



Supplementary Table – Summary of Observational & Epidemiological Data Sources				
Acronym	Baseline Year	Purpose		URL
BRFSS	1984	<i>Behavioral Risk Factor Surveillance System (CDC)</i> . 50 state phone survey of health-related risk behaviors, chronic health conditions, & use of preventive services. 400,000 adult interviews annually; largest continually conducted health survey in the world.	Adult phone interviews	http://www.cdc.gov/brfss/
MEPS	1996	<i>Medical Expenditure Panel Survey (AHRQ)</i> . Survey of US families, medical providers, & employers across US. The Medical Provider Component (MPC) of MEPS is a survey of hospitals, physicians, and other medical facilities that provided care to sampled household members in the course of the survey year. Most complete source of data on the cost and use of healthcare and health insurance coverage. MEPS currently has two major components: the Household Component and the Insurance Component .	Computer-assisted personal interviews using laptop	http://meps.ahrq.gov/mepsweb/
NAMCS	1973	<i>National Ambulatory Medical Care Survey (CDC)</i> . National survey designed to meet the need for objective information about the provision and use of ambulatory medical care services in the US. Findings based on sample of visits to non-federal employed office-based physicians primarily engaged in direct patient care. Data recorded by physician or office staff on encounter form.	Office visit physician interviews & data forms	http://www.cdc.gov/nchs/ahcd.htm
NDI	1973	<i>National Death Index (CDC)</i> . Centralized database state vital statistics deaths collected annually including names of the states in which those deaths occurred, the dates of death, and the corresponding death certificate numbers and cause of death.	State vital statistics database	http://www.cdc.gov/nchs/ndi.htm
NHAMCS		<i>The National Hospital Ambulatory Medical Care Survey (CDC)</i> . Designed to collect data on the utilization and provision of ambulatory care services in hospital emergency and outpatient departments and in ambulatory surgery centers. Specially trained interviewers visit facilities prior to their participation in the survey to explain survey procedures, verify eligibility, develop a sampling plan, and train staff in data collection procedures.	Survey instrument is patient record form	http://www.cdc.gov/nchs/ahcd.htm

Supplementary Table – Summary of Observational & Epidemiological Data Sources				
Acronym	Baseline Year	Purpose		URL
NHANES	early 1960s	<i>National Health and Nutrition Examination Survey (CDC)</i> . A program of studies designed to assess the health and nutritional status of adults and children in the US. Combines interviews and physical examinations. Nationally representative sample of about 10,000 persons each year.	Interviews, physical exams & tests	http://www.cdc.gov/nchs/nhanes/about_nhanes.htm
NHCS		<i>National Hospital Care Survey (CDC)</i> . Integrates inpatient, emergency, outpatient, hospital-based & free-standing ambulatory surgery. Provides trends in healthcare & use of healthcare resources, quality & disparities. Integrates inpatient data formerly collected by the National Hospital Discharge Survey (NHDS) with the emergency department (ED), outpatient department (OPD), hospital-based ambulatory surgery location (ASL), and freestanding ambulatory surgery center (ASC) data collected by the National Hospital Ambulatory Medical Care Survey (NHAMCS). The integration of these 2 surveys along with the collection of personal identifiers (protected health information) will allow linking of care provided to the same patient in the ED, OPD, ASC, and inpatient departments.	Participating hospitals upload data	http://www.cdc.gov/nchs/nhcs.htm
NHDS	1965-2010	<i>The National Hospital Discharge Survey (CDC)</i> . Survey conducted annually from 1965-2010, and was a national probability survey designed to provide information on characteristics of inpatients discharged from non-Federal US short-stay hospitals. Now included in the NHCS. NHDS collected data from a sample of approximately 270,000 inpatient records acquired from a national sample of about 500 hospitals. From 2008 to 2010 the sample size was reduced to 239.	Manual medical transcription hospital records & automated electronic	http://www.cdc.gov/nchs/nhds.htm

Supplementary Table – Summary of Observational & Epidemiological Data Sources				
Acronym	Baseline Year	Purpose		URL
NHHCS	1992-2007	<i>The National Home and Hospice Care Survey (CDC)</i> . Nationally representative US sample surveys designed to provide descriptive information on home health and hospice agencies, their staffs, their services, and their patients. Last fielded by NCHS in 2007. In 2012, NCHS initiated the National Study of Long-Term Care Providers (NSLTCP) The total number of agencies that participated in 2007 NHHCS is 1,036, & data are available on 9,416 current home health patients & hospice discharges from these agencies.	Computer-assisted personal interviewing directors or staff	http://www.cdc.gov/nchs/nhhcs.htm
NHIS	1957	<i>National Health Interview Survey (CDC)</i> . Interviewers visit 35,000–40,000 households across the country and collect data about 75,000–100,000 individuals. US Census Bureau collected disease & risk factor prevalence since 1957.	Household members H&Ps & some tests	http://www.cdc.gov/nchs/nhis.htm
NVSS		<i>National Vital Statistics System (CDC)</i> . Oldest inter-governmental data sharing in public health		http://www.cdc.gov/nchs/nvss.htm
NIS	1988	<i>National Inpatient Sample (HCUP)</i> . Survey developed for Healthcare Cost and Utilization Project (HCUP) ; 20% sample of discharges from participating hospitals. National estimates of healthcare utilization, access, charges, quality & outcomes. representing more than 95% U.S. population. The NIS is the largest publicly available all-payer inpatient healthcare database in US yielding national estimates of hospital inpatient stays. Unweighted, it contains data from >7 million hospital stays each year. Weighted, it estimates > 36 million hospitalizations nationally.	Sample of discharge from all community hospitals allows trend over time	https://www.hcup-us.ahrq.gov/nisoverview.jsp
NNHS	1973	<i>National Nursing Home Survey (CDC)</i> . National sample surveys of nursing homes, residents, & staff. Last fielded by NCHS in 2004. In 2012, NCHS initiated the National Study of Long-Term Care Providers (NSLTCP).		http://www.cdc.gov/nchs/nnhs.htm

Supplementary Table – Summary of Observational & Epidemiological Data Sources				
Acronym	Baseline Year	Purpose		URL
NSLTCP	2012	<p><i>National Study of Long-Term Care Providers (CDC)</i>. Biennial initiative to monitor trends in the major sectors of paid, regulated long-term care services providers. Uses data from surveys of residential care communities,  [PDF - 168 KB] and adult day services centers,  [PDF - 152 KB] & administrative data on home health agencies, nursing homes & hospices. NSLTCP replaces NCHS' previous National Nursing Home Survey, National Home & Hospice Care Survey, & National Survey of Residential Care Facilities.</p>	Mail & web questionnaires to directors	http://www.cdc.gov/nchs/nsltcp/about_nsltcp.htm
YRBSS	1991	<p><i>Youth Risk Behavior Surveillance System (CDC)</i>. Monitors 6 health-risk behaviors - injuries; sexual; alcohol/drug; tobacco; diet/obesity; physical activity in 9th–12th grade students. Students complete the self-administered questionnaire during one class period and record their responses on a computer-scannable questionnaire booklet or separate answer sheet. Allows trends over time.</p>	Self-administered questionnaires in classrooms	http://www.cdc.gov/HealthyYouth/yrbs/index.htm
Material was copied – often verbatim – from websites; H&P: history and physicals				

Study	Year	N	Sites	Focus	Age, y	Ethnicity	URL
ARIC[1] NHLBI	1987	15,792	4	<i>The Atherosclerosis Risk in Communities Study.</i> Designed to investigate causes of atherosclerosis & its clinical outcomes, & variation in CVD risk factors, medical care, and disease by race, gender, location, and date.	45-64	25% AA; 75% EA	http://www2.csc.unc.edu/aric/
CARDIA [2] NHLBI	1985	5,115	4	<i>Coronary Artery Risk Development In Young Adults</i> Examines the development of heart disease in black and white adults.	18-30	52% AA; 58% EA	http://www.cardia.dopm.uab.edu/
CHS[3] NHLBI	1989	5,888	4	<i>The Cardiovascular Health Study.</i> Study of risk factors for CVD in adults 65 years or older.	≥65	16% AA; 84% EA	https://chs-nhlbi.org/
DHS [4] Foundat ion*	2001	6,101	1	The <i>Dallas Heart Study.</i> The DHS was designed to identify new social, behavioral, environmental, genetic, protein, and imaging biomarkers that are risk factors for CVD. Data collected in home & clinic. *Donald W. Reynolds Foundation.	18-65	56% AA; 23% EA; 19% HA; 2% other	http://www.utsouthwestern.edu/edumedia/edufiles/research/center_translational_medicine/dallas_heart_study/dhs-study-overview.pdf
FHS[5-7] NHLBI	1948	15,345	1	<i>The Framingham Heart Study.</i> Identifies the epidemiology and longitudinal development of CVD and CVD risk factors; 3 generations and 2 waves of a multiethnic Omni [8] cohort.	19-70	94% EA; 6% AA/HA/Asian	http://www.framinghamheartstudy.org/
GOCAD ON [9] NHLBI	2000	1,214	7	<i>Genetics of Coronary Artery Disease in Alaska Natives.</i> Investigate the genetic determinants of CVD and its risk factors family study Eskimos from several coastal villages in the Norton Sound region of Western Alaska.	≥18	100% Alaskan Eskimo	http://www.ncbi.nlm.nih.gov/pubmed/16050315

Study	Year	N	Sites	Focus	Age, y	Ethnicity	URL
HCHS/ SOL [10,11] NHLBI	2008	16,000	4	<i>The Hispanic Community Health Study / Study of Latinos.</i> Study the prevalence & development of disease, acculturation, & risk factors that play protective or harmful roles in Cuban, Puerto Rican, Mexican, & Central/South American individuals	18-74	100% Hispanics/Latinos	https://www2.csc.unc.edu/hchs/
JHS [12-14] NHLBI	2000	5,301	1	<i>The Jackson Heart Study.</i> Investigate the causes of CVD in African Americans to learn how to best prevent this group of diseases in the future.	≥21	100% AA	https://www.jacksonheartstudy.org/jhsinfo/
MESA [15] NHLBI	2000	6,814	6	<i>The Multi-Ethnic Study of Atherosclerosis.</i> Study of the characteristics of subclinical CVD and the risk factors that predict progression to clinically overt CVD or progression of the subclinical disease.	45-84	28% AA, 22% HA, 12% Chinese Am	http://www.mesa-nhlbi.org/
REGARDS [16] NINDS	2003	30,239	US	The REasons for Geographic and Racial Differences in Stroke project, observational study of risk factors for stroke, which included a telephone interview followed by an in-home physical exam.	≥45	AA & EA	http://www.regardsstudy.org/about
Strong Heart [17] NHLBI	1989	4,549	4	<i>Strong Heart Study.</i> Estimate CVD mortality & morbidity rates & prevalence of known and suspected CVD risk factors in American Indians. The study population consists of 12 tribes in three geographic areas.	45-74	100% Am. Indians	http://strongheart.ouhsc.edu/
WHI [18] NHLBI	1991	161,808	40	<i>The Women's Health Initiative.</i> Long-term national health study focused on strategies for preventing heart disease, breast & colorectal cancer, & osteoporotic fractures in postmenopausal women.	50-79	Women 82% EA/ AA/Asian RCT+ Observational	https://www.whi.org/about/SitePages/About%20WHI.aspx http://www.nhlbi.nih.gov/whi/
http://www.nhlbi.nih.gov/research/resources/epidemiology-studies .							
CVD=Cardiovascular disease; AA=African American; EA=European American; HA=Hispanic American; RCT=randomized controlled trial.							

1. ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol.* 1989;129:687-702.
2. Friedman GD, Cutter GR, Donahue RP, Hughes GH, Hulley SB, Jacobs DR, Jr., Liu K, Savage PJ. CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol.* 1988;41:1105-1116.
3. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, Kuller LH, Manolio TA, Mittelmark MB, Newman A. The Cardiovascular Health Study: design and rationale. *AnnEpidemiol.* 1991;1:263-276.
4. 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 I. 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.
5. Dawber TR, Meadors GF, Moore FE, Jr. Epidemiological approaches to heart disease: the Framingham Study. *American Journal of Public Health & the Nations Health.* 1951;41:279-281.
6. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol.* 1979;110:281-290.
7. Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, D'Agostino RB, Sr., Fox CS, Larson MG, Murabito JM, O'Donnell CJ, Vasan RS, Wolf PA, Levy D. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol.* 2007;165:1328-1335.
8. Quan SF, Howard BV, Iber C, Kiley JP, Nieto FJ, O'Connor GT, Rapoport DM, Redline S, Robbins J, Samet JM, Wahl PW. The Sleep Heart Health Study: design, rationale, and methods. *Sleep.* 1997;20:1077-1085.
9. Howard BV, Devereux RB, Cole SA, Davidson M, Dyke B, Ebbesson SO, Epstein SE, Robinson DR, Jarvis B, Kaufman DJ, Laston S, MacCluer JW, Okin PM, Roman MJ, Romenesko T, Ruotolo G, Swenson M, Wenger CR, Williams-Blangero S, Zhu J, Saccheus C, Fabsitz RR, Robbins DC. A genetic and epidemiologic study of cardiovascular disease in Alaska natives (GOCADAN): design and methods. *Int J Circumpolar Health.* 2005;64:206-221.
10. Sorlie PD, Aviles-Santa LM, Wassertheil-Smoller S, Kaplan RC, Daviglius ML, Giachello AL, Schneiderman N, Raji L, Talavera G, Allison M, Lavange L, Chambless LE, Heiss G. Design and implementation of the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol.* 2010;20:629-641.

11. Lavange LM, Kalsbeek WD, Sorlie PD, Aviles-Santa LM, Kaplan RC, Barnhart J, Liu K, Giachello A, Lee DJ, Ryan J, Criqui MH, Elder JP. Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol*. 2010;20:642-649.
12. Sempos CT, Bild DE, Manolio TA. Overview of the Jackson Heart Study: a study of cardiovascular diseases in African American men and women. *Am J Med Sci*. 1999;317:142-146.
13. Wyatt SB, Diekelmann N, Henderson F, Andrew ME, Billingsley G, Felder SH, Fuqua S, Jackson PB. A community-driven model of research participation: the Jackson Heart Study Participant Recruitment and Retention Study. *Ethnicity & Disease*. 2003;13:438-455.
14. Wilson JG, Rotimi CN, Ekunwe L, Royal CD, Crump ME, Wyatt SB, Steffes MW, Adeyemo A, Zhou J, Taylor HA, Jr., Jaquish C. Study design for genetic analysis in the Jackson Heart Study. *Ethnicity & Disease*. 2005;15:S6-30-37.
15. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, Greenland P, Jacob DR, Jr., Kronmal R, Liu K, Nelson JC, O'Leary D, Saad MF, Shea S, Szklo M, Tracy RP. Multi-ethnic study of atherosclerosis: objectives and design. *AmJEpidemiol*. 2002;156:871-881.
16. Howard VJ, Cushman M, Pulley L, Gomez CR, Go RC, Prineas RJ, Graham A, Moy CS, Howard G. The reasons for geographic and racial differences in stroke study: objectives and design. *Neuroepidemiology*. 2005;25:135-143.
17. Lee ET, Welty TK, Fabsitz R, Cowan LD, Le NA, Oopik AJ, Cucchiara AJ, Savage PJ, Howard BV. The Strong Heart Study. A study of cardiovascular disease in American Indians: design and methods. *Am J Epidemiol*. 1990;132:1141-1155.
18. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Control ClinTrials*. 1998;19:61-109.