectively. Nonetheless, baseline enrollment was presumably maintained through social networking effects. The HOP monthly email newsletters have an average open rate of 38% with a range from 27% to 56% from the study start date to Dec. 2022. This is much higher than the industry average and the survey reminder within the monthly newsletter typically receives 17.1% of the clicks each month, indicating a substantial level of reader interaction.

Efforts were made and continue to be made to improve survey response rates, such as email reminders about the surveys in general and specific surveys. Each month, one survey is featured in the newsletter to highlight its importance along with a link to the app to complete it. Social media campaigns with both organic posts on HOP's social pages and also ads targeted towards people who have clicked an app download ad or followed HOP include appeals to complete the surveys. Continuous efficiency testing occurred for the HOP ads and organic social media posts. Meta's Ads Experiments feature allows for A/B testing of ads to compare click rates, reach, app downloads, and website views of ads that are randomly assigned to groups within a selected target audience. HOP tested message attributes, such as visuals, wording, ad format, and message frames to improve ad efficiency. An example of such testing was a comparison of a local cue in the ad that either stated, "Oregon residents are invited to join HOP" or "Lane County residents" (tailored to each county the ad was shown in), which found that the localized county ads were 43% cheaper per website view than the "Oregon" ads. Another example was a comparison of a theme focused on the personal benefit of participating (knowing your cancer risk) vs. an ad with identical esthetics, but the text was themed around the altruistic benefit (contributing to cancer research). The personal benefit-themed ad was 37% cheaper per website view than the altruistic benefit ad.

We further measured the effectiveness of additional engagement through newsletters in our population, where the cost per participant in the weeks following an engagement email to those already identified by low-cost lead generation campaigns led to a

**Table 4.** Demographic characteristics of Healthy Oregon Project (HOP) participants

| Participant characteristics  | HOP participants<br>n (%) | Oregon<br>population <sup>2</sup> |
|--|---------------------------|-----------------------------------|
|  | Total N = 34,695          | 3,402,545                         |
| <b>Rural/Urban Status<sup>3</sup></b>                                  |                           |                                   |
| Urban  | 24,079 (72.3%)            | 65%                               |
| Rural  | 9,214 (27.6%)             | 33%                               |
| Unknown/other states   | 474                       | N/A                               |
| Undefined  | 928                       | N/A                               |
| <b>Age Group<sup>1</sup> (median = 44.2;<br/>mean SD = 46.8, 14.4)</b> |                           |                                   |
| 18- < 35   | 7,906 (22.8%)             | 28%                               |
| 35- < 50   | 13,910 (40.1%)            | 25%                               |
| 50- < 65   | 8,025 (23.1%)             | 23%                               |
| ≥65  | 4,854 (14.0%)             | 24%                               |
| <b>Gender</b>  |                           |                                   |
| Female   | 26,111 (75.3%)            | 50.1%                             |
| Male   | 7,945 (22.9%)             | 40.9%                             |
| Unknown or undefined   | 639 (1.8%)                |                                   |
| <b>Race/Ethnicity (n = 35,213)</b>                                     |                           |                                   |
| White  | 28,793 (83.0%)            | 73.5%                             |
| Black or African American  | 411 (1.2%)                | 2.3%                              |
| American Indian or Alaska Native                                       | 469 (1.4%)                | 1.9%                              |
| Asian  | 1,281 (3.7%)              | 5.1%                              |
| Native Hawaiian Pacific Islander                                       | 180 (0.5%)                | 0.5%                              |
| Middle Eastern/North African   | 67 (0.2%)                 | N/A                               |
| Hispanic   | 2,164 (6.2%)              | 14.4%                             |
| Others   | 290 (0.8%)                | 1%                                |
| Unknown or undefined   | 1,108 (3.2%)              | N/A                               |

<sup>1</sup>We included consented participants with valid age information into the descriptive analysis.

<sup>2</sup>Oregon population data were accessed through US Census Bureau, QuickFacts: Oregon. <https://www.census.gov/quickfacts/OR>. The total N shows OR adult population ≥ 18 years old. More refined age distribution data were obtained from <https://censusreporter.org/profiles/04000US41-oregon/> and <https://www.statista.com/statistics/1022743/oregon-population-share-age-group/>.

<sup>3</sup>Urban/rural status was accessed through Oregon Office of Rural Health. <https://www.ohsu.edu/oregon-office-of-rural-health/about-rural-and-frontier-data>.

median 19% decrease in costs for the week after an email newsletter compared to the week before the newsletter (Table 2). No effects were observed between the week after the newsletter and the subsequent two weeks.

The structure of smartphone engagement encourages short and frequent interaction. This approach is not highly conducive to data collection through long- or multi-part surveys, as is often observed in traditional cohort studies. However, by splitting surveys into multiple short and straightforward engagements, participants can complete numerous research surveys when convenient, and it is possible to control the order of survey completion by making the survey available only to participants who have completed other surveys within the app. When possible, the research surveys were

designed to be completed in five minutes, maximizing convenience to the participant and reducing the burden on any one sitting. In total, seven surveys were initially made available, and depending on the survey, 27%–72% of the participants completed each survey. A full breakdown of all participants and the number of completed surveys is presented in Table 3.

### Participant characteristics

The demographics of the in-person cohort and mail-based cohort recruited from the start until December 9, 2022, are presented in Tables 4 and 5. Consistent with the convenience sampling approach, the study population was not randomized. Our analytical cohort included 34,695 participants aged 18–100 years (median = 44.2 years). Distributions for demographic and behavioral factors of HOP Participants are: 72.3% lived in urban areas, 75.3% were females, 83.0% were white, 36.7% were obese, and 28.7% were overweight. Regarding behavior-related factors, 39.1% had smoked at least 100 cigarettes in their lives, and 92.7% had consumed at least 30 alcoholic beverages in their lives. Most participants had inadequate fruit and vegetable intake, with 70.3% having ≤ 1 cup/day of fruit intake and 59.1% having ≤ 1 cup/day of vegetable intake. The majority of participants did not report regular exercise or exercised for less than 1 h/day (56.2%) during weekdays. Most participants reported more than 8 h/day sleep (54.5% during weekdays and 56.1% during weekends).

### The expanding research opportunities and impact of the HOP study

One key feature of HOP is that it can easily provide additional research opportunities. When HOP participants initially consented to the study, they were informed about the potential for receiving invitations to participate in additional studies.

Subsequent studies launched within HOP included studies that have included the collection of additional data and specimens, for example, collection of fecal samples and wristband exposure data to understand the association between the microbiome, environmental chemical exposure, and measures of impulsivity and vigilance (National Institute of Environmental Health Sciences of the National Institutes of Health grant number: P30ES030287); targeted recruitment for partner institution studies, such as a prospective observational cohort study to assess SARS-CoV-2 vaccine effectiveness in children and adults (CDC Contract number:75D30121C12297), and older adult cancer survivors to engage in home monitoring (supported by the Knight Cancer Institute's Cancer Center grant number: P30CA69533).

Of great importance, after the launch of HOP, the NCI Moonshot program funded U01CA232819 to evaluate the cost-effectiveness of broad HBOC and Lynch syndrome screening in the HOP population (NCT04494945). We previously published [21] a breakdown of genetic testing results from the first 13,774 results returned to the HOP participants. Briefly, just over 5% percent of HOP genetic health screenings resulted in a test that required confirmation and genetic counseling. Complete collection of data for NCT04494945 is expected in the summer of 2025.

In our ancillary studies derived from HOP, the response rate is contingent on the particular study. For instance, the microbiome study boasts a recruitment rate of 90% due to targeted invitations sent to a specific group. Conversely, when we have reached out to HOP participants via a single email to offer opportunities to

**Table 5.** Lifestyle characteristics of healthy oregon project (HOP) participants

| Participant characteristics                                  | HOP participants<br>n (%) |
|--|---------------------------|
| BMI (n = 20,234)   |                           |
| Underweight  | 253 (1.3%)                |
| Healthy weight   | 6699 (33.1%)              |
| Overweight   | 5805 (28.7%)              |
| Obese  | 7477 (36.9%)              |
| <b>Cigarette smoking (n = 24,701)<sup>1</sup></b>            |                           |
| Yes  | 9665 (39.1%)              |
| No   | 15,046 (60.9%)            |
| <b>Alcohol intake (n = 24,711)<sup>2</sup></b>               |                           |
| Yes  | 22,918 (92.7%)            |
| No   | 1793 (7.3%)               |
| <b>Current alcohol drinker (n = 22,918)<sup>3</sup></b>      |                           |
| Yes  | 16,579 (72.3%)            |
| No   | 6339 (27.7%)              |
| <b>Ever Had Cancer (n = 20,233)<sup>4</sup></b>              |                           |
| Yes  | 3177 (15.7%)              |
| No   | 17,056 (84.3%)            |
| <b>Vegetables intake (cups/day) (n = 20,233)<sup>5</sup></b> |                           |
| 0  | 510 (2.5%)                |
| 1  | 11,442 (56.6%)            |
| 2  | 5305 (26.2%)              |
| 3  | 2001 (9.9%)               |
| 4  | 975 (4.8%)                |
| <b>Fruits intake (cups/day) (n = 20,233)<sup>6</sup></b>     |                           |
| 0  | 964 (4.8%)                |
| 1  | 13,242 (65.5%)            |
| 2  | 4370 (21.6%)              |
| 3  | 1282 (6.3%)               |
| 4  | 375 (1.9%)                |
| <b>Exercise weekday (hours/day) (n = 20,234)<sup>7</sup></b> |                           |
| None   | 1604 (7.9%)               |
| >0 - ≤ 1   | 9776 (48.3%)              |
| 2  | 6972 (34.5%)              |
| ≥3   | 1882 (9.3%)               |
| <b>Exercise weekend (hours/day) (n = 20,234)<sup>8</sup></b> |                           |
| None   | 1980 (9.8%)               |
| >0 - ≤ 1   | 7735 (38.2%)              |
| 2  | 8330 (41.2%)              |
| ≥3   | 2189 (10.8%)              |
| <b>Sleep weekday (hours/day) (n = 20,234)<sup>9</sup></b>    |                           |
| <8   | 9209 (45.5%)              |
| ≥8   | 11,025 (54.5%)            |

(Continued)

**Table 5.** (Continued)

| Participant characteristics                                | HOP participants<br>n (%) |
|--|---------------------------|
| <b>Sleep weekend (hours/day) (n = 20,234)<sup>10</sup></b> |                           |
| <8   | 8888 (43.9%)              |
| ≥8   | 11,346 (56.1%)            |
| <b>General health (n = 20,234)<sup>11</sup></b>            |                           |
| Excellent  | 1504 (7.4%)               |
| Very good  | 6004 (29.7%)              |
| Good   | 8105 (40.1%)              |
| Fair   | 3818 (18.9%)              |
| Poor   | 803 (4.0%)                |

Original survey questions (The analyses were based on data received from beginning of the study to 12/09/2022).

<sup>1</sup>Smoking: Have you smoked at least 100 cigarettes in your ENTIRE life?.

<sup>2</sup>Alcohol: Have you had at least 30 alcoholic beverages in your ENTIRE life?.

<sup>3</sup>Current alcohol drink: Do you CURRENTLY drink alcoholic beverages (beer, wine, liquor, cocktails, coolers)?.

<sup>4</sup>Ever had cancer: Has a physician ever told you that you have CANCER (not including basal or squamous cell skin cancer or cervical dysplasia)?.

<sup>5</sup>Vegetables: About how many cups of VEGETABLES (other than iceberg lettuce and potatoes) do you eat PER DAY?.

<sup>6</sup>Fruits: About how many cups of FRUIT (including 100% pure fruit juice) do you eat or drink PER DAY?.

<sup>7</sup>Exercise weekday: Exercise/physical activity (hours per day WEEKDAYS M-F).

<sup>8</sup>Exercise weekend: Exercise/physical activity (hours per day WEEKENDS Sat & Sun).

<sup>9</sup>Sleep weekday: Sleeping (hours per day on WEEKDAYS M-F).

<sup>10</sup>Sleep weekend: Sleeping (hours per day WEEKENDS Sat & Sun).

<sup>11</sup>General health: In general, would you say your physical health is.

participate in external studies, the email open rate has ranged between 40–55% with approximately 12–15% of participants clicking through to the external study site, and either completing surveys or completing eligibility to enroll in the study.

## Discussion

The HOP study applied novel recruitment approaches to rapidly engage with a statewide population. Prior to the onset of COVID-19, our primary recruitment methods were in-person community events and the installation of vending machines for pick-up/drop-off genetic screening kits. With the onset of COVID-19, we shifted our recruitment model to a remote approach by marketing the study across the state through multiple social media campaigns, both reaching the broad general population and more specifically targeted efforts at hard-to-reach under-recruited populations. The HOP's goal was to contribute to research that is convenient, accessible, and mutually beneficial.

The participant-driven model based on engagement with the questions utilized in the cohort and the prospect of return of information was widely acceptable. Concerns on social media were primarily focused on genetic health screening in HOP and commonly expressed the idea that genetic testing results might impact health or life insurance. Less commonly, social media concerns focused on intrusiveness, concerns about sharing DNA, and privacy rights. Engagement with HOP staff through social media and direct electronic communications pointed potential participants to web resources, including the HOP website ([www.healthyoregonproject.com](http://www.healthyoregonproject.com)), which informed potential participants on security, confidentiality, and federal or state regulations that



prevent discrimination in health care based on inherited genetic differences.

However, two key challenges remain to be addressed. First, we do not yet have firm measures of how well we will continue to engage with participants or how long the engagement will last. Long-term engagement is necessary to facilitate the secondary research projects that HOP is designed to support. Second, HOP represents a single small catchment area that is demographically quite uniform, as Oregon is among the oldest and least diverse states in the US. Broad centralized national cohorts such as Veterans Affairs' Million Veteran Program and NIH's All of Us are certainly other solutions to the demographic limitations of the HOP. However, we imagine another scenario in which HOP could integrate with an array of similar projects across the US to facilitate expansion or validation cohorts for questions arising from those diverse cohorts. Cohort integration would diversify the population base and democratize access to cohorts for researchers. Finally, these cohorts could create an expanded array of driving projects, and the community of cohorts could evaluate how best to bring the community into research. A network of cohorts utilizing the same platforms and similar consents could engage with 100,000 participants nationwide sharing access to data and research questions broadly.

## Conclusions

The HOP study provides a novel, remote-based, inexpensive recruitment approach to effectively establish a large-scale cohort for population-based studies of cancer. This study provides evidence that remote recruitment approaches can effectively establish large-scale cohorts for population-based cancer studies. The implementation of the study facilitates the collection of extensive survey and biological data into a repository that can be broadly shared and supports the creation of cohorts available for collaborative clinical and translational research.

**Data availability statement.** Our study has not been published elsewhere nor has it been submitted simultaneously for publication elsewhere.

**Acknowledgments.** We thank all the HOP participants who provided information or donated samples. We also thank our patient advocates, Sue Williams, and other patient advocates who provided invaluable advice on how to communicate and support participants who learn that they have inherited cancer syndrome. Additionally, we thank Julia Bainbridge, Ryan Lutz, and Joseph Easley, who assisted in participant recruitment and project management.

**Author contribution.** Z Zhang conducted the descriptive data analyses and wrote the first draft of the manuscript. J. Shannon and P.T. Spellman are the Principal Investigators of the Healthy Oregon Project and obtained funding. K Johnson-Camacho performed administrative, human subject research compliance, technical, and material support for all aspects of the HOP study. A Adey provided technical assay guidance. P Anur and A Mitchell created HOP databases and provided HOP data to researchers. KA Brown and PE Farris are managers of the HOP. C Conrad and R Crist coordinated the study participant recruitment and project management. CA Harrington directed the HOP participants' saliva samples for DNA extraction and genetic sequencing analysis. B O'Roak guided mutation detection in populations. C. Sue Richards led the Knight Diagnostic Lab for next-generation sequencing analyses. LK Marriott developed communication materials and facilitated the initial statewide rollout. V Serrato is the marketing coordinator who coordinates statewide recruitment. A Shafer led the public communication and community outreach efforts of the HOP study. All the authors interpreted the results and critically revised the manuscript for valuable intellectual content.

**Funding statement.** The research reported in this publication was supported by funding from the Cancer Early Detection Advanced Research Center at Oregon Health & Science University's Knight Cancer Institute, the National Cancer Institute of the National Institutes of Health under Award Number P30CA069533 (B. Druker), and 1U01CA232819 (P Spellman, J Shannon). This work was also partially supported by Oregon Clinical and Translational Research Institute, which is supported by the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant Award Number UL1TR002369. The content is the sole responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**Competing interests.** None.

## References

1. Manolio TA, Collins R. Enhancing the feasibility of large Cohort studies. *JAMA*. 2010;304(20):2290–2291. doi: [10.1001/jama.2010.1686](https://doi.org/10.1001/jama.2010.1686).
2. Toledano MB, Smith RB, Brook JP, Douglass M, Elliott P. How to establish and follow up a large prospective cohort study in the 21st century—lessons from UK COSMOS. *PLoS One*. 2015;10(7):e0131521. doi: [10.1371/journal.pone.0131521](https://doi.org/10.1371/journal.pone.0131521).
3. Teague S, Youssef GJ, Macdonald JA, *et al.* Retention strategies in longitudinal cohort studies: a systematic review and meta-analysis. *BMC Med Res Methodol*. 2018;18(1):151. doi: [10.1186/s12874-018-0586-7](https://doi.org/10.1186/s12874-018-0586-7).
4. Corcoran L, Friedenreich CM, McNeely ML, *et al.* A qualitative study examining newly diagnosed breast cancer patients' experiences of participating in the Alberta moving beyond breast cancer (AMBER) prospective cohort study. *BMC Cancer*. 2023;23(1):500. doi: [10.1186/s12885-023-10967-3](https://doi.org/10.1186/s12885-023-10967-3).
5. Dawber TR, Meadors GF, Moore FE Jr. Epidemiological approaches to heart disease: the Framingham study. *Am J Public Health Nations Health*. 1951;41(3):279–281. doi: [10.2105/ajph.41.3.279](https://doi.org/10.2105/ajph.41.3.279).
6. Young FE, Nightingale SL, Temple RA. The preliminary report of the findings of the aspirin component of the ongoing physicians' health study. The FDA perspective on aspirin for the primary prevention of myocardial infarction. *JAMA*. 1988;3(21):3158–3160.
7. The Steering Committee of the Physicians' Health Study Research G. Preliminary report: findings from the aspirin component of the ongoing physicians' health study. *N Engl J Med*. 1988;318(4):262–264. doi: [10.1056/NEJM198801283180431](https://doi.org/10.1056/NEJM198801283180431).
8. Gaziano JM, Sesso HD, Christen WG, *et al.* Multivitamins in the prevention of cancer in men: the physicians' health study II randomized controlled trial. *JAMA*. 2012;308(18):1871–1880. doi: [10.1001/jama.2012.14641](https://doi.org/10.1001/jama.2012.14641).
9. Sesso HD, Christen WG, Bubes V, *et al.* Multivitamins in the prevention of cardiovascular disease in men: the physicians' health study II randomized controlled trial. *JAMA*. 2012;308(17):1751–1760. doi: [10.1001/jama.2012.14805](https://doi.org/10.1001/jama.2012.14805).
10. Bao Y, Bertoia ML, Lenart EB, *et al.* Origin, methods, and evolution of the three nurses' health studies. *Am J Public Health*. 2016;106(9):1573–1581. doi: [10.2105/AJPH.2016.303338](https://doi.org/10.2105/AJPH.2016.303338).
11. Denny JC, Rutter JL, Goldstein DB, *et al.* The “all of us” research program. *N Engl J Med*. 2019;381(7):668–676. doi: [10.1056/NEJMs1809937](https://doi.org/10.1056/NEJMs1809937).
12. Coughlin JW, Martin LM, Zhao D, *et al.* Electronic health record-based recruitment and retention and mobile health app Usage: multisite Cohort study. *J Med Internet Res*. 2022;24(6):e34191. doi: [10.2196/34191](https://doi.org/10.2196/34191).
13. Sudlow C, Gallacher J, Allen N, *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;12(3):e1001779. doi: [10.1371/journal.pmed.1001779](https://doi.org/10.1371/journal.pmed.1001779).
14. Bernstein L, Allen M, Anton-Culver H, *et al.* High breast cancer incidence rates among California teachers: results from the California teachers study (United States). *Cancer Causes Control*. 2002;13(7):625–635. doi: [10.1023/a:1019552126105](https://doi.org/10.1023/a:1019552126105).
15. Horn-Ross PL, Hoggatt KJ, West DW, *et al.* Recent diet and breast cancer risk: the California teachers study (USA). *Cancer Causes Control*. 2002;13(5):407–415. doi: [10.1023/a:1015786030864](https://doi.org/10.1023/a:1015786030864).
16. Holmes JH, Beinlich J, Boland MR, *et al.* Why is the electronic health record so challenging for research and clinical care? *Methods Inf Med*. 2021;60(1-02):32–48. doi: [10.1055/s-0041-1731784](https://doi.org/10.1055/s-0041-1731784).

17. **Feigelson HS, Clarke CL, Van Den Eeden SK, et al.** The kaiser permanente research bank cancer Cohort: a collaborative resource to improve cancer care and survivorship. *BMC Cancer*. 2022;**22**(1):209. doi: [10.1186/s12885-022-09252-6](https://doi.org/10.1186/s12885-022-09252-6).
18. **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**(9):906–913. doi: [10.1038/gim.2015.187](https://doi.org/10.1038/gim.2015.187).
19. **Gaziano JM, Concato J, Brophy M, et al.** Million veteran program: a mega-biobank to study genetic influences on health and disease. *J Clin Epidemiol*. 2016;**70**:214–223. doi: [10.1016/j.jclinepi.2015.09.016](https://doi.org/10.1016/j.jclinepi.2015.09.016).
20. **Spreafico A, Hansen AR, Abdul Razak AR, Bedard PL, Siu LL.** The future of clinical trial design in oncology. *Cancer Discov*. 2021;**11**(4):822–837. doi: [10.1158/2159-8290.CD-20-1301](https://doi.org/10.1158/2159-8290.CD-20-1301).
21. **O'Brien TD, Potter AB, Driscoll CC, et al.** Population screening shows risk of inherited cancer and familial hypercholesterolemia in Oregon. *Am J Hum Genet*. 2023;**3**(8):1249–1265. doi: [10.1016/j.ajhg.2023.06.014](https://doi.org/10.1016/j.ajhg.2023.06.014).
22. **Alvord TW, Marriott LK, Nguyen PT, et al.** Public perception of predictive cancer genetic testing and research in Oregon. *J Genet Couns*. 2020;**29**(2):259–281. doi: [10.1002/jgc4.1262](https://doi.org/10.1002/jgc4.1262).