The Data Collaborative on Health Systems Performance (DC-HSP): Towards Better Healthcare for Cardiovascular Diseases

Policy Brief

PREPARED BY
HEALTH SYSTEMS INNOVATION LAB
AT HARVARD UNIVERSITY

31 October 2023

This page is intentionally left blank.

About this Policy Brief

This Policy Brief, *The Data Collaborative on Health Systems Performance (DC-HSP): Towards Better Healthcare for Cardiovascular Diseases*, was developed under the guidance of Professor Rifat Atun, Professor of Global Health Systems at Harvard University and Director of the Health Systems Innovation Lab, and implemented by a team consisting of Dr. Caroline A. Bulstra (Research Fellow, Health Systems Innovation Lab, Harvard University), Dr. Che L. Reddy (Associate Director, Health Systems Innovation Lab, Harvard University), Kiana Beheshtian (Research Intern, Health Systems Innovation Lab, Harvard University), Dr. Zachary J. Ward (Associate Director and Research Scientist, Health Systems Innovation Lab, Harvard University) and Dr. Jennifer Yeh (Assistant Professor of Pediatrics, Harvard Medical School, Harvard University) and with the support of Dr. Gabriela Borin (Senior Researcher, Health Systems Innovation Lab, Harvard University), Aila Santos Felix de Almeida (Research Intern, Health Systems Innovation Lab, Harvard University), Karen Parente (Research Intern, Health Systems Innovation Lab, Harvard University), Qassi Gaba (Research Scholar, Health Systems Innovation Lab, Harvard University), Yusuf Sarwar (Research Intern, Health Systems Innovation Lab, Harvard University), Zia Shakir (Research Associate, Health Systems Innovation Lab, Harvard University). The results, ideas, and insights in this Policy Brief are drawn from the activities conducted during the first phase of the DC-HSP project, which kicked off in September 2022. We want to express our gratitude to all the local and international experts who participated in this project. The DC-HSP project is supported financially by Novartis Foundation, Microsoft, and Accenture. The funders had no role in study design, analysis, decision to publish, or report preparation of the research output pertaining to this grant.

About the Health Systems Innovation Lab

The Health Systems Innovation Lab at Harvard University is a global research and training facility specializing in developing high-value health systems through strategic innovations.[1] The lab employs research, educational endeavors, innovative initiatives, and knowledge dissemination to collaborate with students, governments, the private sector, international organizations, and civil society. This collaboration aims to advocate for policies and practices while expediting the spread of advancements in healthcare systems. With a diverse

interdisciplinary team of clinicians and academic leaders, along with an international cohort of experts from various disciplines, the Health Systems Innovation Lab is able to harness wide-ranging expertise to examine the factors influencing cardiometabolic diseases, assess healthcare accessibility, evaluate health systems performance related to cardiovascular diseases, and enhance overall population health.[1]

Suggested Citation

Caroline A. Bulstra, Che L. Reddy, Kiana Beheshtian, Zachary J. Ward, Jennifer Yeh, Gabriela Borin, Aila Santos Felix de Almeida, Karen Parente, Qassi Gaba, Yusuf Sarwar, Zia Shakir, Rifat Atun. *The Data Collaborative on Health Systems Performance (DC-HSP): Towards Better Healthcare for Cardiovascular Diseases.* Health Systems Innovation Lab, Harvard University, 2023, Boston, USA.

Date of publication

31 October 2023

Executive Summary

Cardiovascular diseases are the leading cause of global hospital admissions and premature deaths. The burden of known behavioral and cardiometabolic risk factors for cardiovascular diseases, such as obesity, smoking, hypertension, and diabetes mellitus, is rising, especially in low- and middle-income countries. Yet, many health systems lack context-sensitive health system performance data concerning cardiovascular diseases and targeted policies and practices to address the immense burden of cardiovascular diseases. Further, many countries lack robust healthcare delivery systems and equitable access to high-value health services for cardiovascular diseases. High-value health services are effective and efficient to deliver “value for money”, and equitable and responsive to provide “value for many”. There is a major opportunity to use empirical data to assess health system performance in relation to cardiovascular diseases to inform the design, implementation, and scale-up of innovations (i.e., targeted policies, programs, and interventions), and transition to high-value health systems to achieve for enhanced cardiovascular disease outcomes at the population level.

The Data Collaborative on Health Systems Performance (DC-HSP) was established to contribute to the creation of high-value health systems by leveraging the vast amounts of health determinants and health systems data to analyze and measure health systems performance, particularly concerning cardiovascular diseases. The DC-HSP aims to conduct research that informs policy, practice, and innovations to improve population health outcomes. The DC-HSP consists of two phases. Phase I of the project covers three streams of work. The "Data" stream of Phase I encompasses a systematic approach to identifying, mapping, and processing the available data across all world regions and developing the Survey Harmonization and Estimation Platform for Health and Economic Research Data (SHEPHERD). The “Research” stream leverages the harmonized data for data visualization and analysis to inform policy and practice. Finally, the "Translation" stream focuses on the translation and application of research findings to inform policy and practice on addressing cardiovascular diseases by improving health systems performance through policy briefs and consultative groups.

We highlight the major findings of Phase I the DC-HSP: (I) the creation of a global dataset to analyze health system performance for cardiovascular diseases, (II) the creation of a regional dataset that focuses on analyzing health system performance in Latin America and the Caribbean, and (III) the establishment of the DC-HSP Consultative Group to guide on new evidence generation and application. The global dataset includes on 1,046 unique surveys from 219 different countries across Africa, Asia, Europe, the Middle East, Latin America and the Caribbean, North America, and the Western Pacific, encompassing 6.7 million individuals. Although sociodemographic and behavioral risk factor variables are more widely available across surveys and countries, the lack of comprehensive data on cardiovascular disease incidence, mortality, and health systems indicators for cardiovascular diseases is a major concern. The lack of sufficient data on cardiovascular diseases has far-reaching implications for public health and underscores the urgent need for more robust data collection efforts.

The Latin America and Caribbean regions were prioritized during Phase I, as the burden of cardiovascular diseases is rising disproportionally fast in this world region, and, currently, substantial disparities exist in the burden and access to care for cardiovascular diseases. For the Latin America and the Caribbean region, the DC-HSP harmonized 288 unique surveys from 37 countries between 1990 and 2019, with data from over 1.2 million respondents.

The first DC-HSP Consultative Group was held in São Paulo, Brazil. It comprised experts from various fields to orient the focus and policy relevance of the research in cardiovascular diseases in Brazil. The Consultative Group identified three key themes in which new evidence was urgently needed: (I) value generation across the care continuum, (II) robust ecosystems for digital health innovation, and (III) enabling benchmarking and analytics at national and sub-national levels. The Consultative Group emphasized that research should be highly collaborative, equity-focused, interdisciplinary, and embrace the digital transformation influencing health systems in Brazil, and globally – to produce system-level change over the long term.

In summary, Phase I of the DC-HSP project provided five valuable lessons regarding how data could be harnessed better to deliver high-value health services for cardiovascular diseases and improve outcomes at the population level: (I) Leveraging Insights from Policymakers; (II) Streamlining Data Harmonization; (III) Enhancing Performance Evaluation Methods; (IV) Establishing a Cross-Learning Network, and; (V) Identifying Critical Data Gaps. The DC-HSP will use the outputs and assets produced in Phase I to advance a series of data, research, and translation outputs in Phase II.

Table of Contents

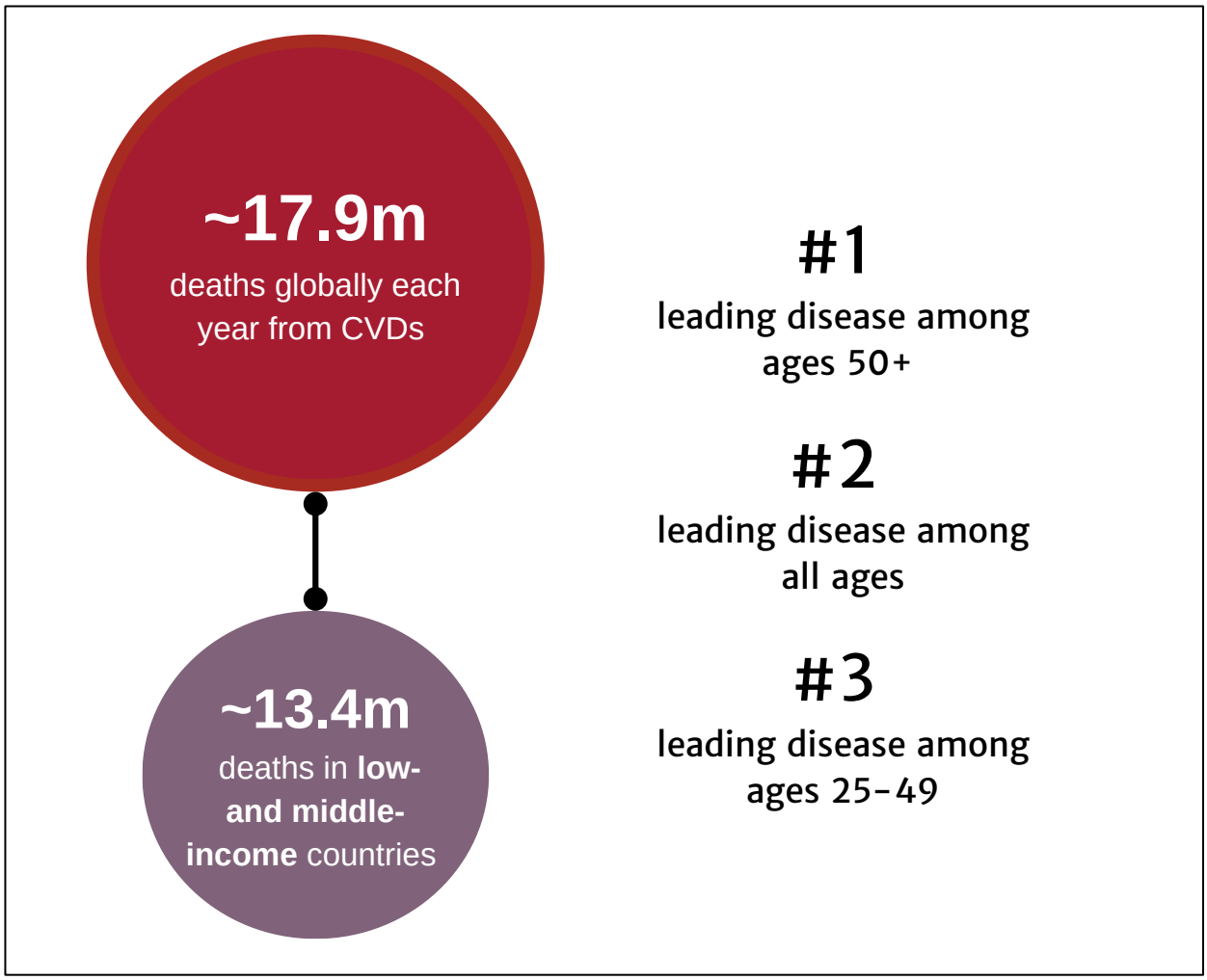
Introduction

This policy brief is organized in four sections. **Section one** discusses what is currently known about cardiovascular diseases from a population health perspective, with a focus on the burden of disease, risk factors, economic impact, and disparities in access and outcomes. **Section two** outlines the major knowledge gaps in improving health systems performance for cardiovascular diseases and how the Data Collaborative on Health Systems Performance (DC-HSP) aims to address these gaps to inform policy, innovation, and practice. **Section three** describes the applied methods for harmonizing and processing the available data used in the DC-HSP project. **Section four** presents the insights emerging from the first phase of the DC-HSP project. The final section, **Section five,** applies these insights to reflect on how policymakers, providers, and payers could contribute to using data more effectively to provide high-value health services and improve health system performance in relation to cardiovascular diseases.

The Rising Burden of Cardiovascular Disease on Health Systems

Cardiovascular diseases, which consist of disorders of the heart and blood vessels, such as ischemic heart disease, cerebrovascular disease, hypertensive heart disease, and other related conditions, are the leading cause of both global hospital admissions and premature deaths in both high-income countries (HIC) and low- and middle-income countries (LMICs).[2] Globally, cardiovascular diseases cause an estimated 17.9 million deaths annually, of which an estimated 13.4 million deaths occur in LMICs.[3] One-third of these deaths affect individuals under 70 years old (**Figure 1**).[2] Three-quarters of cardiovascular disease-related deaths occur in LMICs, with substantial disparities across countries.[5] Addressing behavioral risk factors could prevent a large share of cardiovascular diseases and reduce the burden on health systems, economies, and societies.[6]

Figure 1. Infographic on the importance of cardiovascular diseases (CVD).[3,4,7,8]



Source: Authors

Risk Factors of Cardiovascular Diseases

More than 50% of cardiovascular disease cases can be attributed to modifiable risk factors. Identifying, understanding variation, and managing these risk factors is thus critical to prevention and control efforts.[9,10] The major risk factors for cardiovascular diseases include, among others, high blood pressure, hypercholesterolemia, diabetes mellitus, smoking and secondhand smoke exposure, alcohol consumption, obesity, unhealthy diet, and insufficient physical inactivity.[9,10] High systolic blood pressure remains the primary modifiable risk factor for cardiovascular diseases globally,[7,8] accounting for 10.8 million cardiovascular disease-related deaths overall in 2021, as well as a considerable burden in terms of life years lost and years lived in disability.[7] The population-attributable fraction of high blood pressure

for the individual risk of developing cardiovascular diseases is estimated to be 29% in women and 22% in men, respectively.[11] This means that among 100 individuals with cardiovascular diseases, high blood pressure is responsible for about 29 and 22 cases, respectively. The population-attributable fractions of the most critical risk factors are provided in **Figure 2**. There is substantial regional variation; the regions with the highest rates of cardiovascular disease burden attributable to elevated systolic blood pressure were Central Asia, Central Sub-Saharan Africa, and Eastern Europe.[7] An unhealthy diet—which refers to a diet with insufficient consumption of fruits, vegetables, legumes, whole grains, nuts, seeds, milk, fiber, calcium, omega-3 fatty acids, and polyunsaturated fatty acids, as well as excessive consumption of red and processed meats, sugary beverages, trans-fats, and sodium—contributed towards an estimated 6.58 million cardiovascular disease-related deaths in 2021. The regions with the highest rates of cardiovascular disease burden attributable to dietary risk were Central Asia, Eastern Europe, and Oceania. High body mass index (BMI) contributes to a population-attributable fraction of 8% in both women and men for the individual risk of developing cardiovascular diseases. Furthermore, current smoking contributes to a population-attributable fraction of 7% in women and 11% in men. Major cardiometabolic contributors to cardiovascular diseases are hypercholesterolemia (non-HDL cholesterol attributes 15% in women and 17% in men) and diabetes mellitus (responsible for an estimated 15% of attributable risk in women and 10% in men).[11]

Economic Burden of Cardiovascular Disease

The burden of cardiovascular diseases is a pressing global challenge with a substantial economic impact on individuals, communities, health systems, and society. For instance, in Group of 20 (G20) countries, the direct costs of cardiovascular diseases exceed $600 billion annually, with the United States, Germany, and Japan incurring costs of $244.8 billion, $54.4 billion, and $49.5 billion, respectively, constituting between 11% to 15% of the total health system expenditure in G20 nations.[12]

Economic costs, including direct healthcare expenses and indirect costs such as loss of productivity and human capital, reach $402.2 billion in the United States and $109.6 billion in Japan. Additionally, stroke, a major contributor to cardiovascular disease-related costs,

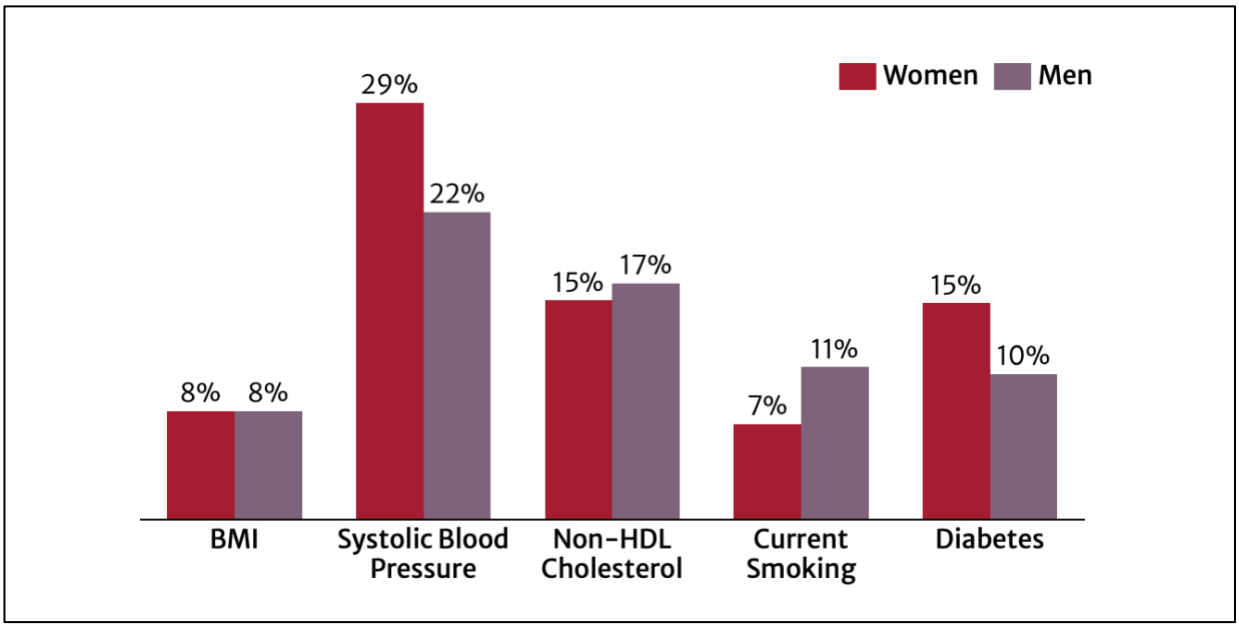
incur annual expenses ranging from $0.9 billion in Australia to $46 billion in the United States, further highlighting the critical need for comprehensive cardiovascular disease management.[12]

Social Determinants of Health and Healthcare Inequalities Concerning Cardiovascular Disease

There are major disparities and inequities in cardiovascular diseases across populations, reflecting differences in risk exposure, unequal access to healthcare services, and varied health outcomes. Disparities differ between and within countries and exist across income, race, geography, and ethnicity, with specific populations facing considerably higher rates of cardiovascular disease incidence, prevalence, and mortality compared to their counterparts. Disparities often stem from a complex interplay of socioeconomic factors, including limited access to quality healthcare, higher rates of poverty, and environmental influences such as exposure to unhealthy living conditions. Inequities, on the other hand, underscore the avoidable and unjust nature of these disparities, highlighting historic and contemporary systemic and structural issues that result in unequal distribution of resources, capabilities, and opportunities, such as individual knowledge related to cardiovascular disease prevention and access to preventive care, diagnostic services, and therapeutic interventions.[13]

In many LMICs, health system performance is inadequate in preventing, diagnosing, treating, and controlling the major cardiovascular risk factors such as hypertension, diabetes, and hypercholesterolemia. Disparities in access to care exist by sex and socioeconomic status. Testing patterns also follow socioeconomic gradients, favoring educated and wealthier individuals even without diagnostic criteria, further suggesting potential misallocation of limited testing resources. Some who meet the criteria are not being screened, while in other settings, individuals undergo screening who do not meet the diagnostic criteria, especially for hypertension.[14] Recent evidence indicates that cardiovascular disease risk factors vary based on genetic- and environmental factors.[15]

Figure 2. Population-attributable fractions of the individual-level risk factors for cardiovascular disease, stratified by women and men.[11]



Source: Authors

The Value of the Data Collaborative on Health Systems Performance (DC-HSP): Generating New Evidence through Existing Data

Health System Performance in Relation to Cardiovascular Diseases: A Critical Opportunity to Improve Health Outcomes at Scale

Well-functioning health systems play a crucial role in improving health by enabling effective, efficient, equitable, and responsive delivery of health services given the available resources in a country.[16,17] Many countries lack robust healthcare delivery systems and equitable access to high-value health services for cardiovascular diseases. High-value health services are effective and efficient to deliver "value for money", and equitably and responsive to provide "value for many".[18,19]

There are major opportunities to improve health systems performance concerning cardiovascular diseases in high-income countries (HICs) and LMICs.[20] For instance, when evaluating health outcomes relative to Gross Domestic Product (GDP), health outcomes in the United States are worse compared to countries that spend relatively less on health compared to the United States, such as the United Kingdom, Singapore, or the Netherlands.[13]

Health systems comprise both public and private institutions and are principally concerned with improving the level and distribution of health in a country while ensuring financial risk protection and user satisfaction. Health systems are open systems that continuously interact with and adapt to contextual factors such as economic, political, ecological, or technological changes.[21] Policymakers can change or modify health system functions (governance and organization, financing, and resource management) to improve the efficiency, effectiveness, equity, and responsiveness of health services for cardiovascular diseases to improve population-level outcomes as scale.

The existing empirical data could be utilized to ascertain how health systems perform in relation to cardiovascular diseases to inform the design, implementation, and scale-up of innovations (i.e., targeted policies, programs, and interventions). Ultimately, such efforts can enable the transition to high-value health systems required to achieve better cardiovascular disease outcomes at the population level.

Overview of the Data Collaborative on Health Systems Performance (DC-HSP)

The DC-HSP, a collaborative that involves the Health Systems Innovation Lab at Harvard University, Novartis Foundation, Microsoft, and Accenture, aims to contribute to providing a new body of evidence needed to create high-value health systems for improving outcomes for cardiovascular diseases.[1] This involves leveraging the vast amounts of data on determinants of health and health systems performance from HICs and LMICs—the majority of which is not publicly accessible—to evaluate and benchmark health systems performance concerning cardiovascular diseases.

The DC-HSP research outputs aim to inform policy, practice, and innovations to improve population health outcomes and create effective, efficient, responsive, and equitable health systems to enhance cardiovascular health. The creation of an integrated dataset with a global scope is essential to advance the next phase of research, innovation, and policy development to understand better health systems performance in different settings.

The Data Collaborative on Health Systems Performance (DC-HSP) Approach

The DC-HSP consists of two phases. Phase I of the project covers three streams of work: "Data," "Research," and "Translation" (**Figure 3**). The "Data" stream encompasses a systematic approach to identifying, mapping, and processing the available data across all world regions and developing a Survey Harmonization and Estimation Platform for Health and Economic Research Data (SHEPHERD). The "Research" stream leveraged the harmonized data for data visualization and analysis to inform policy and practice— utilizing various comparative health systems analysis, decision science, and mathematical modeling methods. Finally, the "Translation" stream focuses on the translation and application of research findings and insights to inform policy and practice on addressing cardiovascular

diseases by improving health systems performance through policy briefs and consultative groups.

The DC-HSP will use the outputs and assets produced in Phase I to advance as a series of data, research, innovation, and translation outputs in Phase II.

Figure 3: The three streams of work comprising the Data Collaborative on Health Systems Performance.

Source: Authors.

What makes the Data Collaborative on Health Systems Performance (DC-HSP) unique?

Collaboratives such as the Global Health and Population Project on Access to Care for Cardiometabolic Diseases (HPACC),[22] Globorisk,[23] Prospective Urban Rural Epidemiology (PURE) study,[24] Global Burden of Disease (GBD),[25] and Non-Communicable Diseases Risk Factor Collaboration (NCD-RisC)[26] have helped to improve our understanding the prevalence and risk factors of cardiovascular diseases. **Appendix Table 1** highlights several critical aspects of these initiatives, including objectives, methods, geographical regions, collected years, data types, variables, study populations, leadership, data access, and the main outputs.

However, to date, studies or collaboratives have yet to focus on understanding and utilizing the breadth of health system data available and analyze these data using a health systems framework to inform policy and practice to improve cardiovascular disease outcomes at the population level. Health Systems Innovation Lab uses a proprietary health systems framework that has been used to analyze the performance of health systems in 50+ nations and inform the design, introduction, and scale-up of major health system reforms.[19,27]

DC-HSP's uniqueness lies in utilizing a wide array of data sources for all the world regions, including individual-level and health system-level data. The study population includes individuals from all ages and sociodemographic contexts. We will employ advanced data analytic methodologies to analyze the large dataset created to generate new insights and evidence on health systems performance aimed at improving health systems performance to address cardiovascular diseases. To increase the utility of our findings, we will develop user-friendly interfaces to display the outputs of our research and create dynamic visualization tools for policy simulation and testing. This approach will help facilitate collaboration across academia, government, private, and non-profit sectors and leverage our networks.

Approach: Applied Methods for Data Harmonization

Approach to Creating the Global Dataset: A Three-Stage Process

Over several months, the DC-HSP team created a global dataset to analyze health system performance in relation to cardiovascular diseases by harmonizing all the suitable surveys identified and prioritized for countries in each of the world regions. The objective was to align variable definitions, measurements, and concepts across the different survey types to enable comparison across surveys, countries, and world regions.

The dataset was developed under the "Data" stream using a three-stage process using a de novo digital environment, SHEPHERD, developed by the Health Systems Innovation Lab. Through this process, the team identified, mapped, and processed all the available data across world regions, with Latin America and the Caribbean region as our first 'use case'.

The different stages are described below. A detailed description of the process is provided in a separate report, "*The Data Collaborative on Health Systems Performance (DC-HSP): Development of the Global Dataset*".

Stage 1: Data Inventory

From the core DC-HSP team of 36, a subset of 30+ global health students, researchers, and experts divided by global regions (Africa, Asia, Europe, Latin America, the Middle East, North America, and the Western Pacific) populated the data inventory. They searched through databases for surveys containing indicators related to cardiovascular disease, outcomes, and health systems interactions.

To populate the data inventory, the Health Systems Innovation Lab research team identified all nationally representative, individual-level, and publicly-accessible surveys with variables of interest related to cardiovascular diseases. The team included surveys that (i) contain

empirical individual-level data or health facility-level data on cardiovascular disease-related risk factors and disease biomarkers (measured or self-reported); (ii) are nationally representative for adults and/or adolescents; (iii) are conducted after 1990, and; (iv) are open-source or accessible upon request. The team prioritized surveys from multi-country programs conducted in low-income and middle-income countries: the Demographic and Health Surveys (DHS) Program,[28] the World Health Organization Stepwise Approach to Non-Communicable Disease (NCD) Risk-Factor Surveillance (STEPS) Surveys,[29] the Global Youth Tobacco Surveys (GYTS),[30] and Global Adult Tobacco Surveys (GATS),[31] the Global School-Based Student Health Surveys (GSHS).[32] Next, we browsed the Global Health Data Exchange (GHDx) database of the Institute for Health Metrics and Evaluation (IHME) for government-led national surveys.[33]

Stage 2: Data Mapping

Data mapping entailed adding each unique survey into the digital data harmonization environment SHEPHERD and defining, examining, and including all relevant variables from the respective codebooks for each identified survey.

Stage 3: Data Processing

Data processing, the final stage, consisted of developing a pooled harmonized dataset in SHEPHERD. The data processing consists of steps to correctly transfer survey metadata and microdata from the raw data into SHEPHERD.

This three-stage process ensured a systematic and replicable process that was followed in each regional team, enabling the pooling of all disparate survey data into a harmonized dataset that allows for analysis across surveys and countries.

SHEPHERD: The Digital Environment Used to Harmonize Survey Data

The DC-HSP project developed the digital data harmonization environment SHEPHERD for data mapping and processing. The objective of developing SHEPHERD is to provide a methodological approach to integrating different surveys and serve as a framework to conduct microsimulation to estimate health outcomes at both population and individual

levels, facilitating policy analyses for prioritizing interventions within the health system. SHEPHERD enables the harmonization and analysis of data from a vast range of sources while maintaining strict privacy standards. When analyzing the harmonized data, the comprehensive modeling approach will enable more robust epidemiological estimations of cardiovascular disease incidence, health system performance, and outcomes.

New Evidence: Insights Emerging from the Data Collaborative on Health System Performance

Phase I of the project, implemented between 1 September 2022 and 31 October 2023, focused on three streams of work: "Data," "Research," and "Translation." In this section, we present the insights emerging from this initial phase of the DC-HSP, highlighting the outcomes of three major outputs: (I) the First meeting of the Consultative Group convened in São Paulo, (II) the harmonized global dataset, and (III) the harmonized regional dataset, which focused on Latin America and the Caribbean.

Insights from the First Consultative Group Meeting in Sao Paulo, Brazil

The first meeting of the Consultative Group was held on 18 May 2023 in São Paulo, Brazil, and organized by the Health Systems Innovation Lab on behalf of the DC-HSP, with support from Fundação Getúlio Vargas (FGV). The Consultative Group comprises experts from various fields as a sounding board to orient the focus and policy relevance of the research in cardiovascular diseases in Brazil. The event titled, "Using the Power of Data Analytics to Deliver High-value Health Services to Improve Cardiovascular Health," was attended by over 30 participants, including senior government officials from Brazil, leading clinicians, researchers, scientists, innovators actively involved in shaping the Brazilian Health System and in the delivery of health services for cardiovascular diseases.

Participants identified six major challenges that currently hinder the provision of high-value health services for cardiovascular diseases with a focus on health system dynamics in Brazil, which include, among others, (I) pervasive disparities in cardiovascular health based on social and racial determinants, (II) inadequate health systems financing, (III) suboptimal resource management, (IV) inefficient coordination across the entire care pathway, (V) fragmented data systems, and (VI) insufficient collaboration within and across health systems.

To strategically address these challenges during Phase II of the DC-HSP project, participants identified 21 research questions, which were reviewed by the Health Systems Innovation Lab and consolidated into three major research themes and 14 research topics that could be systematically researched within the DC-HSP. The three research themes were:

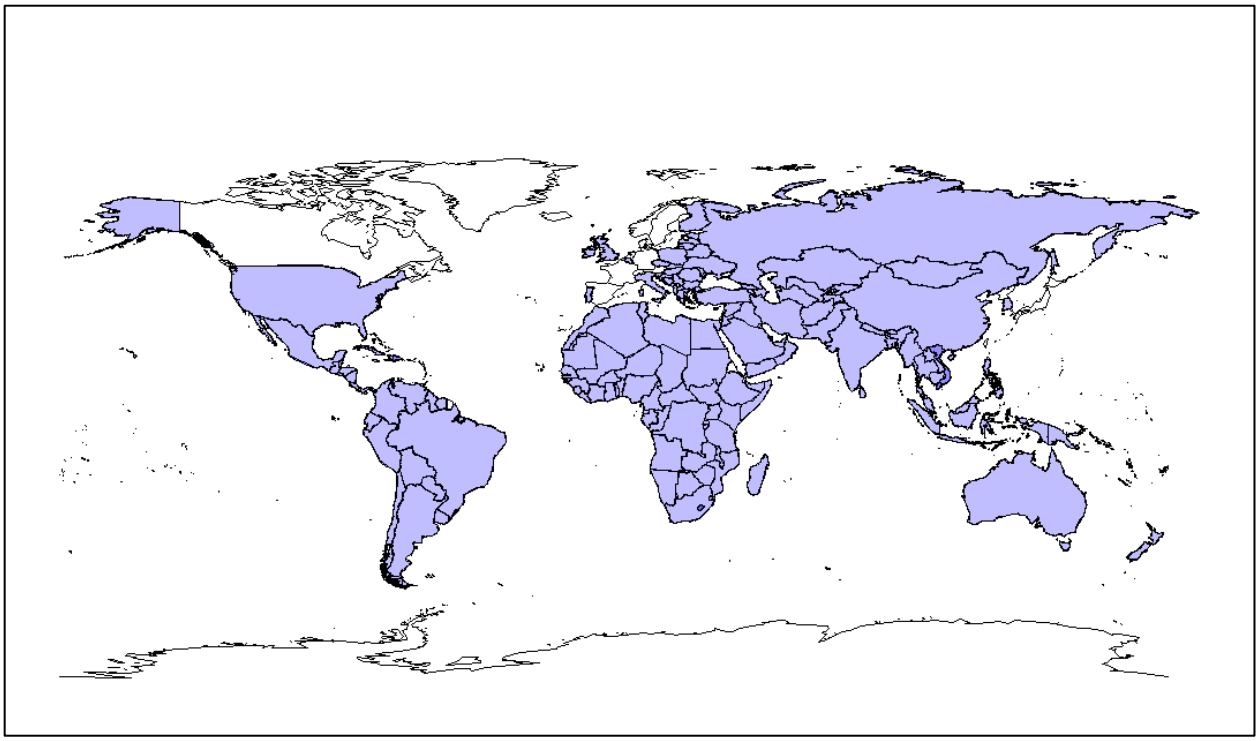
- **Theme 1:** Generating value in care delivery across the continuum of care;
- **Theme 2:** Establishing a robust ecosystem for digital health innovation, and;
- **Theme 3:** Building data systems to enable benchmarking and analytics at national and sub-national levels.

The Consultative Group emphasized that research should be highly collaborative, equity-focused, interdisciplinary, and embrace the digital transformation influencing health systems in Brazil, and globally – to produce system-level change over the long term. Finally, participants emphasized that the research outputs should be used to (I) inform policy engagement; (II) more effective strategic partnerships between the public and private sectors, including academia and civil society, and; (III) build consensus and leadership around cardiovascular health.

Insights from the Harmonized Global Dataset

The global dataset includes 1046 unique surveys from 219 countries across Africa, Asia, Europe, the Middle East, Latin America and the Caribbean, North America, and the Western Pacific. In total, we included 218 Demographic and Health Surveys, 37 Global Adult Tobacco Surveys, 138 Global School-based Health Surveys, 568 Global Youth Tobacco Surveys, 81 Stepwise Approach to Non-Communicable Disease (NCD) Risk-Factor Surveillance Surveys, and 4 surveys from other sources. Other sources included government-led surveys and national cohort studies. The global dataset includes a total of **6,787,201** individuals. **Figure 4** shows all countries for which data is included in the DC-HSP.

Figure 4: Overview of All Countries Included in the Global Data Collaborative on Health Systems Performance (DC-HSP) Dataset (in Purple).



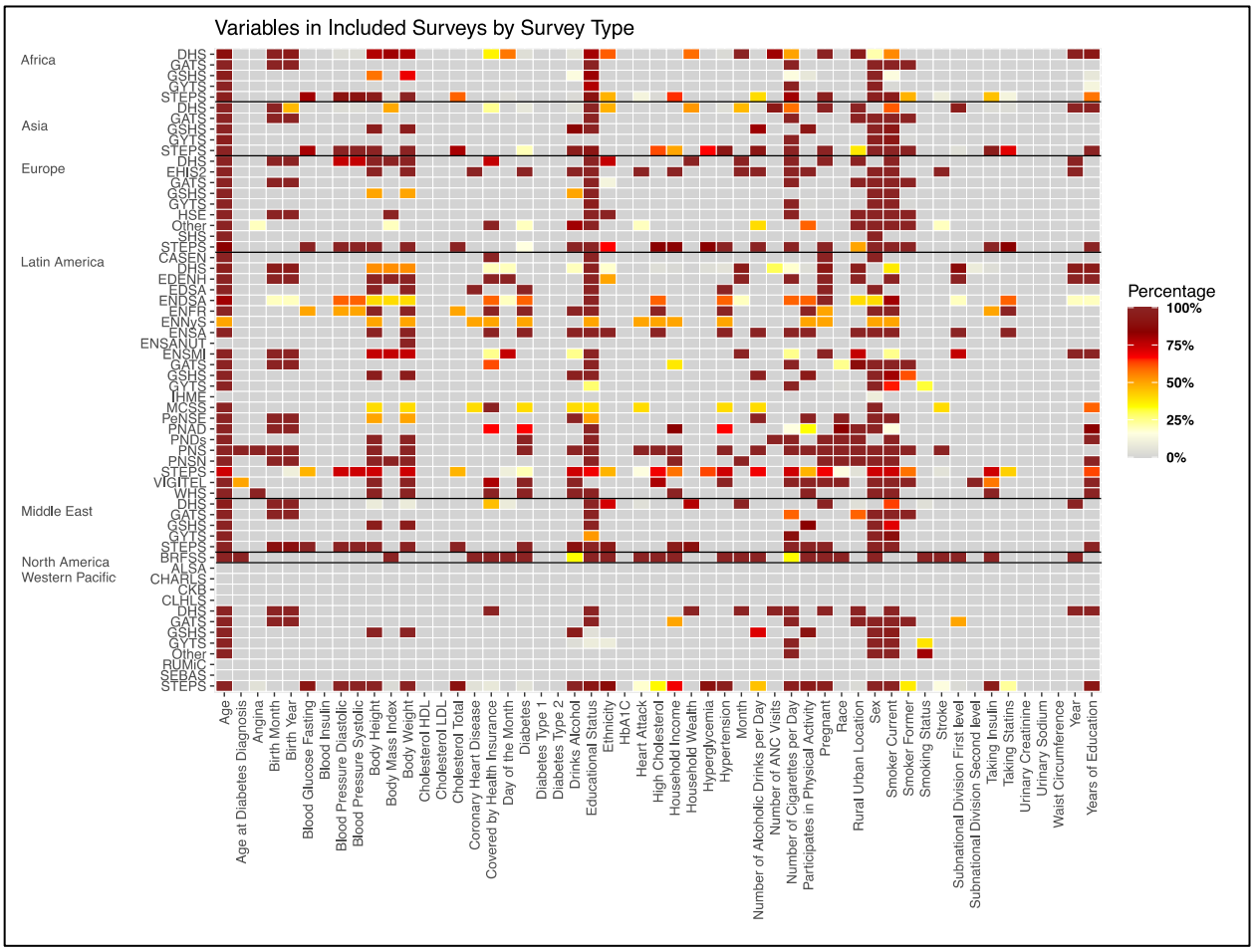
Source: Authors

Figure 5 presents, for each of the different survey types by world region, the percentage of surveys that include the variables of interest, as a percentage of all surveys. Grey indicates that variables are not available. Yellow indicates low variable availability, and dark red indicates high variable availability. Data on age, sex, and educational status were available for most surveys across survey types and regions. Data on race, ethnicity, urban or rural residence, and wealth indicators was available for about half of the surveys. Also, data on the CVD-associated risk factors BMI (or body weight and height) and smoking (including current smoking status, number of cigarettes per day, and former smoking status) were widely available across surveys.

STEPS and GSHS provide most data on body weight, height, and BMI, while data on smoking is most widely available in GYTS and GATS across global regions. Data on cardiometabolic indicators, such as blood pressure, hypertension, blood glucose, and cholesterol, was available in most STEPS surveys across regions, while often lacking in other survey types. Data on diabetes status was less available, while data on blood glucose fasting measurements and taking insulin are more widely available and can be utilized to infer diabetes status in individuals. Data on health insurance coverage was available in DHS surveys and in country-specific surveys such as the EDENH (Brazil) and BRFSS (United States). Data on past coronary heart disease events was widely available in the BRFSS, while lacking in the other included surveys.

A total of 779 surveys include information on two or more indicators of cardiovascular disease-related risk, 59 surveys include information on two or more cardiovascular disease biomarkers, 14 surveys include data on non-communicable disease outcomes, and 46 surveys include information on two or more indicators related to health system interaction.

Figure 5. Variable Availability by World Region and Survey Type for All Surveys Included in the Data Collaborative on Health Systems Performance (DC-HSP) Global Dataset.



Source: Authors

Cardiovascular diseases are the leading cause of mortality worldwide, making it vital to develop a more in-depth understanding of the prevalence, risk factors, and associated demographic characteristics. Although sociodemographic and behavioral risk factor variables are more widely available across surveys and countries, the lack of comprehensive data on cardiovascular disease incidence, mortality, and health systems indicators for cardiovascular diseases is a major concern. The lack of sufficient data on cardiovascular diseases has far-reaching implications for public health and underscores the urgent need for more robust data collection efforts. To address this gap, it is imperative to establish a

comprehensive plan to estimate the burden and risk of cardiovascular diseases, drawing upon demographic and socio-economic profiles and available data on cardiovascular disease risk factors. More information can be found in the report: "*The Data Collaborative on Health Systems Performance (DC-HSP): Development of the Global Dataset*".

Insights from the Harmonized Regional Data for Latin America and the Caribbean

During Phase I, the DC-HSP developed and used SHEPHERD to identify and harmonize available data to inform the development, parameterization, and calibration of the cardiovascular disease microsimulation model. The Latin America and Caribbean regions were prioritized during Phase I, as the burden of cardiovascular diseases is rising disproportionally fast in this world region, and, currently, substantial disparities exist in the burden and access to care for cardiovascular diseases.

We harmonized 288 unique surveys (**Panel 1**) from 37 different countries in Latin America and the Caribbean between 1990 and 2019, with data from over 1.2 million respondents. Surveys were available from the following countries: Anguilla, Argentina, Antigua and Barbuda, Bahamas, Belize, Bolivia, Brazil, Barbados, Chile, Colombia, Costa Rica, Cuba, Curacao, Cayman Islands, Dominica, Dominican Republic, Ecuador, Grenada, Guatemala, Guyana, Honduras, Haiti, Jamaica, Saint Kitts and Nevis, Saint Lucia, Mexico, Montserrat, Nicaragua, Panama, Peru, Paraguay, El Salvador, Suriname, Trinidad and Tobago, Uruguay, Saint Vincent and the Grenadines, Venezuela (**Figure 6**).

Panel 1: Surveys Prioritized in Latin America and the Caribbean with a focus on Brazil

Survey types included for Latin America and the Caribbean were: the Demographic and Health Surveys (DHS) Program,[28] the World Health Organization Stepwise Approach to Non-Communicable Disease (NCD) Risk-Factor Surveillance (STEPS) Surveys,[29] the Global Youth Tobacco Surveys (GYTS),[30] and Global Adult Tobacco Surveys (GATS),[31] the Global School-Based Student Health Surveys (GSHS).[32] In addition to these global surveys, national surveys from Brazil and other countries in Latin America and the Caribbean were included, among which: Pesquisa Nacional por Amostra de Domicílios (PNAD),[34] Pesquisa Nacional de Saúde (PNS),[35] Pesquisa Nacional de Saúde do Escolar (PeNSE),[36] Vigilância de Fatores de Risco e

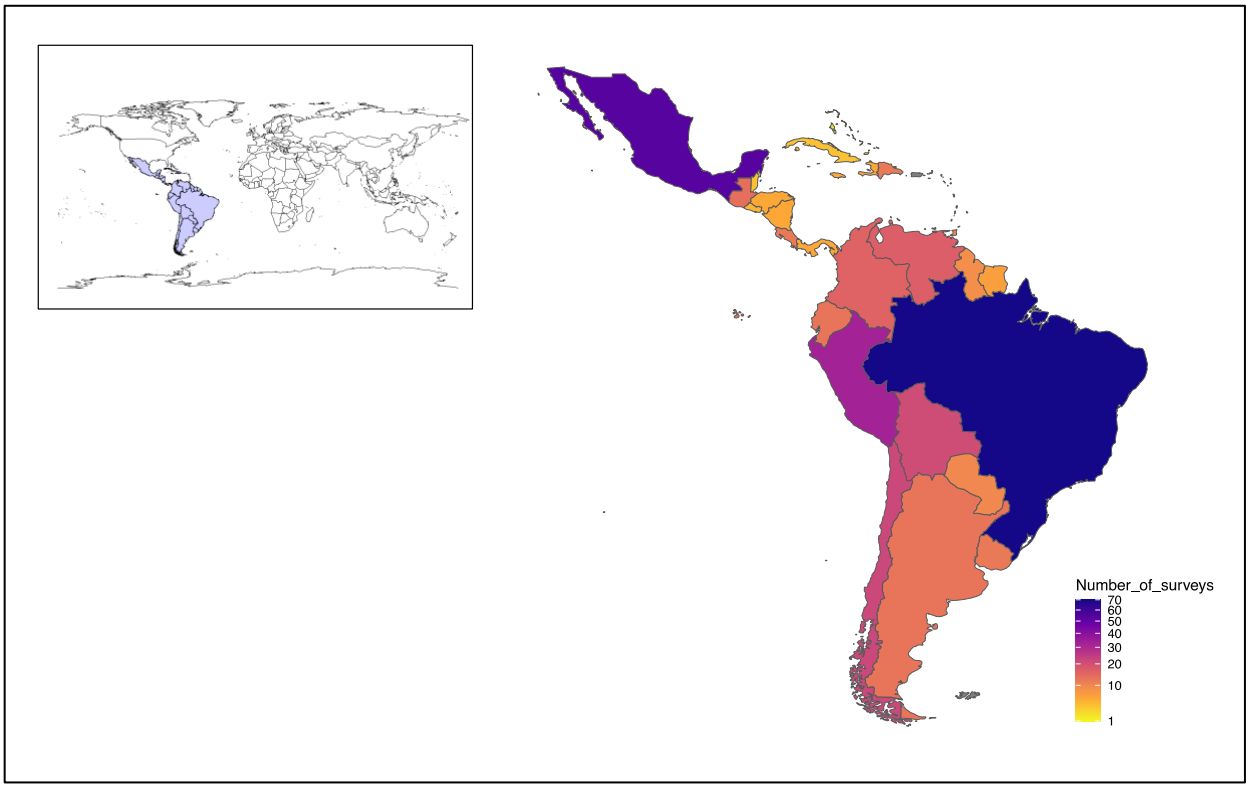
Proteção para Doenças Crônicas por Inquérito Telefônico (Vigitel),[37] and Departamento de Informática do Sistema Único de Saúde (DATASUS).[38]

A total of 210 surveys include information on two or more indicators of cardiovascular disease-related risk, eight surveys include information on two or more cardiovascular disease biomarkers, four surveys include data on non-communicable disease outcomes, and nine surveys include information on two or more indicators related to health system interaction.

Across countries in the Latin America and the Caribbean region, data on age, educational status, and sex was available for the vast majority of surveys across survey types and regions. Data on race, ethnicity, urban or rural residence, and wealth indicators was available for about half of the surveys. Also, data on the CVD-associated risk factors BMI (or body weight and height) and smoking (including current smoking status, number of cigarettes per day, and former smoking status) was widely available across surveys. Similarly, data on participation in physical activity and alcohol-drinking behavior was widely available across surveys. Data on cardiometabolic risk was available for one-third to half of the countries in this region. Most countries lacked data on cardiovascular events, angina, coronary heart disease, and stroke, cross survey types. Ecuador and Guyana are the only countries in the region with data on heart attacks.

During Phase II of the project, DC-HSP will develop a microsimulation-based decision analysis framework, which provides methodological approaches to synthesize evidence from multiple sources to estimate population and individual-level health outcomes and perform policy analyses to prioritize health system interventions.

Figure 6: Overview Of All Countries in Latin America and the Caribbean for which the Data Collaborative on Health Systems Performance (DC-HSP) Included Data (in Purple) and Density of Surveys by Country from Low (in Yellow), Medium (in Orange) to High (in Dark Blue).



Source: Authors

Description of the Sociodemographic Characteristics of the Data Sample

Below, we briefly describe each variable available and present the distributions of the data for the different variables of interest. The reports "*The Data Collaborative on Health Systems Performance (DC-HSP): Development of the Regional Dataset Latin America and the Caribbean*" and "*The Data Collaborative on Health Systems Performance (DC-HSP): Cardiovascular Disease Modelling in Latin America and the Caribbean: Preliminary Insights from Harmonized Empirical Data*" provide a complete overview of the descriptive results. The incidence of cardiovascular diseases often increases with age due to biological factors, independent of risk factor behaviors or other exposures associated with age. While our sample included individuals between 0 and 100 years of age, the majority of observations

were between 18 and 60 years of age. Similarly, behavioral and biological risk factors for cardiovascular diseases often differ by sex. In our sample, approximately 70% of observations are female, and about 30% are male. Education is associated with behavioral risk factors for cardiovascular diseases and factors related to health system access, care-seeking behavior, and treatment adherence. In addition, we include data on urban and rural residences. Living in rural areas is often associated with health system access and quality of care factors and may influence behavioral risk factors.

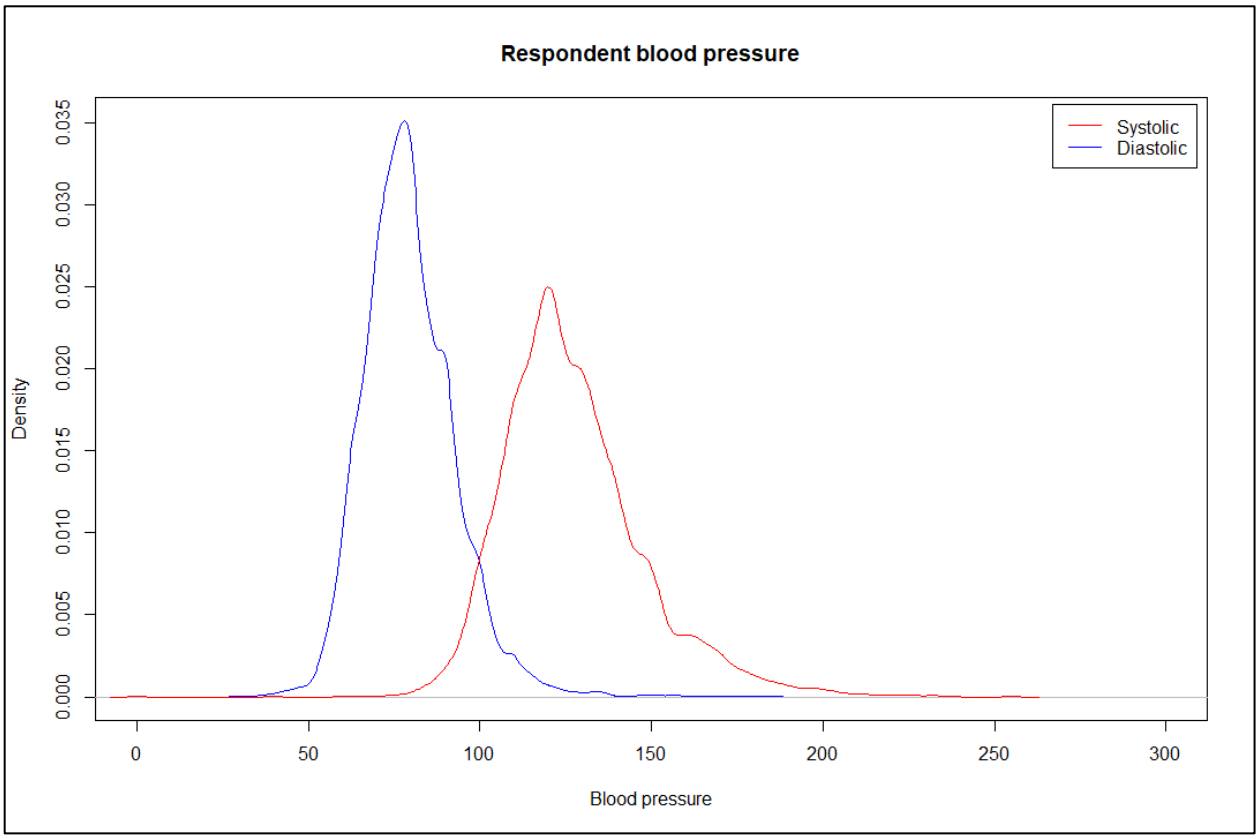
Smoking is a risk factor for cardiovascular diseases and varies by age, time period, sex, educational status, and other factors. Smoking prevalence (unadjusted, unweighted) in our sample is approximately 20%. The prevalence (unadjusted, unweighted) of physical activity in our sample is about 60%. As a measure of adiposity, waist circumference is predictive of cardiovascular disease risk, although this indicator is less frequently measured than other indicators, such as BMI. Similarly, BMI-related cardiovascular disease risks have been well-studied across multiple populations. Although a simple construct, BMI is strongly associated with more sophisticated measures of adiposity, such as DEXA at the individual level, with no significant differences by race/ethnicity.

Description of the Data on Attributable Risk Factors in the Data Sample

Blood pressure is one of the most important attributable risk factors for cardiovascular diseases that can be objectively measured to assess both diagnosed and undiagnosed disease burden of hypertension. **Figure 7** presents the distributions of systolic and diastolic blood pressure in the sample. Cholesterol is also a continuous indicator that can be measured to assess disease risk and burden, although lab capabilities are required. The relative frequency of respondents who responded that they have been diagnosed with high cholesterol is about 30% in our sample. Statin use, which can be used to assess the prevalence of patients being treated for high cholesterol, was lower than 5% in our sample. Furthermore, about 15% of respondents in our sample indicated that they have been told they have hyperglycemia (i.e., high blood glucose or diabetes mellitus). In addition to blood glucose measurements (which may indicate 'normal' results for patients with controlled diabetes), reported insulin use is

another indicator that can be used to assess the prevalence of diagnosed (and treated) diabetes; about 20% of our sample reported to use insulin.

Figure 7. Distributions of Systolic and Diastolic Blood Pressure in the Sample for Latin America and the Caribbean Region. The Y-axis represents the Probability Density of the Blood Pressure Distributions.



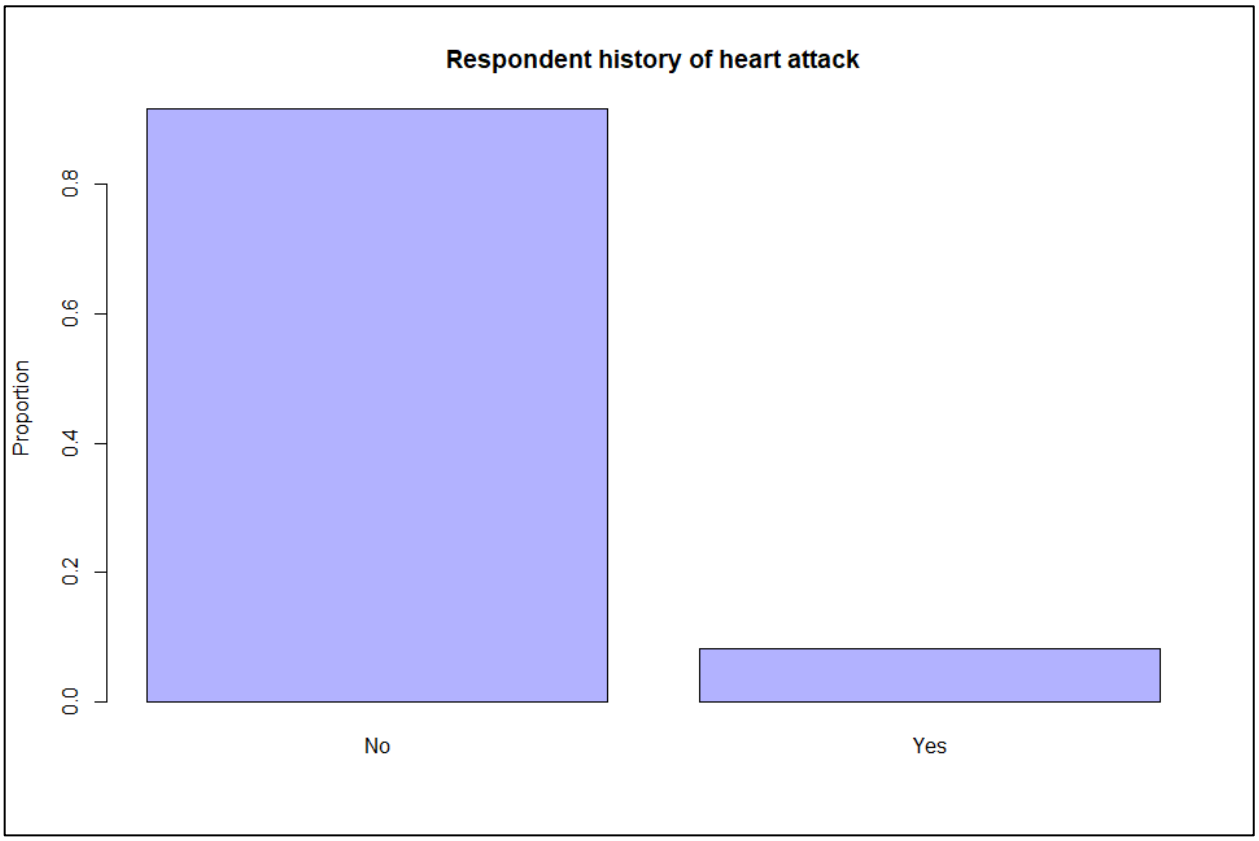
Source: Authors

Description of the Data on Cardiovascular Diseases

Respondents reporting a history of a heart attack is an indicator that can be used to assess the prevalence of myocardial infarction in the population. However, care must be taken as this indicator does not reflect the total incidence of myocardial infarction, as respondents must have survived a previous heart attack and be capable of responding to a survey to answer this question. Because of this issue, most estimates of cardiovascular disease burden are, therefore, prevalence estimates, and cardiovascular disease mortality (both observed

and unobserved) will need to be accounted for to estimate more accurate incidence rates of cardiovascular diseases. **Figure 8** presents the reported history of heart attack among observations in our sample, where about 8% reported to have had a heart attack in the past.

Figure 8. Frequency of Reported History of Heart Attacks in the Sample for the Latin America and the Caribbean Region.



Source: Authors

Health System Performance Indicators

Although not cardiovascular disease-specific, health insurance coverage is a useful proxy indicator for access to care, which impacts the detection probabilities of cardiovascular diseases and chronic disease management. In the sample for Latin America and the Caribbean, about 40% of the respondents reported to have health insurance, compared to 55% who reported not to be covered by any health insurance. Similarly, we include data on antenatal care, which is not cardiovascular disease-specific but is a commonly used proxy indicator to assess access to routine primary care in different contexts.

Using the Evidence Generated by the Data Collaborative on Health Systems Performance to Inform Practice: Looking Ahead

Phase I of the DC-HSP project provided five valuable lessons and recommendations concerning how data could be harnessed better to deliver high-value health services for cardiovascular diseases and improve outcomes at the population level.

(I) **Leverage Perspectives from Policymakers:** Leverage insights from the consultative group in Brazil to identify critical questions that policymakers should address. These questions will serve as a foundation for transitioning to high-value health services for cardiovascular diseases.

(II) **Streamline and Fast-track Data Harmonization:** Utilize standardized measurements, data collection, and reporting practices to facilitate the harmonization process. Emphasize the importance of digital information systems, integrated data management, and enhanced interoperability for effective data utilization.

(III) **Enhance Performance Benchmarking:** Standardize performance evaluation methods to bridge the information-to-policy gap. Develop evidence-based approaches for standardized health outcome and cost measurement reporting and analytics, utilizing available data.

(IV) **Establish a Cardiovascular Diseases Cross-Learning Network:** Facilitate knowledge exchange and cross-learning among countries to share insights and best practices in harnessing data for improvement in health services.

(V) **Identify Critical Health System Performance Data Gaps:** Conduct an assessment to identify critical areas where additional data is needed to understand cardiovascular

diseases comprehensively. Prioritize data collection efforts in these identified areas to fill existing gaps and enhance the overall effectiveness of health services.

In Phase II of the project, we aim to build on the data identification, mapping, and processing efforts to estimate cardiovascular disease incidence, detection, and mortality for countries and populations that currently lack reliable numbers. The next steps of such efforts will involve the development and calibration of a structural microsimulation of cardiovascular disease. Once developed and calibrated, the model can be used to project cardiovascular disease indicators into the future for different groups of interest, and can also be used for policy analysis to evaluate the potential impact of various intervention strategies, as well as the impact of health system performance on cardiovascular disease outcomes. Incorporating both upstream risk factors, such as individual demographic and socio-economic characteristics, and downstream health system factors, such as quality of care, provides essential context that may influence the comparative and cost-effectiveness of potential interventions. A major strength of this approach is the utilization of harmonized empirical data, accounting for observed differences across and within countries, and the synthesis of knowledge from other sources into one integrated analytic framework that can provide evidence-based health system guidance to policy-makers in different contexts.

We encourage policymakers to adopt systems to structurally measure cardiovascular disease incidence, detection, and mortality—promoting and implementing strategic initiatives to collect and analyze data related to cardiovascular diseases that consider sociodemographic and lifestyle factors as well as health systems performance indicators. This data-driven approach will not only enhance our understanding of the distribution and determinants of cardiovascular disease occurrence and health system performance, but will also facilitate the development of targeted interventions and policies to mitigate the impact of cardiovascular diseases on public health. By addressing the scarcity of data on cardiovascular diseases and promoting evidence-based strategies, we can make substantial progress in preventing and managing this widespread health issue.

The DC-HSP will use the outputs and assets produced in Phase I to advance a series of data, research, and translation outputs in Phase II. These efforts will include, among others, extending the health systems performance dataset in scale and scope, using the dataset to build a cardiovascular disease microsimulation model, contributing to the design of an intervention study to introduce a "high-value" health system solution for cardiovascular diseases in Brazil, developing an interactive health systems performance dashboard for policymakers and the public, to help facilitate the development of targeted innovations to address cardiovascular diseases in country health systems. By implementing these recommendations, the DC-HSP project can optimize the use of data for delivering high-value health services, improving outcomes for cardiovascular diseases, and fostering collaborative learning across nations.

References

1 Health Systems Innovation Lab | Harvard University. https://www.hsph.harvard.edu/health-systems-innovation-lab/ (accessed Aug 27, 2023).

2 Fuchs FD, Whelton PK. High Blood Pressure and Cardiovascular Disease. *Hypertension* 2020; **75**: 285–92.

3 Cardiovascular diseases. https://www.who.int/health-topics/cardiovascular-diseases (accessed Oct 9, 2023).

4 American Heart Association | To be a relentless force for a world of longer, healthier lives. www.heart.org. https://www.heart.org/en/ (accessed Oct 9, 2023).

5 Deng P, Fu Y, Chen M, Wang D, Si L. Temporal trends in inequalities of the burden of cardiovascular disease across 186 countries and territories. *Int J Equity Health* 2023; **22**: 164.

6 Cardiovascular diseases (CVDs). https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (accessed Oct 9, 2023).

7 Roth GA, Mensah GA, Johnson CO, *et al.* Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019. *J Am Coll Cardiol* 2020; **76**: 2982–3021.

8 New Report Tracks Latest Trends in Global Cardiovascular Health. Am. Coll. Cardiol. https://www.acc.org/About-ACC/Press-Releases/2022/12/12/19/20/http%3a%2f%2fwww.acc.org%2fAbout-ACC%2fPress-Releases%2f2022%2f12%2f12%2f19%2f20%2fNew-Report-Tracks-Latest-Trends-in-Global-Cardiovascular-Health (accessed Oct 9, 2023).

9 Heart Disease and Stroke | CDC. 2022; published online Sept 8. https://www.cdc.gov/chronicdisease/resources/publications/factsheets/heart-disease-stroke.htm (accessed Oct 30, 2023).

10 Kannel WB, McGee DL. Diabetes and cardiovascular risk factors: the Framingham study. *Circulation* 1979; **59**: 8–13.

11 Global Effect of Modifiable Risk Factors on Cardiovascular Disease and Mortality. *N Engl J Med* 2023; **389**: 1273–85.

12 Health Systems Innovation Lab at Harvard University, Rifat A. The State of Cardiovascular Disease in G20+ Countries. *HPHR J* 2022; **HSIL**. DOI:10.54111/0001/HSIL/cvdg20.

13 How Does the US Healthcare System Compare to Other Countries? https://www.pgpf.org/blog/2023/07/how-does-the-us-healthcare-system-compare-to-other-countries (accessed Aug 27, 2023).

14 Wurie HR, Cappuccio FP. Cardiovascular disease in low- and middle-income countries: an urgent priority. *Ethn Health* 2012; **17**: 543–50.

15 Anand S, Bradshaw C, Prabhakaran D. Prevention and management of CVD in LMICs: why do ethnicity, culture, and context matter? *BMC Med* 2020; **18**: 7.

16 World-Heart-Vision-2030.pdf. https://world-heart-federation.org/wp-content/uploads/World-Heart-Vision-2030.pdf (accessed Oct 9, 2023).

17 Piñeiro DJ, Codato E, Mwangi J, Eiselé J-L, Narula J. Accelerated reduction in global cardiovascular disease is essential to achieve the Sustainable Development Goals. *Nat Rev Cardiol* 2023; **20**: 577–8.

18 Atun R, Moore G. Building a High-Value Health System. Oxford University Press, 2021 DOI:10.1093/oso/9780197528549.001.0001.

19 Reddy CL, Garcia Ruiz J, Atun R. Transitioning to High-Value Health Systems in G20+ Countries. 2023 https://www.hsph.harvard.edu/health-systems-innovation-lab/hvhs/ (accessed Oct 30, 2023).

20 Kruk ME, Gage AD, Arsenault C, *et al.* High-quality health systems in the Sustainable Development Goals era: time for a revolution. *Lancet Glob Health* 2018; **6**: e1196–252.

21 Atun R. Health systems, systems thinking and innovation. *Health Policy Plan* 2012; **27**: iv4–8.

22 HPACC. HPACC. https://www.hpaccproject.org (accessed Nov 15, 2023).

23 Homepage | Globorisk. http://www.globorisk.org/ (accessed Feb 9, 2023).

24 Teo K, Chow CK, Vaz M, Rangarajan S, Yusuf S, PURE Investigators-Writing Group. The Prospective Urban Rural Epidemiology (PURE) study: examining the impact of societal influences on chronic noncommunicable diseases in low-, middle-, and high-income countries. *Am Heart J* 2009; **158**: 1-7.e1.

25 Global Burden of Disease (GBD). https://www.healthdata.org/research-analysis/gbd (accessed Nov 15, 2023).

26 Homepage > NCD-RisC. https://ncdrisc.org/ (accessed Feb 9, 2023).

27 Atun R, de Jongh T, Secci F, Ohiri K, Adeyi O. Integration of targeted health interventions into health systems: a conceptual framework for analysis. *Health Policy Plan* 2010; **25**: 104–11.

28 The DHS Program - Quality information to plan, monitor and improve population, health, and nutrition programs. https://dhsprogram.com/ (accessed Feb 9, 2023).

29 STEPwise approach to NCD risk factor surveillance (STEPS). https://www.who.int/teams/noncommunicable-diseases/surveillance/systems-tools/steps (accessed Feb 9, 2023).

30 Global Youth Tobacco Survey. https://www.who.int/teams/noncommunicable-diseases/surveillance/systems-tools/global-youth-tobacco-survey (accessed Feb 9, 2023).

31 Global Adult Tobacco Survey. https://www.who.int/teams/noncommunicable-diseases/surveillance/systems-tools/global-adult-tobacco-survey (accessed Nov 15, 2023).

32 Resource bank for adolescent health. https://www.who.int/teams/maternal-newborn-child-adolescent-health-and-ageing/adolescent-and-young-adult-health/resource-bank-for-adolescent-health (accessed Feb 9, 2023).

33 Global Health Data Exchange | GHDx. https://ghdx.healthdata.org/ (accessed Feb 9, 2023).

34 IBGE. Volume Brasil | IBGE. PNAD - Pesqui. Nac. Por Amostra Domicílios. https://www.ibge.gov.br/estatisticas/sociais/populacao/9127-pesquisa-nacional-por-amostra-de-domicilios.html (accessed Nov 15, 2023).

35 IBGE. 2019 | IBGE. PNS - Pesqui. Nac. Saúde. https://www.ibge.gov.br/estatisticas/sociais/saude/9160-pesquisa-nacional-de-saude.html (accessed Nov 15, 2023).

36 IBGE. Edições | IBGE. PeNSE - Pesqui. Nac. Saúde Esc. https://www.ibge.gov.br/estatisticas/sociais/educacao/9134-pesquisa-nacional-de-saude-do-escolar.html (accessed Nov 15, 2023).

37 Vigitel. Minist. Saúde. https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/v/vigitel/vigitel (accessed Nov 15, 2023).

38 DATASUS – Ministério da Saúde. https://datasus.saude.gov.br/ (accessed Nov 15, 2023).

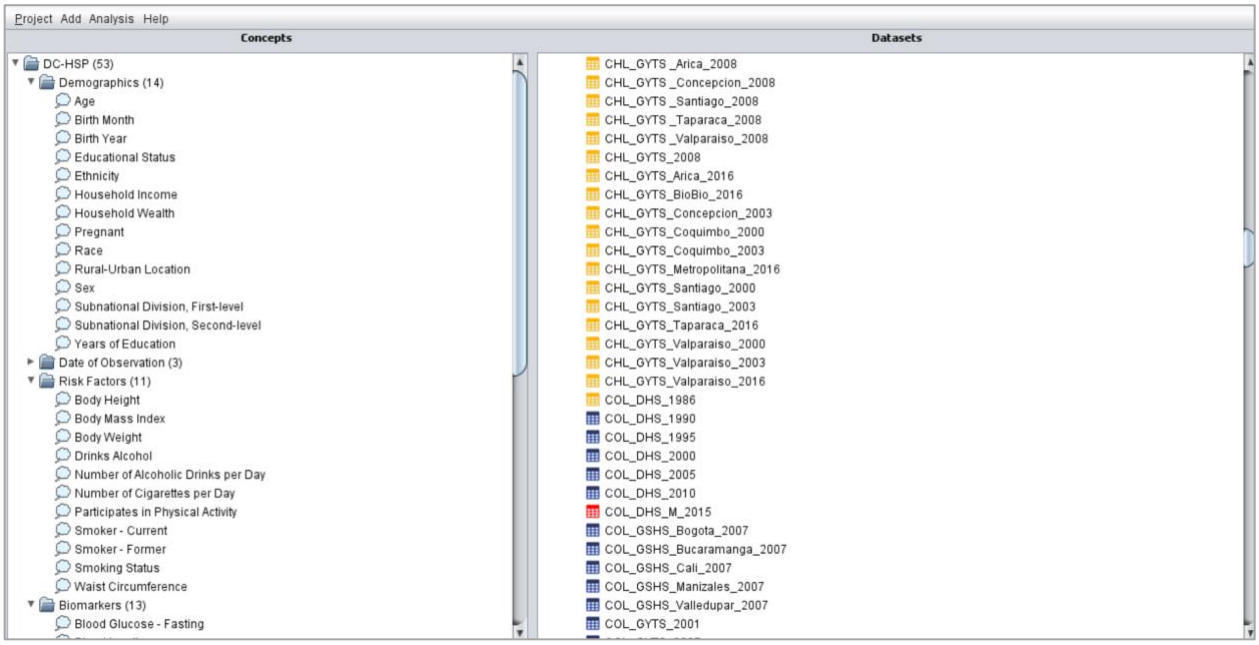
Appendix

Appendix Table 1. Overview of DC-HSP and other existing collaboratives related to cardiovascular diseases (CVDs).

Initiatives	Data Collaborative on Health Systems Performance (DC-HSP)	Globorisk	Global Health and Population Project on Access to Care for Cardiometabolic Diseases (HPACC)	IHME Global Burden of Disease	NCD-RisC	The Prospective Urban Rural Epidemiology (PURE) study
Objective	Leverage previously fragmented health systems data to analyze and measure performance, particularly in relation to cardiovascular diseases, to inform policy, practice, and innovations aimed at improving population health outcomes.	Develop a country-specific CVD-risk calculator and assess the correlation between the WHO and Globorisk 10-year CVD risk prediction models	Conduct high-impact analyses related to cardiometabolic diseases and health systems, inform policymakers on healthcare gaps, and implement primary prevention strategies	Understand the changing health challenges facing countries and provide estimates for the Global Burden of Disease (GBD) and other health scenarios to inform global health policy and decision-making.	Hosts data for variables listed below. Give range to a wide variety of data visualizations.	Examine the relationship between societal influences and prevalence of risk factors, chronic noncommunicable diseases, and societal determinants.
Applied Methods	Empirical Studies, Modelling, Data Science Techniques, Cloud Computing, Interactive Data Visualization	Modelling, Tool designed to predict CVD risk, Compares laboratory-based and office-based data	Empirical studies	Utilizes mathematical modelling techniques based on diverse secondary data sources from over 200 countries and territories	Uses a Bayesian hierarchical model to estimate trends	Community-based sampling and recruitment strategy where urban and rural communities were selected within countries
Regions Covered	Africa, Asia, Europe, Latin America, the Middle East, North America, the Western Pacific	LMICs, 182 countries	First LMICs, now global scope	Global- 204 countries	Global -196 counties	Global: 17 countries categorized into 4 LIC, 10 MIC, and 3 HIC study samples
Years Covered	1990-present	2006 or later	2005 or later	1990-present	1990-present	2002-2030
Types of Data Included	National and regional surveys, including individual-level surveys with or without biomarkers (e.g., DHS, STEPS, WHO and European Health Interview Surveys), national health and nutrition examination surveys (e.g., US NHANES, Korea NHANES), household consumption and expenditure surveys, health facility surveys (SARA and SPA), multi-indicator cluster surveys, educational	Data comprised of a wide range of health examination surveys such as STEPS, Chinese CHNS, Post-Monica in the Czech Republic, Iranian NCDSS, Jamaican JHLS, ENSA in Mexico, KNHANES, Spanish ENRICA surveys and NHANES	Nationally representative population-based surveys. 57 WHO STEPS surveys, 20 other surveys including DHS SAGE and Gateway 2 Global Aging	Observational study – uses surveys, censuses, cohort studies and admin data	More than 3300 multi-country and national population surveys, including WHO STEPS surveys	National World Bank reports, census data, and surveys. At the community level, data is obtained using tools such as the EPOCH instrument and Neighborhood Walkability Scale, along with the aggregation of responses from individual and household questionnaires

	attainment surveys, living standards measurement surveys, health and welfare surveys, environmental surveys, and health and retirement surveys					
Variables	Smoking, BMI (Body Mass Index), diet, physical activity, alcohol consumption, CVD-risk outcomes, blood pressure, blood glucose levels, cholesterol levels, disease prevalence, health system interaction, access to primary care	Age, sex, smoking, BP, Diabetes, total cholesterol	Age, sex, BP, diabetes, lipid biomarkers, health service utilization	Mortality and disability across countries, time, age, and sex	BMI, height, diabetes, blood pressure	Tobacco environment, social and economic, nutrition/food policy, physical activity environment
Study Populations	No restrictions	>40 years or older, males and females	≥15 years or older, males and females	Not specified	Not specified	35 to 70 years and their household members
Leaders and Data Access	Harvard Health Systems Innovation Lab	Harvard T.H. Chan School of Public Health researchers	Global team, multinational collaboration	University of Washington	Imperial College London	Population Health Research Institute (PHRI)
Main Outputs	Policy brief, data dashboard, scientific publications	Office-based model, publications	35 peer reviewed articles; generated evidence for WHO; crafted evidence-based policies.	Publications, used to inform health policies	Scientific publications	Scientific publications

Appendix Figure 1: Overview of a selection of the mapped variables (left-side panel) and a selection of the surveys (right-side panel) in SHEPHERD.

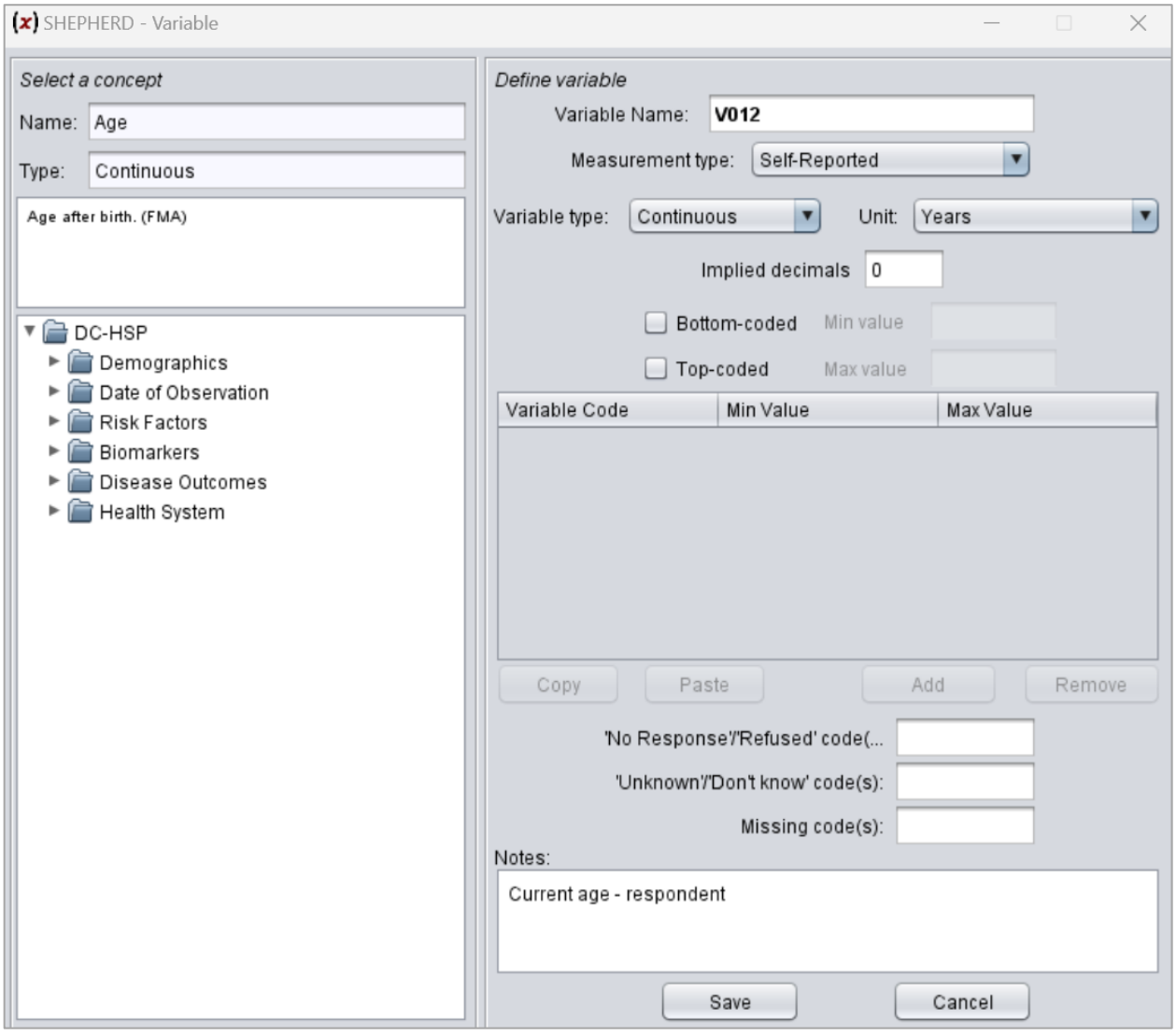


Source: SHEPHERD

Appendix Figure 2: Overview of how a survey is mapped in SHEPHERD, with the Demographic and Health Survey (DHS) from Colombia from 1995 as an example.

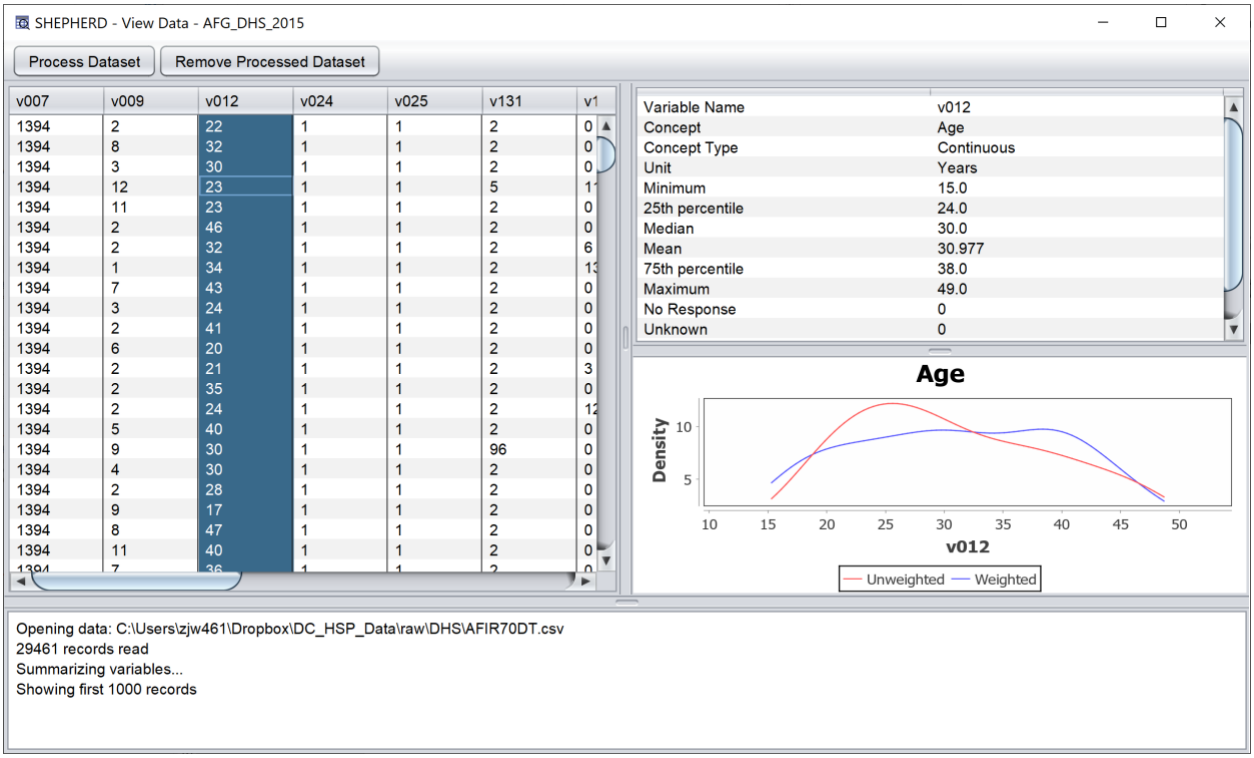
Source: SHEPHERD

Appendix Figure 3: Overview of how a variable is added for a survey in SHEPHERD, with age as an example.



Source: SHEPHERD

Appendix Figure 4: Overview of how variables for each survey are checked before processing surveys in SHEPHERD, with age as an example.



Source: SHEPHERD