

Data Release Notes

Name of the dataset	GRID3 COD - Health Areas v1.0
Name of the file	GRID3_COD_health_areas_v1.0.gpkg
Date of data release	2024-05-10
File format	OGC Geopackage
Dataset version	1.0
Abstract	<p>This document outlines the methodology and data sources used during the production of the GRID3 COD - Health Areas v1.0 dataset. The dataset consists of health area boundaries with name, location, health zone, and attributes for ten provinces in the Democratic Republic of the Congo (COD). Limitations and use constraints are provided.</p> <p>This operational dataset has not been fully validated by government officials or ministries.</p>
Dataset citation	Center for International Earth Science Information Network (CIESIN), Columbia University and Ministère de la Santé Publique, Hygiène et Prévention, Democratic Republic of the Congo, 2024. GRID3 COD - Health Areas v1.0. New York: GRID3. https://doi.org/10.7916/v2bv-4h83 . Accessed <DAY MONTH YEAR>.
Terms of use	<p>Users are free to use, copy, distribute, transmit, and adapt the work for commercial and non-commercial purposes, without restriction, as long as clear attribution of the source is provided.</p> <p>Copyright 2024. The Trustees of Columbia University in the City of New York.</p>
Data license	The data and accompanying document are licensed under a Creative Commons Attribution 4.0 International License, CC BY 4.0 (http://creativecommons.org/licenses/by/4.0) and specified in legal code (http://creativecommons.org/licenses/by/4.0/legalcode).
Contacts and data queries	The authors of this dataset appreciate feedback regarding the data, including suggestions, discovery of errors, difficulties in using the data, and format preferences. For dataset-related questions, please send an email to: info@ciesin.columbia.edu

I. Input datasets

To create this dataset, CIESIN developed a consistent data schema and methodology to harmonize the data from ten different provinces collected as three different province groups. The province groups are legacy groupings corresponding to the order in which the original data was collected.

Province group 1: Haut-Katanga, Kasai, Kasai-Oriental, Kinshasa, and Lomami

Province group 2: Haut-Lomami and Tanganyika

Province group 3: Ituri and Kwilu

Province group 4: Maniema

The input data includes data from all four province groups. The collection process for each of these groups is detailed below by province group.

Province Group 1: Haut-Katanga, Kasai, Kasai-Oriental, Kinshasa, and Lomami

From January to July 2021 with the support of provincial and national health authorities, local healthcare workers (“head nurses”, “health zone management staff”, and “head doctors of the health zones”) and GRID3 GIS specialists (“mappers” and “provincial coordinators”) engaged in a participatory mapping process in Haut-Katanga, Kasai, Kasai-Oriental, Kinshasa, and Lomami. This mapping process occurred at the level of the health zone (an operational unit made up of approximately 15-20 health areas).

Mappers were deployed to health zones in teams of two for approximately nine days where they trained the health area head nurses on data collection using the Geospatial Tracking System (GTS), an Open Data Kit (ODK)-based smartphone application¹. The head nurses routinely work in their respective areas and have a good understanding of the health facilities, settlements, and points of interest (POIs, such as schools and religious centers) present.

While the head nurses collected data in their health areas, the mappers worked with the health zone management team to validate and modify data from the field. After all data were collected, cleaned, and integrated into final geospatial layers, the health zone head doctor validated the preliminary data. From July 2021 through October 2022, the mappers and provincial coordinators worked with CIESIN staff to consolidate the data (spelling, gaps and overlaps, topology, etc). This data was used to produce basemaps at the health area-level and shared back with every health zone and province for the second round of validation.

From November through January 2022, the in-country GIS team worked with CIESIN staff to integrate these corrections into a final geodatabase.

This work was done with the participation and supervision of the Direction du Système National d'Information Sanitaire (DSNIS). The Agence Nationale d'Ingénierie Clinique, de l'Information et de l'Informatique de Santé (ANICiS) also played an important role in the area of data governance.

Province Group 2: Haut-Lomami and Tanganyika

The *Haut-Lomami and Tanganyika Health Catchment Area Boundaries* data was originally created through an extensive fieldwork exercise from July to September 2019 with additional data added from the PNLIP.

Phase 1: 2019 Field data collection

With the support of provincial and national health authorities, local healthcare workers (“head nurses”, “health zone management staff”, and “head doctors of the health zones”) and GRID3 GIS specialists (“mappers” and “provincial coordinators”) engaged in a participatory mapping process in Haut-Lomami and Tanganyika from July to September 2019. This mapping process occurred at the health zone level (an operational unit made up of approximately 15-20 health areas).

Mappers were deployed to the health zones, where they organized participatory mapping meetings with local healthcare workers. They also trained head nurses to collect data on settlements, health facilities, and other points of interest in their respective health areas using an ODK-based smartphone application. Mappers then used this information to delineate health area boundaries, using previously existing data to guide this effort. This data was then sent back to CIESIN for additional quality checks. The sources used to delineate the boundaries were:

- ESP-UCLA, Public health school (École de Santé Publique) of the University of California, Los Angeles. This data was received prior to fieldwork.
- MSP/CIESIN, DRC’s Ministry of Public Health (Ministère de Santé Publique) and the Center for International Earth Science Information Systems)

Phase 2: Integration of data from the PNL

In October 2021, GRID3 received access to a large settlement point dataset collected during a bednet distribution campaign by IMA World Health and owned by PNL. This dataset was combined and consolidated with previous GRID3 settlement data for Haut-Lomami and Tanganyika. In total, the PNL data added more than 3,500 new points to the settlement data. The health zone and health area boundaries were then refined based on the health area and health zone attributes of the PNL point clusters, locations of the points, and the GRID3 friction surface layer.

Table 1: Settlement Data Sources

Source	Description
PNLP	Programme National de Lutte contre le Paludisme (National Malaria Elimination Programme) of the DRC
MSP/CIESIN	DRC’s Ministry of Public Health (Ministère de Santé Publique) and the Center for International Earth Science Information Systems)

Province Group 3: Ituri and Kwilu

The GRID3 Ituri and Kwilu Health Catchment Area Boundaries were created by using settlement point data from DRC’s National Malaria Elimination Program (Programme National de la Lutte contre le Paludisme or PNL). Health area boundaries were verified with local health teams through in-person missions in both provinces.

GRID3 received data from the PNLP (National Program for the Fight Against Malaria) covering Ituri and Kwilu in October 2021 and August 2022, respectively¹. This data consisted of household-level GPS points with village, health area, and health zone attributes. It had been collected by IMA World Health, an implementing partner of the PNLP, during province-wide bed net distribution campaigns. In Ituri, over 1,165,000 household points were received from a bed net distribution campaign in June 2021. In Kwilu, over 1,191,000 household points were received from a bed net distribution campaign in August 2022. The PNLP data was combined with a small number of settlement points from other organizations that also contained health zone and health area attribute data.

CIESIN used the following process to delineate preliminary health area boundaries based on the PNLP data:

- Assessment of already existing health area boundaries. Available point data available (settlements, health facilities, other points of interest available with health area attributes collected independently) with health area attributes were overlaid with the boundaries to determine the percent of match and select the best layer available, in concert with the local health authorities. In Kwilu, the 2018 data produced by UCLA on behalf of the PNLTHA - and already integrated into the DSNIS national database - was considered the best available data. In Ituri, a combination of recent WHO and OSM boundary work was used as a reference.
- Creation of a friction surface layer. Several geospatial datasets - elevation, topography, land cover, water bodies (rivers, streams, permanent bodies of water, lakes, etc.), and road infrastructure (primary, secondary, tertiary road networks) were combined to construct a surface (raster or grid, at ~100 m x 100m resolution) with a travel cost associated with it - which quantifies the amount of time, or how easy it is to travel from one point to another. Rivers, permanent bodies of water, or cliffs were used as physical barriers and considered obstacles to be circumvented.
- Inclusion of the best available health areas to the friction layer. In addition to natural and man-made barriers, the best available health area boundaries were added to the friction surface to avoid modifying or redrawing existing boundaries. If extensive work had already been conducted and validated in an area, and only a few points fell outside of the original boundaries - no changes were made to the boundaries, as the amount of data available did not justify modifications. Conversely, if a sufficient quantity of recent and reliable data points fell outside current health area limits, these were adjusted accordingly.
- Creation of village catchment areas. The granularity of the data collected during the ITN distribution campaign of the PNLP (door-to-door distribution of bednets) at the provincial scale allowed us to create buffers around each survey location in order to generate village catchment areas (settlement contours).
- Creation of health areas. Each village catchment area was dissolved into health area boundaries, based on their health area attribute. An additional visual inspection identified and corrected minor issues and ensured the boundaries were topologically correct.
- Match health areas to the master list. The final step was to ensure that the data inputs from the geo-referenced household survey were in line with the official health area lists.

¹ PNLP-IMA World HealthSANRU-GRID3 Official Agreement for the use of data (01/09/2021)

- Map the preliminary boundary data. Health zone boundaries along with other contextual data (settlements, health facilities, points of interest, etc) were then produced and printed in a large format (A0).
- Present them to the local health team to edit and refine them. These preliminary health area boundaries (pre-alpha) were then presented to local health officials (Médecins Chefs de Zone, Infirmiers Superviseurs) to make sure all the settlements and health facilities did fall within their respective boundaries - and/or to make the necessary adjustments to the boundaries when necessary.

In Kwilu, the data was verified through a mission organized in October 2022. Two GRID3 mappers traveled to Kikwit and Bandudu, respectively, to present the cartographic improvements driven by the use of the PNL data and to verify that the boundaries were correctly delineated. The majority of this work was carried out in collaboration with antenne-level and provincial level staff, rather than working directly with health zone staff.

In Ituri, the data was verified through a mission organized in October 2022. Four GRID3 mappers traveled to Bunia and Aru, respectively, to validate the health area boundary data in all health zones. During this mission, the GRID3 mappers arranged for the transportation of local health teams to the antenna, where they validated the preliminary layers and suggested improvements where they were needed.

Province Group 3: Maniema

Table 2. Data Sources

Name	Data type/ format	Release year	Input data year	Resolution (if applicable, in meters)
Fieldwork data collected by the Kinshasa School of Public Health in the Democratic Republic of the Congo (KSPH, DRC)	Spatial points and qualitative feedback	2024	2024	N/A
Pre-Distribution Registration Survey (PDRS) from the National Malaria Control Programme (PNLP) collected as part of the anti-malaria campaigns in the Democratic Republic of the Congo	Polygons derived from spatial points	2023	2023	N/A

The Maniema data was created through an extensive fieldwork data collection conducted by the Kinshasa School of Public Health (Ecole de Santé Publique de Kinshasa, ESPK) spanning from October 2023 to January 2024, and supplemented with additional data from the Pre-Distribution Registration Survey (PDRS) from the National Malaria Control Programme (PNLP) collected as part of the anti-malaria campaigns in the Democratic Republic of the Congo.

Phase 1: Integration of data from the PNL

In October 2021, CIESIN received access to a large settlement point dataset collected during a bednet distribution campaign by IMA World Health and owned by the PNL. This dataset was explored, cleaned, and matched against health area and health zone lists within DRC's DHIS2. This attribute information was used to produce a preliminary, draft boundary to be validated with the help of local and provincial health authorities.

Phase 2: 2023-2024 Field data collection

The Kinshasa School of Public Health (ESPK), with the support of provincial and national health authorities, local healthcare workers (i.e. head nurses, health zone management staff, and head doctors of the health zones) and GRID3 GIS specialists engaged in data collection from October 2023 to January 2024. GRID3 GIS specialists were deployed to each health zone, and liaised with local authorities and local health workers to validate the list of health areas within each health zone and validate (or make corrections) to the preliminary draft boundary produced in Phase 1 using ESRI software. Corrections were sent back to CIESIN for processing and quality checks, and validated against point-data attributes collected as part of the same fieldwork.

II. Methodology

Data processing, integration, and harmonization scripts have been developed and updated by CIESIN as new data becomes available. A data schema prioritizing information preservation was developed to include fields common to all or most of the datasets in all four province groups. Common generalized values were created for all fields and the original values in each dataset were mapped to these. The data schema enabled the consolidation of data from two previously published health areas datasets in province groups 1 & 2 (Haut-Katanga, Haut-Lomami, Kasai, Kasai-Oriental, Kinshasa, Lomami and Tanganyika) and the unpublished data from province groups 3 & 4 covering health areas in the provinces of Ituri, Kwilu and Maniema.

III. Dataset Description

The *GRID3 COD - Health Areas v1.0* dataset consists of health area spatial polygon data. The data are available for download in OGC Geopackage format contained in a zip file.

This data release supersedes the following datasets:

- [GRID3 COD - Health Areas: five provinces v1.0](#)
- [GRID3 COD - Health Areas: two provinces v1.0](#)

The codebook for health areas is shown below.

Table 3: Codebook for Health Areas

Variable Names	Values example	Type	Definition
OBJECTID	1	numeric	Software-generated unique code
pays	République démocratique du Congo	text	Country name French official UN member state spelling
iso3	COD	text	ISO3 code
province	Lomami	text	DHIS2 province name
prov_uid	F9w3VW1cQmb	text	The identifier of the province in the DHIS2
antenne	Kabinda	text	Antenne name
zonesante	Kabinda	text	DHIS2 health zone name
zs_uid	NR5cDy2jhiJ	text	The identifier of the health zone in the DHIS2
airesante	Kemayi	text	DHIS2 health area name
as_uid	BuOpOXDJJpy	text	The identifier of the health area in the DHIS2
date	2022	text	Year of data collection or last edit/modification
source	UCLA	text	Institution or project providing input data for this dataset
edit_par	GRID3	text	Editor of the original data
grid3id	as_F6FD8B8CFD	text	Internal GRID3 ID

IV. Known Data Limitations and Disclaimer

The spatial accuracy of the health areas data is dependent on both the accuracy of the input data collected in the field as well as on the correctness of the edits made to the collected data throughout the validation process. In general, it was assumed that the field collected data were the most accurate data available. Temporal mismatches exist among the point datasets and the satellite imagery used to perform quality checks. Spelling mistakes and/or mismatches may have occurred due to colloquial variations on how data points are referred to in the field.

The scope of the GRID3 Mapping For Health project included fieldwork and validation for Haut-Katanga, Kasai, Kasai-Oriental, Kinshasa, and Lomami provinces and 115 health zones.

This operational dataset has not been fully validated by government officials or ministries.

CIESIN, Columbia University, and its co-authors follow procedures designed to ensure that data disseminated by the project are of reasonable quality. If, despite these procedures, users encounter apparent errors or misstatements in the data, they should contact CIESIN, info@ciesin.columbia.edu

CIESIN, Columbia University, its co-authors, and their sponsors do not guarantee the accuracy, reliability, or completeness of any data provided. We provide these data without warranty of any kind whatsoever, either expressed or implied, and shall not be liable for incidental, consequential, or special damages arising out of the use of any data provided.

V. Acknowledgments

CIESIN and its co-authors thank the following institutions that provided input data and/or assistance with data production:

Acasus, Switzerland
Agence Nationale d'Ingénierie Clinique et du Numérique de la Santé (ANICNS) , DRC
Bill & Melinda Gates Foundation, USA
Bluesquare, Belgium
Bureau Central du Recensement (BCR), DRC
Caritas, USA
Centers for Disease Control and Prevention (CDC), USA
Direction d'Etudes et Planification (DEP), DRC
Direction des Soins de Santé Primaires (DSSP), DRC
Division du Système National d'Informations Sanitaires (DSNIS), DRC
Division Provinciale de la Santé (DPS) de Kinshasa, Kwilu, Kasai, Kasai-Oriental, Lomami, Haut-Lomami, Tanganyika, Haut-Katanga, and Ituri
Kinshasa, and Lomami, DRC
Ecole de Santé Publique de Kinshasa (ESPK), DRC
Gavi, the Vaccine Alliance, Switzerland
Geospatial Evaluation and Observation Lab (geoLab), College of William & Mary, USA
Global Affairs Canada (GAC), Canada
Global Good, USA
IMA World Health, DRC
Initiative Régionale de Documentation et d'Accompagnement Communautaire au Développement (IDRAC Sarl), DRC
International Federation of Red Cross and Red Crescent Societies (IFRC), Switzerland
International Medical Corps (IMC), USA
Médecins Sans Frontières (MSF), Switzerland
Ministère de l'Environnement et Développement Durable (MEDD), DRC
Ministère de la Santé publique, Hygiène et Prévention, DRC
Novel-T, Switzerland
Open Street Map (OSM), DRC
PATH, USA
Programme Elargi de Vaccination (PEV), DRC
Programme National de Lutte contre le Paludisme (PNLP), DRC
Référentiel Géographique Commun (RGC), DRC
Soins de Santé Primaires en Milieu Rural (SANRU), DRC
The International Organization for Migration (IOM), DRC
United Nations Children Fund (UNICEF), USA
United Nations Development Programme (UNDP), USA
United Nations Office for Project Services (UNOPS), Denmark and DRC CO

United Nations Office for the Coordination of Humanitarian Affairs (OCHA), USA
United Nations Organization Stabilization Mission in the Democratic Republic of the
Congo (MONUCSO), DRC
University of California, Los Angeles (UCLA) DRC Health Research and Training
Program, USA
VillageReach, USA
World Health Organization (WHO), Switzerland (HQ), Brazzaville (AFRO), Kinshasa (CO)
World Resources Institute (WRI), USA

Funding for the development and dissemination of this dataset was provided by GRID3 Inc, the Bill & Melinda Gates Foundation, and Gavi, the Vaccine Alliance.