

Design and outcomes of the Patient Centered Outcomes Research Institute coronary heart disease cohort study

Christianne L. Roumie^{a,b,c,d,*}, Niral J. Patel^{a,c,d}, Daniel Muñoz^{e,f}, Justin Bachmann^{e,f}, Ashton Stahl^c, Ryan Case^c, Cardella Leak^c, Russell Rothman^{a,c,d}, Sunil Kripalani^{a,c,d}

^a Division of General Internal Medicine and Public Health, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, United States

^b Veterans Health Administration-Tennessee Valley Healthcare System Geriatric Research Education Clinical Center (GRECC), HSR&D Center, Nashville, TN, United States

^c Center for Health Services Research, Vanderbilt University Medical Center, Nashville, TN, United States

^d Center for Clinical Quality and Implementation Research, Vanderbilt University Medical Center, Nashville, TN, United States

^e Division of Cardiovascular Medicine, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, United States

^f Vanderbilt Translational and Clinical Cardiovascular Research Center (VTRACC), Vanderbilt University Medical Center, Nashville, TN, United States

ABSTRACT

Background: The Patient Centered Outcomes Research Institute (PCORI) established Clinical Data Research Networks (CDRNs) to support pragmatic research. The objective was to electronically identify, recruit, and survey coronary heart disease (CHD) patients and describe their characteristics, health status, and willingness to participate in future research.

Methods: We developed a computable phenotype and assembled CHD patients 30 years or older and had visits or hospitalizations between 2009 and 2015. A sample of patients was surveyed between August 2014 and September 2015. Survey administration included the following methods: face-to-face, telephone, paper or web portal. Survey items covered broad domains including: health literacy and numeracy, and socio-demographics, physical and mental health, health behaviors, access to medical care, and willingness to participate in future research.

Results: Of 5517 approached patients, 2605 completed the survey. Participants were mostly white (~88%), male (68%) and had a median age of 69 years (interquartile range [IQR] 61–76 years). Most respondents' health literacy and numeracy were adequate (83.2% and 84.3%, respectively). Only 4% of respondents reported that their overall health or physical health was excellent. The majority (~58%) reported that their health was good or very good, while 40% reported that their general and physical health were fair or poor. The majority reported that their quality of life was good to excellent (81%). Limitations in physical health and function were common, including often/always having fatigue (25%), pain (38.7%), or sleep difficulty (19.7%). A patient sample (n = 1936) was provided with a trial summary which would randomize their aspirin dose; and 63% reported that they would consider participating.

Conclusion: Many patients with CHD had limitations in physical health. However, the majority reported a good or excellent quality of life.

1. Introduction

Over the past 30 years, there have been numerous randomized clinical trials to evaluate the impact of medications, procedural therapies, and management strategies on outcomes in patients with cardiovascular disease. In general, these trials have relied on traditional methods for identification and recruitment of patients using research nurses, which can be resource-intensive. In an era of big data and

pragmatic clinical trials, more efficient and cost-effective approaches are needed [1]. The construction of cardiovascular disease cohorts from administrative systems could facilitate identification of a potentially large segment of patients willing to participate in future research. Furthermore, future studies could focus enrollment on populations often underrecruited in cardiovascular clinical trials; for example, women and minorities. However, broad definitions of coronary heart disease are often not validated in large administrative databases [2–5].

* Corresponding author. Nashville VA Medical Center, 1310 24th Ave South GRECC, Nashville, TN 37212, United States.

E-mail addresses: christianne.roumie@vanderbilt.edu (C.L. Roumie), niral.patel@vanderbilt.edu (N.J. Patel), daniel.munoz@vanderbilt.edu (D. Muñoz), justin.m.bachmann@vanderbilt.edu (J. Bachmann), astahl256@gmail.com (A. Stahl), ryan.case@pop.belmont.edu (R. Case), leak0620@comcast.net (C. Leak), russell.rothman@vanderbilt.edu (R. Rothman), sunil.kripalani@vanderbilt.edu (S. Kripalani).

<https://doi.org/10.1016/j.conctc.2018.03.001>

Received 22 September 2017; Received in revised form 16 February 2018; Accepted 8 March 2018

Available online 09 March 2018

2451-8654/ Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Abbreviations

ADAPTABLE trial	Aspirin Dosing: A Patient-centric Trial Assessing Benefits and Long-Term Effectiveness
CDRN	Clinical Data Research Networks
CHD	Coronary heart disease
EHR	Electronic Health Record
ICD 9 CM	International Classification of Diseases, Ninth Revision, Clinical Modification
PCORI	Patient Centered Outcomes Research Institute

PHCS-2	Perceived Health Competence Scale
PROMIS	Patient Reported Outcomes Measurement Information System
RA	Research Assistant
RD	Research Derivative
REDCap	Research Electronic Data Capture
U.S.	United States
VHAN	Vanderbilt Health Affiliated Network
VUMC	Vanderbilt University Medical Center

Efficient recruitment of participants is paramount to the success of clinical research and is of particular importance in pragmatic clinical trials, including the ongoing Patient-Centered Outcomes Research Institute (PCORI) funded ADAPTABLE trial (Aspirin Dosing: A Patient-centric Trial Assessing Benefits and Long-Term Effectiveness) [6]. Thus, starting with a broad cardiovascular disease population seen in clinical practice is likely to be an effective strategy in engaging a broader, more diverse pool of research participants. To this end, PCORI funded 13 Clinical Data Research Networks (CDRNs) to build infrastructure for comparative effectiveness research and pragmatic clinical trials. Each CDRN was tasked with developing a cohort for a common condition, a rare condition, and one for weight-related research [7–9]. Here we describe the identification, recruitment, and enrollment of the Mid-South CDRN common condition cohort of patients with Coronary Heart Disease (CHD). We present their sociodemographic characteristics, health status using patient-reported outcome measures, and willingness to participate in research.

2. Methods

2.1. Setting

At the time of this study, the Mid-South CDRN was comprised of Vanderbilt University Medical Center (VUMC), the Vanderbilt Health Affiliated Network (VHAN), and Greenway Prime Research Network. VUMC is a tertiary care academic medical center in Nashville, TN, which includes large cardiology and primary care practices. The VHAN is a clinically integrated network which includes more than 40 hospitals and 300 ambulatory practices, with an estimated reach of more than 3 million patients in the Mid-South area. Greenway Health provides electronic health record (EHR) and practice management software to more than 2000 sites around the country. For the present CHD cohort, only VUMC and nearby VHAN clinical sites were used.

2.2. CHD computable phenotype

We used the VUMC Research Derivative (RD) to develop a computable phenotype to identify a population of patients with CHD between January 2009 and 2014. The RD is one component of the PCORI Mid-South common data model database and is composed of clinical and related data derived from VUMC's enterprise data warehouse and restructured for research [10]. The RD includes data from the EHR, ORMIS (Operating Room Management Information System), scheduling systems, and medication prescribing and administration. Data types include inpatient and outpatient encounters, clinical notes and documentation, nursing records, medication data, laboratory data, and vital signs. Data may be structured (ICD-9 CM codes [11]), semi-structured (laboratory tests and results), or unstructured (patient summaries and physician progress reports). The medical record number, other person identifiers, and dates are preserved within the database. Patient vital status is derived from data available through the Social Security Death Index.

CHD computable phenotype definitions were developed using an

iterative process, and a sample of 50 charts was reviewed by 2 physicians until consensus was achieved on the presence of coronary disease. The computable phenotype identified patients with CHD, on the basis of outpatient or inpatient billing codes. Patients fit the phenotype by fulfilling either an outpatient case definition for coronary heart disease (case type 1) or having a revascularization procedure (case type 2). Case type 1 was defined as having two outpatient visits on separate days for prior myocardial infarction (ICD 9-CM diagnosis code 410.*; 412.*; 429.7*) or obstructive coronary artery disease (411.*; 413.*; 414.*; V45.81; V45.82). For case type 2, a revascularization procedure for CHD was defined as having one inpatient or outpatient procedure code for coronary artery bypass or percutaneous transluminal coronary angioplasty (CPT codes 33140; 33533–36; 33510–23; 33530; 33533–33536; 92920–92921; 92924–92925; 92928–92929; 92933–92934; 92937–92938; 92941; 92943–92944; 92980–82; 92984; 92995–6; 92974 or ICD-9 CM Procedure code: 36.01; 36.02; 36.03; 36.05; 36.09; 36.10–36.19). We then validated the final above algorithm. Research assistants abstracted 470 charts. The recorded diagnosis of CHD in the patient record or discharge summary was considered the reference standard for validation. The positive predictive value of the final above definition was 98.5% (192 true positives/195 algorithm positive) and the sensitivity was 94.6% (192 true positives/203 coronary disease positive patients) [12].

2.3. Study sample

To identify a pool of potentially eligible patients, we applied the phenotype to patient records from VUMC, as well as patients seen by Vanderbilt cardiologists at nearby VHAN sites whose data were available in the RD. To increase the likelihood that identified patients would have accurate contact information on file and feel a greater sense of engagement with the medical center, we limited the search to patients with inpatient or outpatient clinical encounters within the last 5 years (January 2009 through June 2014). Two updates were performed to capture additional newly diagnosed CHD patients. Thus, the end search dates were modified to be through December 2014 and April 2015. We restricted the sample to patients aged 30 years or older, to exclude patients likely to have non-traditional (non-atherosclerotic) coronary conditions, including patients with congenital heart disease. We also excluded patients who had an unknown date of birth or sex, were receiving hospice care, or who had a recorded date of death at the time of the search. The institutional review board of Vanderbilt University approved this study. Study participants were enrolled between August 2014 and September 2015. All surveyed participants provided informed consent and were offered \$10 for survey completion.

2.4. Study procedures

Patients who met the computable phenotype definition of CHD and had any form of contact with the VUMC health system were recruited for survey participation. Research assistants further excluded patients if the medical record reported impaired cognition (dementia or severe psychiatric illness), legal blindness, significant hearing loss, if the

patient was unable to communicate in English, or, if after contact, the patient's illness severity rendered them unable to take the survey. The primary method of recruitment used by the 3 research assistants was face-to-face contact at scheduled cardiology or primary care visits. Research assistants (RA) approached a convenience sample of patients who had been identified by the CHD phenotype for inclusion. Additional methods included mailed letters with instructions on accessing the survey via web portal, mailed paper survey, and email with a link to the survey web portal. Telephone recruitment was used for patients who were missed at clinic visits. These different approaches were used to reach a broader pool of patients, including those who received care in satellite cardiology clinics, and to accommodate patient preference for method of survey completion. Additionally, patients could be referred for participation by another research team that was conducting a healthy weight survey (weight cohort). Patients could have been contacted multiple times using more than one method.

For face-to-face recruitment, trained staff reviewed VUMC cardiology clinic appointment schedules and approached patients with a CHD diagnosis. Patients who preferred not to complete the survey while in clinic were given the option to complete it via phone, mail, or through a link to the web portal by an email.

Among patients who did not have an upcoming visit or who received care at an off-site practice, we used two mail-based approaches. One involved sending patients a recruitment letter that informed them of the study and provided information about how to complete the survey via web portal or telephone. The second approach involved mailing a paper survey with a stamped return envelope.

Patients with an account to the patient portal, MyHealthAtVanderbilt [13], received a general opt-in email to be contacted directly about research opportunities. Those who agreed were enrolled in MyResearchAtVanderbilt and each eligible patient was offered study participation if they fulfilled the CHD phenotype.

Patients who completed the survey by phone were typically those whom RAs had attempted to meet at a clinic visit, but missed for logistical reasons. RAs also performed follow-up telephone calls for recruitment at least once after mailing the recruitment letter or survey, or sending an email invitation. Patients had the option to complete the survey by phone at that time.

2.5. Survey measures

Depending on the mode of recruitment, surveys were administered

to patients in-person on paper, in-person via tablet computer that accessed the web portal, by telephone (verbally), self-administered on paper, or self-administered via a web portal. In general, the survey took between 15 and 30 min to complete. The web-based survey was designed and distributed only in English using REDCap (Research Electronic Data Capture), a secure application for research [14].

Table 1 describes the survey domains. Many survey items were chosen based on the PCORI common data model and work of the Patient-Reported Outcomes Task Force, which recommended items for use across research networks. Others items for inclusion were selected by the research team. The survey was reviewed during a patient engagement studio for clarity and content.

2.6. Health status, behaviors, and habits

Health status of participants was measured using several items from the Patient Reported Outcomes Measurement Information System (PROMIS) Global Health Scale, which asks about general, physical, and mental health, and quality of life [15]. These items are scored on a 5-point scale ranging from excellent to poor. Self-reported problems during the last 7 days related to fatigue, depressive symptoms, pain, and sleep were also assessed. The 2-item Perceived Health Competence Scale (PHCS-2) summarized patients' ability to manage their health and problem-solve [16].

Medication adherence was queried with a single item that normalized the occurrence of missed medication doses by asking "how many days in the past week did you miss taking one or more of your prescription medicines?" Additional questions assessed aspirin and anticoagulant use.

A detailed measure of physical activity was obtained by using the 7 item short International Physical Activity Questionnaire [17]. This questionnaire has four domains: leisure time physical activity, domestic and gardening (yard) activities, work-related physical activity, and transport-related physical activity. Physical activity was scored on three levels: low (no activity), moderate (3 + days at vigorous intensity, 5 + days at moderate intensity, or 5 + days combining walking, moderate and vigorous intensity), or high (3 + days of vigorous intensity or 7 + days of combining walking, moderate and vigorous intensity).

Tobacco use was measured using 11 items from the National Adult Tobacco Survey [18]. These items included past and current smoking status, number of cigarettes per day, and the use of smokeless tobacco

Table 1
Construct domains and scales utilized in the Coronary Heart Disease Cohort.

Domains	Scales	Domains/description
Health Status, Social Support and Health Behaviors	<ul style="list-style-type: none"> Adapted Patient Reported Outcomes Measurement System (PROMIS) Perceived Health Competence and Social Support Medication adherence International Physical Activity Questionnaire National Adult Tobacco Survey Alcohol Use Disorder Inventory Test National Health and Nutrition Examination Survey National Health Interview Survey 	<ul style="list-style-type: none"> Global Health Physical health Mental health Ability to do things for my health Difficult to find effective solutions for my health How often you can count on someone for emotional support Medication Adherence Exercise frequency and duration Smoking Status Alcohol intake Diet Ability to receive medical care Reasons for delays in care Desire to participate in health related research studies
Access to Medical Care		<ul style="list-style-type: none"> Race, ethnicity Employment status Education achieved Household income
Future Research Participation		Confidence with reading related to health
Socio-demographics		Confidence with calculating numbers related to health
Health Literacy/Numeracy	<ul style="list-style-type: none"> Brief Health Literacy Scale Subjective Numeracy Scale 	

and electronic cigarettes. Alcohol consumption was assessed using 2 items from the Alcohol Use Disorder Inventory Test [19], which is scored and categorized into four risk strata: low risk (scores 0–7), risky or hazardous (scores 8–15), harmful (scores 16–19), or high risk (scores 20+). Healthy diet was assessed with the single item Healthy Eating Index-2010 from the National Health and Nutrition Examination Survey [20]. Individuals were asked: “In general, how healthy is your overall diet?” The Healthy Eating Index-2010 is a reliable and valid measure that distinguishes the healthfulness of diets [21] and yields comparable results to longer assessments of 24-h dietary recall [22]. Responses occurred on a 5-point scale from 1 “poor” to 5 “excellent.”

2.7. Access to medical care

Access to medical care was assessed using 8 items from the National Health Interview Survey [23], which asked about satisfaction with medical care and common reasons patients might choose to delay getting healthcare.

2.8. Research interest and trust

A set of items asked participants about their prior participation in research, interest in participating in future research, and preferred mode of contact. Questions asked about specific types of studies (survey, behavioral study, blood draw, medication trial, overnight hospital stay). Patients were also asked to rate their trust in different sources of medical information, ranging from magazines to physicians

and health agencies. In preparation for launch of the ADAPTABLE trial, patients were given a brief description of the study and asked if they would be interested in participating and willing to change aspirin dose if it was ok with their doctor.

2.9. Socio-demographics, health literacy and numeracy

The survey also included 7 items on socio-demographic characteristics including number of people in the household, marital status, level of education, household income, gender, employment, and race and ethnicity. Health literacy was assessed using the 3-item Brief Health Literacy Screen [24] which asks patients to rate their confidence level with reading health information, completing medical forms, and learning new information. Each item is scored on a 5-point scale. Adequate health literacy was defined as a score of > 9 [25]. Numeracy was measured using a shortened 3-item version of the Subjective Numeracy Scale [26], which asks patients about their use of numerical data and calculations. Each item is scored on a 6-point scale, yielding a scale range of 6–18 [27]. Adequate subjective numeracy was defined as 10 or higher.

2.10. Statistical analysis

Using descriptive statistics, proportions, means, standard deviations, and ranges, we examined characteristics of participants and non-participants. To be included in the analysis, respondents should have completed $\geq 70\%$ of the survey questions. For the sample that

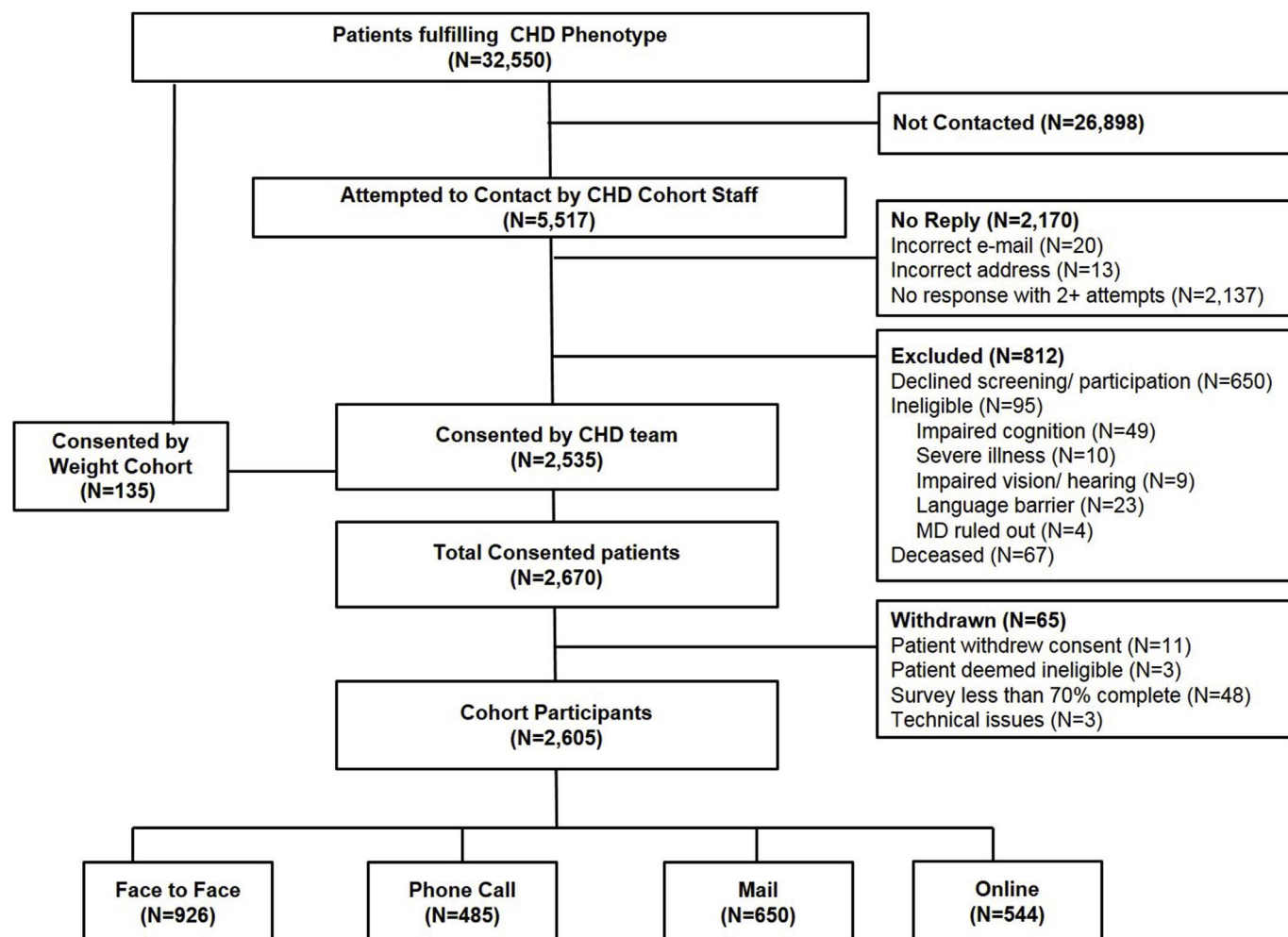


Fig. 1. Flow of participants.

completed the survey, we report participant responses to their self-reported health collected using the PROMIS measures and examined their distributions in relation to socio-demographics using Pearson's chi-squared comparisons of proportions. We also report patient willingness to participate in a trial about optimal aspirin dose.

3. Results

3.1. Identification of participants

A total of 32,550 patients met the CHD phenotype criteria. Because of staffing limitations, the study staff attempted to contact a sample of 5517 patients, and the remainder was not contacted. Patients were contacted through face-to-face (convenience sample because of already scheduled visits), email (entire sample of patients with an available email) and mail contact (random sample of the remaining). The weight cohort consented 135 CHD patients. Of those attempted, 2170 (39%) could not be reached, and 812 were excluded because they declined screening or participation, were ineligible, or deceased. Of the 2670 patients who consented, 65 (2.4%) subsequently withdrew or were withdrawn from the sample for < 70% survey completion. Thus, 2605 participants remained in the CHD survey cohort with the following enrollment distribution: face-to-face (n = 926, 35.5%); phone call (n = 485, 18.6%); mail (n = 650, 25.0%); or online (n = 544, 20.9%) (Fig. 1).

3.2. Participant characteristics

Patient characteristics, clinical data, and comorbidities extracted from the Research Derivative are shown in Table 2. Survey participants and non-participants were predominantly reflective of the patients seen in VUMC portion of the MidSouth CDRN. Patients were white (85%), males (68%), with median age of 69 years (interquartile range [IQR] 61–76 years). Survey participants had more healthcare utilization than non-participants (median 17 visits vs. 8 visits in past 2 years); therefore, they more often had co-morbidities and medications recorded in their electronic health record. Almost 90% of the survey participants had aspirin recorded in their electronic health record and 40% had a record of other antiplatelet agents (clopidogrel, ticlopidine, aspirin/dipyridimide, dipyridimide alone, prasugrel, or ticagrelor). The proportion of survey respondents with adequate health literacy (83.2%) and subjective numeracy (84.3%) was relatively high, but approximately 1 in 6 participants had limited skills in these areas.

3.3. Health status

Only about 4% of respondents reported that their overall health or physical health was excellent. The majority (about 58%) reported that their overall health was good or very good, while about 40% reported that their general and physical health were fair or poor. More patients, however, reported that their quality of life was good to excellent (81%). Overall, among respondents, 23% reported that their mental health was excellent and 45% reported that in the past 7 days they had not experienced depressive symptoms. Limitations in physical health and function were commonly reported, including often or always having fatigue (25%), pain (38.7%), or sleep difficulty (19.7%) (Fig. 2). With regard to perceived health competence, 83.2% of patients reported that they were “able to do things for my health as well as most other people.” Despite this, 27.5% also reported that they strongly agreed with the statement “it is difficult for me to find effective solutions for my health problems.”

There were differences in health status by patient socio-demographics. Patients who were older than 65 more often reported excellent/very good general health (63% vs. 57%, prevalence difference [PD] 6%; 95% CI [2.3%, 10.3% p = 0.002]) and physical health (63% vs. 53%, PD 9.6%; 95% CI [5.6%, 13.6% p < 0.001]) compared

to patients who were 45–64 years old. Women more often reported poorer general health (47% vs. 35%, PD 12%; 95% CI [7.9%, 16% p < 0.001]), physical health (46% vs. 36%, PD 10%; 95% CI [5.9%, 14% p < 0.001]), and quality of life (23% vs. 17%, PD 6%; 95% CI [2.6%, 9.3% p = 0.0003]) compared to men. Participants who reported being unable to work reported poor quality of life compared to those respondents who were retired from work (47% vs. 16%, PD 31%; 95% CI [26.2%, 36.7% p < 0.001]). There were no major differences in general health, physical health or quality of life between participants who had more versus fewer healthcare visits. Patients who were hospitalized in the last two years reported poorer general health, physical health, and quality of life (Table 3).

3.4. Future participation in research

Of the 2605 respondents, 32.8% stated that they would be very interested in future survey and behavioral research, 32.1% would be willing to donate a sample of blood and 10.6% were willing to participate in research if it involved an overnight hospital stay.

A sample (N = 1936) of patients were provided with a brief summary of the ADAPTABLE trial (including randomization of aspirin dose)

Table 2
Patient characteristics extracted from the Research Derivative.

	Participants N = 2605	Non-Participants ^a N = 29,946
Age, years median (IQR)	68.9 (61.3–75.9)	68.7 (60.4–76.6)
Sex, n (%)		
Male	1780 (68.3)	20,116 (67.2)
Race, n (%)		
White	2310 (88.7)	25,329 (84.6)
Black	191 (7.3)	1,891 (6.3)
Other	38 (1.5)	466 (1.6)
Missing	66 (2.5)	2,260 (7.5)
Health Literacy (n (%) adequate)	2168 (83.2%)	–
Subjective Numeracy (n (%) adequate)	2196 (84.3%)	–
Number of Visits in past 2 years, median (IQR)	17.0 (8.0–34.0)	8.0 (4.0–18.0)
Hospitalized in past 2 years, n (%)	1304 (50.1)	10,352 (34.6)
Clinical Variables median (IQR)		
Body Mass Index, (kg/m²)	29.5 (26.2–33.6)	28.8 (25.5–32.9)
N with measure (%)	2491 (95.6)	20,008 (66.8)
Systolic Blood Pressure, (mmHg)	126 (119–135)	127 (118–137)
Diastolic Blood Pressure, (mmHg)	70 (65–76)	70 (64–77)
N with measure (%)	2573 (98.8)	23,984 (80.0)
HbA1c, (%)	6.3 (5.7–7.2)	6.2 (5.6–7.3)
N with measure (%)	1051 (40.3)	7201 (24.0)
Low Density Lipoprotein, mg/dL	81.0 (64.0–100.0)	83.0 (66.0–106.0)
N with measure (%)	1706 (65.5)	10,381 (34.7)
Co-morbidities^b, N (%)		
Hypertension	1988 (76.3)	14,900 (49.8)
Type 2 Diabetes	989 (38.0)	9942 (33.2)
Hyperlipidemia	2051 (78.7)	14,108 (47.1)
Congestive Heart Failure	589 (22.6)	4417 (14.7)
Depression	164 (6.3)	1043 (3.5)
Smoking	208 (8.0)	2130 (7.1)
Medications, N (%)		
ACEI/ARBs	1955 (75.0)	16,753 (55.9)
Beta Blockers	1998 (76.7)	17,724 (59.2)
Anticoagulants	1036 (39.8)	8716 (29.1)
Non Aspirin Antiplatelet Agents	1072 (41.2)	10,187 (34.0)
Aspirin	2332 (89.5)	20,792 (69.4)
Statins	2331 (89.5)	20,415 (68.2)

IQR = Interquartile range; ACEI = angiotensin-converting enzyme inhibitor; ARB = angiotensin-receptor blocker.

^a Non-participants include those who were approached and declined and those who were not approached.

^b Each co-morbid condition was defined as present if there was 1 specified inpatient code, 2 specified outpatient codes separated by 30 days, 1 specified procedure code, or prescription for a medication defining that comorbid condition in the 730 days (2 years) prior to the search query.

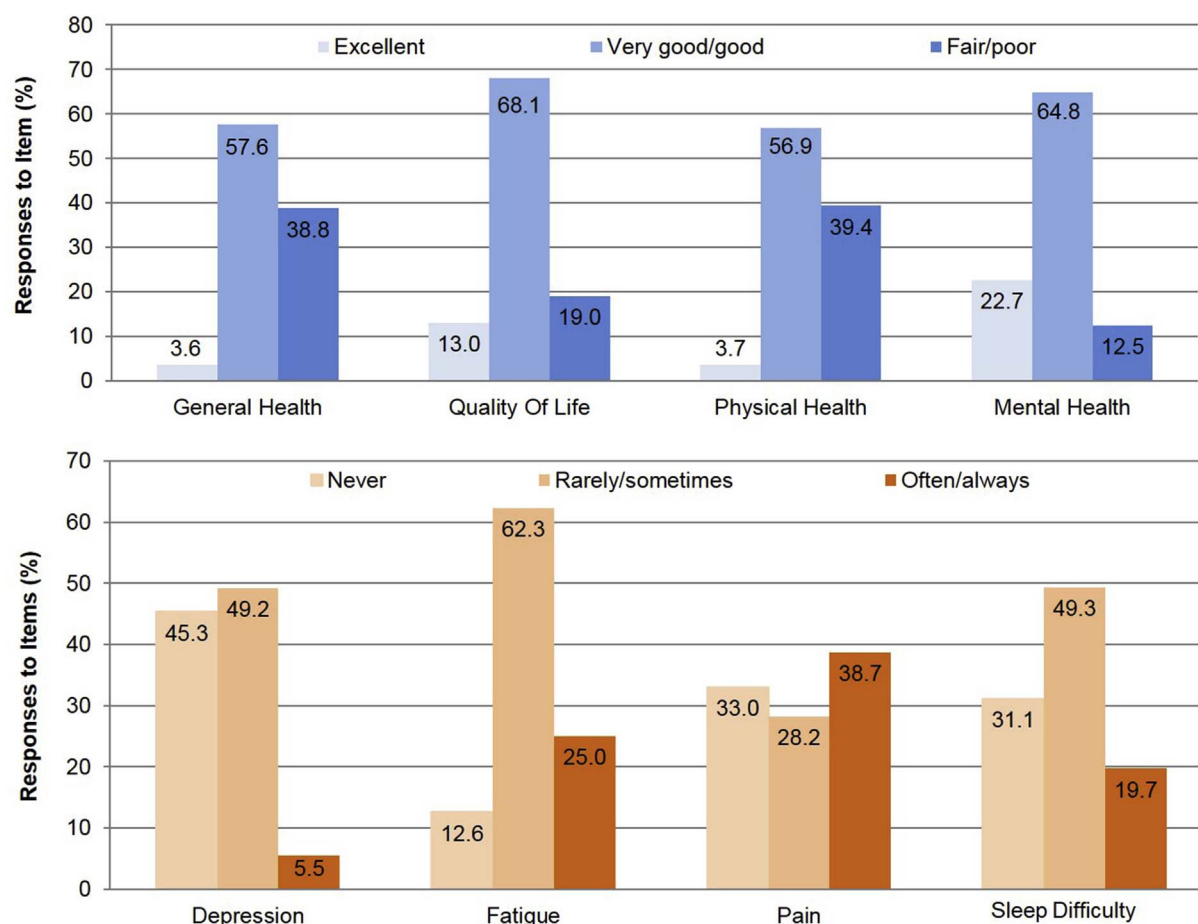


Fig. 2. Responses by participants on PROMIS domains in general health, depression fatigue, pain and sleep difficulty.

and asked if they would consider participating if it was “okay with your doctor.” Most respondents said that they would consider participating ($n = 1223$ [63.2%]) and 15.6% reported that they weren't sure ($n = 302$). When asked about their willingness to change their aspirin dose, most participants reported that they would be willing to change ($n = 1251$ [64.6%]).

4. Discussion

A major focus of PCORI has been identifying strategies to accomplish effective and efficient recruitment of large populations of patients for participation in research, including pragmatic clinical trials. The identification of willing patients from clinical practice can be a challenge. Use of administrative datasets for common chronic conditions is becoming an increasingly popular strategy. In this study, our team was successful in approaching and engaging a large number of patients with CHD, many of whom were older and had a high comorbidity burden. The results of this survey provide insights into the “personome” [28] of a large community based cohort with CHD – that is, their socio-demographic, lifestyle, psychosocial, and functional characteristics. These patients were receptive to future research participation and a sizable majority of patients reported that they were potentially willing to participate in a pragmatic trial to test aspirin dosing effectiveness.

Clinical questions posed in effectiveness research and facing patients and clinicians can be difficult to answer. Traditional clinical trials are often criticized for being selective, while observational cohorts are limited in their ability to draw causality because of issues such as confounding by disease severity and indication. Identification and inclusion of a study population which is representative of the typical patients who seek care for a condition can help strengthen the

generalizability of lessons derived from the results of future research studies.

By recruiting from a pool of patients who recently utilized our health care system, we achieved a high participation rate and enriched the sample with patients who often had multiple documented co-morbidities. More than 75% of patients had documented hypertension, almost 26% had a history of heart failure and 50% of the patient sample had been hospitalized in the past two years. Patients are often recipients of care based on extrapolation of clinical trial results or on the basis of expert opinion. PCORI seeks to recruit and evaluate important clinical questions among the general population, reflective of people who seek care in real world practice and who are often limited in their recruitment for traditional clinical trials because of their age, gender or co-morbidities [29,30]. The proportion of women with CHD recruited into this cohort was approximately 32%. Landmark studies of CHD including COURAGE [31] and the BARI 2D trial [32] only enrolled 15% and 29% of females, respectively. In contrast, only 7% of patients self-identified as black, which is less than the general population and likely reflective of the population seeking care within our medical system.

Another important finding from this project, is that we observed differences in self-reported health status by age, sex and employment status. Despite the relatively limited physical health of our participants, the majority reported good quality of life. It has been described that patients with chronic diseases and older patients often use social comparison to compare themselves with others in relation to an outcome. Specifically, downward comparisons (comparing one's situation to another person's situation that is worse than one's own) and response shifting is a common strategy used in accommodating to one's chronic illness [33]. Given the cross sectional nature of this study it is difficult to ascertain if quality of life among respondents changed as a method of

Table 3
Item response by patient sociodemographics.

Characteristics	Global General Health N = 2590 respondents		Global Physical Health N = 2586 respondents		Quality of Life N = 2595 respondents	
	Excellent/Very Good/Good N = 1584	Fair/Poor N = 1006	Excellent/Very Good/Good N = 1567	Fair/Poor N = 1019	Excellent/Very Good/Good N = 2103	Fair/Poor N = 492
Age, years (%)						
< 44 (N = 46)	28 (60.9)	18 (39.1)	28 (60.9)	18 (39.1)	37 (80.4)	9 (19.6)
45–54 (N = 210)	120 (57.1)	90 (42.9)	108 (52.2)	99 (47.8)	150 (71.8)	59 (28.2)
55–64 (N = 610)	342 (56.6)	262 (43.4)	329 (54.3)	277 (45.7)	473 (77.5)	134 (22.1)
65–74 (N = 955)	606 (63.7)	346 (36.3)	611 (64.4)	338 (35.6)	795 (83.4)	158 (16.6)
75+ (N = 784)	488 (62.7)	290 (37.3)	491 (63.1)	287 (36.9)	648 (83.1)	132 (16.9)
Sex (%)						
Male (N = 1780)	1149 (65.0)	620 (35.0)	1131 (64.0)	636 (36.0)	1469 (82.8)	306 (17.2)
Female (N = 825)	435 (53.0)	386 (47.0)	436 (53.2)	383 (46.8)	634 (77.3)	186 (22.7)
Race (%)						
White (N = 2310)	1423 (62.0)	874 (38.0)	1406 (61.3)	889 (38.7)	1870 (81.3)	430 (18.7)
Black (N = 191)	92 (48.2)	99 (51.8)	92 (48.4)	98 (51.6)	149 (78.0)	42 (22.0)
Other (N = 38)	26 (68.4)	12 (31.6)	26 (70.3)	11 (29.7)	32 (84.2)	6 (15.8)
Missing (N = 66)	43 (67.2)	21 (32.8)	43 (67.2)	21 (32.8)	52 (78.8)	14 (21.2)
Employment^a (%)						
Employed (N = 705)	531 (75.9)	169 (24.1)	510 (73.0)	189 (27.0)	639 (90.6)	66 (9.4)
Retired (N = 1411)	899 (64.0)	505 (36.0)	923 (65.8)	480 (34.2)	1184 (84.3)	220 (15.7)
Unable to work (N = 401)	105 (26.3)	295 (73.8)	87 (21.9)	310 (78.1)	210 (52.8)	188 (47.2)
Unemployed/Homemaker/Student (N = 73)	46 (63.9)	26 (36.1)	43 (58.9)	30 (41.1)	61 (83.6)	12 (16.4)
Number of Visits in past 2 years (%)						
≤ 8 (N = 1355)	827 (61.4)	519 (38.6)	820 (61.0)	525 (39.0)	1086 (80.4)	264 (19.6)
> 8 (N = 1250)	757 (60.9)	487 (39.1)	747 (60.2)	494 (39.8)	1017 (81.7)	228 (18.3)
Hospitalized in past 2 years (%)						
Yes (N = 1304)	722 (55.7)	575 (44.3)	714 (55.2)	579 (44.8)	1019 (78.4)	280 (21.6)
No (N = 1301)	862 (66.7)	431 (33.3)	853 (66.0)	440 (34.0)	1084 (83.7)	212 (16.4)

^a Row percent shown in table may not add to 100% because some participants have missing responses. There were 15 participants missing employment status and missing response to global health, physical health, or quality of life items.

accommodating to CHD.

We report a number of lessons learned that can influence future studies and patient recruitment. First, patients can effectively be recruited for research study participation using large administrative data that targets a condition of interest. When developing the phenotype, we tried to include general data elements that could be “exported” to other CDRNs and utilized the PCORI common data model. Second, it is important to update vital status in the recruitment database. As we attempted to contact patients who had been identified by the phenotype, we learned that incorporation of social security death index into our data warehouse has a “lag” which resulted in sending invitation letters to a number of participants who had recently died. Third, despite chronic illnesses and differences in health status, our patients also indicated a willingness to participate in future research. In anticipation of the ADAPTABLE study, PCORI’s first attempt to leverage nationwide clinical data research network infrastructure in the performance of a large pragmatic clinical trial, our CHD cohort has identified a large population who seem open and amenable to the idea of participating in either survey based research or research that involves changing their aspirin dose if it were “ok with their doctor.”

While our study has many strengths, there are also a number of limitations to consider. First, because of the available staffing resources, we reached out only to a convenience sample of patients who were followed in cardiology or primary care practices. Thus only 17% of the eligible pool of 32,550 patients was recruited. The trends and associations observed in our results may not be representative of the overall group of patients with CHD in the Vanderbilt health system, perhaps because of chance or because of unintended selection bias. Furthermore, we did not compare those who were true non-responders to the survey from those who were not approached; thus, there could also be differences within these groups that are difficult to separate. Second, the lessons that we have learned through the conduct of our phenotype development and survey deployment are derived from a single center and may not generalize to other regions of the country.

Third, while the preponderance of respondents did report a positive attitude to research and willingness to participate in a clinical trial of aspirin dosing, we are currently enrolling in this trial and find that actual willingness to participate may be lower. This may be due to the concept of participation in a hypothetical versus actual study.

5. Conclusion

In our CHD cohort study, despite participants’ self-report of poorer physical health, they nonetheless reported very high levels of perceived health competence and quality of life. There were no differences in general health status by race and previous healthcare utilization, while there were differences by age, gender and employment status. Given that many patients report that they would be interested in future research participation, our results provide a guide and platform from which future research can work to recruit a mix of participants who are representative of the patients who seek care in our healthcare system.

Ethics approval and consent to participate

The institutional review board of Vanderbilt University approved this study. All surveyed participants provided informed consent and were offered \$10 for survey completion.

Consent for publication

Not applicable.

Availability of data

The datasets used and analyzed during the current study are available from the corresponding author with a request in writing.

Competing interests/Disclosures

None.

Funding

This project was funded by research grants from the Mid-South Clinical Data Research Network - Patient Centered Outcomes Research Institute (PCORI) (R-1306-04869 and 1501-26498), and in part by the National Center for Advancing Translational Sciences (2 UL1 TR000445-06). Dr. Roumie was supported in part by Center for Diabetes Translation Research P30DK092986. Dr. Bachmann was supported by grant number K12HS022990 from the Agency for Healthcare Research and Quality. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

Author contributions

Design (Roumie, Muñoz, Bachmann, Rothman, Kripalani).

Conduct/data collection (Stahl, Case, Leak, Kripalani).

Analysis (Roumie, Patel).

Drafting manuscript (Roumie).

Critical revision of manuscript (Roumie, Patel, Muñoz, Bachmann, Stahl, Case, Leak, Rothman, Kripalani).

Dr. Roumie and Mrs. Patel had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Acknowledgements

We would like to thank the participants of this study for their contributions.

References

- [1] I. Ford, J. Norrie, Pragmatic trials, *N. Engl. J. Med.* 375 (5) (2016) 454–463.
- [2] S.L. Cutrona, S. Toh, A. Iyer, et al., Validation of acute myocardial infarction in the Food and Drug Administration's Mini-Sentinel program, *Pharmacoepidemiol. Drug Saf.* 22 (1) (2013) 40–54.
- [3] S. Toh, M.A. Baker, J.S. Brown, C. Kornegay, R. Platt, Rapid assessment of cardiovascular risk among users of smoking cessation drugs within the US Food and Drug Administration's Mini-Sentinel program, *JAMA Intern. Med.* 173 (9) (2013) 817–819.
- [4] G.L. Grunau, S. Sheps, E.M. Goldner, P.A. Ratner, Specific comorbidity risk adjustment was a better predictor of 5-year acute myocardial infarction mortality than general methods, *J. Clin. Epidemiol.* 59 (3) (2006) 274–280.
- [5] H. Quan, V. Sundararajan, P. Halfon, et al., Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data, *Med. Care* 43 (11) (2005) 1130–1139.
- [6] ADAPTABLE trial, <http://theaspirinstudy.org/>, (2016).
- [7] R. Fleurence, J.V. Selby, K. Odom-Walker, et al., How the Patient-Centered Outcomes Research Institute is engaging patients and others in shaping its research agenda, *Health Aff.* 32 (2) (2013) 393–400.
- [8] J.V. Selby, A.C. Beal, L. Frank, The Patient-Centered Outcomes Research Institute (PCORI) national priorities for research and initial research agenda, *Jama* 307 (15) (2012) 1583–1584.
- [9] J.V. Selby, S.H. Lipstein, PCORI at 3 years—progress, lessons, and plans, *N. Engl. J. Med.* 370 (7) (2014) 592–595.
- [10] I. Danciu, J.D. Cowan, M. Basford, et al., Secondary use of clinical data: the Vanderbilt approach, *J. Biomed. Inf.* 52 (2014) 28–35.
- [11] International Classification of Diseases, Ninth Revision, Clinical Modification, Public Health Service, US Dept of Health and Human Services, Washington, DC, 1988.
- [12] C. Roumie, J. Shirey-Rice, S. Kripalani, MidSouth CDRN - coronary heart disease algorithm, PheKB (2014), <https://phekb.org/phenotype/midsouth-cdrn-coronary-heart-disease-algorithm>.
- [13] C.Y. Osborn, S.T. Rosenbloom, S.P. Stenner, et al., MyHealthAtVanderbilt: policies and procedures governing patient portal functionality, *J. Am. Med. Inf. Assoc.* 18 (Suppl 1) (2011) i18–23.
- [14] P.A. Harris, R. Taylor, R. Thielke, J. Payne, N. Gonzalez, J.G. Conde, Research Electronic Data Capture (REDCap) - a metadata-driven methodology and workflow process for providing translational research informatics support, *J. Biomed. Inf.* 42 (2) (2009) 377–381.
- [15] R.D. Hays, J.B. Bjorner, D.A. Revicki, K.L. Spritzer, D. Cella, Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items, *Qual. Life Res.* 18 (7) (2009) 873–880.
- [16] M.S. Smith, K.A. Wallston, C.A. Smith, The development and validation of the perceived health competence scale, *Health Educ. Res.* 10 (1) (1995) 51–64.
- [17] M. Hagströmer, P. Oja, M. Sjöström, The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity, *Publ. Health Nutr.* 9 (06) (2006) 755–762.
- [18] CDC, CfDCaP, (2011).
- [19] J.B. Saunders, O.G. Aasland, T.F. Babor, J.R. De La Fuente, M. Grant, Development of the Alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful Alcohol consumption-II, *Addiction* 88 (6) (1993) 791–804.
- [20] J.N. Variyam, B.-H. Lin, NHANES: development of a flexible consumer behavior survey module, *Faseb. J.* 21 (5) (2007) A53.
- [21] P.M. Guenther, S.I. Kirkpatrick, J. Reedy, et al., The Healthy Eating Index-2010 is a valid and reliable measure of diet quality according to the 2010 Dietary Guidelines for Americans, *J. Nutr.* 144 (3) (2014 Mar), <http://dx.doi.org/10.3945/jn.113.183079>.
- [22] T.K. Adjoian, M.J. Firestone, D. Eisenhower, S.Y. Stella, Validation of self-rated overall diet quality by Healthy Eating Index-2010 score among New York City adults, *Prevent. Med.* 2016 (3) (2013) 127–131.
- [23] S.L. Botman, Design and Estimation for the National Health Interview Survey, 1995–2004, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics, 2000.
- [24] L.D. Chew, K.A. Bradley, E.J. Boyko, Brief questions to identify patients with inadequate health literacy, *Fam. Med.* 36 (8) (2004) 588–594.
- [25] L.D. Chew, J.M. Griffin, M.R. Partin, et al., Validation of screening questions for limited health literacy in a large VA outpatient population, *J. Gen. Intern. Med.* 23 (5) (2008) 561–566.
- [26] A. Fagerlin, B.J. Zikmund-Fisher, P.A. Ubel, A. Jankovic, H.A. Derry, D.M. Smith, Measuring numeracy without a math test: development of the Subjective Numeracy Scale, *Med. Decis. Making* 27 (5) (2007) 672–680.
- [27] C.D. McNaughton, K.L. Cavanaugh, S. Kripalani, R.L. Rothman, K.A. Wallston, Validation of a short, 3-item version of the subjective numeracy scale, *Med. Decis. Making* 35 (8) (2015) 932–936.
- [28] R.C. Ziegelstein, Personomics, *JAMA Intern. Med.* 175 (6) (2015) 888–889.
- [29] I. Ridda, R. Lindley, R.C. MacIntyre, The challenges of clinical trials in the exclusion zone: the case of the frail elderly, *Australas. J. Ageing* 27 (2) (2008) 61–66.
- [30] M.D. Witham, M.E. McMurdo, How to get older people included in clinical studies, *Drugs Aging* 24 (3) (2007) 187–196.
- [31] W.E. Boden, R.A. O'Rourke, K.K. Teo, et al., Optimal medical therapy with or without PCI for stable coronary disease, *N. Engl. J. Med.* 356 (15) (2007) 1503–1516.
- [32] R.L. Frye, P. August, M.M. Brooks, et al., A randomized trial of therapies for type 2 diabetes and coronary artery disease, *N. Engl. J. Med.* 360 (24) (2009) 2503–2515.
- [33] M.A. Sprangers, C.E. Schwartz, Integrating response shift into health-related quality of life research: a theoretical model, *Soc. Sci. Med.* 48 (11) (1999) 1507–1515.