

Health information systems data for decision-making: case study in three cities on current practices and opportunities

Sarah Rajkumar^{1,2} · Florence Secula^{1,2} · Daniel Cobos^{1,2} · Anna Socha^{1,2} · Johannes Boch³ · Sarah des Rosiers³ · Theresa Reiker³ · Joseph Barboza⁴ · Karim Seck³ · Mariana Silveira⁵ · Thuy Nguyen⁶ · Peter Steinmann^{1,2}

Received: 23 January 2024 / Accepted: 9 August 2024

Published online: 22 August 2024

© The Author(s) 2024 [OPEN](#)

Abstract

A functional and reliable Health Information System (HIS) is vital for data-based decision-making in public health. Here we describe the assessment of data processes and general HIS principles by adapting a global approach to three cities. The assessments supported the data strategy of the CARDIO4cities initiative in each city aiming to improve urban population health by increasing the use of cardiovascular disease (CVD) data to inform decision-making. We aimed to explore data collection processes and entities, data availability and quality as well as data ownership and sharing regarding a set of identified key performance indicators (KPIs). KPIs were based on a global theory of change (ToC) and a global evaluation and indicator framework and were tailored to each location. By first assessing existing sources and processes regarding data, recommendations for changes and improvements are sure to build on current circumstances. To map the data, existing data collection, analysis and storage processes were investigated. A flow chart was created to visualize the data pathways and challenges for each system and findings were compared across cities to document differences and similarities. Data quality and interoperability of various separate systems were the most prominent challenges for all HISs. The observed data quality issues originated from incorrect, missing and incomplete data and were connected to the misunderstanding of indicators, incomplete data input forms or the lack of a systematic data routine in the workflow. Harmonization of the HISs to ensure interoperability can facilitate data collection and analysis of health data and can provide a solid basis for health management decision-making. Based on the presented HIS cases, we recommend to examine, map and verify current processes when conducting a HIS assessment, to visualize findings and to gauge the interest of government entities to ensure political support.

1 Introduction

Health information is one of the six building blocks of a comprehensive health system according to the World Health Organization (WHO) framework for action [1]. A health information system (HIS) refers to a paper-based or electronic system designed to capture, transmit and manage healthcare data. The Global Observatory for eHealth 2015 survey reported that 66% of 125 surveyed countries have a HIS policy or strategy in place [2]. A well-maintained, functional and reliable HIS contains complete and trustworthy data that can be readily analysed and used for data-based decision-making and the development and implementation of sustainable health interventions. The timely availability of accurate public health data is therefore vital [3]. Unfortunately, routine health data is often outdated, inconsistent and incomplete

✉ Sarah Rajkumar, sarah.rajkumar@swisstph.ch | ¹Swiss Tropical and Public Health Institute, Kreuzstrasse 2, 4123 Allschwil, Switzerland. ²University of Basel, 4003 Basel, Switzerland. ³Novartis Foundation, Basel, Switzerland. ⁴IntraHealth, Dakar-Fann Dakar, Sénégal. ⁵Instituto Tellus, São Paulo, Brazil. ⁶PATH, Ho Chi Minh City, Vietnam.



[4]. Generally speaking, quality data should come from a reliable source, be accessible to relevant stakeholders, be delivered timely, and be relevant and consistent, accurate and complete [5]. However, several pre-requisites for using data to inform policy decisions such as the access to information, the skill to interpret and use that information adequately as well as the relevance to consider local evidence may be insufficient [6]. To ensure data quality and the use of evidence, sufficient data literacy as well as enhanced technical tools are needed [7].

Particularly in the context of the world's leading cause of death, cardiovascular disease (CVD), evidence-based actions are crucial to address the burden at scale. CVDs killed approximately 17.9 million people in 2019 with three quarters of these deaths occurring in low and middle-income countries (LMICs) [8–10].

Within its broad strategy to improve cardiovascular population health, the CARDIO4Cities approach aspires to catalyse the use of CVD data to inform health system decision-making and integrate data-generating digital innovations into existing HISs [11]. The CARDIO4Cities approach builds on six pillars: quality of **C**are, early **A**ccess, policy **R**eform, **D**ata and digital technology, **I**ntersectoral collaboration, and local **O**wnership. From 2018 to 2020, CARDIO4Cities validated its approach and demonstrated increases in hypertension control after 1–2 years of implementation in São Paulo (from 12.3 to 31.2%), Dakar (from 6.7 to 19.4%) and Ulaanbaatar (from 3.1 to 19.7%) [12]. This resulted in an estimated averted 3.3–12.8% of strokes and 3.0–12.0% of coronary heart disease (CHD) events [13]. In 2021, the approach added the measurement of further CVD target risk factors such as diabetes type II (blood sugar), dyslipidaemia (LDL-cholesterol) and obesity (BMI).

CARDIO4Cities applies solutions tailored to local needs alongside rigorous monitoring and evaluation (M&E) activities that generate data to inform data-based decision-making. Before improved approaches can be integrated into an HIS, it is key to map, understand and analyse current processes regarding data collection, storage and quality control [14]. Where the HIS contain several parallel sub-systems, their interoperability needs to be explored. In view of applying the CARDIO4Cities approach in various settings worldwide, a rapid assessment of the baseline conditions regarding HIS and CVD data flows was conducted in three cities: Dakar, Senegal; São Paulo, Brazil and Ho Chi Minh City (HCMC), Vietnam to get an overview of existing HIS and CVD data flows and identify gaps and challenges. The three locations were chosen as previous initiatives on CVDs and urban health had been implemented there by the same funding organization, strong partnerships had been established, and they are geographically diverse.

This manuscript outlines a global approach of assessing data processes and HIS principles. It describes the local adaptation of this methodology to the three selected cities and the findings of the analysis. The discussion contextualizes the findings within minimum structures needed to monitor and evaluate the cascade of care of the CV risk factor hypertension (i.e., screening, diagnosis, treatment and control) under the perspective of sustainable implementation of data-based decision-making.

2 Methods

The CARDIO4Cities initiative established a global theory of change (ToC) and a global evaluation and indicator framework to monitor progress, outcomes and impact [15] Fig 1. This provided the starting point for adapted indicator lists including key performance indicators (KPI) tailored to each location. KPIs included CV risk factors such as hypertension, high blood sugar (blood glucose or HbA1c), high low-density lipoprotein (LDL) cholesterol, and high body mass index (BMI).

In this context, the assessment of the local HIS aimed to explore data collection processes and entities, data availability and quality as well as data ownership and sharing regarding the identified KPIs. Further objectives of characterizing the HIS were (i) to establish a common understanding with local partners on KPIs and the feasibility of collecting them in the existing HIS, (ii) to explore the potential for integration of different sub-systems and (iii) to inform the development of an evaluation framework.

In each location, relevant stakeholders from public and private institutions were identified. To carry out data mapping, existing data collection, analysis and storage processes were investigated through desk research and stakeholder interviews and by conducting field visits to health facilities. Aggregated and individual-level data sources were mapped considering the KPIs. Those sources comprised paper-based and electronic data from health facilities, health systems data, and data collected by private entities. Each data source was further described with properties such as the collecting entity, the frequency of collection, details on reporting (frequency, recipient) as well as who owns the data and who has access to it.

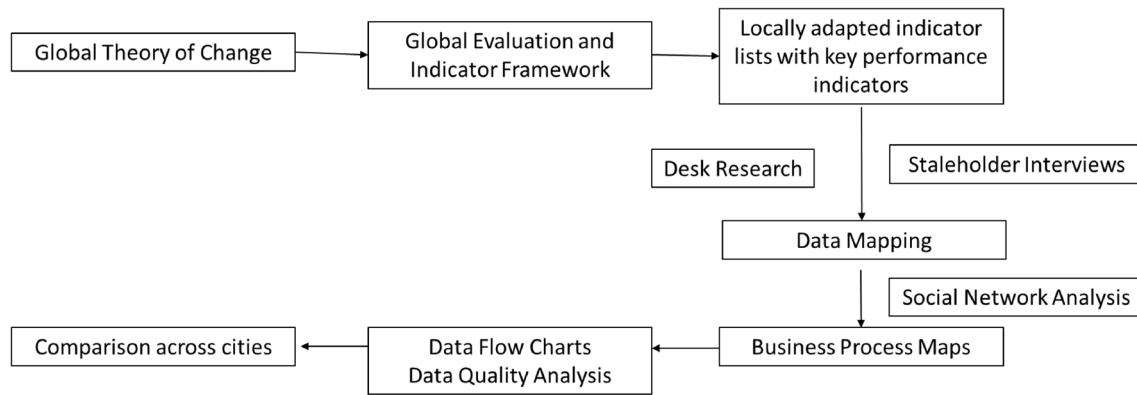


Fig. 1 Steps taken in the performance of this assessment

In Dakar and São Paulo, business process maps were produced through a social network analysis using the Bizagi software. They provided additional insights on hypertension data-related processes by identifying relevant local data systems and helped to further refine the routine data collection flowchart in these locations.

In all locations, a flow chart was created to visualize the data pathways from collection to analysis, storage, and use. Data gaps were identified with reference to the local evaluation framework. Data quality was assessed using a standardized data quality checklist that looks at accessibility, timeliness, relevance, consistency, accuracy and completeness. For each factor, between 1 and 6 specific indicators were created. Each indicator received a score from 0 to 4 (0 = No answer/Not applicable (NA), 1 = Not present, needs to be developed, 2 = Needs a lot of strengthening, 3 = Needs some strengthening, 4 = Satisfactory, no action needed). Representatives from institutions owning, processing or hosting these data were consulted at various stages to fill any knowledge gaps and verify the gathered information. Based on these investigations and the flow charts, challenges on each system level (community, regional, national) were identified and discussed with local stakeholders.

Findings were compared across cities to document differences and similarities that would indicate minimal requirements for tracking cascade of care data of CV risk factors longitudinally and to assess requirements for the sustainable application of data-based decision-making, for example with the support of digital tools.

In each location, a local partner organization supported and advised on the iterative process of conducting the assessment and carried out networking activities with stakeholders: IntraHealth in Dakar, Instituto Tellus in São Paulo, and PATH in HCMC.

3 Results

3.1 Dakar, Senegal

3.1.1 Background and status of HIS

DHIS2 (District Health Information Software 2), which was introduced in Senegal in 2010 to collect data on infectious diseases, has been the official national Routine Health Information System (RHIS) since 2016 [16, 17]. DHIS2 is a free, open source health information management system that is extensively used in LMICs for the collection of data on infectious diseases [18]. In 2018, IntraHealth started to collect individual level data on hypertension. Before that, no NCD data was collected in a systematic way in Senegal. In 2021, the MoH decided to integrate NCD (hypertension, diabetes and several other chronic diseases) data collection into DHIS2. It developed its own set of NCD indicators (including hypertension indicators) and set up the technical processes of DHIS2 to enable data entry and management.

3.1.2 Data collection and governance

Blood pressure data collection using DHIS2 was piloted jointly by the MoH and the Cardio4Dakar team in a selected number of health facilities of Dakar during the last quarter of 2021. The MoH also set up a parallel data collection

system and appointed NCD focal points in district health management teams to systematically analyse the data system in collaboration with the local partner organization of CARDIO4Cities. For both Q1 and Q2 2022, a comparison between DHIS2 data and “actively collected” data showed substantial underreporting of cases in DHIS2 although reporting increased over time, suggesting a learning curve for providers. The use of hypertension data for decision-making has remained limited to formulating health-facility specific recommendations for improvement of performance.

The flow chart maps the NCD data collection system following integration into DHIS2 (2022 onwards) (Fig. 2). A list of the most relevant data sources on CV risk factors informing the cascade of care in Dakar is given in Table 1.

3.1.3 Data flow and quality

The primary sources for filling DHIS2 data are paper-based registries with several intermediate aggregation steps. Health workers screen patients for hypertension and enter results into paper-based consultation registries at health facility level. Hypertension data for each indicator noted in the consultation registries is tallied in monthly paper-based reports. The tallied results are eventually reported by indicator in a paper-based monthly SNIS (National Sanitary Service, Service National de l'Information Sanitaire) report, which is then entered into the DHIS2 system. The main issues regarding data quality are accuracy due to the many intermediary steps and completeness.

3.2 São Paulo, Brazil

3.2.1 Background and status of HIS

Public health in São Paulo is largely managed by social organizations (Organização Social de saúde, OSS), that are mandated by and contractually bound to the São Paulo Municipal Health Authorities (SMS-SP). Each OSS has its own HIS. The social network analysis showed that the HIS landscape in São Paulo is extensive and highly fragmented and that there are no communication channels or interlinkages between the different sub-systems used in parallel. To promote interoperability between the systems, the eSaúdeSP application, a digital HIS platform implemented by the SMS-SP, that consolidates all personal health information for an individual user, is being developed locally. eSaúdeSP is part of a bigger project called *Avança Saúde* that aims to modernize the health systems and is funded by the International Development Bank (IDB) in partnership with the São Paulo Municipal Health Department. eSaúdeSP includes an electronic platform that offers patients access to personalized health services and information, and a web-based platform where health professionals can offer remote health assistance. This web-based platform already hosts two tools developed by the CARDIO initiative in São Paulo, namely the cardiovascular risk stratification application and the self-care pact form (Plano de Autocuidado Pactuado—PAP). The SMS-SP in collaboration with the local partner organization of CARDIO4Cities, established a solution to accelerate detection of cardiovascular risk factors. By integrating interoperability, this solution also constitutes a first integration of clinical indicators, outcome values and health system performance indicators.

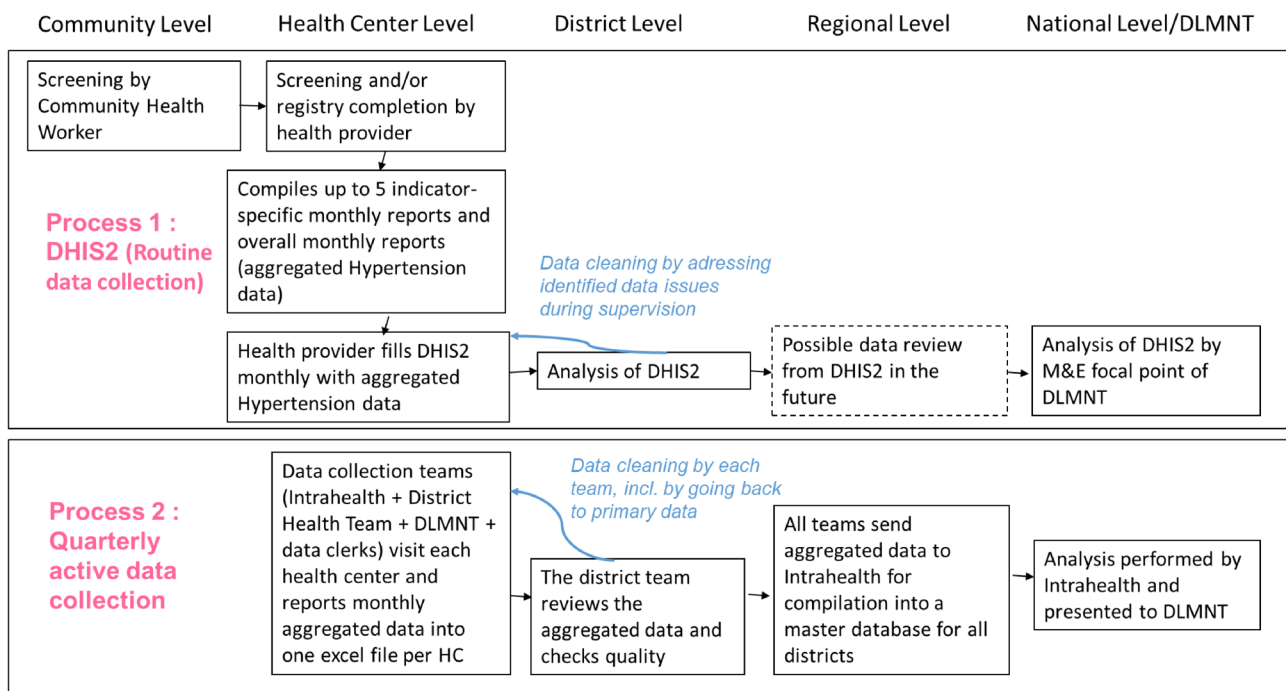
3.2.2 Data collection and governance

Public and private data sources on NCD risk factors that were either directly accessed by the local partner institution or through the Ministry of Health or the São Paulo Municipal Health Department were considered for the flow chart (Fig. 3). They include paper-based medical records and digital applications from primary health centers as well as digital inpatient, mortality and outpatient information systems that are maintained by hospitals. The majority of the data is collected, transmitted and stored electronically, for example in excel files. An integrated healthcare management system called SIGA used by public health care facilities organizes and controls the flow of patients through different levels of the health system. Frequency of reporting and exact definition of indicators of interest (hypertension, blood sugar, LDL-cholesterol, BMI), for example absolute numbers, rates or frequencies, vary as well depending on the system. Different types of software and data formats are used. Data collected by public health facilities and OSS is reported to the federal Ministry of Health or the São Paulo Municipal Health Department. A list of the most relevant data sources and of who owns them is given in Table 2.

Table 1 Overview of most relevant CVD data sources discovered in the research, Dakar

Data sources	Entity collecting	Data collected	Data format/tools	Frequency	Reporting	Data management/access
SNIS / Primary level	Health centre	Individual patient data (eg. demographics, BP level, BMI)	Patient files + Consultation registries	Monthly	Data compiled in to a monthly report and transferred to district level + monthly aggregated data input into DHIS2	SNIS—primary data in health centres; aggregate data in DHIS2
SNIS/DHIS2	Health centre	Aggregated Hypertension/ CVD data by health centre, according to MoH indicators*	Digital	Monthly		SNIS
SNIS	Medical region of Dakar	Aggregated hypertension/ CVD data by health centre, according to MoH indicators	Report	Quarterly		SNIS
Active data collection on hypertension (Cardio-led)	Team (IntraHealth + district health teams + data clerks)	Aggregated hypertension/ CVD data by health centre, according to MoH indicators	Digital (one excel by Health centre) + Excel database (all health centers)	Quarterly	DLMNT	DLMNT
Hospital HIS	Hospitals	Individual patient data (demographics, BP level, BMI, mortality, etc.)	Digital/paper	Monthly	Monthly activity report	Hospital
Laboratories	Laboratories	Test results	Digital/paper		Public labs report, private labs do not report	Laboratories

BMI body mass index, DLMNT Division de la lutte contre les maladies non transmissibles, NCD department of MSAS; SNIS: Service National de l'Information Sanitaire, National Health Information Service



CHW: Community Health Worker; DLMNT: Division de la lutte contre les maladies non transmissibles, NCD department of MSAS

Fig. 2 Routine health data flow in Dakar, Senegal, from 2022 onwards, valid July 2023. CHW Community Health Worker, DLMNT Division de la lutte contre les maladies non transmissibles, NCD department of MSAS

3.2.3 Data flow and quality

The social network analysis listed potentially relevant local data systems, but none of them were interacting or sharing data. A “data lake” is being created within eSaúdeSP with the objective to connect the data across the systems. A data lake is a centralized repository designed to store, process, and secure large amounts of structured, semistructured, and unstructured data [19]. This project was still too preliminary to be reviewed for the present publication. While data was considered to be accurate, consistency was rather low due to several data sources, data was difficult to access, not delivered in time and incomplete. The data mapping exercise found relevant indicators in the local HIS, but some databases had not been updated in the last 8+ years and were therefore not considered useful.

3.3 Ho Chi Minh City, Vietnam

3.3.1 Background and status of HIS

In HCMC, the HCMC Center for Disease Control (HCDC) collects, reviews, and archives NCD data from health facilities at district and provincial level that are under HCDC management. The HCDC reports the data to the HCMC Department of Health (DoH) and simultaneously to the Pasteur Institute. The HCMC Pasteur Institute is a public institution that gives technical guidance on disease control, conducts research and outbreak investigations and trains health staff. It aggregates NCD data from 20 provincial Centres for Disease Control (CDC) and delivers it to the General Department of Preventive Medicine (GDPM) at the MoH, which uses it for monitoring and guidance on NCD prevention and control. The GDPM is a department of the MoH, which specializes on implementing legal regulations in the field of preventive medicine.

3.3.2 Data collection and governance

CVD data on hypertension, blood sugar, LDL-cholesterol and BMI are all at least partially collected in an aggregated format using templates provided by the HCDC. Screening and treatment data on CVDs are collected in Commune Health Stations

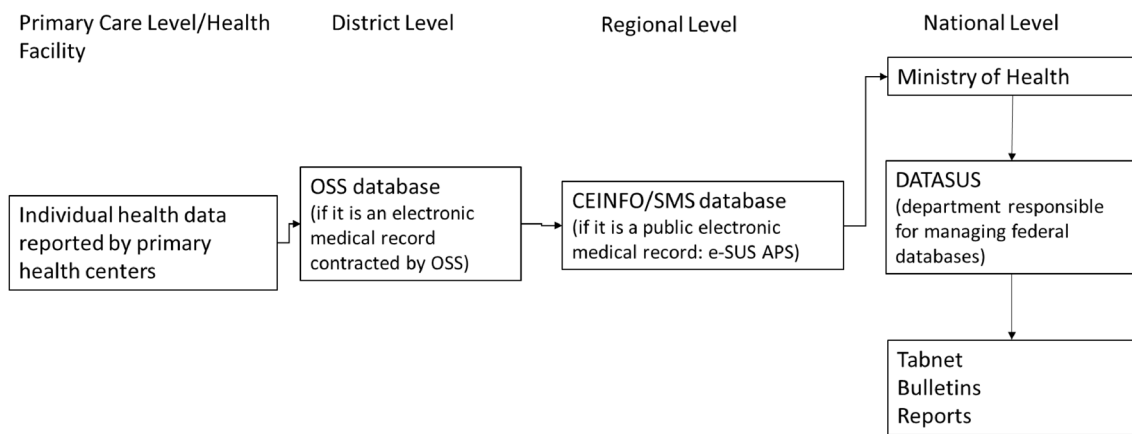
Table 2 Overview of most relevant NCD data sources discovered in the research, São Paulo

Tool provided by	Data sources	Entity collecting	Data collected	Data format/tools	Frequency
City of São Paulo	Medical records (paper)	Primary health center	Constant sample of individual patient data (eg, BP level, diagnosis)	Paper medical record	Monthly (IQVIA)
	Risk stratification app	Primary health center, to be integrated into eSaúde	Cardiovascular risk score	Digital, restricted (Mobile application)	Anytime the risk stratification app is used
	SIH (Hospital Information System)	Public Hospitals/Establishments that perform hospitalizations	Collection of individual and organization level information related to hospital admissions	Digital (Public Website)	Anytime there is a hospitalization
	SIM (Mortality Information System)	Public Hospitals/Establishments	Death declaration and cause of death (incl. ICD codes)	Death certificates, online public website	Anytime there is a death
	SIA (Outpatient Information System)	Outpatient Production Bulletin (Ministry of Health—Federal)	Information related to care provided by outpatient clinics financed by the SUS	Medical records and form when notification is needed to SIA	Anytime relevant event takes place
	SIGA (Integrated Healthcare Management System)	Primary health center and Public Hospitals/Establishments	Organization and control the flow of patients, regulating access to health services (eg scheduling appointment/exams) and recording care across levels	Different types of format, depends on the public health facility and managing "social organizations"	Anytime patient appointments and exams are scheduled
	Pact for self-care (PAP)	Primary health center	Patient goals, and Usage of PAP	PAP card (paper/ digital)	Anytime a patient fills out a PAP card
	TABNET São Paulo	Secretaria Municipal da Saúde	Data across various health information system from SUS including SIH and SIM	Online public website	Monthly to quarterly
	eSaúde app	Primary health center and Public Health Establishments; Patient, Avançada Saude Program	Collects patient information directly and from various data systems (pregnancy data currently but expanding)	Mobile application and website	Anytime data is entered by patient or health care provider
	City of São Paulo/CSO (Civil Society Organization)—Santa Marcelina	Santa Marcelina data system	Primary health center funded by Santa Marcelina	Individual level patient data on hypertension	Excel file

Table 2 (continued)

Tool provided by	Data sources	Entity collecting	Data collected	Data format/tools	Frequency
Tellus	Screening corner	Primary health center, to be integrated into eSaude	Individual patient data (demographics, blood pressure level, BMI, etc.)	Screening corner	Anytime a screening corner at the health facility is visited
	Tellus internal control	Tellus	Data on implementation activities, eg outreach, HCP training	Excel	Monthly
	Tellus management dash-board*	Health unit manager and Tomara via survey/inter-view + automatically via medbem database	Monitoring the implementation of the Cuidando de Todos Protocol via: Screening corner data; Use of dashboard	Google data studio, google forms	Monthly
UMANE	UMANE observatory dash-board	UMANE	Use of dashboard; Primary health center management of patients Vigitel—Ministry of Health	Interactive dashboard	Yearly

BMI body mass index, *HCP* healthcare professionals, */CD* international classification of disease, *SiH* hospital information system, *SiM* mortality information system, *SiA* outpatient information system, *SIGA* integrated healthcare management system, *SUS* unified health system (sistema único de saúde), *CSO* civil society organization

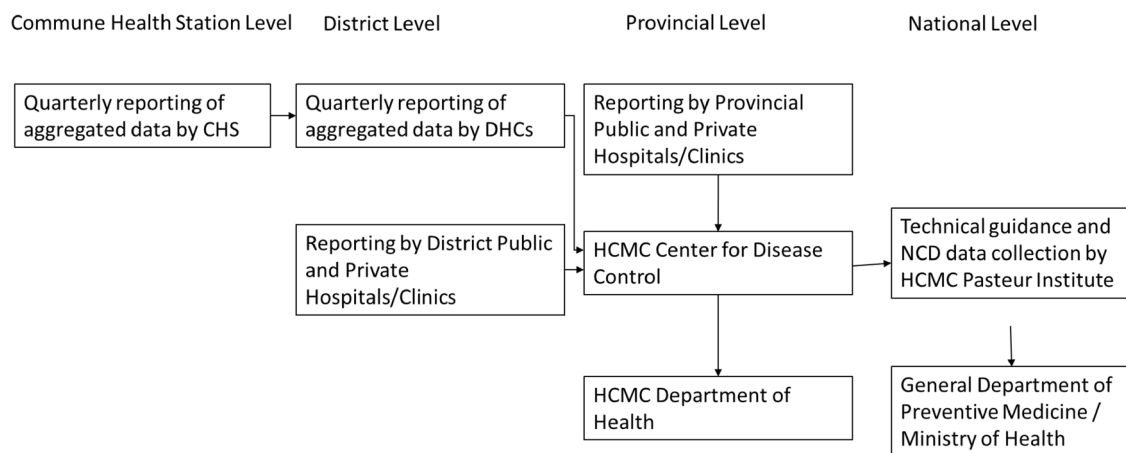


APS: Primary Care (Atenção Primária), OSS: social health organizations (organizações sociais de saúde), CEINFO: epidemiology and information coordination (coordenação de epidemiologia e informação); SMS: municipal health department (secretaria municipal de saúde)

Fig. 3 Health data flow in São Paulo, before applying the CARDIO approach, valid until the implementation of eSaúde (planned for 2024) and the Screening Corner (April 2023). APS Primary Care (Atenção Primária), OSS social health organizations (organizações sociais de saúde), CEINFO epidemiology and information coordination (coordenação de epidemiologia e informação), SMS municipal health department (secretaria municipal de saúde)

(CHS) and reported in an aggregated format to the district health centres (DHC) on a quarterly basis for reviewing, merging with their service data and reporting to the HCDC. Public and private hospitals and clinics at provincial and district levels also report their data to the HCDC including screening, diagnosis, treatment and control data (Fig. 4). However, reporting of CVD data from hospitals and clinics to the HCDC relies on the willingness of the staff. The CHSs, DHCs, and the hospitals/clinics own the data they generate. A list of the most relevant NCD data sources in HCMC is given in Table 3.

The HCDC is mandated by the government to carry out regular data collection and supervision visits in all CHSs. An official dispatch issued by the GDPM provides guidance for NCD prevention and control including screening, diagnosis, treatment, and mortality.



CHS: commune health station; DHC: district health centre; HCMC: Ho Chi Minh City; NCD: non-communicable diseases

Fig. 4 Health data flow in Ho Chi Minh City, Vietnam, valid July 2023. CHS commune health station, DHC district health centre, HCMC Ho Chi Minh City, NCD non-communicable diseases

Table 3 Overview of most relevant CVD data sources discovered in the research, HCMC

Data sources	Entity collecting	Data collected	Data format/tools	Frequency	Reporting	Data management/access
Electronic Health Records EHR	Health workers, interns		Electronic, individual level	Retrospective data, updated any time a health facility is visited	No data auto-sharing between CHSs and DHs EHR and EMR are not connected yet	CHS, DHC (access)
CHS HIS	Health workers	BMI, Blood pressure, Blood sugar, and other lab tests if performed	Entry mask individual level	Blood pressure is captured every time a health facility is visited. Health information of hypertension/DM patients are captured on a needs basis	Quarterly to DHC in an aggregated format	CHS (mgmt) CHS, DHC (access)
CHS logbook	Health workers	BMI, Blood pressure, blood sugar, and other lab tests if performed	Excel Individual level	Anytime a health facility is visited (except BMI is only captured in the first visit) Health information of hypertension/DM patients are captured on a needs basis	Quarterly to DHC, aggregated	CHS
DHC and DH HIS	Health workers	BMI, lab test results, blood pressure preventive and treatment data	Individual level	Anytime a health facility is visited	Quarterly to HCDC in an aggregated format if DHs/DHCs deploy any NCD prevention and control activities	DHC, DH (Mgmt)

BMI body mass index, CHS commune health station, DHC district health centre, DH department of health, HCDC Ho Chi Minh City centre for disease control, HIS health information system, NCD non-communicable diseases

3.3.3 Data flow and quality

Depending on the district, a distinct paper-based form or software is used for data collection of individual patient data. For reporting, data is aggregated at the level of the health facility and passed on to the HCDC in a standardized format, thereby harmonizing HIS formats by requiring the health workers of each facility to aggregate the data from their HIS into the required format. This process is error-prone as there are no guidelines and no regular supervision. Data quality was not systematically assessed in HCMC as no data was collected there.

3.4 Comparative summary

Table 4 provides an overview of current processes, needs and challenges regarding routine data collection in the three cities based on the HIS assessment. It shows which of the key indicators outlined by CARDIO are not routinely collected by the current system. Challenges included for example incomplete data sources, redundant indicators, fragmented systems, overlapping processes and poor data quality. Although the situation is different in each location, there are similarities, for example, data on LDL-cholesterol and BMI are unavailable key indicators everywhere. This comparison served as a basis to inform data collection activities.

4 Discussion

This manuscript presents an approach of assessing the most relevant data processes and HIS related to NCD risk factors in three cities. The similarities between the three settings were limited. In all three locations, the focus of currently existing NCD data collection activities were on hypertension with a secondary emphasis on diabetes. Data quality was usually low for a variety of reasons such as multiple, often manual processing steps, incomplete data sources or parallel systems with differing indicator definitions.

4.1 Comparison across cities

In Dakar, the MoH has adopted an active role in promoting data collection to monitor disease indicators over time, although initially this pertained mainly to infectious disease indicators. Using lessons learnt from the data collection on hypertension indicators, the government actively conducted NCD data integration into the existing DHIS2 system. Thus, in Dakar, the initiative has been promoting and supporting an ongoing process that is led by the government and aims to integrate NCD data collection into a centralized system. This represents an ideal setting for establishing a sustainable, uniform HIS that is led by the authorities and supported by the CARDIO4Cities initiative and through the local partner organization. A possibility for data harmonization is the recently introduced e-tracker which may, in the future, allow electronic hypertension patient files to be established nationwide (personal communication by local partner). The e-tracker is the DHIS2 app for individual-level or case-based transactional data. It supports data collection, case monitoring and follow-up, analysis and reporting within the DHIS2 system.

In São Paulo, different public and private stakeholders are involved at all levels of the health system. Many different data collection systems for CVD data exist, but no interoperability layers or communication channels could be identified. Statements by local experts emphasized that the commitment and support of the SMS-SP is needed to integrate and streamline the different systems. The highly fragmented governance of the different sub-systems, each managed by a different social organization, makes this integration particularly challenging. ESaúdeSP shows great potential for serving as an interoperable tool to integrate different HISs as one of its elements is a clinical repository that collects patient information directly from various data systems in real-time. ESaúdeSP has already been widely adopted to manage the COVID-19 and other vaccination programs, and is the main HIS for neonatal care in the municipality. Its integrated functionality of telehealth appointments that allow patients to access medical care remotely from a smartphone application is expected to drive adoption of the eSaúdeSP platform even further (personal communication by local partner).

In HCMC, the HIS constitutes of several parallel paper-based and electronic systems. At the CHS level, different formats for data collection through HIS, logbooks, and reporting templates cause challenges for data reconciliation. Depending on the district, a distinct paper-based form or software is used for data collection of individual patient data. Collection of aggregated data is organized and carried out by HCDC staff, making government support indispensable when it comes to harmonization. This centralized system with the HCDC driving and controlling all major decisions seems predestined

Table 4 Overview of routine health information system and CARDIO data collection, including future data strategy and key challenges

	Dakar	São Paulo	Ho Chi Minh City
Most relevant source for routine data collection on CV risk factors	DHIS2	SIGA (partial)	Aggregated data reported to district, regional, and national level
Key challenges regarding routine data collection	Low quality of hypertension data in the centralized health information system	Numerous and incomplete data sources Parallel data collection systems	Data aggregation at health facility level is highly error-prone
Missing key indicators	LDL-cholesterol and BMI	LDL-cholesterol and BMI	LDL-cholesterol and BMI
Key challenges regarding interoperability and sustainability	High staff turnover creating ongoing training needs for data reporting	Fragmented health information system: inconsistent indicator definitions and reporting processes	Fragmented health information system: error-prone data aggregation for reporting

to implement a uniform electronic data collection tool across all districts. Aligning forms and software would lead to a better overview, easier data flow and standardized data exchange as well as ensuring comparability of data.

4.2 Key challenges: data quality and interoperability

To enable usage of an HIS as an integral tool to inform policy or population health decision-making requires different elements to be in place [3]. Data quality and interoperability are the ones that stand out most when comparing across the three case studies.

4.2.1 Data quality

Even with routine data collection systems in place in Dakar and HCMC, data quality was generally considered to be low. Quality issues originated from incorrect, missing and underreported data and are connected to the misunderstanding of indicators, incomplete data input forms or the lack of a systematic data routine in the workflow. In some cases the primary sources were paper-based registries which require several aggregation steps. Even if data was aggregated at the level of the health facility and passed on in a standardized format, manual aggregation led to errors resulting in lack of trust in the data.

The introduction of a short-term parallel approach to data collection and standardized aggregation may be an opportunity to address data quality issues. However, these parallel systems for comparison should be short-term only and used to enable sustainable long-term solutions through a capacity building and change process.

Capacity building in combination with establishing clear data governance and accountability is crucial to improve data quality [7, 20]. High staff turnover may lead to loss of skills and knowledge, making regular refresher trainings indispensable [21]. Concrete and practical approaches for data collection, aggregation and reporting as well as data quality control should be included in trainings. Clear governance and accountability for data quality needs to be reinforced. This may include regular supervision visits, assignment of specific responsibilities to the team members, with clear visibility of impact of data quality and a joint understanding and overview over the importance for the health system [14]. Additional human and financial resources need to be considered when adding responsibilities to already stretched health staff. It may also be necessary to adjust job descriptions or add incentives to perform additional tasks. Replacing paper-based processes with digital technologies can help to improve data quality by avoiding manual processing steps and media disruption, which both may cause errors, delays and additional costs. The implementation of an electronic innovation/software needs to be closely coordinated with existing work practices [18]. A transition period where paper-based and electronic processes are used in parallel may increase the workload, requiring more resources [14]. Additionally, electronic processes may introduce new threats to patient safety through security leaks, which must be considered during development and implementation [14].

4.2.2 Interoperability

According to the Healthcare Information and Management System Society, interoperability is the ability of different information systems, devices and applications to access, exchange, integrate and cooperatively use data in a coordinated manner [22]. Fragmented systems have often emerged in response to a variety of concrete needs without being rooted in an overarching long-term strategic design leading to limited interoperability. Indicator definitions and data collection processes can differ. Harmonization of the HIS may provide a solid basis for health management decision-making and can support improvements in patient and health system outcomes [4]. It may also include the integration into one interoperability layer. Different electronic databases can be linked and the data collated [4, 20]. Harmonization facilitates data collection and analysis of health data [4], and is critical to getting an overview of population health for decision-making in an efficient manner [20]. One single interoperability layer prevents information gaps and reduces redundant data collection and costs [23]. Fast Healthcare Interoperability Resources (FHIR) may help standardize data to facilitate interoperability [24]. It would also be advisable to develop a holistic approach to achieve data interoperability linking high quality data from heterogeneous sources, similar to the mechanism that has previously been described for data from IoT medical devices [25]. Regarding data access and ownership, it is key to define who needs the data for decision-making. This also asks for communication and collaboration between different actors of the public and private sector, depending on the nature of local HIS providers [26].

There are some limitations to this work worth mentioning: No data was collected in HCMC and therefore data quality was not systematically assessed there. Due to the wide fragmentation of data sources in São Paulo, the list provided may be incomplete. Including stakeholders from diverse backgrounds (for example health workers, governmental and non-governmental organizations involved in the health sector, digital health specialists) helps to create a list that is as complete as possible. Applying this approach in another location may yield different results from the cities mentioned here as HIS are very diverse. Minimizing assumptions in the beginning is necessary to avoid preconceived opinions and miss out on unexpected findings.

We started by defining common, clear and mutually agreed upon indicators that were adapted to local contexts [15]. They served as references for benchmarking the documented HIS in terms of data sources and format, collection and reporting processes as well as ownership and storage. A platform such as the newly developed CrowdHEALTH that uses big data management mechanisms may help to analyze the potential for interoperability between systems and provide solutions for data harmonization [27]. To lay the foundation for sustainable health interventions such as emerging digital technologies and design new health policies, the PolicyCLOUD, an extensible framework that allows the exploitation and management of policy-relevant data flows may serve as a supporting tool [28].

5 Conclusions and recommendations

Based on the presented HIS case studies, we recommend some basic principles for future investigations of HIS in the context of population health and data-based decision making. We direct our recommendations at various potentially interested target groups: for example scientists in the field of health systems research or digital transformation, NGOs working on local health systems or on NCDs, or policy makers and governmental institutions interested in refining their NCD strategies and looking to perform a baseline assessment of their current HIS.

Local, real-life knowledge—The local context needs to be examined and current processes mapped and verified in an iterative approach. Key to understanding the HIS in any setting is an extensive and solid network of contacts among involved local stakeholders. Not only are local idiosyncrasies often not obvious to external observers but the actual situation may not correspond to the official set-up. Local collaborators are needed to contextualize and verify any assembled information.

Visualizing findings—To visualize interconnections in fragmented HIS brings transparency and a possibility to discuss with stakeholders. Interoperability layers should be explored to facilitate data exchange and lower costs. Using Process Mapping Tools such as Bizagi is helpful for visualizing complex systems as well as for identifying interoperability layers and bottlenecks.

Political support—The support of government entities is needed for system integration and driving data-based decision making. Gauging the interest of responsible government agencies early on is advantageous for a thorough assessment, especially if structures are centralized. Readiness for aligning to methods already in place to strengthen the HIS is important for a successful collaboration.

Concluding from the evidence presented here, optimizing the use of data for decision-making in health systems remains a case-by-case exercise. A harmonized HIS providing high quality data that is used by authorities for decision-making may seem like the gold standard in any context, but the extent to which this is feasible and desirable as well as the steps taken to reach this goal can differ greatly from place to place.

Acknowledgements We thank Sabrina Simon and Marta Palmeirim for their support in data collection and interpretation.

Author contributions SR wrote the manuscript; FS, DC and AS contributed to the writing, tables and figures and revised the manuscript; JoBa, KS, MS, TN contributed to the data collection and revised the manuscript; JoBo, SdR, TR, PS revised the manuscript. All authors have approved the final manuscript.

Funding Open access funding provided by University of Basel. This study was funded by the Novartis Foundation.

Data availability All data supporting the findings of this study are available within the paper.

Code availability Not applicable.

Declarations

Ethics approval and consent to participate This study does not involve human subjects data. Consequently, according to Swiss law, ethical approval is not deemed necessary for this study.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. WHO. Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies. Geneva: World Health Organization; 2010.
2. WHO. Global diffusion of eHealth: making universal health coverage achievable. Report of the third global survey on eHealth. Geneva: World Health Organization; 2016.
3. WHO. Reinforcing health information systems. World Health Organization Regional Office for the Eastern Mediterranean. Geneva: World Health Organization; 2014.
4. Schmidt BM, Colvin CJ, Hohlfeld A, Leon N. Definitions, components and processes of data harmonisation in healthcare: a scoping review. *BMC Med Inform Decis Mak.* 2020;20:222. <https://doi.org/10.1186/s12911-020-01218-7>.
5. Mumbere O, Kopi L. How to conduct a data quality assessment (DQA): an aid memoir for a COR/AOR. USAID; 2012.
6. Leon N, Balakrishna Y, Hohlfeld A, Odendaal WA, Schmidt BM, Zweigenthal V, et al. Routine Health Information System (RHIS) improvements for strengthened health system management. *Cochrane Database Syst Rev.* 2020;8:CD012012. <https://doi.org/10.1002/14651858.CD012012.pub2>.
7. Lemma S, Janson A, Persson LA, Wickremasinghe D, Kallestal C. Improving quality and use of routine health information system data in low- and middle-income countries: a scoping review. *PLoS ONE.* 2020;15:e0239683. <https://doi.org/10.1371/journal.pone.0239683>.
8. Roth GA, Mensah GA, Fuster V. The global burden of cardiovascular diseases and risks: a compass for global action. *J Am Coll Cardiol.* 2020;76:2980–1. <https://doi.org/10.1016/j.jacc.2020.11.021>.
9. WHO. Cardiovascular diseases (CVDs) key facts. Geneva: World Health Organization; 2023.
10. Vaduganathan M, Mensah GA, Turco JV, Fuster V, Roth GA. The global burden of cardiovascular diseases and risk: a compass for future health. *J Am Coll Cardiol.* 2022;80:2361–71. <https://doi.org/10.1016/j.jacc.2022.11.005>.
11. Aerts A, Boufford J. A new whole-of-city strategy for addressing cardiovascular population health. *Cities & Health.* 2021. <https://doi.org/10.1080/23748834.2021.1979774>.
12. Boch J, Venkitachalam L, Santana A, Jones O, Reiker T, Rosiers SD, et al. Implementing a multisector public-private partnership to improve urban hypertension management in low-and middle- income countries. *BMC Public Health.* 2022;22:2379. <https://doi.org/10.1186/s12889-022-14833-y>.
13. Reiker T, Des Rosiers S, Boch J, Partha G, Venkitachalam L, Santana A, et al. Population health impact and economic evaluation of the CARDIO4Cities approach to improve urban hypertension management. *PLOS Glob Public Health.* 2023;3:e0001480. <https://doi.org/10.1371/journal.pgph.0001480>.
14. Cresswell KM, Bates DW, Sheikh A. Ten key considerations for the successful implementation and adoption of large-scale health information technology. *J Am Med Inform Assoc.* 2013;20:e9–13. <https://doi.org/10.1136/amiajnl-2013-001684>.
15. Saric J, Aerts A, Barboza J, Boch J, Cobos D, Rajkumar S, et al. CARDIO4Cities: from global evaluation framework to local monitoring in Dakar and São Paulo. submitted to *BMC Public Health.* 2024.
16. Nielsen P. Advancing Health Information Systems: Experiences from Implementing DHIS 2 in Africa; 2012.
17. Sané I. Utilisation à grande échelle du logiciel Dhis2 : Le Sénégal dans l'ère révolutionnaire de la gestion de l'information sanitaire. Osiris; 2016.
18. Byrne E, Sæbø JI. Routine use of DHIS2 data: a scoping review. *BMC Health Serv Res.* 2022;22:1234. <https://doi.org/10.1186/s12913-022-08598-8>.
19. Google Cloud. What is a Data Lake?: Google.
20. Stansfield S, Walsh J, Prata N, Evans T. Information to Improve Decision Making for Health. 2006.
21. Jones C, Gates M. The costs and benefits of nurse turnover: a business case for nurse retention. *Online J Issues Nurs.* 2007. <https://doi.org/10.3912/OJIN.Vol12No03Man04>.
22. HIMSS. Interoperability in healthcare. Chicago: Healthcare Information and Management System Society; 2023.
23. Sligo J, Gauld R, Roberts V, Villa L. A literature review for large-scale health information system project planning, implementation and evaluation. *Int J Med Inform.* 2017;97:86–97. <https://doi.org/10.1016/j.ijmedinf.2016.09.007>.

24. Vorisek CN, Lehne M, Klopfenstein SAI, Mayer PJ, Bartschke A, Haese T, et al. Fast Healthcare Interoperability Resources (FHIR) for interoperability in health research: systematic review. *JMIR Med Inform.* 2022;10:e35724. <https://doi.org/10.2196/35724>.
25. Mavrogiorgou A, Kiourtis A, Perakis K, Pitsios S, Kyriazis D. IoT in healthcare: achieving interoperability of high-quality data acquired by IoT medical devices. *Sensors.* 2019. <https://doi.org/10.3390/s19091978>.
26. Nabyonga-Orem J. Monitoring Sustainable Development Goal 3: how ready are the health information systems in low-income and middle-income countries? *BMJ Glob Health.* 2017;2:e000433. <https://doi.org/10.1136/bmjgh-2017-000433>.
27. Kyriazis D, Autexier S, Boniface M, Engen V, Jimenez-Peris R, Jordan B, et al. The CrowdHEALTH project and the Hollistic health records: collective wisdom driving public health policies. *Acta Inform Med.* 2019;27:369–73. <https://doi.org/10.5455/aim.2019.27.369-373>.
28. Biran O, Feder O, Moatti Y, Kiourtis A, Kyriazis D, Manias G, et al. PolicyCLOUD: a prototype of a cloud serverless ecosystem for policy analytics. *Data Policy.* 2022;4:e44. <https://doi.org/10.1017/dap.2022.32>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.