# **WP8: SALURBAL case study**

**Discipline-specific or cross-discipline area**

Urban Health is considered a domain of multiple disciplines involved in understanding how urban environment contributes to shape population health and the social inequalities related to health disparities among city dwellers.

For example, the role of cities as drivers of human activity and wellbeing requires urban planners’ perspectives on how physical environment and services in cities are designed and planned. Social scientist’s knowledge is also necessary in urban health to bring insights on people’s behaviors and interactions that contributes to their development and wellbeing. Epidemiologists’ viewpoints contribute to the documentation of the burden of disease in urban settings and the factors associated with those population disease patterns.

Therefore, urban health requires the interconnection of knowledges and practices from different disciplines to depict the complexity of urban systems. This poses several challenges in the study of urban health such as the need of a common vocabulary (i.e., how we define urban areas, informal settlements); the use of data from different sources (i.e., place based and spatial data, health registries); the creation of data that could be comparable across urban areas and over time (i.e accounting for differences between cities and within cities over time), among others (Quistberg 2019).

These challenges are related to the principles of FAIR (Findable, Accessible, Interoperable, and Reusable) and CARE (Collective benefit, Authority to control, Responsibility, and Ethics) data policies and practices (Wilkinson 2016; Russo Carroll 2021).

**Organisation coordinating disciplinary or cross-disciplinary activities**.

The SALURBAL project (SALud URBana en America Latina, Urban Health for Latin American cities) is a five-year project based at the Urban Health Collaborative - Drexel University and with partners throughout Latin America and in the United States. SALURBAL aims to study how urban environments and urban policies impact the health of residents from almost 370 cities in 11 Latin American countries. Its findings inform policies and interventions to create healthier, more equitable, and more sustainable cities world-wide.

SALURBAL specific aims are

1. To quantify the contributions of city and neighbourhood-level factors to differences in levels of health and health inequalities among and within cities.
2. To evaluate the health and environmental impact of city and neighbourhood-level policies and interventions by capitalizing on natural experiments.
3. To employ systems thinking and formal systems simulation models to better understand the dynamic relations between the urban environment, health and environmental sustainability and identify the plausible impacts of selected policies.
4. To engage with the scientific community, the public, and policy makers to disseminate findings and translate them into policies and interventions.

The Urban Health Collaborative (UHC) is a research center that is part of the Dornsife School of Public Health at Drexel University in Philadelphia, USA. The UHC is Directed by Dean Diez Roux and Co-Directed by Dornsife Associate Professor Gina Lovasi. The organizational structure of the UHC includes an overall administrative core as well as a research and data core, a training core, and a policy and community engagement core. The UHC Data and Research Core compiles and analyzes data on health in cities locally, nationally, and internationally. It also provides statistical consultation on a range of analytical issues.

This case study will engage Drexel- based researchers working for the SALURBAL project. However, the full SALURBAL consortium is involved in the integration of data and the expertise detailed below.

**How the case study will support international cooperation on the FAIRness of both data & digital objects in the discipline proposed.**

To study drivers of urban health, health equity, and environmental sustainability, the SALURBAL project created - and continue expanding and improving- a data platform that flexibly linked various type of data and allows between and within- city comparisons (Quistberg 2019).

During this process the SALURBAL project acquired expertise in:

* + identifying and defining the geographical units and subunits that constitutes the universe of cities included in the study
  + collecting, processing, and harmonizing health data (from national vital statistics registries and national health surveys)
  + creating and updating data on the physical and social environment in cities
  + integrating all available information within a multilevel data structure that allowed definition and measurement of constructs at different geographical levels and over time

These activities represent practical examples on data management and stewardship that could support FAIR and CARE principles on data provenance and contribute to promote best practices in data sharing and use. It could serve as valuable lessons in the process of creating similar data platforms for urban health research and practice, particularly in highly urban or rapidly urbanizing regions of the Global South, where integration of data is still an incipient process (Gatzweiler 2020).

Moreover, lessons learned from the SALURBAL experience on data compilation, processing, and integration could be translated to urban data systems beyond research and academia, supporting the establishment of guidelines and recommendations in the development and implementation of information systems oriented to climate crisis and further epidemic preparedness in urban settings (Gatzweiler 2020)

**What are the current initiatives and best practices, globally, within the target scientific discipline, and to what extent can work with this field facilitate the exchange of best practices across disciplines.**

Cities are considered the primary contributors to global environmental change and human development, being at the center of leading mitigation and adaptation strategies that could promote human health along with environmental sustainability (Liu 2020)

Urban health is therefore directly connected to other fields involved in studying the interlinkages between human health, urbanization, and environmental degradation -such as Planetary Health- as well as those disciplines focusing on governance system transformations that integrate social, economic, and environmental policies that could contribute to equitable and sustainable solutions for humans and the shared environment (Pineo 2020).

Given the transdisciplinary approach of Urban Health, challenges faced within this field are also common to other areas. In the same way, solutions proposed from the Urban Health perspective could also promote advancement beyond its discipline.

A recent publication that emerged from a knowledge interchange experience between researchers in Urban and Planetary Health, identified knowledge sharing, capacity building, and trust-building as three critical principles to foster effective collaboration across and beyond disciplines and to pursuing valuable solutions for complex problems in cities (Pineo 2020).

Many stakeholders from different disciplines and levels of governance are involved in the creation, maintenance and dissemination of data used for Urban Health research. Some of the challenges related to knowledge sharing are:

* data acquisition and use
* availability of data at spatial scales and levels of aggregation necessary for studying problems in cities
* integration of data from different sources using different levels of spatial reference depending on the definition of urban setting used
* disparities in quality and completeness of data over time and across geographies
* adequate standardisation processes to make data comparable (across cities within and between countries and regions)

The use of complex data requires also that researchers and other actors could work across disciplines, contexts and sectors recognizing the diversity of knowledge systems that pertains to each field while generating common grounds as the base of collaborative work. Understanding the process involved in the generation of data (data lifecycle) as well as the acknowledgment of the different jargon or operational languages used in data are part of the capacities necessary to develop within and beyond Urban Health for the correct interpretation and interaction with data in research and evidence translation. These capacities are also essential part in the creation of accountability and bridging the diversity of backgrounds, knowledge, and experience between actors from different sectors in order to promote trust- building among them.

**Provide a detailed description of the case study that will be implemented. Include details on both domain-specific and domain-independent recommendations in FAIR practice (from the Research Data Alliance [RDA], the Committee on Data of the International Science Council [CODATA], or from other relevant organisations).**

SALURBAL is a five-year project funded since 2017 by Wellcome Trust Fund, as part of the initiative ‘Our Planet, Our Health’. SALURBAL have worked in the collection, integration, and use of urban health data providing, for the first time, information on the heterogeneity within and between cities in almost 370 cities of 100,000 residents or more in 11 Latin American countries (Argentina, Brazil, Chile, Colombia, Costa Rica, Guatemala, El Salvador, Mexico, Nicaragua, Panama, and Peru).

Through this work, SALURBAL has addressed many of the challenges exposed in previous sections.

* The team has systematized a process for city- definition and operationalization that integrated multiple ways in which a city can be delimited: based on country-specific definitions of metropolitan areas, based on administrative boundaries linked to censuses, or based on urban footprints assessed through data-driven efforts (Diez Roux 2018).
* The various and integrated city definitions permitted the creation of a data structure with several levels of aggregation and spatial definitions that allowed the incorporation of data from different sources, making it shareable across several cores and disciplines which enabled the examination of research questions depicting the complexity of urban health problems (Quistberg 2019).
* The management of issues related to data access, quality, and completeness during the process of data harmonization derived in the development of procedures and standards that are systematically documented. This is an ongoing work that intends to bring accountability on the data lifecycle, granting data replicability and establishing recommendations for best use of this data.

SALURBAL experience on balancing aspirations with practical realities have shown that it still possible to generate evidence and potential solutions to complex urban health issues while working with imperfect data. Moreover, the insights provided by rigorous analyses with imperfect data contribute to show the gaps and needs in data use and management that ultimately helps to further improve data life cycle.

This expertise gained by SALURBAL team could support the implementation of FAIR (Findable, Accessible, Interoperable, Reusable) data principles within and beyond the Urban Health field towards an Open Science (Wilkinson 2016; Cox 2020). Actions developed by SALURBAL on city- definition and data integration are practical examples of interoperability (‘I’): the generation of consensus on city terminologies and other vocabularies across domains in the project could help other data systems in cities to adopt similar languages, facilitating in the future the use of data for research and action. The elaboration and documentation of standard procedures for data collection, processing and harmonization contributes to the accessibility and reuse (‘A’, ‘R’) of data, as this brings transparency and authentication in data use, and could stimulate replicability of these processes using similar data in other parts of the world.

Moreover, since SALURBAL gathers urban health data from a vast part of the Latin American region, data documentation made at SALURBAL could contribute to initiate a dialogue and reflection about CARE (Collective benefit, Authority to control, Responsibility, and Ethics) principles and practice among stakeholders involved in data lifecycle (Russo Carroll 2021), bringing a perspective of transparency, integrity, equity and ethics towards minority and vulnerable population in the implementation of open data in Latin America and other parts of the Global South.

Some of the potential contributions of SALURBAL to the FAIR and CARE principles have been already materialized in previous collaborations with CODATA and other stakeholders.

A recent publication has detailed the initial city definition and data integration process (*Quistberg 2019*). This has been also disseminated through a [data brief](https://drexel.edu/~/media/Files/lac/Data/Data%20Brief_ENG_FINAL_web.ashx?la=en) that was presented in several SALURBAL activities, including the [Urban Health Workshop for Latin América and the Caribbean](https://council.science/events/first-latin-american-and-caribbean-urban-health-workshop/), in which CODATA Director was one of the main partners. In a similar way, SALURBAL and CODATA envision to create a documentation on the process of National Health Surveys harmonization that include recommendations of FAIR practice and could guide governments carrying out these surveys in future data collection.

The ethical dimensions of data harmonization and the strategies for responding to them in relation to data provenance and integrity have been introduced in previous [workshop](https://vimeo.com/user91439529/review/482787542/7c134b8fd2?sort=lastUserActionEventDate&direction=desc) at the FAIR Convergence Symposium 2020 and in a [podcast](https://crdf.org.in/podcast/data-knowledge-action-for-urban-systems) within the Data- Action Knowledge System Series at CEPT Research and Development foundation, in collaboration with CODATA Collaborator Dr Theresa Anderson. SALURBAL and CODATA are interested in continuing and expanding this discussion process through the creation of second workshop and the elaboration of a document that integrates the experiences shared by its participants that would lead to the creation of guidelines and training in best practices for data provenance.

**Please provide a brief overview of how you would develop, pilot and possibly deploy interoperability standards and guidelines for increasing FAIRness in your scientific disciplines as part of the work of this project**

In this case study, SALURBAL and CODATA teams seek to further enrich the application of FAIR principles by evaluating the challenges of abstraction throughout the lifecycles of data collection and analysis through the lens of the CARE principles. By foregrounding concerns about the vulnerability of data to misinterpretation, misuse, and misappropriation, the team seeks to integrate FAIR and CARE principles in ways that can be applied to large-scale cross-domain projects. Building on work shaped through previous workshop, we will start with core concepts related to the ecology of visible and invisible work described by Star and Strauss (1999) and further developed by Dr Theresa Anderson (2020). While no data is inherently either visible or invisible, degrees of visibility can be identified in any given context to create awareness of ways relationships between the visible and invisible negotiate how and what we “see” as data at any point in time.

This CODATA-SALURBAL collaboration would pilot a process for enhancing empathic and ethical ways of thinking about analysing data in urban health data systems and projects. Shedding light on bias that may occur during the process of data collection and harmonization promotes an enriched notion of FAIR approaches. The ultimate intention will be the translation of the lessons learnt in the pilot into online training, demonstrating ways data cleaning and data harmonisation can become analytical tools to uncover data ‘absences’ and assess their implications. By offering practical examples on cross-domain challenges faced over the course of the SALURBAL project, it is expected to set initial guidelines for FAIR practice across researchers, local governments and stakeholders involved in global regulations.

**Provide some details on potential assessment and evaluation methodologies to appraise FAIRness within this discipline, that could be further developed or emerge as part of the activity proposed.**

Previous and proposed activities by SALURBAL and CODATA contribute to the improvement of data use and management within complex, transdisciplinary projects. Cross-domain data projects worldwide face challenges associated with data provenance in relation to the processes of sharing and harmonising the data collected that might potentially - though unintentionally - contribute to inequity and bias. By examining the ethical dimensions of data harmonisation and strategies for responding to the challenges these dimensions present, our project seeks to provide guidance that can be applied in a range of situations.

Surfacing considerations of who or what might be missing, misrepresented, or underrepresented in our data builds assurances to help look after the most vulnerable members of a community. This case study would use a framework of keystone practices for working with community data (Anderson 2020) as the starting point for assessing these practices.

The framework and evaluation methodologies shaped through this project will involve data compilations linked to cities, which can have far-reaching benefits for research, policy, and information systems. We believe that the approach we are taking is particularly relevant in the rapid urbanizing regions of the Global South where the decolonization of data practices is especially needed. It is also aligned with some initiatives taking place in Europe, such as  [CitieS Health](https://citieshealth.eu/about/) or [The Institute for Future Cities](http://ifuturecities.com/) which encourage citizens participation in the collection of urban data.

**Please outline the expected impact on your discipline as a result of this activity. What will be the added value?**

An enhanced awareness on the life cycle and the implementation of FAIR principles in cross-domain challenges related to data use and share may have implications at many levels:

* For research: understanding data provenance can have significant impact for the evidence generated and the further agenda of research. Implementation of FAIR and CARE principles in data use could improve the translation of evidence into decision making.
* For local governance: In emerging cities where infrastructure for data systems is particularly acute, the recognition of the different instances and actors involved in cultures of registration, for example, could mobilize efforts to improve processes of collection and registration that make data more valuable and useful for decision-making.
* For global regulations: discussion of FAIR and CARE principles across disciplines and geographies is a way to move forward processes toward Open Science connecting initiatives on data management and stewardship across borders while raising awareness to the way cultures of counting can contribute to downstream discrimination and bias in data globalization.

Furthermore, as a transdisciplinary project, this case study has far-reaching value for a range of contexts and disciplines beyond Urban Health. The health of our cities is inextricably intertwined with the health of our planet. We foresee opportunities to engage actively with geoethics initiatives bringing community perspectives of geospatial data more deliberately into consideration about the complexities of monitoring, managing, and improving the health of people in our cities.

**What do you foresee as concrete outputs and deliverables of this activity? (Please include estimated timing of delivery over 24-month period)**

**Deliverable 1. Urban Health Data - Guidelines and Recommendations (By Month 12)**

1. Publication documenting the cross-domain challenges in the harmonization of National Health Surveys across Latin American countries that led to the establishment of a set of guidelines and recommendations for FAIR practices in data harmonization, use and sharing, as well as in data provenance, that could guide governments carrying out these surveys in future data collection*.*

**Deliverable 2.** **Urban Health Data - Learning and Training (By Month 18 and Month 22)**

1. Implementation of a workshop discussing the FAIR and CARE dimensions in data provenance, that also allows mapping, gathering, and connecting researchers and data initiatives working with urban data systems across regions. Potential participants are [ARISE project](https://www.ariseconsortium.org/) (Accountability in Urban Health) working on urban health in countries from South East Asia, and [CitieS Health](https://citieshealth.eu/about/) and [The Institute for Future Cities](http://ifuturecities.com/) working with cities within Europe and in collaboration with other countries in Africa and Latin America. (**M18**)
2. Development of training materials shaped through the workshop (**Deliverable 2a**) and documentation guidelines l (**Deliverable 1a**) co-designed with the potential collaborators mentioned above along with other stakeholders that could be identified throughout the pilot project. This activity intends to establish a community of practice that can  share knowledge  and learning**(M22)**

**References**

Anderson, Theresa Dirndorfer (2020) "Keystone Practices to Enable Smart Cities to Flourish." In *Geography Research Forum*, vol. 40, pp. 171-192. <https://grf.bgu.ac.il/index.php/GRF/article/view/603>

Cox SJD et al. (2021) Ten simple rules for making a vocabulary FAIR. PLoS Comput Biol 17(6): e1009041. <https://doi.org/10.1371/journal.pcbi.1009041>

Diez Roux A, Slesinski C et al. (2019) A Novel International Partnership for Actionable Evidence on Urban Health in Latin America: LAC-Urban Health and SALURBAL. Global Challenges 2019, 3, 1800013. <https://doi.org/10.1002/gch2.201800013>

Gatzweiler F. (2021) Data for Intelligent Urban Systems / Data-Knowledge-Action Systems. [Under review]

Liu J, Gatzweiler F et al. (2021) Co-creating FAIR data for healthy and resilient urban development. [Under review]

Pineo H, Audia C et al. (2021) Building a Methodological Foundation for Impactful Urban Planetary Health Science. J Urban Health 2021, 98:442–452. <https://doi.org/10.1007/s11524-020-00463-5>

Quistberg D , Diez Roux A et al. (2019) Building a Data Platform for Cross-Country Urban Health Studies: the SALURBAL Study. J Urban Health 2019- 96:311–337. <https://doi.org/10.1007/s11524-018-00326-0>

Russo Carroll S et al. (2021) Operationalizing the CARE and FAIR Principles for Indigenous data futures. Nature -Scientific Data. 2021 8:108. <https://doi.org/10.1038/s41597-021-00892-0>

Star, S.L., & Strauss, A. (1999) Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work Computer -Supported Cooperative Work, 8: 9-30.

Wilkinson M et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. Nature- Scientific Data. 2016 -3:160018 <https://doi.org/10.1038/sdata.2016.18>