Am J Transl Res 2012;4(4):458-470
www.ajtr.org /ISSN:1943-8141/AJTR1209002

Original Article

The Measurement to Understand Reclassification of Disease of Cabarrus/Kannapolis (MURDOCK) Study Community Registry and Biorepository

Sayanti Bhattacharya[1], Ashley A Dunham[1], Melissa A Cornish[1], Victoria A Christian[1], Geoffrey S Ginsburg[2], Jessica D Tenenbaum[1], Meredith L Nahm[3], Marie Lynn Miranda[4], Robert M Califf[3], Rowena J Dolor[5*], L Kristin Newby[5]

*[1]Duke Translational Research Institute, Duke University, Durham, NC, USA; [2]Institute for Genome Science and Policy, Duke University, Durham, NC, USA; [3]Duke Translational Medicine Institute, Duke University, Durham, NC, USA; [4]University of Michigan, Ann Arbor, MI, USA; [5]Duke Clinical Research Institute, Duke University Medical Center, Durham, NC, USA. *Present affiliation: Vanderbilt University, Nashville, TN, USA.*

Received September 13, 2012; accepted October 5, 2012; Epub October 10, 2012; Published October 30, 2012

Abstract: Current understanding of chronic diseases is based on crude clinical characterization, imaging studies, and laboratory testing that has evolved over decades. The Measurement to Understand Reclassification of Disease of Cabarrus/Kannapolis (MURDOCK) Study is a multi-tiered, longitudinal study designed to enable classification of chronic diseases using clinically annotated biospecimen collections, -omic technologies, electronic health records, and standard epidemiological methods. We expect that detailed molecular classification will improve mechanistic understanding of chronic diseases, augmenting discovery and testing of new treatments, and allowing refined selection of prevention and treatment strategies. The MURDOCK Study Community Registry and Biorepository will serve as a bridge for validation of initial exploratory studies, a platform for future prospective studies in targeted populations, and a resource of both data (analytical and clinical) and samples for cross-registry meta-analyses and comparative population studies. Participation of local health care providers and the Cabarrus County/Kannapolis, NC, community will facilitate future medical research and provide the opportunity to educate and inform the public about genomic research, actively engaging them in shaping the future of medical discovery and treatment of chronic diseases. We present the rationale and study design for the MURDOCK Community Registry and Biorepository and baseline characteristics of the first 6000 participants.

Keywords: Disease reclassification, community registry, biorepository

Introduction

The Framingham Heart Study first quantitatively characterized universally recognized risk factors for heart disease [1]. However, when initiated in 1948, the study could not anticipate the clinical or mechanistic diversity of these seemingly discrete risk factors, the range of influence of each on health and illness, or the complexity of clinical responses to attempts to modulate them. In fact, the Framingham clinical risk model had only modest ability to discriminate who would or would not suffer a future myocardial infarction or die from heart disease (c-index in men, 0.69; c-index in women, 0.72) [2, 3], highlighting the need for more refined classification of disease state and risk using molecular profiling. Subsequently, through the Framingham Offspring Study and Gen III cohorts, investigators have advanced linking population registries with biospecimen collection to foster molecular epidemiological research (http://www.framinghamheartstudy.org/).

Many population-based biorepositories of varying size have been created [4-12], each with its strengths and limitations. One of the largest is the UK Biobank, an effort to enroll 500,000 population representative participants in the United Kingdom [13, 14]. However, it is restricted by age (40-69 years) and is representative of a unique health system [13]. The Fram-

ingham cohorts have limited racial and ethnic diversity [12], while several other registries are focused primarily on cardiovascular disease [9-11].

The MURDOCK Community Registry and Biorepository will complement these ongoing initiatives. Its location in the fast-growing southeastern United States will bolster future studies of disease heterogeneity, prevention, and treatment in diverse populations. Its absence of age restrictions will allow insight into healthy aging and disease progression across age groups, selective studies of age group-specific diseases, and studies prior to onset of chronic illnesses plaguing older populations. The study is also distinguished by its methodological enhancements, including geospatial mapping of enrollees for social and environmental correlations with molecular profiles, clinical disease states, and treatment; collection of diverse sample types, patient-reported health status and lifestyle information, and yearly follow-up; and access to a population for which electronic health records (EHRs) are commonly available. Finally, MURDOCK methodology includes recruitment of a representative sample nested within volunteer recruitment.

The MURDOCK Community Registry and Biorepository is designed so that clinical data and results of -omics analyses can be integrated with other clinical and genomic databases, creating opportunities for large-scale meta-epidemiological-genomic investigations necessary to advance global healthcare. For example, similar to the electronic Medical Records & Genomics Network (eMERGE) concept [15], using common survey elements, accepted data standards, and access to EHRs, the MURDOCK Community Registry and Biorepository could contribute to consortia models without requiring additional recruitment or sample collection.

Methods/Design

The MURDOCK Study was funded by a gift to Duke University from the David H. Murdock Institute for Business and Culture. The overarching aims and infrastructure of the MURDOCK Study have been described previously [16]. The MURDOCK Community Registry and Biorepository is a component of this study, which is organized into "Horizons" (**Figure 1**). The MURDOCK Community Registry and Biorepository (Horizon 1.5) provides a bridge between initial molecular discovery studies in Horizon 1, prospective studies and validation studies in Horizon 2, and broad-based, collaborative meta-epidemiological-genomic studies envisioned for Horizon 3. Creating the registry and biorepository in parallel with Horizon 1 was intended to reduce the time and expense required to establish longitudinal cohorts with archived clinical data, biospecimens, and accrued follow-up time that could be used in validation studies and new prospective studies.

The MURDOCK Community Registry and Biorepository coordinating center, in Kannapolis, NC, sits adjacent to the NC Research Campus and Core Laboratory, home to -omics technology platforms and advanced imaging and informatics capabilities. Biospecimens from registry participants are stored locally at the LabCorp Biobank, a 40,000-square foot facility with significant sample storage capacity.

Participants

The MURDOCK Community Registry and Biorepository will include approximately 50,000 adult residents of Cabarrus County and/or Kannapolis, NC, representing approximately one-third of the local adult population. Participants must (1) be ≥18 years old and residents of Cabarrus County and/or Kannapolis, NC (defined by zip codes) for ≥6 months per year, (2) provide written informed consent (or consent by proxy from a legal guardian/caregiver), and (3) participate in all five components listed in **Box 1**. There are no exclusion criteria. For specific studies, populations may be recruited from surrounding areas and may include children and/or adolescents.

The MURDOCK Community Registry and Biorepository and related ancillary studies are or must be approved by the institutional review boards of both Duke University Medical Center (Durham, NC) and Carolinas HealthCare System (Charlotte, NC).

Recruitment and enrollment strategies

Participant recruitment began in February 2009. Community engagement activities are used to educate the local population about the MURDOCK Study. Community and health care advisory boards and an advisory board of epidemiology experts support the registry and biore-

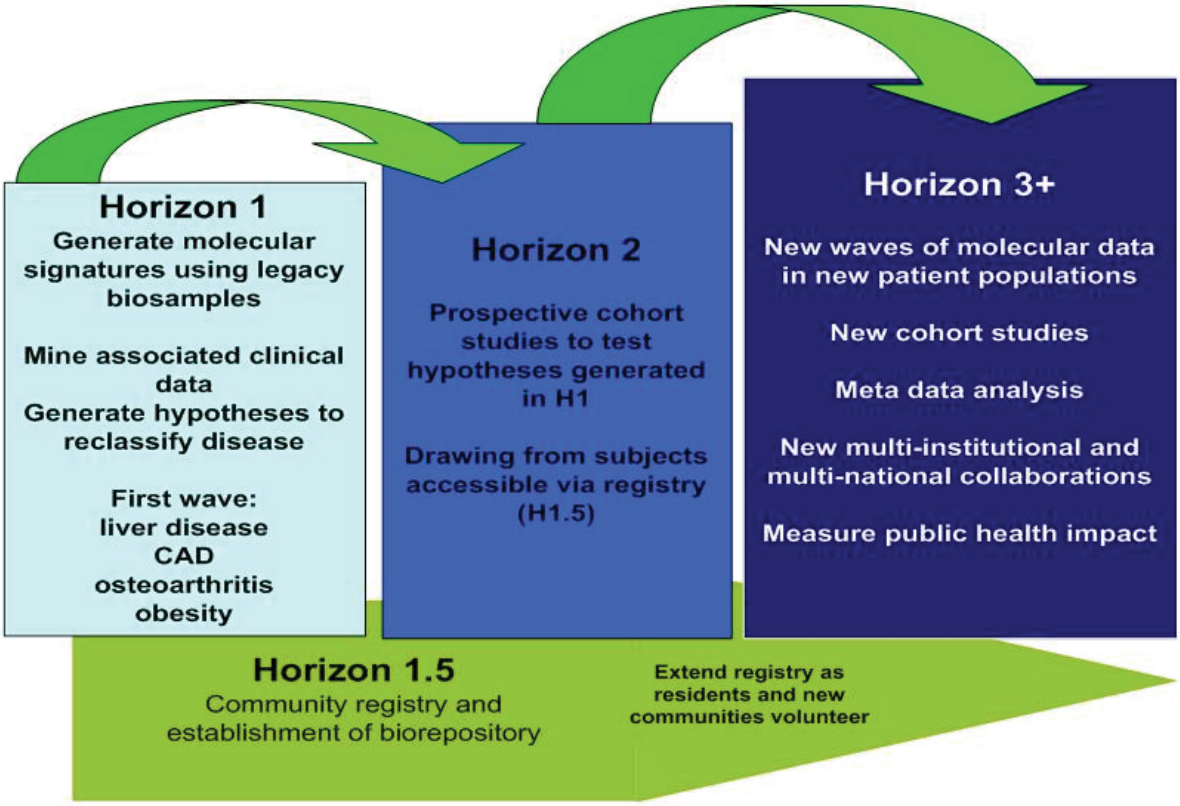

Figure 1. The MURDOCK Study.

Box 1. Core components of MURDOCK Community Registry and Biorepository participation

Use of contact information and data provided in the Participant Registry Questionnaire (and Memory and Cognitive Questionnaire, if applicable) for epidemiological research
Annual follow-up, including direct contact, to update health status and medical events
Release of medical information, including but not limited to demographics, problem lists, medications, social and family history, and results of laboratory and other testing modalities from paper or electronic health records
Blood and urine collection and storage for future biomedical research
Contact up to four times yearly to request participation in additional research studies

pository efforts. To facilitate recruitment of minority populations, study materials have been translated into Spanish and can be administered by bilingual coordinators.

Multiple methods to facilitate enrollment are used [5, 17, 18], and several general procedures govern enrollment. After reviewing materials, participants provide informed consent, complete the Participant Registry Questionnaire, and have blood pressure, heart rate, and waist circumference measured. Blood and urine samples are either obtained on-site or at a MURDOCK Study-designated laboratory within 30 days.

Two general recruitment methods are used: open enrollment of adult volunteers (target n = 35,000) and representative sample enrollment (target n = 15,000). Open enrollment encourages community participation and trust in the growing scientific and epidemiological work rep-

Table 1. Targeted ancillary studies

Disease/demographic	Targeted population
Multiple sclerosis	Multiple sclerosis patients from Mecklenburg County, NC; Greenville, SC; Cabarrus County/Kannapolis, NC
Centenarians	Individuals ≥100 years old; no geographic restrictions
Memory health	Individuals ≥55 years old from Cabarrus County/Kannapolis
Physical functioning	1000 individuals ≥30 years old from Cabarrus County/Kannapolis
Severe acne vulgaris	250 patients ≥12 years old with prior severe acne vulgaris treated with isotretinoin, from Cabarrus County/Kannapolis

resented by the MURDOCK Study. It augments the sample size available for future studies with lower cost and greater efficiency than possible with a similarly sized, completely random population sampling design. Quarterly monitoring ensures that recruitment goals are met and that major demographic imbalances relative to the population of Cabarrus County/Kannapolis do not occur. This nonrandom recruitment phase allows for targeted enrollment and collection of data elements in specific disease areas or demographics that may be of low prevalence and otherwise underrepresented in open enrollment or the planned representative sample (see examples in **Table 1**).

Interested individuals may initiate enrollment via the Web (www.murdock-study.org), by mailing interest cards, or by contacting MURDOCK Study staff by phone or e-mail; they are then enrolled at MURDOCK Study sites, public events, or their homes (for more details, see Supplementary Appendix 1).

The other recruitment method used, random population-based sampling, can eliminate biases arising from failure to reflect population demographics in the study cohort [19]. For example, failure to consider nonrepresentative participation in genome-wide association studies has resulted in false-positive and false-negative associations [20].

A population representative sample of 15,000 individuals (30% of anticipated enrollment) was prospectively determined and represents approximately 11.5% of the adult Cabarrus County/Kannapolis, NC, population. According to the U.S. Census Bureau, Cabarrus County is 71% non-Hispanic whites, 15% blacks, and 9% Hispanics; therefore, this strategy should identify at least 1000 individuals from these key demographic groups. Also, a representative sample of this size should include approximately 1000 participants prior to onset of chronic diseases and individuals within most major chronic disease categories at baseline.

Population representative sample recruitment will use an address-based sampling frame stratified by zip code. Addresses will be cross-matched with other sources to yield telephone numbers and e-mail addresses to increase response rates [21, 22]. All adult individuals living in selected households will be encouraged to enroll. Enrollment procedures, data and sample collection, and longitudinal follow-up will be identical to those for open enrollment participants. Population representative sample recruitment is expected to begin in January 2013 after 6 months of intensive community engagement and will continue for approximately 36 months.

Longitudinal follow-up and retention

All participants are consented for annual follow-up, which is collected via paper form, telephone, or the Web (as of summer 2012). Nonresponders receive two mailed reminders, then three phone calls or e-mails, after which no further attempt is made until the next annual follow-up. Participants who miss two annual follow-ups and for whom there are no EHR data are considered lost-to-follow-up. Physician support is enlisted to contact lost-to-follow-up participants, and Social Security Death Index and annual National Death Index searches are used to update vital status. Currently, the median duration of follow-up is 1.6 years (interquartile range 0.9-2.3), no participants are lost-to-follow-up, and 38 individuals have withdrawn consent for future contact.

To maintain engagement and minimize attrition,

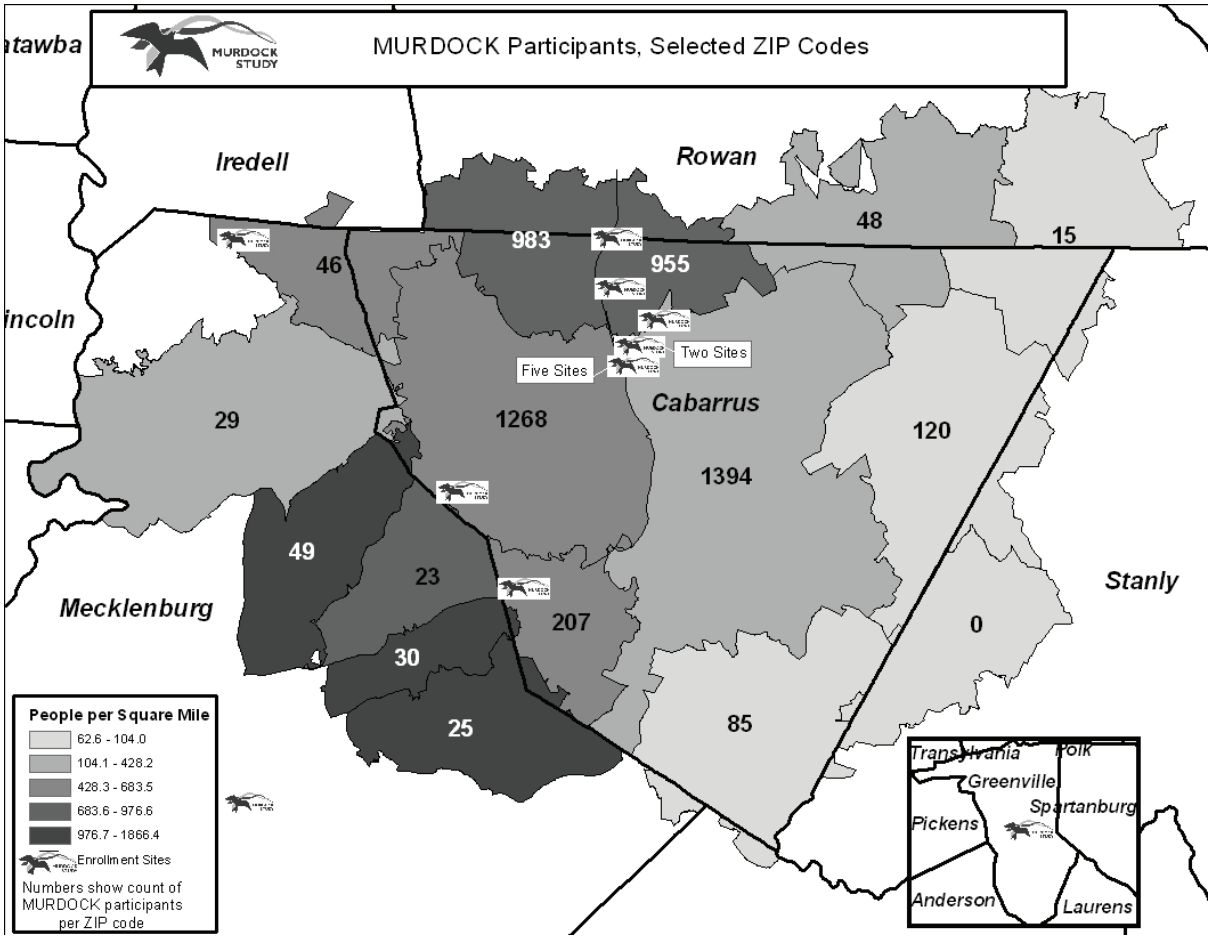

Figure 2. Geocoded enrollment by zip code.

participants receive quarterly paper newsletters, monthly e-mail updates, and birthday and holiday cards. Newsletters and the study website provide updates on study progress, new projects, general results, and contact information. Community events are staged to provide forums for questions and presentations of research.

Data collection

The Participant Registry Questionnaire collects self-reported height, weight, general demographics, patient-reported health status (focused PROMIS questions [23]), clinical illnesses, family medical history, alcohol and tobacco consumption, the Stanford Brief Activity Survey [24], and general dietary questions. MURDOCK staff members measure waist circumference, blood pressure, and heart rate at baseline using standard techniques. Data and contact information are entered into the database via a Web interface and managed by the Duke Translational Medicine Institute/Duke Center for Health Informatics in a secure relational database management system with access controls. Data processing fidelity is assessed annually; the most recent audit revealed 46 errors per 10,000 fields.

Most participant data originate from self-reports. A random sample of registry self-reported data from baseline and follow-up questionnaires will be confirmed by re-interview and compared against medical record information to assess two-way discrepancy rates. The MURDOCK Study takes place in the context of a local Beacon community and a developing state health informatics exchange that may obviate point-to-point data transfers [25, 26]. EHR data transfer is anticipated to begin in late 2012.

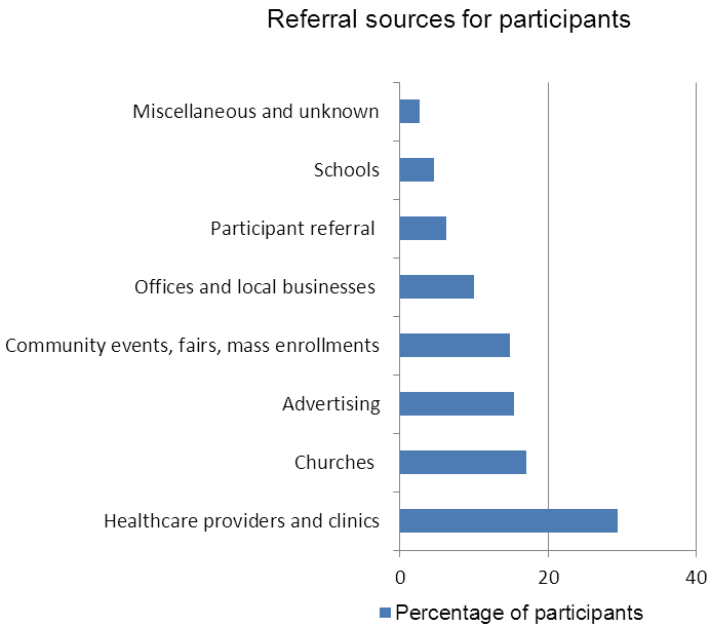

Figure 3. Participant referral sources.

Geospatial coding of participants

Geocoding assigns each participant specific map coordinates using the address provided at enrollment. Other publicly available information (e.g., location of schools, parks and recreation facilities, bus lines, health care providers, fast-food restaurants, and environmental features) can be layered onto the map alongside participant demographic or clinical features. By enabling examination of the geographic distribution of participants by demographic, clinical, and social/economic characteristics, geocoding supports evaluation of social and environmental correlates of health and illnesses as well as molecular profiles. Geocoding is also used to guide recruitment activities and cohort selection for future clinical, population, environmental, and -omics projects. The National Center for Geospatial Medicine has geocoded the addresses of 5430 of the first 5935 participants (91.4%) (**Figure 2**); geocoding will be updated as enrollment occurs.

Collection and storage of biospecimens

All participants provide blood and urine samples that are kept in long-term storage at -80°C for future research. Fifty ml of blood is obtained for DNA (three 2-ml purple top EDTA tubes), plasma (one 6-ml and one 10 -ml purple top EDTA tube), serum (two 10-ml red top tubes), and RNA (three 2.5-ml PAXgene RNA tubes). Urine specimens are aliquoted immediately without further preparation into four 10-ml cryovials. Details of sample collection, processing, and biobanking are provided in Supplementary Appendix 1 online.

Results

Baseline characteristics of the first 6000 participants

There are 15 active recruitment sites in Cabarrus County/Kannapolis, and two sites recruiting multiple sclerosis patients outside this region. Approximately 225 participants are enrolled per month. Baseline characteristics for the first 6000 participants are presented here, but recruitment is ongoing. Over 8000 participants are now enrolled, and >270,000 sample aliquots have been biobanked. Details of the MURDOCK Study, including publications and presentations, can be found at www.murdock-study.org.

Figure 3 shows referral sources for the first 6000 participants. In the third quarter of 2011, >25% of referrals came from word of mouth, confirming increasing community awareness of the study. **Figure 2** displays geocoded enrollment distribution relative to local zip code population densities and locations of MURDOCK recruitment sites, demonstrating pockets of the county that are missed with current recruitment efforts; this information will guide future efforts.

Table 2 shows baseline characteristics of the first 6000 MURDOCK Registry and Biorepository participants compared with 2010 U.S. Census statistics for Cabarrus County, if available. Continuous characteristics are presented as median (interquartile range) and discrete variables as percentages. Descriptive data were generated using SAS software Version 9.2 (SAS Institute, Cary, NC); no statistical testing was performed. Compared with Cabarrus County population statistics, initial participants were older, and men were underrepresented, regardless of race or ethnicity (**Tables 2** and **3**). Recognizing

Table 2. Baseline characteristics of initial 6000 participants (compared with population data when available)

	MURDOCK Study (n=6000)	Cabarrus County population[a] (n=125,533)
Demographics		
Age (years)	54.2 (42.1–65.1)	36.7
Female sex	64.3	50.9
Race/ethnicity		
White	78.3	75.4
African American	12.4	15.3
Hispanic	5.4	9.4
Native American/Alaskan native	0.5	0.4
Asian	0.6	2.0
Hawaiian/Pacific islander	0.05	0.0
More than one race	1.9	2.1
Baseline examination		
Systolic blood pressure (mm Hg)	125 (113–138)	
Diastolic blood pressure (mm Hg)	75 (68–83)	
Heart rate (beats/minute)	71 (64–79)	
Weight (kg)	80 (68–96)	
Height (cm)	168 (160–175)	
Body mass index (kg/m^2)	28 (24–33)	
Waist circumference (cm)	95 (84–106)	
Medical history		
Cigarette smoking		
Current smokers	10.4	15.4
Ever smokers	41.0	32.8
Obesity	27.1	25.0
Diabetes mellitus	15.8	9.0
Hypertension	38.3	29.0
Hypercholesterolemia	39.9	32.0
Thyroid disease	12.3	
Cardiovascular disease		
Coronary artery disease	5.4	3.5
Atrial fibrillation	4.1	
Myocardial infarction or angina	5.2	4.4
Congestive heart failure	2.4	
Implantable defibrillator/pacemaker	1.0	
Neurological/psychiatric disorders		
Alzheimer's disease	0.4	
Depression	23.3	
Other mental illness	3.0	
Stroke	2.9	3.0
Multiple sclerosis	2.8	
Gastrointestinal disease		
Crohn's disease/ulcerative colitis	1.7	
Liver disease	1.7	
Kidney disease	2.0	
Cancer		
Breast	2.6	
Lung	0.4	
Prostate	1.8	
Colon	0.9	
Cervical	1.0	

Melanoma	2.5	
Skin (non-melanoma)	11.5	
Others	3.3	
Pulmonary disease		
Asthma	12.2	12.6
Emphysema/chronic obstructive pulmonary disease	3.6	
Bone and joint disease		
Osteoarthritis	18.7	
Rheumatoid arthritis	7.3	
Other autoimmune disease	4.7	
Osteoporosis/osteopenia	10.3	
Gout	4.8	
Family history		
Diabetes mellitus	41.0	
Coronary artery disease	50.3	
Mental illness	23.5	
Dementia	15.7	
Cancer	47.4	
Socioeconomic		
Education		
<High school diploma	7.2	14.8
High school diploma or equivalent	21.0	30.5
Some college or associate degree	37.2	32.5
Bachelor's degree	21.7	22.2[b]
Master's or higher professional degree	12.9	
Median household income		$54,274
<$10,000	5.1	
$10,000–29,999	16.0	
$30,000–49,999	18.1	
$50,000–69,999	15.8	
$70,000–89,999	12.4	
$≥90,000	23.0	
Marital status		
Married	65.8	55.4[c]
Divorced	10.9	9.7
Separated	2.8	3.1
Widowed	7.1	5.9
Never married	11.4	25.9
Domestic partner	2.1	Not available
Diet (servings/day)		
Fruits/vegetables	2 (1–3)	
Milk/dairy foods	2 (1–2)	
Protein	2 (2–3)	
Sweets	1 (1–2)	
Caffeinated drinks	1 (0–2)	
Sweetened beverages	0 (0–1)	

Continuous variables reported as median (interquartile range); all other values are percentages except median household income for Cabarrus County. [a]Estimates from U.S. Census Bureau Cabarrus County 2010 reports. (http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk) and Behavioral Risk Factor Surveillance System 2010, Centers for Disease Control and Prevention (http://www.epi.state.nc.us/SCHS/brfss/2010/caba/topics.html#pagetop). Age reported as median, but interquartile range not available. [b]Bachelor's degree or higher. [c]Rates for age >15 years.

Table 3. Risk factors, self-reported cardiovascular diseases, and baseline measurements by race and sex

	White		African American		Hispanic		Asian	
	Female (n=2967)	Male (n=1724)	Female (n=495)	Male (n=247)	Female (n=235)	Male (n=91)	Female (n=20)	Male (n=18)
Risk factors								
Hypertension	34.9	41.2	53.9	59.1	19.6	20.9	15.0	5.6
Diabetes	12.9	17.3	25.6	23.6	14.5	13.2	10.0	16.7
Obesity	32.2	20.0	36.0	15.0	15.3	7.7	5.0	11.1
Hypercholesterolemia	40.3	44.9	36.2	44.5	21.3	17.6	10.0	16.7
Smoking	39.4	51.6	31.5	48.6	13.2	46.2	5.0	22.2
Family history of cardiovascular disease	56.5	50.3	42.7	36.4	24.3	23.1	20.0	33.3
Cardiac diseases								
Coronary artery disease	2.9	11.7	2.2	5.3	0.4	2.2	0	0
Atrial fibrillation	3.8	6.4	1.6	4.5	0	0	0	0
Myocardial infarction	3.6	9.0	3.4	6.5	0.4	3.3	0	5.6
Congestive heart failure	2.1	2.7	2.4	7.7	0.4	1.1	0	0
Pacemaker placement	0.6	1.6	0.8	2.5	0.4	0	0	0
Stroke	2.6	3.7	2.4	5.3	0.9	0	0	0
Baseline measurements								
Age (years)	55.4 (43.9–65.4)	57.4 (44.3–67.5)	52.6 (41.6–62.1)	53.9 (44.4–61.7)	38.6 (31.0–46.1)	40.1 (30.0–48.0)	41.1 (33.9–50.4)	42.6 (38.3–45.3)
Height (cm)	165 (160–170)	175 (165–183)	168 (160–175)	175 (165–180)	158 (155–162)	168 (165–173)	159 (152–168)	174 (170–178)
Weight (kg)	73 (63–86)	89 (80–102)	84 (71–104)	98 (82–111)	71 (61–80)	82 (73–91)	55 (51–67)	76 (68–86)
Body mass index (kg/m^2)	27 (24–32)	28 (25–32)	31 (27–38)	30 (26–34)	29 (25–33)	29 (26–32)	23 (21–26)	25 (22–29)
Systolic blood pressure (mm Hg)	121 (110–135)	129 (118–141)	126 (113–141)	135 (122–150)	112 (102–122)	123 (114–132)	113 (101– 130)	124 (112–134)
Diastolic blood pressure (mm Hg)	74 (67–82)	76 (69–84)	79 (71–87)	81 (74–89)	71 (64–79)	75 (67–87)	70 (63–78)	75 (68–80)
Waist circumference (cm)	89 (79–102)	101 (93–112)	97 (87–110)	100 (92–114)	88 (80–95)	95 (88–103)	80 (68–82)	92 (84–100)

Continuous variables are presented as median (interquartile range). All other values are percentages.

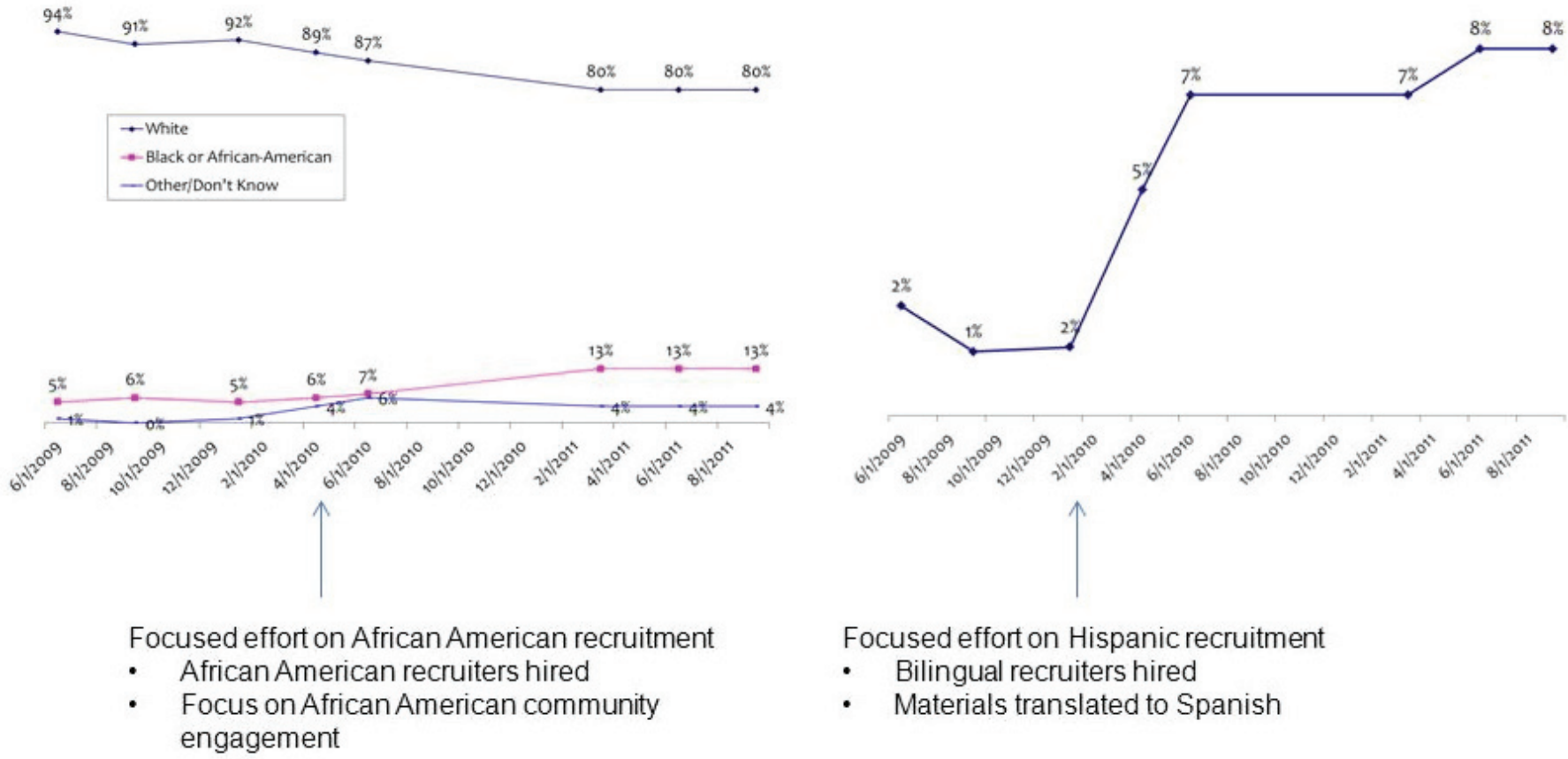

Figure 4. African American and Hispanic recruitment patterns.

this imbalance, recruitment strategies to increase male enrollment were implemented: opening recruitment to and actively targeting men who work for Cabarrus County/Kannapolis public services (e.g., fire and police) and health care providers who work in but live outside of recruitment zip codes; sponsoring sporting events; and adding enrollment sites to medical practices focused primarily on men. Random sample recruitment will also ensure that a subset of the registry is representative of the population.

In large part due to implementation of specific recruitment efforts (**Figure 4**), we had nearly achieved Cabarrus County population representation by race/ethnicity through the first 6000 participants (78.3% white, 12.4% African American, and 5.4% Hispanic).

Relative to the Cabarrus County population, the MURDOCK cohort had fewer individuals who had not completed high school (7.2% vs. 14.8%) and more individuals who had at least some college or higher education (71.8% vs. 54.7%). The estimated median household income for Cabarrus County was $54,274. Only 15.8% of MURDOCK participants reported a median household income in the $50,000-$69,000 range; 39.2% had lower and 35.4% had higher median household incomes. Reflecting an older cohort than the Cabarrus County population, nearly two-thirds were married, >7% were widowed, and only 11.4% were never married.

In regard to health and health behaviors, we considered the American Heart Association's (AHA's) seven factors to improve population health: not smoking, body mass index (BMI) <25 kg/m^2, ≥150 minutes of moderate-intensity physical activity or 75 minutes of vigorous-intensity activity (or a combination) weekly, healthy diet, total cholesterol <200 mg/dL, blood pressure <120/80 mm Hg, and fasting glucose <100 mg/dL [27, 28]. Excluding the healthy diet component for which we do not collect equivalent data elements, only 4.7% of our population met all of the other six goals. However, 99.8% achieved one or more of the goals, and the median number achieved was 4.0 (3.0-4.0). Of diet factors we collect, the median number of servings of fruits and vegetables per day was two (4.5 recommended), but the median number of sugar-sweetened beverages was well within the AHA recommendation.

Twenty-seven percent of participants self-reported obesity, but clinical measurements at enrollment revealed a median BMI of 28 (24-33); thus, nearly half the population had some degree of obesity. Given these observations, it is sobering that 18.2% of participants reported they were inactive, 34.5% did only light-intensity activity (for more details, see Supplementary Appendix 2 online), and few (15%) regularly checked nutrition labels on food.

Regarding major chronic medical illnesses, nearly 19% of participants reported osteoarthritis, and 5% had gout. Also, 23% reported depression, and 3% had other mental illnesses. Over 10% of enrollees were current smokers, and 41% reported past smoking; 3.6% reported chronic obstructive pulmonary disease, and 0.4% had a history of lung cancer.

Diabetes was reported by 16%, 38% had hypertension, 40% hypercholesterolemia, and 5% coronary artery disease (CAD). Except for CAD, which was similar in frequency to the Cabarrus County population, self-reported rates of these chronic diseases in the MURDOCK cohort were greater than the overall population. Among both sexes, African Americans had higher rates of hypertension and diabetes compared with whites, but smoking and atrial fibrillation were more frequent among whites (**Table 3**). Among men, but not women, African Americans more frequently reported heart failure and stroke than the overall Cabarrus County population, but obesity, prior CAD or myocardial infarction, and atrial fibrillation were more common among whites.

Discussion

The MURDOCK Community Registry and Biorepository will support molecular reclassification of chronic diseases through detailed clinical characterization and biospecimen collection for future targeted and exploratory research using multiple -omic technology platforms. Open enrollment coupled with a population representative sample will efficiently recruit a sample size large enough to explore molecular underpinnings of chronic diseases across the spectrum of age, race, and ethnicity, with the ability to generalize to the population of Cabarrus County. The clinical and sample databases will support rapid identification and recruitment of individuals with selected characteristics (clinical and/or

molecular) for participation in future prospective clinical trials and pharmacoepidemiological or pharmacogenomic studies. In conjunction with parallel efforts at Duke and beyond, it is envisioned that the clinical and molecular data generated will be available for large-scale meta-epidemiological-genomic investigations necessary to advance understanding of health and disease. In particular, the ability to combine demographically and environmentally diverse populations will enhance opportunities to explore interrelations of factors such as diet, environment, and race with molecular profiles. These opportunities would not be possible within an isolated registry/biorepository effort and will yield sample sizes adequate to address genomic questions that might be unattainable in a single study.

The MURDOCK Community Registry and Biorepository has limitations. The focused population is relatively stable with limited migration over time. Though useful for this effort and generalizability of results over time, this creates challenges for genetic and genomic research due to relatedness among participants. Relatedness can be managed either with up-front data collection or downstream at the sample analytic phase using genome-wide approaches to identify relatedness genetically, along with generalized estimating equations or other statistical methods to account for clustering of effects. Downstream approaches will be relied on when needed.

Conclusions

The MURDOCK Community Registry and Biorepository are positioned as an enduring scientific resource that can integrate with ongoing efforts in multiple local, national, and international arenas. By maintaining close contact with similar resources elsewhere, the MURDOCK Study will be positioned to make major contributions to collaborative initiatives and meta-epidemiological-genomic research agendas that may emerge. For further information about the MURDOCK Community Registry and Biorepository and inquiries about potential collaborations, investigators are encouraged to visit the study website (www.murdock-study.org) or contact the following individuals: L Kristin Newby, Co-Principal Investigator, MURDOCK Community Registry and Biorepository (Tel: 919-668-8805; E-mail: kristin.newby@duke.edu), or Ashley A Dunham, Project Leader, MURDOCK Community Registry and Biorepository (Tel: 919-257-1789; E-mail: ashley.dunham@duke.edu).

Acknowledgements

This work was supported by a gift to Duke University from the David H. Murdock Institute for Business and Culture and grant 1UL1 RR024128-01 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH) and NIH Roadmap for Medical Research; its contents are solely the responsibility of the authors and do not necessarily represent the official view of NCRR or NIH.

Address correspondence to: Dr. L Kristin Newby, Duke Clinical Research Institute, P.O. Box 17969, Durham, NC 27715-7969. Tel: 919-668-8805; Fax: 919-668-7056; E-mail: kristin.newby@duke.edu

References

[1] Dawber TR, Meadors GF, Moore FE Jr. Epidemiological approaches to heart disease: the Framingham Study. Am J Public Health Nations Health 1951; 41: 279-286.

[2] Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998; 97: 1837-1847.

[3] Ohman EM, Granger CB, Harrington RA, Lee KL. Risk stratification and therapeutic decision making in acute coronary syndromes. JAMA 2000; 284: 876-878.

[4] Drolet BC, Johnson KB. Categorizing the world of registries. J Biomed Inform 2008; 41: 1009-1020.

[5] McCarty CA, Wilke RA, Giampietro PF, Wesbrook SD, Caldwell MD. Marshfield Clinic Personalized Medicine Research Project (PMRP): design, methods and recruitment for a large population-based biobank. Personalized Med 2005; 2: 49-79.

[6] Ginsburg GS, Burke TW, Febbo P. Centralized biorepositories for genetic and genomic research. JAMA 2008; 299: 1359-1361.

[7] Ollier W, Sprosen T, Peakman T. UK Biobank: from concept to reality. Pharmacogenomics 2005; 6: 639-646.

[8] Gerhard GS, Langer RD, Carey DJ, Stewart WF. Electronic medical records in genomic medicine practice and research. In: Ginsburg GS, Willard HF, editors. Essentials of Genomic and Personalized Medicine. Amsterdam, the Netherlands: Elsevier; 2010; pp: 142-150.

[9] The ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and

objectives. Am J Epidemiol 1989; 129: 687-702.
[10] Connolly DC, Oxman HA, Nobrega FT, Kurland LT, Kennedy MA, Elveback LR. Coronary heart disease in residents of Rochester, Minnesota, 1950-1975. I. Background and study design. Mayo Clin Proc 1981; 56: 661-664.
[11] Victor RG, Haley RW, Willett DL, Peshock RM, Vaeth PC, Leonard D, Basit M, Cooper RS, Iannacchione VG, Visscher WA, Staab JM, Hobbs HH; Dallas Heart Study Investigators. The Dallas Heart Study: a population-based probability sample for the multidisciplinary study of ethnic differences in cardiovascular health. Am J Cardiol 2004; 93: 1473-1480.
[12] Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study: design and preliminary data. Prev Med 1975; 4: 518-525.
[13] Barbour V. UK Biobank: a project in search of a protocol? Lancet 2003; 361: 1734-1738.
[14] Petersen A. Securing our genetic health: engendering trust in UK Biobank. Sociol Health Illn 2005; 27: 271-292.
[15] McCarty CA, Chisholm RL, Chute CG, Kullo IJ, Jarvik GP, Larson EB, Li R, Masys DR, Ritchie MD, Roden DM, Struewing JP, Wolf WA; eMERGE Team. The eMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. BMC Med Genomics 2011; 4: 13.
[16] Tenenbaum JD, Christian V, Cornish MA, Dolor R, Dunham AA, Ginsburg GS, Kraus VB, McCarthy JJ, McHutchison JG, Nahm ML, Newby LK, Svetkey LP, Udayakumar K, Califf RM. The MURDOCK study: A long-term initiative for disease reclassification through advanced biomarker discovery and integration with electronic health records. Am J Transl Res 2012; 4: 291-301.
[17] Hays J, Hunt JR, Hubbell FA, Anderson GL, Limacher M, Allen C, Rossouw JE. The Women's Health Initiative recruitment methods and results. Ann Epidemiol 2003; 13: S18-S77.
[18] Brown DR, Fouad MN, Basen-Engquist K, Tortolero-Luna G. Recruitment and retention of minority women in cancer screening, prevention, and treatment trials. Ann Epidemiol 2000; 10: S13-S21.
[19] Lunetta KL. Genetic association studies. Circulation 2008; 118: 96-101.
[20] Pearson TA, Manolio TA. How to interpret a genome-wide association study. JAMA 2008; 299: 1335-1344.
[21] Link MW, Battaglia MP, Frankel MR, Osborn L, Mokdad AH. Address-based versus random-digit-dial surveys: comparison of key health and risk indicators. Am J Epidemiol 2006; 164: 1019-1025.
[22] Link MW, Battaglia MP, Frankel MR, Osborn L, Mokdad AH. A comparison of address-based sampling (ABS) versus random-digit dialing (RDD) for general population surveys. Public Opinion Q 2008; 72: 6-27.
[23] DeWalt DA, Rothrock N, Yount S, Stone AA; PROMIS Cooperative Group. Evaluation of item candidates: the PROMIS qualitative item review. Med Care 2007; 45: S12-S21.
[24] Taylor-Piliae RE, Norton LC, Haskell WL, Mahbouda MH, Fair JM, Iribarren C, Hlatky MA, Go AS, Fortmann SP. Validation of a new brief physical activity survey among men and women aged 60-69 years. Am J Epidemiol 2006; 164: 598-606.
[25] U.S. Department of Health and Human Services, The Office of the National Coordinator for Health Information Technology. Improving health through health information technology. Beacon Community Program website. http://healthit.hhs.gov/portal/server.pt/community/healthit_hhs_gov__onc_beacon_community_program__improving_health_through_health_it/1805. Updated February 21, 2012. Accessed May 29, 2012.
[26] U.S. Department of Health and Human Services, The Office of the National Coordinator for Health Information Technology. State Health Information Exchange Cooperative Agreement Program website. http://healthit.hhs.gov/portal/server.pt/community/healthit_hhs_gov__state_health_information_exchange_program/1488. Updated April 11, 2012. Accessed May 29, 2012.
[27] Sacco RL. The new American Heart Association 2020 goal: achieving ideal cardiovascular health. J Cardiovasc Med (Hagerstown) 2011; 12: 255-257.
[28] Lloyd-Jones DM, Hong Y, Labarthe D, Mozaffarian D, Appel LJ, Van Horn L, Greenlund K, Daniels S, Nichol G, Tomaselli GF, Arnett DK, Fonarow GC, Ho PM, Lauer MS, Masoudi FA, Robertson RM, Roger V, Schwamm LH, Sorlie P, Yancy CW, Rosamond WD; American Heart Association Strategic Planning Task Force and Statistics Committee. Defining and setting national goals for cardiovascular health promotion and disease reduction: the American Heart Association's strategic Impact Goal through 2020 and beyond. Circulation 2010; 121: 586-613.