

DataSuite

CD2030 DataSuite is the Countdown to 2030 data analytics platform for extracting, checking, adjusting and analysing routine health facility data.

2026-10-29
datasuite.damurka.com

Contents

Part I · Get started 7

1 Introduction 8

2 DataSuite documentation 9

3 Setting up CD2030 DataSuite 11

3.1 Download and install for Windows 12

3.2 Requirements for DataSuite 16

3.3 Portable mode 19

4 Navigating the CD2030 DataSuite interface 21

4.1 Menus and shortcuts 27

4.2 Language and themes 30

4.3 Extensions 36

Part II · Methodology 41

5 Framework overview 42

6 Analysis setup 44

6.1 Country input facility data set 46

6.2 Checks on the uploaded data 48

6.3 Household survey results 50

6.4 Global estimates 52

6.5 National survey parameters 54

7 Data quality assessment 57

7.1 Completeness of monthly facility reporting 59

7.2 Missing monthly values 62

7.3 Extreme outliers 64

7.4 Consistency of annual reporting 67

7.5 Annual data quality score 72

8 Data adjustment 75

8.1 Adjusting for completeness of reporting 80

8.2 Low reporting rates 84

8.3 Extreme outliers 87

8.4 Missing values 90

8.5 Removing years 93

9 Denominator assessment and selection 95

9.1 Quality of the DHIS2 population projections 98

9.2 Facility data derived denominators 103

9.3 Selection of the best performing denominator 108

10 National coverage 115

10.1 Antenatal care 118

10.2 Delivery care 122

10.3 Immunization 126

10.4 Family planning 132

10.5 Comparison with surveys and targets 134

10.6 Continuum of care 137

11 Bayesian modeling approach 143

11.1 Estimates and short-term forecasts 145

11.2 How survey data are used 148

11.3 How routine data are used 150

12 Equity assessment (equiplots) 153

12.1 Inequality by wealth quintile 156

12.2 Inequality by maternal education 160

12.3 Inequality by area of residence 163

13 Subnational analysis 165

13.1 Subnational coverage 168

13.2 Subnational inequality (MADM) 172

13.3 Global coverage targets 175

14 Institutional mortality 179

14.1 Institutional maternal mortality ratio (iMMR) 182

14.2 Institutional stillbirth rate (iSBR) 185

14.3 Neonatal deaths before discharge 189

14.4 Completeness of facility death reporting 193

14.5 Ratio of stillbirths to maternal deaths 198

14.6 Outliers in mortality data 201

15 Curative health services utilization for sick children 204

15.1 Quality of outpatient and inpatient data 207

15.2 Outpatient service utilization 210

15.3 Inpatient service utilization 215

15.4 Case fatality rate 219

15.5 MCH preventive vs curative index 222

16 Health system performance 225

16.1 Health systems inputs 228

16.2 Health systems outputs by inputs 235

16.3 Private sector and RMNCAH-N services 242

17 Planning for dissemination, decisions and action 245

17.1 Identifying use cases and audiences 248
17.2 Follow-on activities 250
17.3 Information products 252

Part III · DataSuite Apps 255

18 DataSuite applications 256

19 RMNCAH and vaccine analyses 258

20 Leveraging AI for Analysis and Interpretation 260

21 Data Extractor 269

21.1 DHIS2 login 273
21.2 Mapping indicators 277
21.3 Downloading data 283

22 RMNCAH and Vaxx apps 289

22.1 Load Data 295

23 Data Quality Assessment 306

23.1 Reporting Rate 309
23.2 Outlier Detection 312
23.3 Data Missingness 315
23.4 Internal Consistency 317
23.5 Overall Data Quality Score 320

24 Data Adjustment 322

24.1 Data Adjustment 324
24.2 Data Adjustment Changes 331

25 Denominators 333

25.1 Denominator Assessment 335
25.2 Denominator Selection 338

26 National Analysis 342

26.1 National Coverage 345
26.2 Coverage Target 349
26.3 National Inequality 352
26.4 Equity Assessment 355

27 Sub-National Analysis 359

27.1 Sub-National Coverage 362
27.2 Sub-National Inequality 365
27.3 Coverage Target 368

28 Advanced Analysis 371

28.1 Bayesian Analysis 373

28.2 Continuum of Care 376
28.3 FP Projection 379
29 Mortality 381
29.1 Institutional Mortality 384
29.2 Mortality Mapping 387
29.3 Mortality Completeness 389
30 Service Utilization and Health System Performance 391
30.1 Service Utilization DQA 393
30.2 National Utilization 395
30.3 Sub-National Utilization 398
30.4 MCH Preventive vs Curative Index 401
30.5 National Health System 403
30.6 Sub-National Health System 405
30.7 PHC Performance 407
30.8 Public/Private Share 409
31 Exports and Ask AI 411
32 Reports 415
33 Pooled App overview 424
33.1 Preparing caches 428
33.2 Using the app 431
33.3 Exporting results 439
Part IV · Developer resources 443
34 Developer resources 444
35 Building apps for DataSuite 445
36 CacheConnection reference 452
37 Data dictionary 483
Part V · FAQs 489
38 FAQs 490
Appendix 495
Appendix A · Glossary 496
Appendix B · Abbreviations 506
Appendix C · Indicator definitions 509

PART I

Get started

Install DataSuite, learn the Countdown method and use the apps that analyse routine health facility data.

CHAPTER 1

Introduction

Analyse routine health facility data on your computer

DataSuite is the data analytics platform of the **Countdown to 2030** initiative. It takes routine data from DHIS2 (p. 498), checks and adjusts it, and compares it with household surveys to track women's, children's and adolescents' health, nationally and by region.

Built for data analysts, researchers and public health professionals. Free and built on open source.

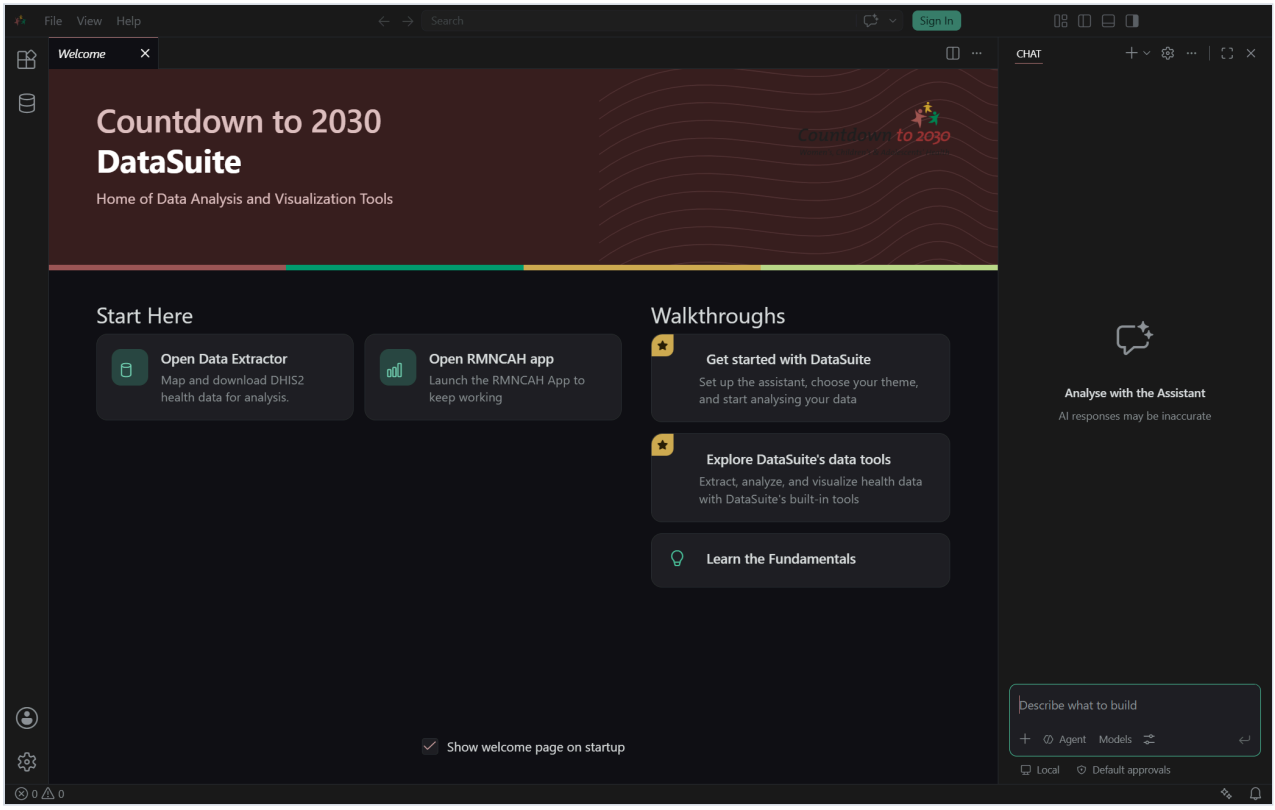
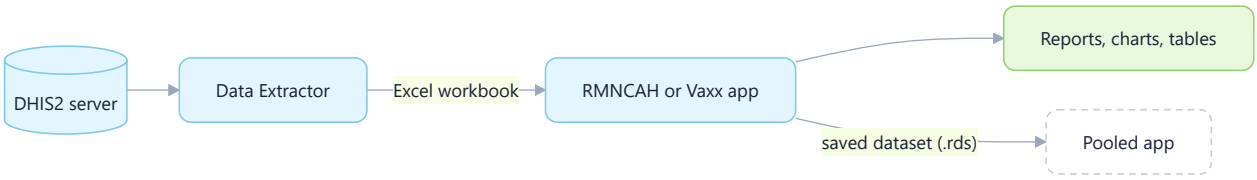


Figure 1.1 — *The DataSuite Welcome page, with Open Data Extractor and Open RMNCAH app*

How it works

Data moves from one app to the next. Each app saves a file that the next one opens.



CHAPTER 2

DataSuite documentation

Install DataSuite, learn the Countdown method and use the apps that analyse routine health facility data.

Analyse your health facility data, step by step

DataSuite brings the Countdown 2030 analysis to your computer. Get the data out of DHIS2 (p. 498), check and adjust it, then estimate coverage and share the results.

Learn about the first steps *(p. 21)*

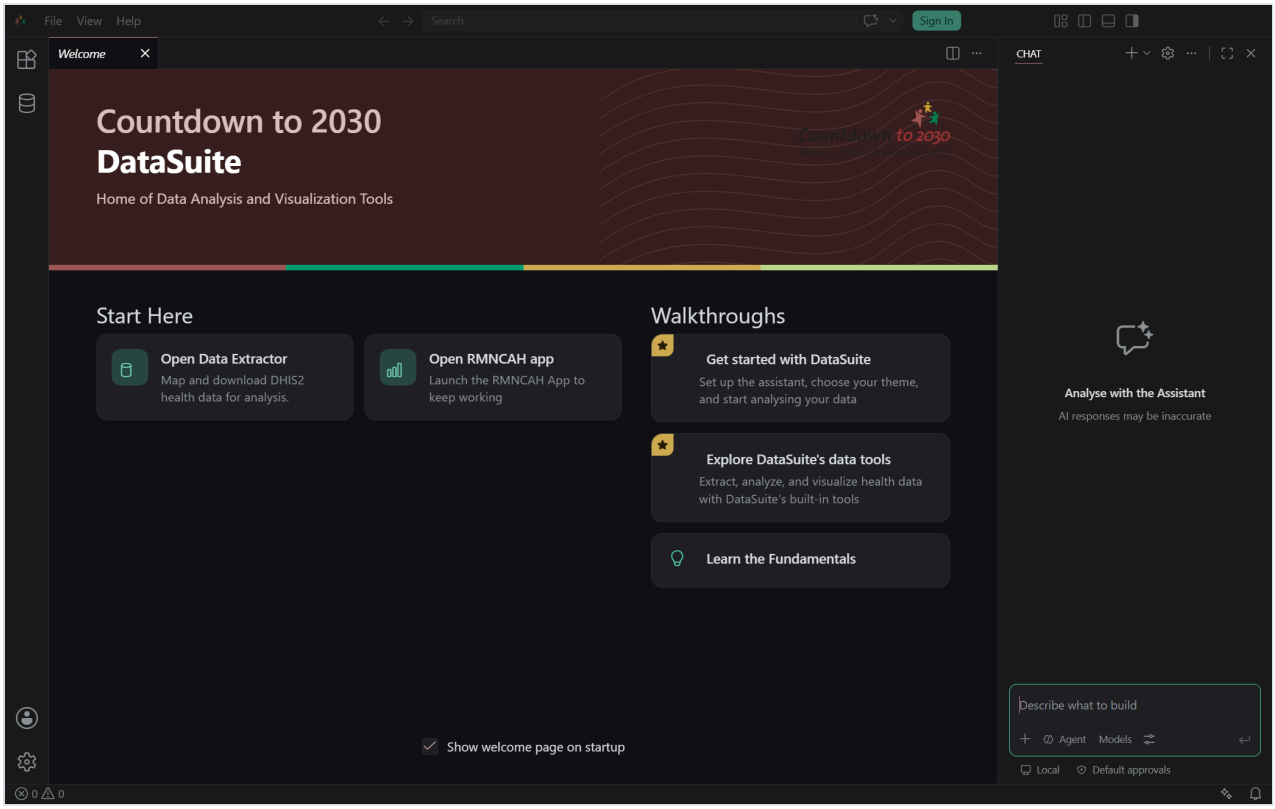


Figure 2.1 — *The DataSuite Welcome page, with Open Data Extractor and Open RMNCAH app under Start Here*

Methodology

The Countdown methodology is the analysis itself: each step, how it is done and how to read the results.

Look things up

What an indicator, a term or an abbreviation means, and how the two analyses differ.

Privacy

By default, DataSuite auto-updates to new versions, and does not collect usage data or crash report information.

CHAPTER 3

Setting up CD2030 DataSuite

Setup steps for CD2030 DataSuite: review the system requirements, download and install the app, then get started with the user interface.

CD2030 DataSuite runs on 64-bit Windows 10 and 11, and on 64-bit Linux (see Requirements (p. 16)). It is a lightweight platform, allowing you to install it or use a portable version and begin your data operations in a matter of minutes.

To get started, follow these step-by-step setup instructions:

1. **Review the system requirements (p. 16)**
   - To ensure your hardware and computer configuration are supported.
2. **Download and install CD2030 DataSuite for your platform**
   - Select the correct installer <https://datasuite.damurka.com/en/downloads#installer-options> option for your PC.
   - Install (p. 12) the DataSuite app.
3. **Go to Getting started with DataSuite (p. 21)**
   - To get a glimpse of the user interface (UI) and key features of CD2030 DataSuite.

Next steps

Once you have installed CD2030 DataSuite, you can explore the following pages to learn more about the platform and its applications:

- Platform overview (p. 21) - Understand the core architecture.
- DHIS2 Data Extractor (p. 269) - Explore the Data Extractor.
- RMNCAH App (p. 289) - Learn how to navigate and use the RMNCAH App.
- Vaxx App (p. 289) - Learn how to use the Vaxx App.

FAQs

Having challenges downloading and installing CD2030 DataSuite? Check out the commonly asked questions (p. 490).

3.1 Download and install for Windows

Step-by-step instructions to download and install CD2030 DataSuite on Windows, with the installer or the ZIP file, from licence agreement to finish.

Before you install, check the requirements (p. 16). DataSuite does not include R, which its apps need: install R 4.1 or newer (p. 16) too.

To download and install CD2030 DataSuite on Windows, follow these steps:

1. **Download and install DataSuite**
   - Select the installer option that is appropriate for your use <https://datasuite.damurka.com/en/downloads#installer-options>.
2. **Once the download is complete, open the folder and run the installer**

Install DataSuite on Windows

Use the Windows installer

1. **Download the DataSuite installer <https://datasuite.damurka.com/en/downloads> for Windows**
2. **Once it is downloaded, run the installer (`DataSuiteUserSetup-{version}.exe`)**
   If prompted by User Account Control (UAC), click **Yes** to allow the app to make changes to your device.

③ Accept the license agreement

Read through the End-User License Agreement (EULA). Select **I accept the agreement** and click **Next**.

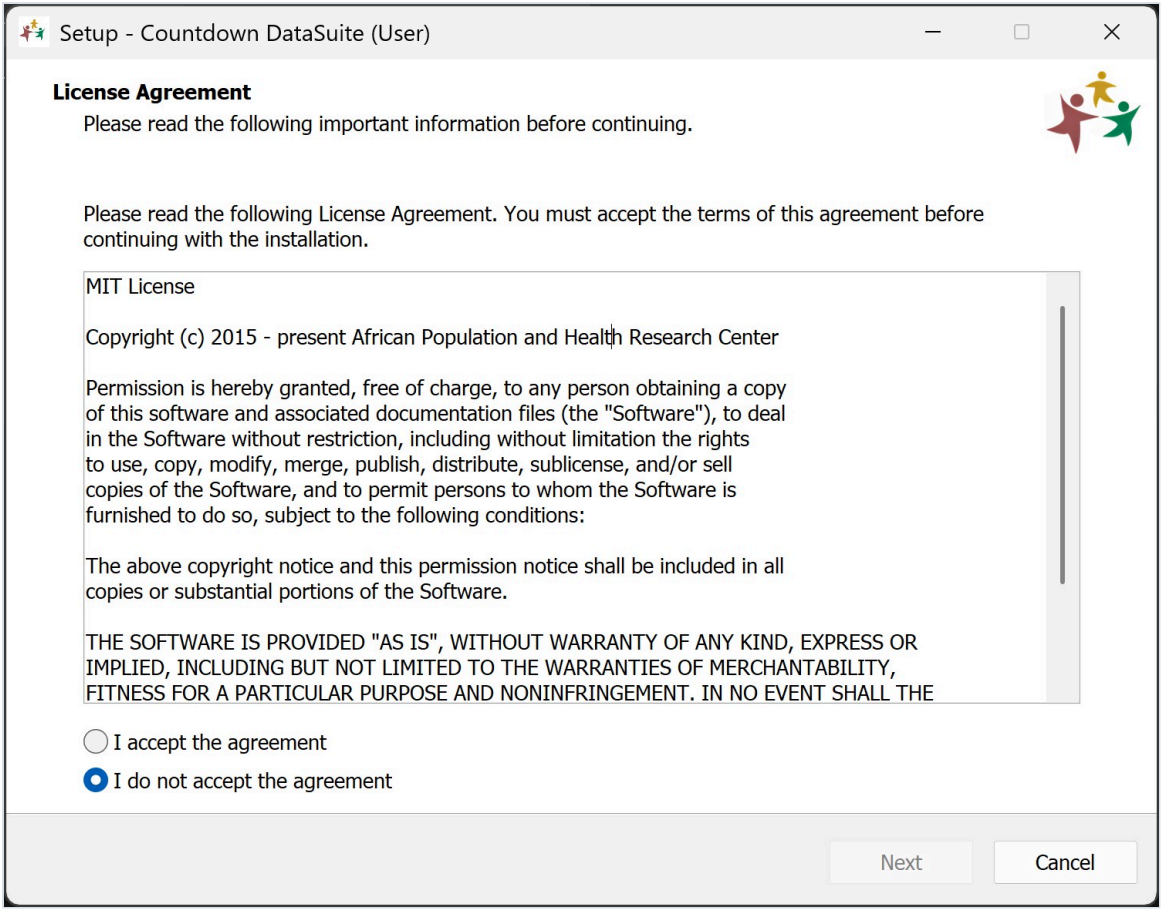


Figure 3.1 — *License agreement screen of the DataSuite installer*

④ Choose the installation location

Choose the destination folder where you want DataSuite to be installed. By default, DataSuite is installed under `C:\Users\{Username}\AppData\Local\Programs\DataSuite`. Click **Next** to proceed.

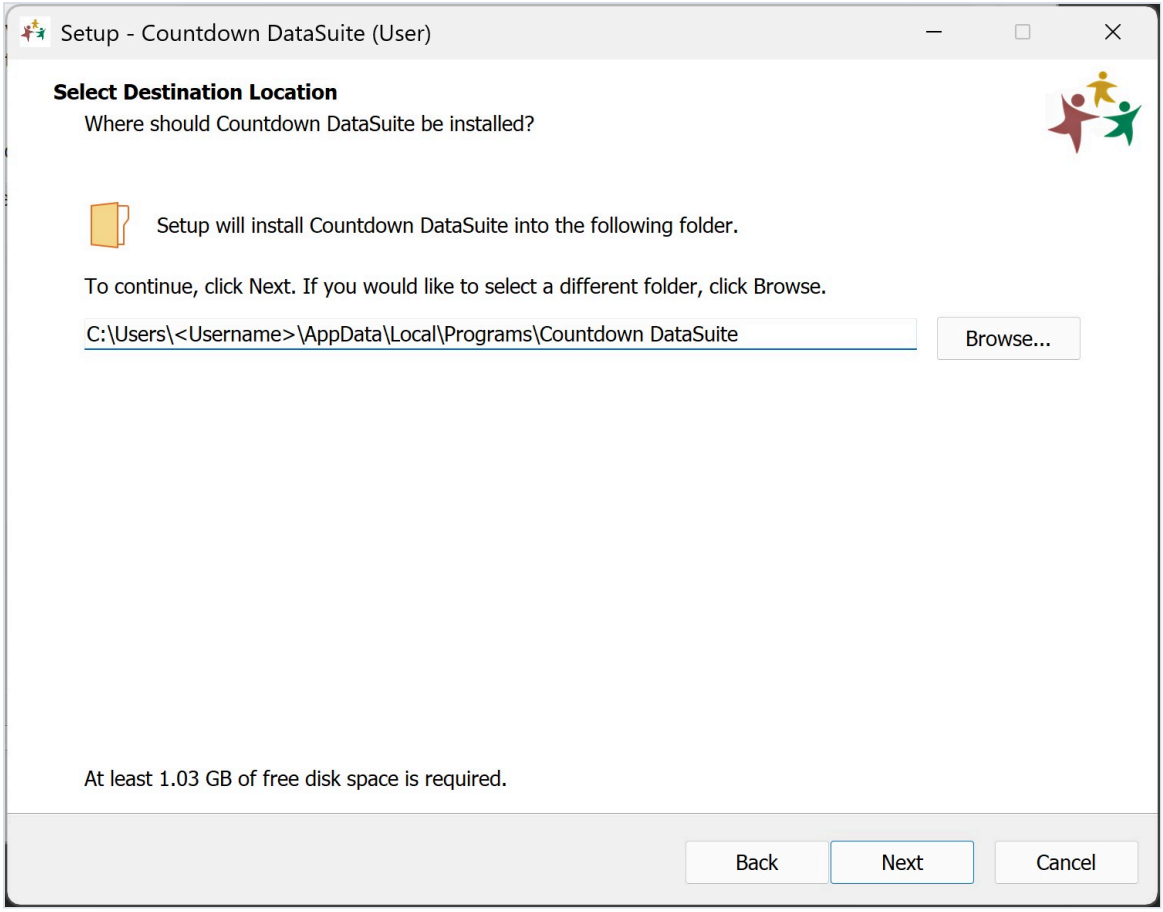


Figure 3.2 — *Installation location screen of the DataSuite installer*

⑤ Install and finish

Click **Install** to begin the process. Once the progress bar completes, click **Finish** to exit the setup wizard and launch DataSuite.

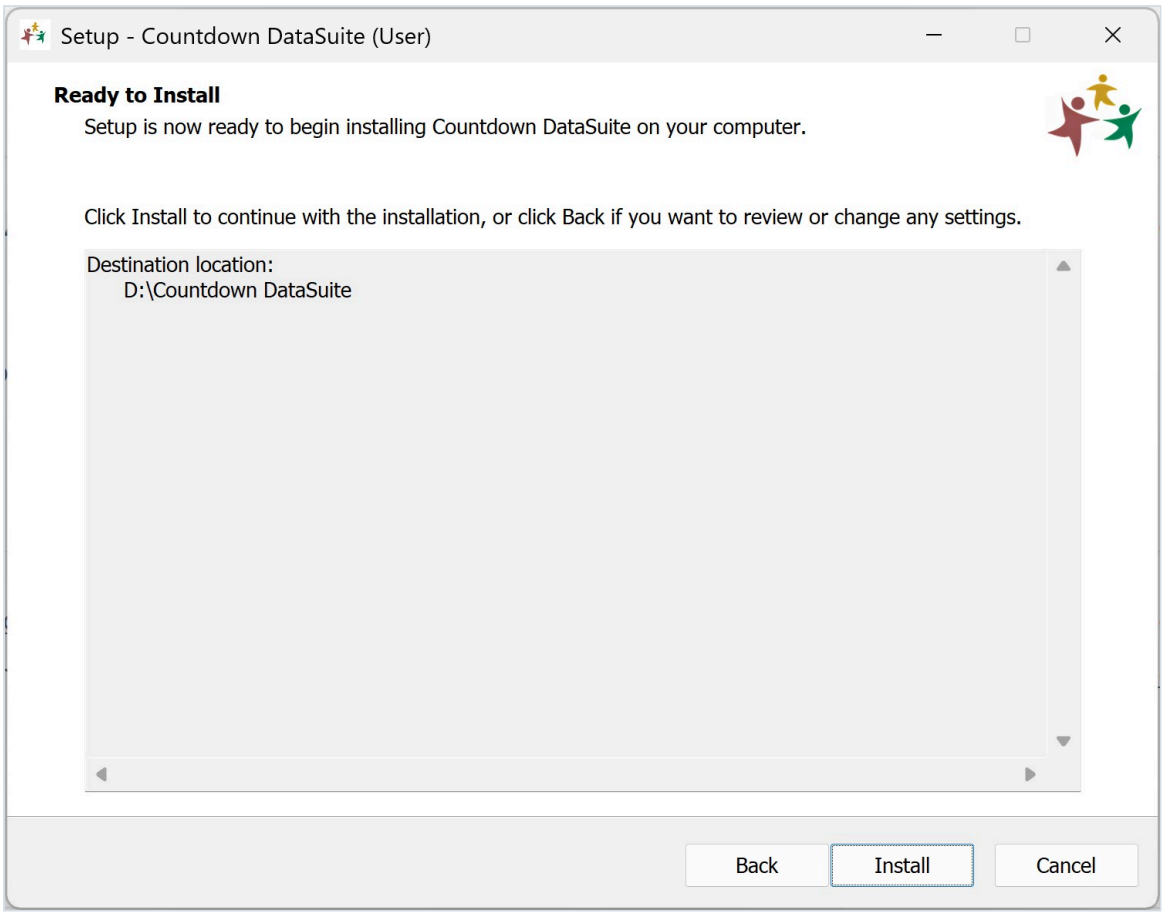


Figure 3.3 — *Ready to install screen of the DataSuite installer*

Use the ZIP file

1. Download the DataSuite ZIP archive <https://datasuite.damurka.com/en/downloads>.
2. Extract the ZIP archive, and run DataSuite from there.

3.2 Requirements for DataSuite

What CD2030 DataSuite needs: hardware, 64-bit Windows 10 and 11 or Linux, and the software it does not include: R for the apps, Python or a licensed Stata for notebooks, Word, PowerPoint or LibreOffice for PDFs.

Hardware

CD2030 DataSuite is a small download (< 250 MB) and has a disk footprint of < 1 GB. DataSuite is lightweight and should easily run on today's hardware.

We recommend:

- 1.6 GHz or faster processor
- 1 GB of RAM

The RMNCAH, Vaxx and Pooled apps download their R packages the first time you open them: allow up to about 1 GB more disk space for them.

Platforms

DataSuite is supported on the following platforms:

- Windows 10 and 11 (64-bit)
- Linux (64-bit), a distribution with glibc 2.39 or newer, such as Ubuntu 24.04, Debian 13 or Fedora 40 (DataSuite 1.1.3 and later): see Linux <https://datasuite.damurka.com/en/downloads#linux>
- macOS 14 (Sonoma) or later, on Apple silicon or Intel: see macOS <https://datasuite.damurka.com/en/downloads#macos>

Software DataSuite does not include

DataSuite does not include R, Python or Stata: install the ones you need yourself, before or after DataSuite. The apps need only R; Python and Stata are needed only for notebooks in those languages.

Software	Needed for	Version	Get it
R	The RMNCAH, Vaxx and Pooled apps, R notebooks and the AI assistant's analyses	4.1 or newer (a recent release is best)	CRAN <https://cloud.r-project.org/> (on Windows: **Download R for Windows** > **base**)
Python	Python notebooks	3.10 or newer	python.org <https://www.python.org/downloads/>
Stata, licensed	Stata notebooks	Stata 17 or newer, StataNow included, any edition (MP, SE or BE)	stata.com <https://www.stata.com/>
Microsoft Word, PowerPoint or LibreOffice	A report's PDF: Word or LibreOffice for a document, PowerPoint or LibreOffice for a slide deck	A recent version	LibreOffice <https://www.libreoffice.org/download/download-libreoffice/> (free)

You also need an internet connection the first time you open an app (to download its R packages), for the Data Extractor (p. 269) (to reach your DHIS2 server) and for the AI assistant (p. 260) (signed in with a DataSuite account).

R

DataSuite finds R itself: the setting `datasuite.rHome` if you set it, then the `R_HOME` variable, R on the `PATH`, the Windows registry and the usual install folders (such as `C:\Program Files\R`). On Linux, install R with your package manager, for example `sudo apt install r-base`.

When no R is found, or it is older than 4.1, DataSuite says so when it starts and when you open an app: "DataSuite needs R 4.1.0 or newer to run its apps and analyses, and none was found." Choose **Download R** to get it, **Choose R Folder…** if R is installed where DataSuite does not look, then **Check Again**. You can check at any time with the command **DataSuite: Check R Installation**.

You do not install any R package yourself: DataSuite installs what the apps need into its own library and leaves your R library as it is.

Python

DataSuite finds Python itself: the setting `datasuite.pythonPath` if you set it, then the Windows `py` launcher, Python on the `PATH` and the usual install folders. Python needs its shared library, which the python.org installers include. The notebooks' Python packages go into DataSuite's own environment, not into your Python.

Without a usable Python, a Python notebook does not run and says: "Python was not found. Python apps need Python 3.10 or newer: install it from https://www.python.org, or set `datasuite.pythonPath` to the Python to use."

Stata

DataSuite uses the Stata in the setting `datasuite.stataHome` (the folder Stata is installed in, for example `C:\Program Files\StataNow19`). Without it, it looks in the `STATA_HOME` variable, the Windows registry, the `PATH` and the usual install folders, and uses a licensed Stata, the newest first. A Stata without its licence (`stata.lic`), or older than 17, cannot be used.

Without one, a Stata notebook does not start and says: "Stata notebooks need Stata 17 or newer, licensed, on this computer". Set `datasuite.stataHome` if DataSuite does not find your Stata or picks the wrong one.

Word, PowerPoint or LibreOffice

The app makes a report's Word (`.docx`) and PowerPoint (`.pptx`) files itself: only the PDF needs one of these programs. Without one, a slide deck's PDF cannot be made ("Making a PDF of slides needs Microsoft PowerPoint or LibreOffice"), and a document's PDF is printed from the page, so its pages can differ from the Word file. See Formats and downloads (p. 422).

Next steps

- Download and install DataSuite for Windows (p. 12)
- Install DataSuite on Linux <https://datasuite.damurka.com/en/downloads#linux>

- Commonly asked questions (p. 490)

3.3 Portable mode

How to run DataSuite in portable mode from the Windows ZIP archive by creating a data folder, so all its data lives next to the app and can be moved around.

DataSuite supports portable mode <https://en.wikipedia.org/wiki/Portable_application>. This mode enables all data created and maintained by DataSuite to live near itself, so it can be moved around across environments.

Portable mode is supported on the ZIP download for Windows. See the downloads page <https://datasuite.damurka.com/en/downloads> to find the correct `.zip` file for your platform.

> Do not attempt to configure portable mode on an installation from the **Windows User or System installers**. Portable mode is only supported on the Windows ZIP (`.zip`) archive. Note also that the Windows ZIP archive does not support auto update.

Enable portable mode

Windows

After unzipping the DataSuite download, create a `data` folder within DataSuite's folder:

```
|- DataSuite-win32-x64-1.0.0
|   |- Suite.exe (or code executable)
|   |- data
|   |- bin
|   |  |- code
|   |  |- ...
|   |- ...
```

From then on, the `data` folder will be used to contain all DataSuite data, including session state, preferences, etc.

The `data` folder can be moved to other DataSuite installations. This is useful for updating your portable DataSuite version, in which case you can move the `data` folder to a newer extracted version of DataSuite.

Update portable DataSuite

On **Windows**, you can update DataSuite by copying the `data` folder over to a more recent version of DataSuite.

Migrate to portable mode

You can also migrate an existing installation to portable mode.

Windows

1. Download the DataSuite <https://datasuite.damurka.com/en/downloads> ZIP distribution for your platform.
2. Create the `data` folder as above.
3. Copy the user data directory `DataSuite` to `data` and rename it to `user-data` :

- **Windows**: `%APPDATA%\DataSuite`

As an example, here's the desired outcome on **Windows**:

```
|- DataSuite-win32-x64-1.84.2
|   |- DataSuite.exe (or suite executable)
|   |- data
|   |   |- user-data
|   |   |   |- ...
|   |- ...
```

TMP directory

By default, the `TMP` directory is still the system one even in portable mode, since no state is kept there. If you want to also have your TMP directory within your portable directory, you can create an empty `tmp` directory inside the `data` folder. As long as a `tmp` directory exists, it will be used for TMP data.

CHAPTER 4

Navigating the CD2030 DataSuite interface

A tour of the DataSuite window for first-time users - the first-start window, the Welcome page and its apps, the activity bar, the chat, the panel and the status bar - and what is new.

Find your way around DataSuite

DataSuite opens on its **Welcome** page. From there you open the apps, the Data Extractor and the AI assistant. This page walks through each part of the window.

Learn about the menus and shortcuts *(p. 27)*

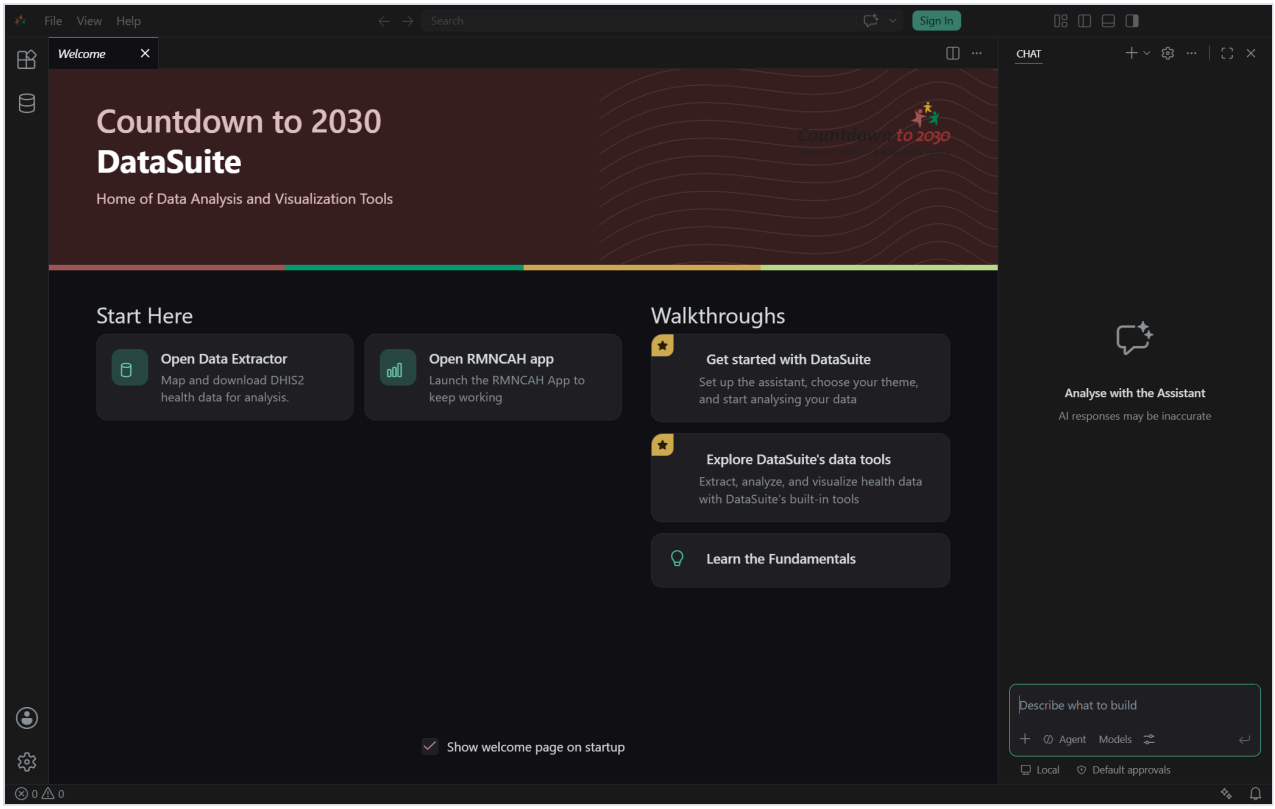


Figure 4.1 — *The DataSuite window: the Welcome page with Start Here and Walkthroughs, and the chat on the right*

Before you start, download <https://datasuite.damurka.com/en/downloads> and install (p. 12) DataSuite, and install R (p. 16). The apps need R.

What is new

- **Extensions.** The apps come as extensions from DataSuite's own registry, and you can add more. See Extensions (p. 36).
- **Language.** DataSuite is in English, French and Portuguese. See Change the display language (p. 30).

- **A first-start window** asks you to sign in and to choose a theme.
- **The Countdown icon** lists the published Countdown datasets your account can open.

The first start

The first time you open DataSuite, a **Welcome to DataSuite** window opens. It has three steps.

1. **Sign in.** Click **Sign in to DataSuite** to use the AI assistant and the Countdown datasets. Or click **Continue without Signing In**. You can sign in later.
2. **Make It Yours.** Under **Color Theme**, click a theme: **Countdown Dark**, **Dark High Contrast**, **Countdown Light** or **Light High Contrast**. Click **Continue**.
3. **Build with AI Agents.** This step shows what the chat can do. Click **Get Started**.

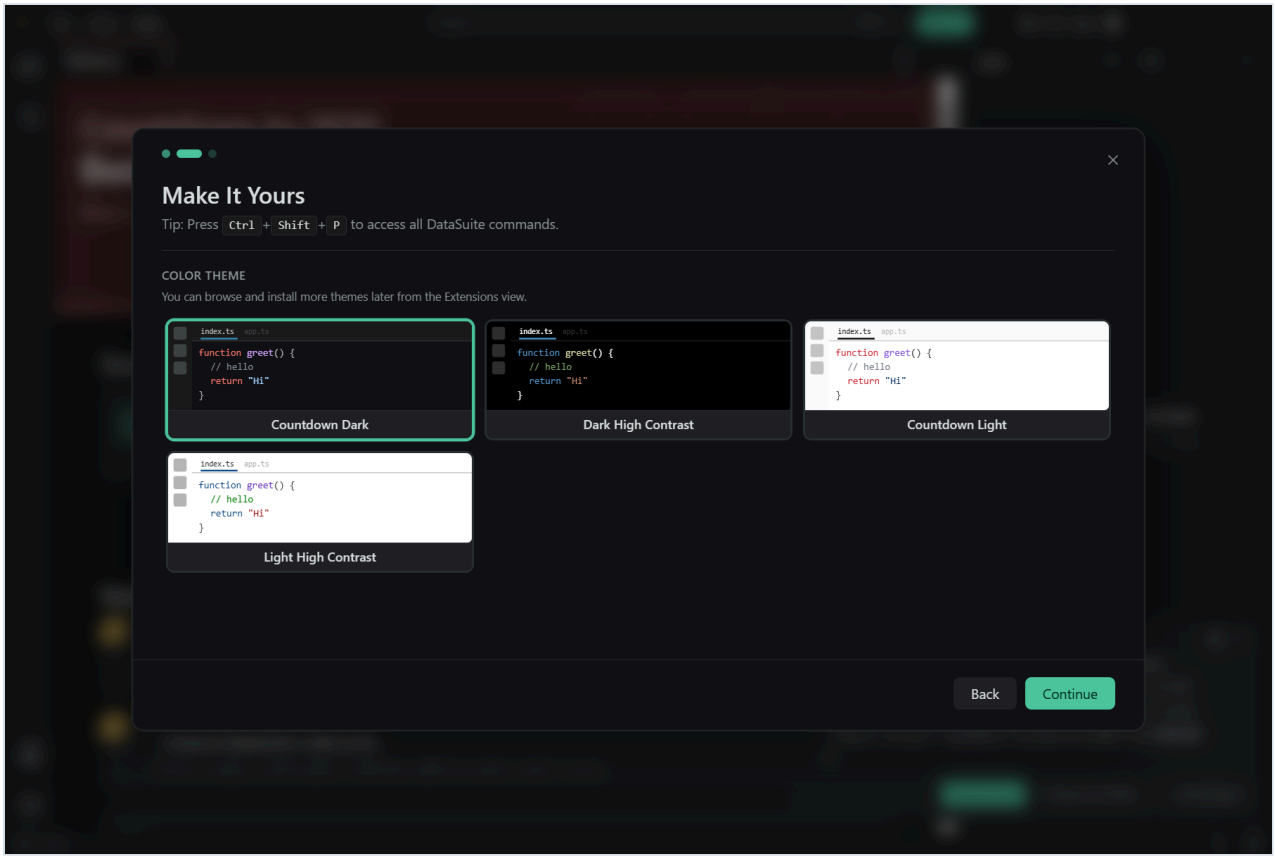
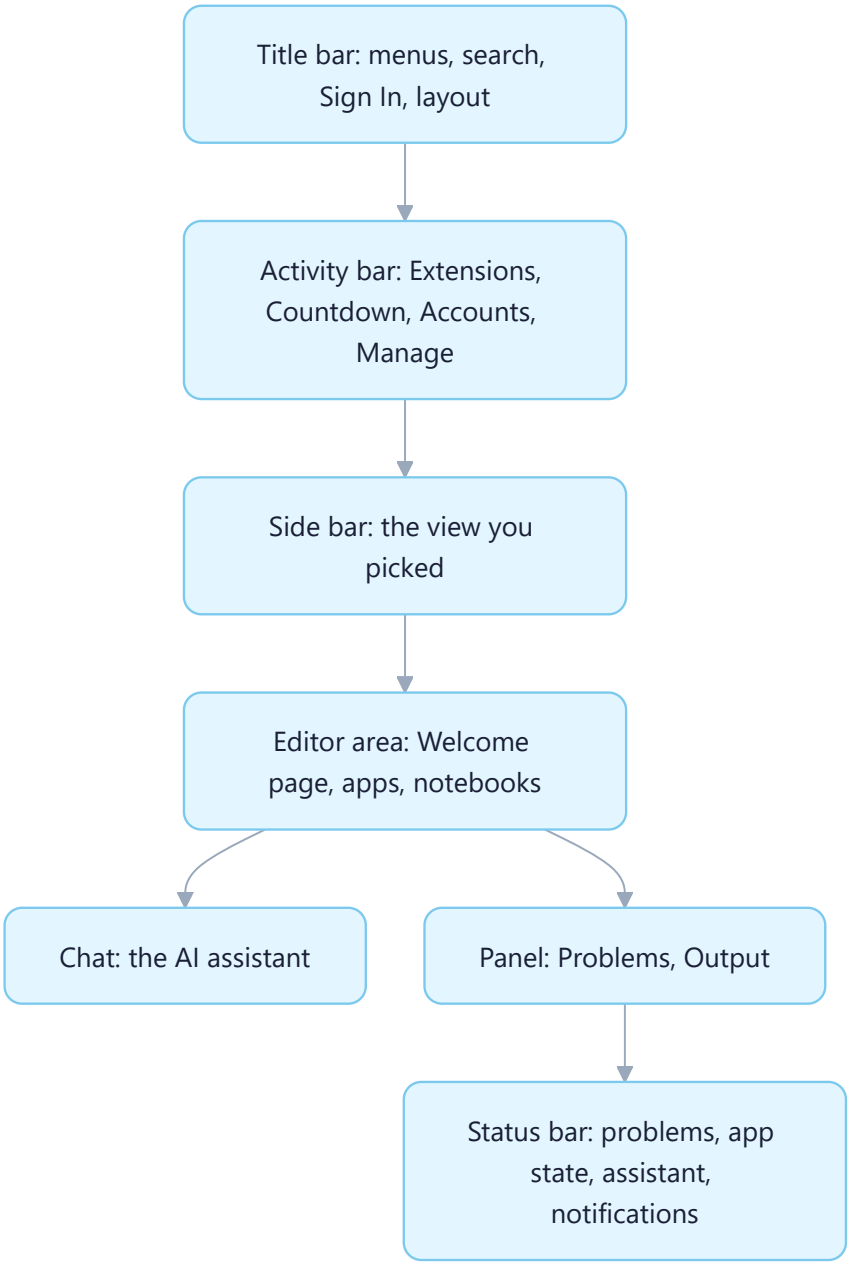


Figure 4.2 — *The Make It Yours step: four colour themes to choose from*

You can change all of this later. Close the window with × to skip it.

The parts of the window



① Title bar

At the top left are three menus: **File**, **View** and **Help**. See Menus and shortcuts (p. 27).

In the middle is the **Search** box. Type a file name to open the file. Type `>` to run a command.

On the right are **Sign In** and four layout buttons: **Customize Layout...**, **Toggle Primary Side Bar** (`Ctrl+B`), **Toggle Panel** (`Ctrl+J`) and **Toggle Secondary Side Bar** (`Ctrl+Alt+B`). The chat icon next to the search box is **Toggle Chat**.

② Activity bar

The column of icons on the left. Click an icon to open its view in the side bar. Click it again to close the side bar.

Icon	Opens
Extensions (`Ctrl+Shift+X`)	The extensions you have and those you can add. See Extensions (p. 36).
Countdown	The **Datasets** view: the published Countdown datasets your account can see. Download one, or open it in RMNCAH or Vaxx. Sign in first.
Accounts (at the bottom)	**Sign in to DataSuite…**, then your account once you are signed in.
Manage (the gear at the bottom)	**Command Palette…**, **Profiles**, **Settings**, **Extensions**, **Keyboard Shortcuts** and **Themes**.

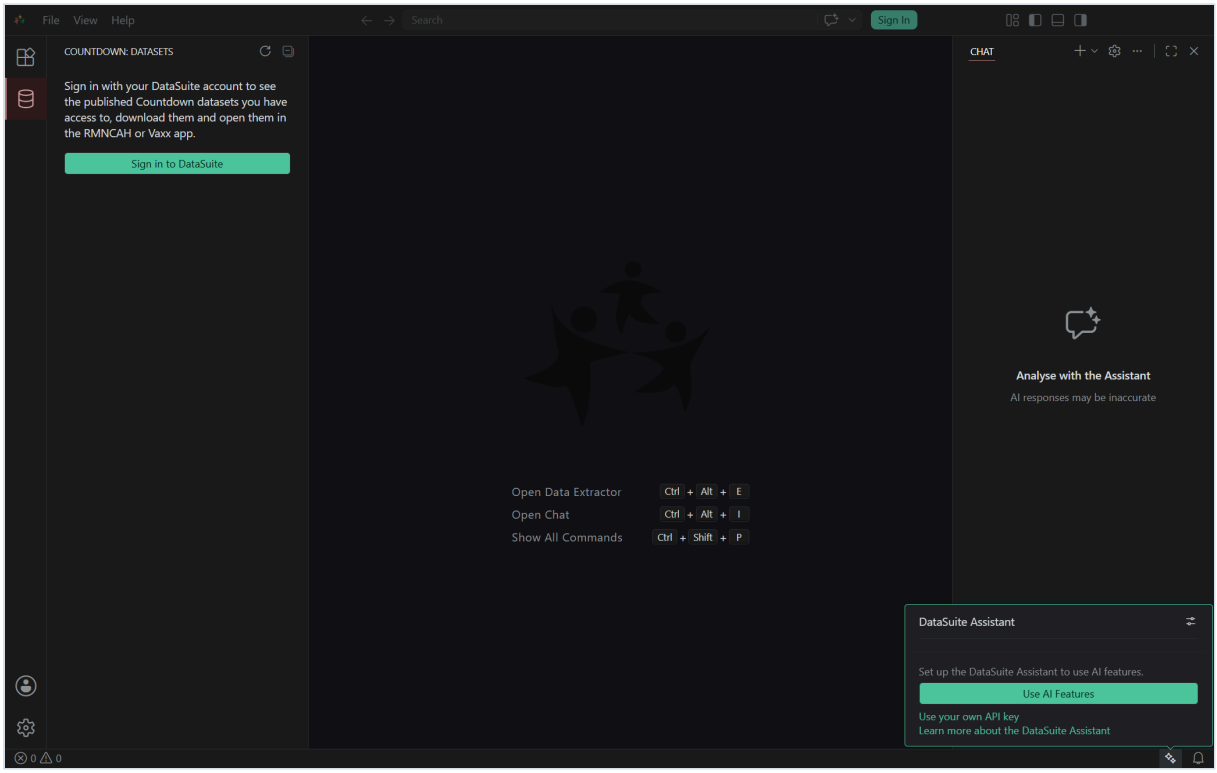


Figure 4.3 — *The Countdown Datasets view asks you to sign in; the DataSuite Assistant pop-up is open at the bottom right*

③ Welcome page

DataSuite opens on the **Welcome** page. It has two parts.

- **Start Here**: buttons that open the Data Extractor (p. 269) and the RMNCAH app.
- **Walkthroughs**: short guides. **Get started with DataSuite** sets up the assistant and your theme. **Explore DataSuite's data tools** shows the extractor, the apps and R in the chat. **Learn the Fundamentals** covers settings, extensions and shortcuts.

To open the page again, choose **Help > Welcome**. Clear **Show welcome page on startup** at the bottom if you do not want it at each start.

When no file or app is open, the editor area shows three shortcuts: **Open Data Extractor** (`Ctrl+Alt+E`), **Open Chat** (`Ctrl+Alt+I`) and **Show All Commands** (`Ctrl+Shift+P`).

④ Open Data Extractor

Opens the Data Extractor. It connects to DHIS2 (p. 498), maps indicators and downloads routine health facility data. Click **Open Data Extractor** on the Welcome page, choose **File > Open Data Extractor**, or press `Ctrl+Alt+E`. See Data Extractor (p. 269).

⑤ Open the apps

Open RMNCAH app on the Welcome page asks for a data file and opens the RMNCAH app on it. The **File** menu has all three apps: **Open Pooled...**, **Open RMNCAH...** and **Open Vaxx...**. Each app opens in its own tab. See Apps overview (p. 256) and Open the app (p. 295).

Once you have opened an app on a file, **File > Open Recent** lists it. Pick it to open the same app on the same file again.

⑥ Editor area and notebooks

Apps, notebooks, reports and settings open as tabs here. Drag a tab to the side to see two at once.

Notebooks run R, Python or Stata. Choose the kernel at the top right of the notebook: **R** (DataSuite's R, with the app packages), **Python** (DataSuite's Python environment) or **Stata** (Stata 17 or newer, on this computer). The **Variables** view lists what is in memory. Double-click a data frame to see it in the **Data Viewer**. See Requirements (p. 16) for what Python and Stata need.

⑦ Chat

The **Chat** opens on the right, in the secondary side bar. Type a question under **Analyse with the Assistant** and press `Enter`. Open it with **View > Chat** (`Ctrl+Alt+I`) or **Toggle Chat** in the title bar. The assistant needs a DataSuite sign-in or your own API key. See AI assistant (p. 267).

⑧ Panel

The panel opens under the editor area with `Ctrl+J`. **Problems** (`Ctrl+Shift+M`) lists errors and warnings. **Output** (`Ctrl+Shift+U`) shows the logs, for example an app's R session.

⑨ Status bar

The bar at the bottom of the window.

- **Left**: the number of errors and warnings. Click it to open **Problems**.
- **Shiny app**: when an app tab is active, its name and state. Green means it is running. Red means it has stopped. Click it to reload the page, restart the R session or open the app in your browser.
- **Right**: the **DataSuite Assistant status** icon. Click it to set up the assistant: **Use AI Features** or **Use your own API key**. The bell shows your **Notifications**.

Next steps

- Change the language or the colours: Language and themes (p. 30).
- Add tools: Extensions (p. 36).
- Get your data: Data Extractor (p. 269).
- If something does not work: Troubleshooting (p. 490).

4.1 Menus and shortcuts

The DataSuite menus (File, View, Help, Accounts and Manage), the commands and keyboard shortcuts people use most, update checks, and where DataSuite saves your settings.

DataSuite keeps its menus short. This page lists what is in each one, the shortcuts worth learning, and where your settings are saved. For a tour of the window, see Navigating the interface (p. 21).

The Command Palette

Every action in DataSuite is a command. Press `Ctrl+Shift+P` (or **View > Command Palette...**), type a few words and press `Enter`. For example, type `display language` or `color theme`.

The **Search** box in the title bar does the same when you type `>` first.

Menus

File

Item	What it does
New Window (`Ctrl+Shift+N`)	Opens a second DataSuite window.
Open Pooled..., **Open RMNCAH...**, **Open Vaxx...**	Asks for a data file and opens that app on it. See Apps overview (p. 256).
Open Data Extractor	Opens the Data Extractor (p. 269).
Open Recent	The apps and files you opened before. It appears once you have opened one.
Preferences	**Profiles**, **Settings** (`Ctrl+,`), **Color Theme** (`Ctrl+K Ctrl+T`) and **Display Language**. See Language and themes (p. 30).
Close Editor (`Ctrl+F4`)	Closes the active tab.
Close Window (`Alt+F4`), **Exit**	Closes the window, or DataSuite.

View

Item	What it does
Command Palette... (`Ctrl+Shift+P`)	Runs any command.
Open View...	Opens a view by its name.
Appearance	Shows or hides parts of the window, full screen and zoom.
Editor Layout	Splits the editor area to see several tabs at once.
Extensions (`Ctrl+Shift+X`)	Opens the Extensions view. See Extensions (p. 36).
Chat (`Ctrl+Alt+I`)	Opens the AI assistant. See AI assistant (p. 267).
Browser (`Ctrl+Alt+/`)	Opens a web page inside DataSuite.
Problems (`Ctrl+Shift+M`), **Output** (`Ctrl+Shift+U`)	Opens the panel on errors, or on the logs.

Help

Item	What it does
Welcome	Opens the Welcome page.

Show All Commands (`Ctrl+Shift+P`)	Same as the Command Palette.
Documentation	Opens this site.
Open Walkthrough...	Opens one of the Welcome page guides.
Show Release Notes	What changed in this version. See Release notes <https://datasuite.damurka.com/en/release-notes>.
Ask @datasuite	Asks the assistant how to do something in DataSuite.
Video Tutorials	Opens the videos.
Search Feature Requests, **Report Issue**	Tells the DataSuite team what you need or what went wrong.
View License, **Privacy Statement**	The licence and the privacy statement.
Toggle Developer Tools, **Open Process Explorer**	For finding faults.
About	The DataSuite version. Give it when you report a problem.

Accounts and Manage

The two icons at the bottom of the activity bar.

- **Accounts**: **Sign in to DataSuite...**. One sign-in serves the AI assistant and the Countdown datasets. Once you are signed in, your account and **Sign Out** are here.
- **Manage** (the gear): **Command Palette...**, **Profiles**, **Settings**, **Extensions**, **Keyboard Shortcuts** and **Themes**. **Check for Updates...** is here too.

Updates

DataSuite looks for a new version by itself and tells you when one is ready. To look now, open **Manage** (the gear) and choose **Check for Updates...**. When an update is downloaded, the same menu shows **Restart to Update (1)**.

The setting **Update: Mode** (`update.mode`) controls this. Set it to `none` to stop update checks, or to `manual` to check only when you ask. The apps' extensions update by themselves too. See Extensions (p. 38).

Keyboard shortcuts

Shortcut	What it does
`Ctrl+Shift+P`	Command Palette
`Ctrl+Alt+E`	Open Data Extractor
`Ctrl+Alt+I`	Open the chat
`Ctrl+Shift+X`	Extensions view
`Ctrl+K Ctrl+T`	Color Theme
`Ctrl+,`	Settings
`Ctrl+B`	Show or hide the side bar
`Ctrl+Alt+B`	Show or hide the chat side bar

`Ctrl+J`	Show or hide the panel
`Ctrl+F4`	Close the active tab

To change a shortcut, open **Manage > Keyboard Shortcuts** (`Ctrl+K Ctrl+S`), find the command and double-click it.

Where your settings are saved

Open **File > Preferences > Settings** (`Ctrl+,`) to change a setting. Search for it by name. DataSuite saves what you change in a file:

What	Windows	Linux
Settings, such as the colour theme	`%APPDATA%\DataSuite\User\settings.json`	`~/.config/DataSuite/User/settings.json`
Display language (`locale`)	`%USERPROFILE%\.datasuite\argv.json`	`~/.datasuite/argv.json`
Extensions you added	`%USERPROFILE%\.datasuite\extensions`	`~/.datasuite/extensions`

In portable mode (p. 19), all of this lives in the `data` folder next to DataSuite. To reset DataSuite, see Uninstall <https://datasuite.damurka.com/en/get-started/install/uninstall#clean-uninstall>.

4.2 Language and themes

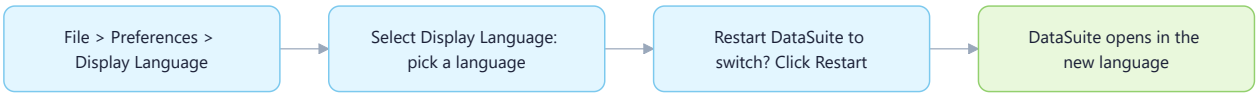
How to use DataSuite in English, French or Portuguese with its language packs, how to choose a light, dark or high contrast colour theme, and where DataSuite saves both choices.

DataSuite comes in three languages and six colour themes. Both are yours to change at any time. Your choices are kept for the next start.

Change the display language

DataSuite is in **English**, **French** (`français`) and **Portuguese** (`português (África)`, as written in Africa). The French and Portuguese **language packs** come with DataSuite, so you do not need the internet to switch.

At first start, DataSuite uses your computer's language if it is French or Portuguese. Otherwise it starts in English.



1. **Open the language list**
   Choose **File > Preferences > Display Language**. Or press `Ctrl+Shift+P` and run **Configure Display Language**.

② Pick a language

The list **Select Display Language** shows the installed languages: **English (en)**, **français (fr)** and **português (África) (pt)**. The current one is marked **(Current)**.

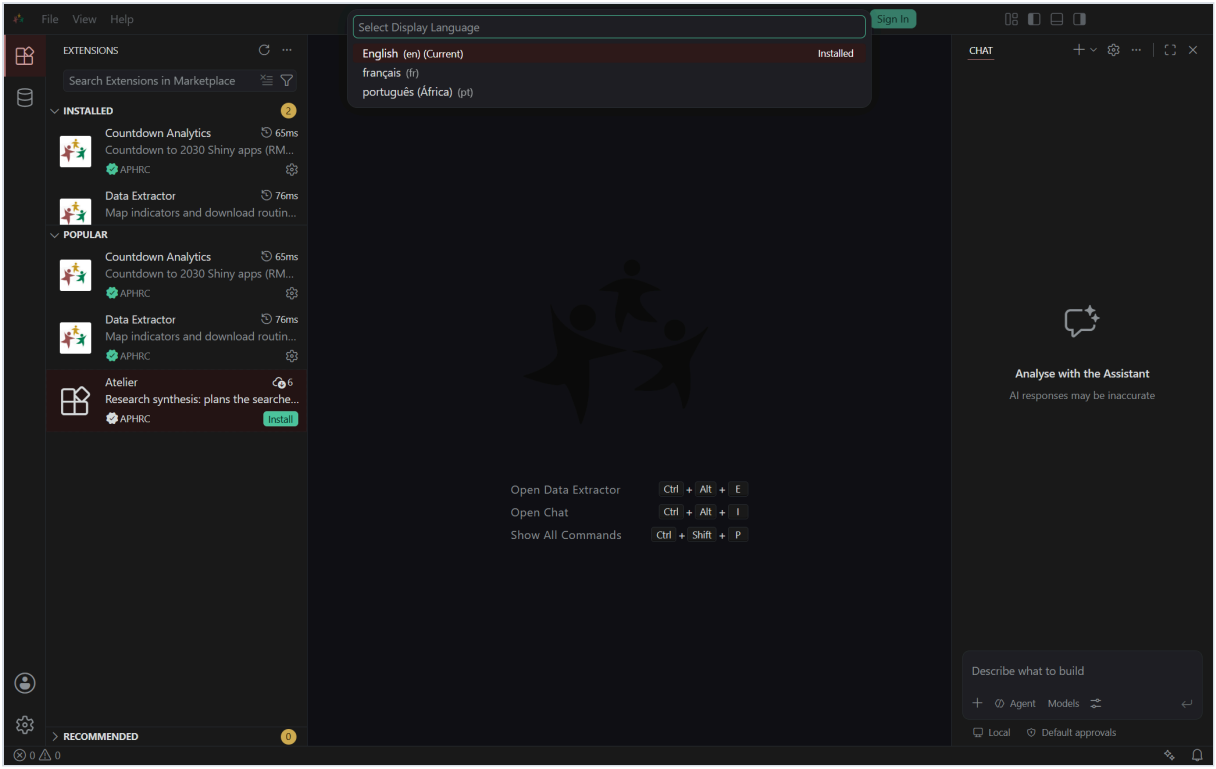


Figure 4.4 — *The Select Display Language list with English, français and português (África)*

③ Restart

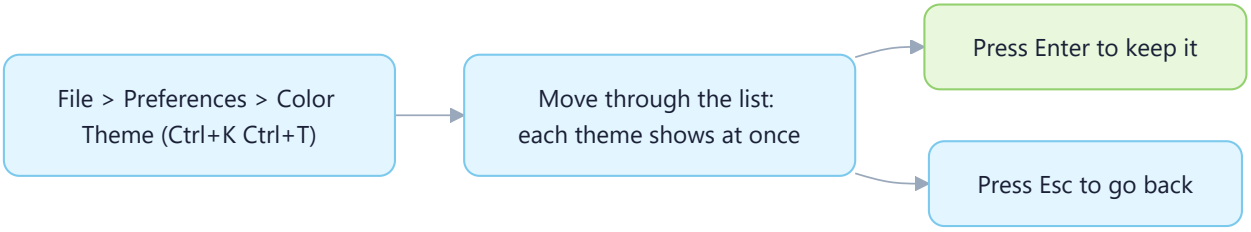
DataSuite asks **Restart Countdown DataSuite to switch to ...?** Click **Restart**. DataSuite closes and opens again in the new language.

The menus, the Welcome page, the settings and the apps follow the display language. A few new labels may still be in English until the next language pack.

To go back to your computer's language, run **Clear Display Language Preference** from the Command Palette and restart.

Change the colour theme

A colour theme sets the colours of the whole window. DataSuite starts with **DataSuite Analysis Dark**.



1. **Open the theme list**
   Choose **File > Preferences > Color Theme**, or **Manage > Themes > Color Theme**, or press `Ctrl+K Ctrl+T`.

② Try the themes

Move up and down the list with the arrow keys. DataSuite shows each theme as you move.

DARK

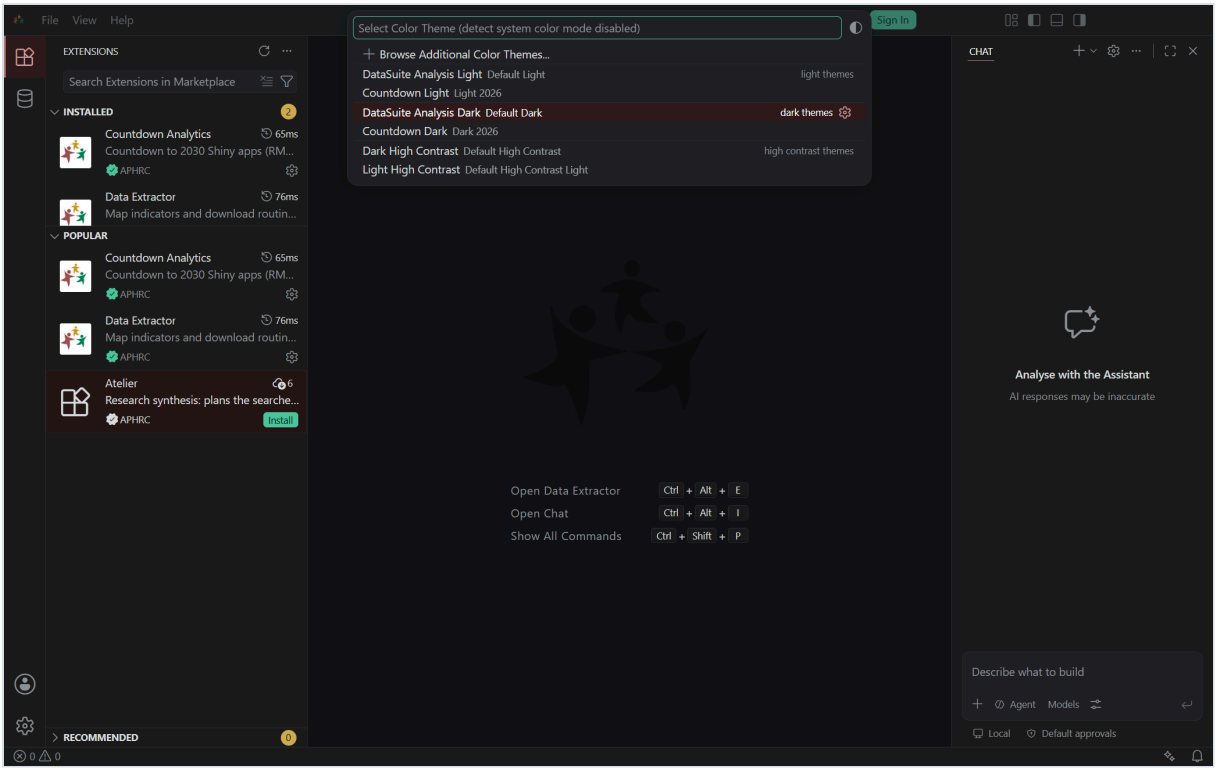


Figure 4.5 — *The Select Color Theme list with the dark theme in use*

LIGHT

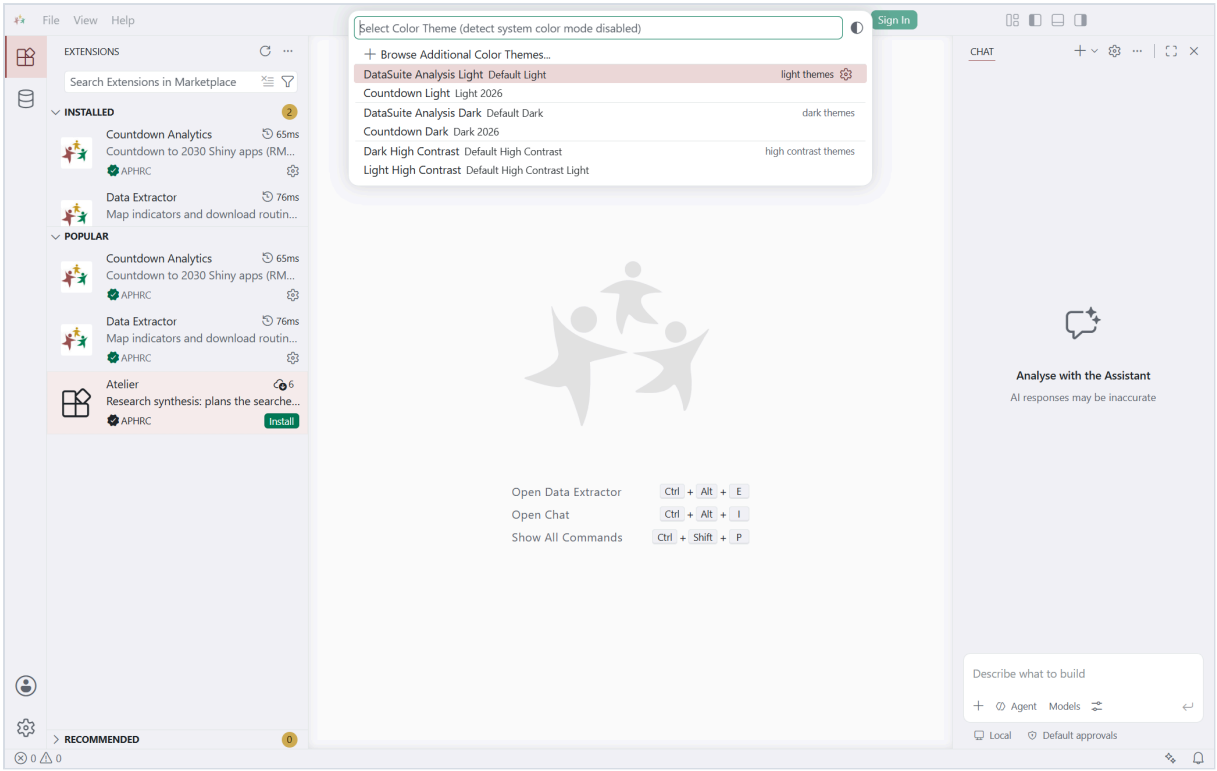


Figure 4.6 — *The same list with DataSuite Analysis Light shown*

③ Keep it

Press `Enter` to keep the theme. Press `Esc` to go back to the theme you had.

The themes that come with DataSuite:

Theme	Kind	Notes
DataSuite Analysis Light	Light	The default light theme.
Countdown Light	Light	The Countdown colours on a light background.
DataSuite Analysis Dark	Dark	The default theme.
Countdown Dark	Dark	The Countdown colours on a dark background.
Dark High Contrast	High contrast	Strong contrast for low vision or bright rooms.
Light High Contrast	High contrast	The same, on white.

The **Get started with DataSuite** walkthrough on the Welcome page also has a **Choose your theme** step, with **Browse Color Themes**.

More themes

Browse Additional Color Themes... at the top of the list searches the DataSuite extension registry for themes. See Extensions (p. 36).

DataSuite has no file icon theme of its own. **Manage > Themes > File Icon Theme** lists any you install.

Follow your computer's light or dark mode

The half-filled circle at the right of the theme list turns on **detect system color mode**. DataSuite then uses your light theme when Windows is light and your dark theme when it is dark. The settings **Preferred Light Color Theme** (`workbench.preferredLightColorTheme`) and **Preferred Dark Color Theme** (`workbench.preferredDarkColorTheme`) choose which.

Where your choices are saved

Choice	Saved in	Setting
Display language	`%USERPROFILE%\.datasuite\argv.json` (Linux: `~/.datasuite/argv.json`)	`"locale": "fr"`
Colour theme	`%APPDATA%\DataSuite\User\settings.json` (Linux: `~/.config/DataSuite/User/settings.json`)	`"workbench.colorTheme"`

In portable mode (p. 19), both are in the `data` folder next to DataSuite. See also Where your settings are saved (p. 29).

4.3 Extensions

What DataSuite extensions are and which ones exist, and how to find, install, update, turn off and remove them, from DataSuite's own registry or from a file when there is no internet.

Extensions add apps and tools to DataSuite. The RMNCAH, Vaxx and Pooled apps and the Data Extractor are extensions. DataSuite installs them for you and keeps them up to date. You can add more from the **DataSuite registry**.

Where extensions come from

DataSuite gets its extensions from its own registry at datasuite.damurka.com/registry <https://datasuite.damurka.com/registry>. It does not use the Visual Studio Marketplace. Extensions from the DataSuite team are published by **APHRC** and carry a verified mark.

You can browse the registry in a web browser too. Its categories include **Shiny Apps**, **Python Apps** and **Data Extractors**.

The DataSuite extensions

Extension	Comes with DataSuite	What it adds
Countdown Analytics	Yes, updates itself	The **RMNCAH**, **Vaxx** and **Pooled** apps; the **Countdown** icon with the **Datasets** view; what the assistant knows about the Countdown method and your dataset. See Apps overview (p. 256).
Data Extractor	Yes, updates itself	**Open Data Extractor**: map indicators and download routine data from DHIS2 (p. 498) to Excel. See Data Extractor (p. 269).
DataSuite Assistant	Yes, updates itself	The AI assistant in the chat. See AI assistant (p. 267).
French and Portuguese language packs	Yes	DataSuite in French and Portuguese. See Language and themes (p. 30).
Countdown Themes	Yes	The Countdown and DataSuite Analysis colour themes.
R, Python and Stata Language Basics	Yes	Colours and brackets in R, Python and Stata code and do-files.
Atelier	No, install it	Research synthesis: it searches PubMed, Europe PMC, OpenAlex and Semantic Scholar, ranks the papers and writes a cited summary with DataSuite's AI.

The extensions that come with DataSuite cannot be uninstalled. You can turn them off.

Open the Extensions view

Click **Extensions** in the activity bar, or press `Ctrl+Shift+X`, or choose **View > Extensions**.

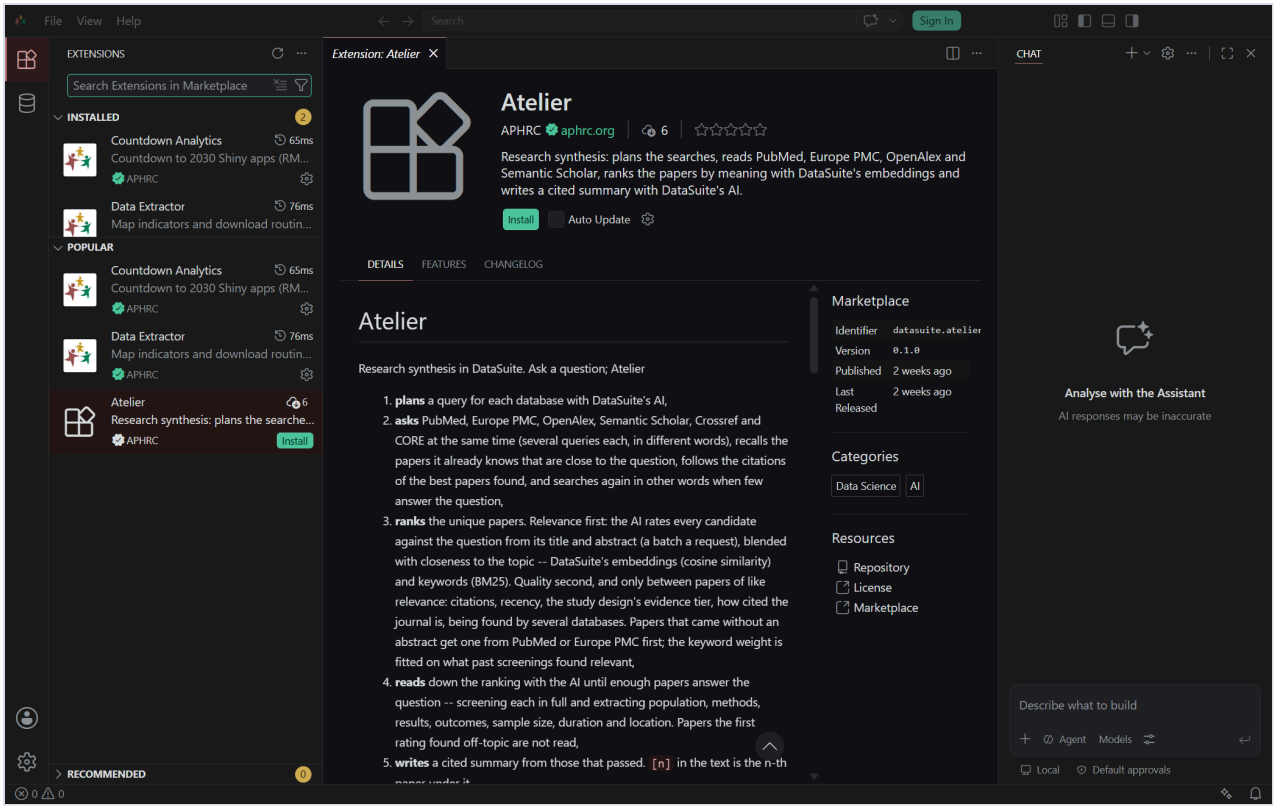


Figure 4.7 — *The Extensions view: Installed and Popular lists, with the Atelier extension page open*

The view has three lists:

- **Installed**: the extensions you have, with Countdown Analytics and the Data Extractor.
- **Popular**: the most installed extensions in the registry.
- **Recommended**: extensions suggested for what you have open.

Click an extension to open its page: **Details**, **Features** and **Changelog**. The changelog says what each version changed.

Install an extension



① Search

Type a name or a word in **Search Extensions in Marketplace**. The list shows what the registry has.

② Install

Click **Install** on the extension, in the list or on its page. DataSuite downloads it and turns it on.

③ Use it

The extension now appears under **Installed**. An app extension adds its app to the **File** menu, and sometimes to **Start Here** on the Welcome page. If DataSuite asks to reload or restart, do so.

Only install extensions you trust. An extension can read your files and run programs.

Update extensions

DataSuite updates your extensions by itself. Countdown Analytics, the Data Extractor and the DataSuite Assistant update too, so app fixes do not wait for a new DataSuite version.

- To check now, open the **…** menu at the top of the Extensions view and choose **Check for Extension Updates**.
- An extension with a new version shows an **Update** button. Click it.
- On an extension page, clear **Auto Update** to keep it at its version.

To update DataSuite itself, see Updates (p. 28).

Turn an extension off or remove it

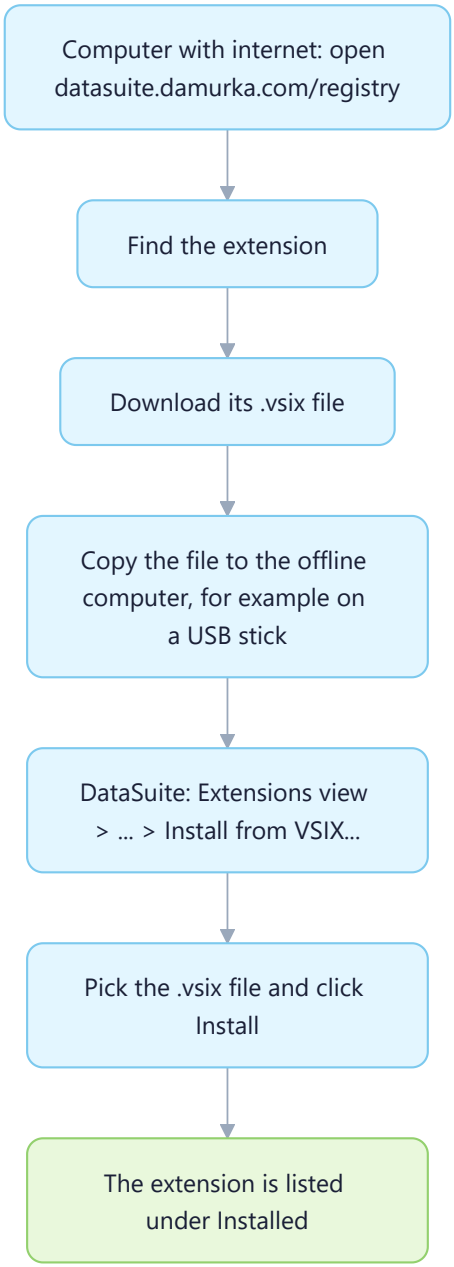
Click the gear on the extension in the list, or open its page.

- **Disable** turns it off and keeps it. **Enable** turns it on again.
- **Disable (Workspace)** turns it off only for the folder you have open.
- **Uninstall** removes it. The extensions that come with DataSuite have no **Uninstall**: disable them instead.

DataSuite may ask you to reload to finish.

Without internet

DataSuite works offline. The extensions that come with it are already installed, and so are the language packs. To add an extension on a computer without internet, bring it as a `.vsix` file.



1. **Download the file**

   On a computer with internet, open datasuite.damurka.com/registry <https://datasuite.damurka.com/registry>. Find the extension and download its `.vsix` file.

2. **Install it**

   On the offline computer, open the Extensions view. Open the **…** menu at the top and choose **Install from VSIX…**. Or run **Extensions: Install from VSIX…** from the Command Palette. Pick the file and click **Install**.

An extension installed this way does not update itself while you are offline. Install the newer `.vsix` the same way.

An app extension still needs R, and the R packages of its app. DataSuite installs those the first time the app opens, which needs the internet. See Requirements (p. 16).

Where extensions are saved

Extensions you install are in `%USERPROFILE%\.datasuite\extensions` (Linux: `~/.datasuite/extensions`). The extensions that come with DataSuite are in its install folder. In portable mode (p. 19), extensions are in the `data` folder next to DataSuite.

PART II

Methodology

Overview of the CD2030 DataSuite analytical framework: from data quality and adjustment to denominators, coverage and subnational inequality.

CHAPTER 5

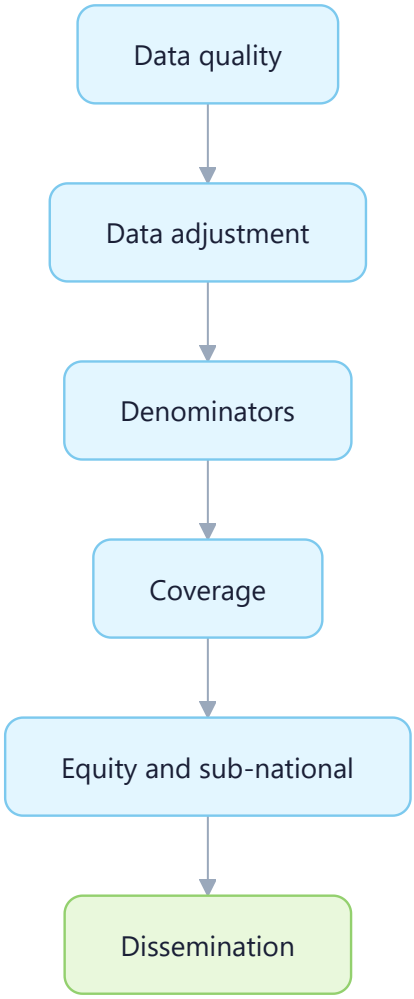
Framework overview

Overview of the CD2030 DataSuite analytical framework: from data quality and adjustment to denominators, coverage and subnational inequality.

From raw facility data to coverage you can defend

Facility data are plentiful but often unreliable; surveys are reliable but rare and expensive. The framework checks and corrects the facility data, calibrates them against surveys, and turns them into coverage trends you can trust.

Learn about the end-to-end workflow (p. 43)



End-to-end workflow

The workflow follows a logical sequence: assess data quality, correct for incompleteness and errors, select evidence-based denominators, calculate coverage, and examine inequality patterns at national and geographic levels.

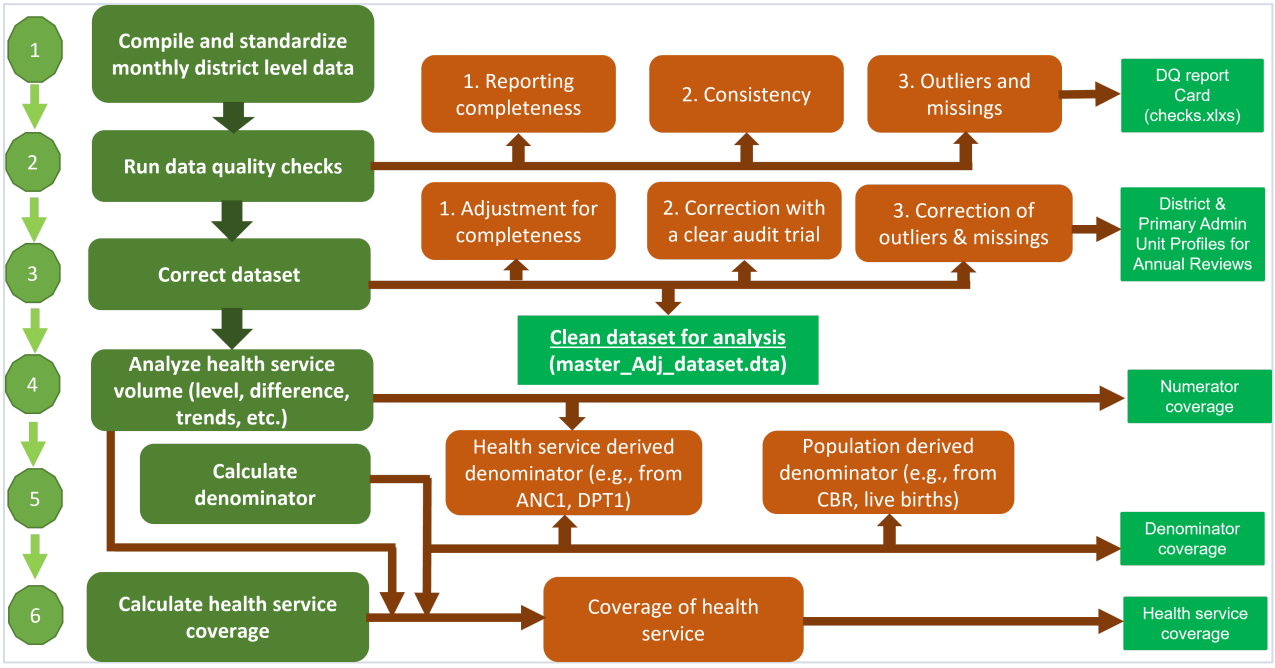


Figure 5.1 — *Countdown analysis workflow*

The framework follows a practical sequence:

1. Load and configure (p. 44) facility and survey data.
2. Calibrate key national benchmark rates (p. 54).
3. Assess data quality (p. 57) before estimating coverage.
4. Adjust (p. 75) incomplete, unstable, or unreliable data.
5. Select the best-performing denominator (p. 95).
6. Generate and interpret **coverage** (p. 115) (including **Bayesian coverage estimation** (p. 143)), **equity** (p. 153), **sub-national** (p. 165), **mortality** (p. 179), **service utilization** (p. 204) and **health system performance** (p. 225) results, in that order.
7. Plan **dissemination, decisions and action** (p. 245).

In the apps, the same sequence is followed page by page: see Analysis steps (p. 291) (RMNCAH app).

The final purpose of the framework is practical use. Findings should inform country review processes, planning discussions, and technical working groups so that the analysis leads to concrete RMNCAH action rather than standing alone as a report.

CHAPTER 6

Analysis setup

How to set up the analysis: the facility data workbook from HMIS/DHIS2 and the checks it must pass, the household survey files, the global reference estimates, and the national survey parameters, and how they are loaded.

The analysis setup prepares the data inputs of the analysis: the country's health facility data (by month, district and indicator, from DHIS2 (p. 498)), the household survey (p. 500) results, the global estimates of population and other indicators, and the national survey parameters. Once loaded and checked, they form the master facility crude data set on which every later step works.

It is the first step of the analysis. The next step is the Data quality assessment (p. 57).

At a glance

- **Answers:** what data, reference datasets and national parameters does the analysis need, and how are they loaded?
- **Uses:** the Excel workbook generated from HMIS (p. 499)/DHIS2 exports (see the DHIS2 data indicator definitions (p. 509)), the optional survey files and shapefile, the embedded reference datasets, and the national survey parameters.
- **Produces:** a working folder with the companion `.rds` file, the reference datasets matched to the country, and the national parameters used in the data quality assessment (p. 57) and for denominator selection (p. 95).
- **In the apps:** Getting started (p. 295) (the **Load Data** wizard).

Why

Routinely reported health facility data are an important data source for health indicators at facility and population level. Household surveys, mostly DHS or MICS, are the primary source of population-based coverage statistics, which can be disaggregated by socioeconomic characteristics, place of residence, subnational area, age and gender. Global estimates of population and immunization coverage complete them. Both kinds of data are critical and need to be considered in conjunction with each other, so all of them are loaded, checked and matched to the same country and regions before the analysis starts.

How the parts fit together

The facility data workbook is required; everything else is matched to it. The workbook must first pass the blocking checks. The survey files and a custom shapefile are optional, but their region names must then be matched to the workbook's. The global estimates are chosen automatically from the

workbook's country. The national survey parameters are always required: they are pre-filled from the national survey file when one is uploaded, and must be checked.

Doing it in the app

1. **Prepare a working folder**
   Create a dedicated folder and put the Excel workbook in it: see Working folder (p. 47).
2. **Load the data**
   Open the app on the workbook (see Getting started (p. 295)). The **Upload Data** step shows the loaded file and lets you replace the built-in reference data.
3. **Pass the checks**
   The **Data Quality** step runs the checks on the uploaded data (p. 48). Fix any blocking problem in the workbook and load it again.
4. **Add the survey files and the shapefile (optional)**
   Upload the survey files (p. 50) and, if needed, a custom shapefile; then match their region names in **Map Survey** and **Map Shapefile**.
5. **Confirm the national rates**
   On **National Rates**, check every national survey parameter (p. 54).
6. **Finish loading**
   Press **Finish Loading Data**. The app saves the dataset and shows **Ready for analysis**.

Points to consider

- Are all the years and districts you want to analyse in the workbook?
- Are the national survey parameters from the most recent national survey, and validated?
- Do the survey and shapefile region names match the dataset's?

Common pitfalls

- Loading a changed workbook while its saved `.rds` copy is still in the folder: the app opens the saved copy instead. Delete the copy to load the workbook afresh.
- Accepting the pre-filled national parameters without checking them: they feed the expected ratios of the data quality assessment and the denominators.
- Mixing several countries or export batches in one working folder.

6.1 Country input facility data set

The workbook of monthly health facility data by district and indicator, exported from HMIS/DHIS2, and how it is loaded.

The country input facility data set is the country's DHIS2 (p. 498) data by month, district and indicator, in an Excel workbook. The app reads it, checks it and saves it as the master facility crude data set, the starting point of the data quality assessment (p. 57) and every later step.

At a glance

- **Holds:** the monthly service numbers, reporting completeness, population and administrative data of every district.
- **Format:** an Excel workbook (`.xlsx` or `.xls`) from the Data Extractor (p. 269) the first time; the saved `.rds` afterwards.
- **In the app:** the **Upload Data** step of Getting started (p. 298).

Why

Routinely reported health facility data are an important data source for health indicators at facility and population level. The data are reported by health facilities on events such as immunizations given, or live births attended. They give the numerators of the coverage statistics, for the country and for every region and district, year after year.

What

The workbook, in the format produced by the Data Extractor (p. 269), must contain these sheets:

Sheet	What it holds
`Admin_data`	The country, its first administrative level (Admin 1 (p. 496)) and districts, with the number of facilities and health workers (Admin data (p. 496))
`Population_data`	Population, births and target groups by district and year (Population data (p. 502))
`Reporting_completeness`	Monthly reports expected and received, and the reporting rate, for each reporting form (Reporting completeness data (p. 503))
`Service_data` (at least one)	Monthly numbers of services, such as antenatal care visits, deliveries, vaccinations and outpatient visits (Service data 1 (p. 503), Service data 2 (p. 503), Service data 3 (p. 503))

Each indicator is defined in the DHIS2 data indicator definitions (p. 509).

The app accepts these files for the facility data:

Format	When to use it
Excel workbook: `.xlsx` or `.xls`	The first time: the workbook with the required sheets.
Stata dataset: `.dta`	A single, already merged dataset.
A dataset already saved by the app: `.rds`	Every later time: it opens without the checks.

Working folder

Prepare a dedicated working folder before loading the data. It is where you keep the input files for the app, and where the app creates the saved data file used for faster loads.

- Choose a clear folder name that you can easily remember and reference, preferably on your computer.
- Place the Excel workbook in that folder before starting the app.
- Keep any related documentation, data or country-specific files together, so the working folder is self-contained.

> **Best practice**
>
> - Use one working folder per country or DHIS2 export batch.
> - Keep the original Excel source file and the generated `.rds` file together.
> - Do not rename the `.rds` file: the app manages it.

Loading the data

For the first load, use the Excel workbook generated from the HMIS/DHIS2 exports. For later loads, use the saved file, which the app creates in the working folder when the workbook is loaded.

The saved .rds file

- When you load an Excel workbook for the first time, the app converts it and saves a companion `.rds` file (p. 503) in the same working folder.
- On later loads, the app looks for that `.rds` file and uses it, which is much faster.
- When you open the same Excel or Stata file again, the saved copy is loaded instead. Delete it if you changed the workbook and want to load it afresh.

The first load may therefore take a little longer, while the workbook is read and transformed.

Interpretation

- **Does the workbook cover every year and district you want to analyse?** Missing years or districts cannot be added later in the app.
- **Did you change the workbook since the last load?** Then delete the saved `.rds` copy first, or the app opens the old data.
- **Is the dataset for the right analysis?** A dataset saved by the RMNCAH app cannot be opened in the Vaxx app, and the other way round.

After the setup

When loaded, the workbook goes through the checks on the uploaded data (p. 48). Once they pass and loading is finished, the data quality assessment can start.

6.2 Checks on the uploaded data

The checks the facility data workbook must pass when it is loaded: blocking, informational and mapping checks.

When the facility data are loaded, the app checks that the workbook is complete and consistent before the analysis can go on. These checks are about the structure of the file (sheets, columns, names, months), not about the quality of the numbers, which the data quality assessment (p. 57) examines afterwards.

> **At a glance**
>
> - **Checks:** that the sheets, columns, country, districts and months of the workbook can be read and match.
> - **Good result:** every blocking check (p. 497) passes.
> - **In the app:** the **Data Quality** step of Getting started (p. 299).

Why

The analysis adds up monthly numbers by district, year and indicator, and matches them with population and reporting data and with survey regions. A missing sheet, a district spelled differently in two sheets, or a missing month would silently produce wrong totals. Checking them at load time stops the analysis before it starts on broken data.

What

There are three kinds of check.

Kind	Checks	What happens if one fails
Blocking	Admin, Population and Service sheets found; country, admin-1 and district columns present in the Admin sheet; exactly one country in the Admin sheet, recognised by name; districts match between the Admin sheet and every other sheet; district names consistent across years; no missing months (every district-year has all 12); every month name recognised (English or French)	The analysis cannot continue until the workbook is corrected and loaded again
Informational	Estimated live births not lower than facility (reported institutional) live births, with a 10% margin; population data not pasted into service data; every indicator has some data; month names labelled consistently across sheets	A warning; the analysis can continue
Mapping	Survey or shapefile region names that do not match the dataset	The names are matched to the dataset's region names in the mapping steps (see Matching region names (p. 51))

Interpretation

- **Did a blocking check fail?** Read the failed check, correct the workbook (the source file) and load it again. Delete the saved `.rds` copy first, or the app reopens the old data.
- **Are estimated live births lower than the live births reported by facilities?** The population projections in the workbook are probably too low; this will matter when denominators are assessed in Denominator assessment and selection (p. 95).
- **Does an indicator have no data at all?** It may not be collected in the country, or its column may be empty in the export. It will be missing from every later output.
- **Are there region name mismatches?** Match them in the mapping steps before finishing the setup.

After the setup

Once the checks pass and loading is finished, the data quality of the numbers themselves is assessed in the Data quality assessment (p. 57): completeness of reporting, missing values, extreme outliers and consistency.

6.3 Household survey results

The optional survey data files and shapefile, and how their region names are matched to the facility data.

Household survey results, such as those of a DHS or MICS, give population-based coverage nationally, by region and by socioeconomic group. They can be loaded as survey data files, with a shapefile (p. 503) for maps. Both are optional, but when they are used their region names must match the facility data.

> **At a glance**
>
> - **Holds:** survey coverage results nationally, by region, by area of residence, by maternal education and by wealth quintile.
> - **Format:** up to five Stata files (`.dta`); a shapefile (`.shp` , `.dbf` , `.shx` , `.prj`) for maps.
> - **In the app:** the **Survey Files**, **Shapefile**, **Map Survey** and **Map Shapefile** steps of Getting started (p. 300).

Why

Household surveys (p. 500), mostly modules of DHS or MICS, or special coverage surveys, are the primary source of population-based coverage statistics, which can be disaggregated by socioeconomic characteristics, place of residence (urban/rural), subnational area, age and gender. Facility data and surveys are both critical pieces of information and need to be considered in conjunction with each other: the survey results are the benchmark against which the facility-based coverage is compared, and the source of the equity analysis.

What

There are five survey data files (`.dta`), each optional:

File	Survey coverage results
National data	For the whole country
Regional data	By region (admin 1)
Area data	By area of residence (urban/rural)
Maternal education data	By the mother's level of education
Wealth index quintile data	By household wealth quintile

Together they power the equity analysis (the Equity Assessment), the survey comparison lines and the subnational comparisons. Without them, the app keeps its built-in reference data (see Global estimates (p. 52)). A national survey file also pre-fills the national survey parameters (p. 54).

A shapefile is used for maps. A standard shapefile for the country is built in; a custom one can be used instead, by choosing the column that holds the region (admin-1) names.

Matching region names

When the survey files or a custom shapefile name their regions differently from the dataset, their region (admin-1) names must be matched to those in the dataset. Names are matched automatically where possible; the analyst reviews the matches, corrects them and saves them, at least once, even if nothing needs correcting.

Interpretation

- **Are the survey files from the most recent survey?** Survey comparisons and the equity analysis are only as recent as the files.
- **Does every survey region have a match in the dataset?** A region left unmatched is missing from the subnational comparisons and maps.
- **Did regions change since the survey (split or merged)?** Then a one-to-one match is not possible for those regions; note it when comparing.

After the setup

The survey results are used in the comparisons of Coverage (p. 115), in the equity analysis and, through the national survey parameters, in the Data quality assessment (p. 57) and Denominator assessment and selection (p. 95).

6.4 Global estimates

The reference datasets built into the app (UN estimates, survey data, WUENIC, UN mortality data, FPET) and how they are matched to the country.

Besides the facility data, the analysis uses global estimates of population and other indicators: population and birth projections, immunization coverage estimates, mortality and family planning estimates, and survey results. They are built into the app as reference datasets and chosen automatically for the country whose facility data are loaded.

> **At a glance**
>
> - **Holds:** UN estimates, survey data, WUENIC estimates, UN mortality data and FPET data for the country.
> - **Format:** built in; the UN, WUENIC (and, in RMNCAH analyses, UN mortality) estimates can be replaced with your own `.dta` files.
> - **In the app:** the **Additional reference data** card of the **Upload Data** step in Getting started (p. 298).

Why

The facility data alone cannot say whether their population projections are plausible, or how facility-based coverage compares with other estimates. Global estimates give that comparison: the UN projections are used to assess the population projections in the facility data and as one of the denominator options, and the WHO and UNICEF estimates of immunization coverage are compared with the annual coverage estimated from the facility data.

What

Reference dataset	What it holds
UN estimates (p. 504)	UN population projections: total population, births and target groups by year
Survey data	Coverage results of the country's household surveys
WUENIC (p. 505)	The WHO and UNICEF estimates of national immunization coverage, produced annually with countries from all data sources
UN mortality data	UN estimates of mortality rates
FPET (p. 499) data	Family planning estimates from the Family Planning Estimation Tool

Because these datasets are named after the country, the app matches them automatically to the country of the uploaded facility data: there is nothing to choose.

You can replace the built-in UN estimates and WUENIC estimates (and, in RMNCAH analyses, the UN mortality estimates) with your own `.dta` files, for example a more recent release. **Restore default** returns to the built-in data.

Interpretation

- **Is the country recognised?** The reference datasets are matched through the country name in the `Admin_data` sheet; an unrecognised name stops the loading (see Checks on the uploaded data (p. 48)).
- **Are the built-in estimates recent enough?** If a more recent UN or WUENIC release is available, upload it.

After the setup

The global estimates are used in Denominator assessment and selection (p. 95) (UN projections) and in Coverage (p. 115) (comparison with WUENIC and the surveys).

6.5 National survey parameters

The national survey coverage values and demographic rates the analysis needs to check the data and derive denominators.

The national survey parameters are a short list of national values: coverage of key services from the most recent household survey (p. 500) (such as DHS or MICS), and rates of pregnancy loss (p. 503), twin births (p. 504), stillbirth (p. 504) and infant mortality. They are used to compute expected ratios and denominators from the facility data, and are applied uniformly across subnational units.

At a glance

- **Holds:** survey coverage of antenatal care, delivery and vaccination services, five demographic rates, and two survey years.
- **Required:** every field; pre-filled from the national survey file or the country's survey data, to be checked.
- **In the app:** the **National Rates** step of Getting started (p. 301).

Why

These parameters support the selection of denominators for service coverage. The denominators can be computed from the facility data by making assumptions: for example, if ANC1 (p. 496) coverage in the survey is 95% and the number of ANC1 visits in the facility data for the year is 100,000, the number of pregnant women is 100,000 ÷ 0.95 = 105,263. Live births are then obtained from pregnancies with the rates of pregnancy loss, stillbirths and twin births, and infants eligible for vaccination with neonatal mortality. Country-specific values are preferred where available.

The survey coverage of ANC1, penta1 and penta3 (p. 502) also sets the expected ratios of the consistency checks in the data quality assessment (p. 67).

What

Parameter	Definition
ANC1 survey (%)	Percentage of women aged 15–49 with a live birth who received at least one antenatal care visit during their last pregnancy. (Example: if 940 out of 1,000 women had at least one ANC visit, ANC1 coverage is 94%.)
ANC4 (p. 496) (%)	Percentage of women aged 15–49 with a live birth who received four or more antenatal care visits during their last pregnancy. (Example: if 680 out of 1,000 women had four or more ANC visits, ANC4 coverage is 68%.)
Pregnancy loss (proportion)	Proportion of pregnancies that ended in miscarriage, abortion or stillbirth out of all reported pregnancy outcomes. (Example: if 80 of 1,000 pregnancies ended in loss, the proportion is 0.08.)
Twin births (proportion)	Proportion of live births that are twins out of all live births. (Example: if 30 out of 1,500 live births were twins, i.e. 15 twin pairs, the proportion is 0.02.)
Neonatal mortality (p. 502) (proportion)	Proportion of live births that die within the first 28 days of life. (Example: if 20 out of 1,000 live births died within the first 28 days, the proportion is 0.02.)

Post-neonatal mortality (p. 503) (proportion)	Proportion of live births that die between 28 days and 11 months of life. (Example: if 15 out of 1,000 live births died between 28 days and 11 months, the proportion is 0.015.)
Stillbirth rate (proportion)	Proportion of pregnancies that ended in stillbirth out of all reported pregnancy outcomes. (Example: if 10 out of 1,000 pregnancies ended in stillbirth, the proportion is 0.01.)
Institutional delivery (p. 500) (%)	Percentage of live births that occurred in a health facility. (Example: if 850 out of 1,000 live births were in a health facility, institutional delivery coverage is 85%.)
BCG (p. 496) coverage (%)	Percentage of infants aged 12–23 months who received the BCG (Bacillus Calmette-Guérin) vaccine. (Example: if 950 out of 1,000 infants received BCG, BCG coverage is 95%.)
Penta1 coverage (%)	Percentage of infants aged 12–23 months who received the first dose of pentavalent vaccine. (Example: if 920 out of 1,000 infants received the first dose, penta1 coverage is 92%.)
Penta3 coverage (%)	Percentage of infants aged 12–23 months who received the third dose of pentavalent vaccine. (Example: if 850 out of 1,000 infants received the third dose, penta3 coverage is 85%.)
Measles 1 (p. 501) coverage (%)	Percentage of infants aged 12–23 months who received the first dose of measles vaccine. (Example: if 900 out of 1,000 infants received the first dose, measles 1 coverage is 90%.)
C-section rate (%)	Percentage of live births that were delivered by caesarean section. (Example: if 150 out of 1,000 live births were by caesarean section, the rate is 15%.)
Low birth weight (p. 501) (%)	Percentage of live births with a low birth weight (less than 2,500 grams). (Example: if 100 out of 1,000 live births had a low birth weight, the percentage is 10%.)
Recent survey year	The calendar year in which the survey was conducted and from which the vaccine coverage values (BCG, penta1, penta3, measles 1) are drawn. (Example: if the survey was conducted in 2022, that is the survey year.)
Survey data start year	The calendar year from which the survey data are used: it sets the period for which the survey data are relevant. Ideally the two most recent surveys are used. (Example: if a country has surveys in 2008, 2013, 2018 and 2023, the survey data start year is 2018, the two most recent.)

For vaccines, ANC4, C-section rate and low birth weight are not asked for; the survey coverage of OPV1 and OPV3 (p. 502) is asked for instead. See RMNCAH and vaccine analyses (p. 258).

Settings

Setting	Default	What it means
Pre-filled values	From the country's survey data, or the uploaded national survey file	The starting values; update them if you have more recent, validated values.
Starting twin and pregnancy loss rates	0.015 and 0.03	A new dataset starts with a twin rate of 1.5% and a pregnancy loss rate of 3%; neonatal, post-neonatal mortality and stillbirth rates come from the survey.
Units	Percentages and proportions	Coverage values are percentages (0 to 100). Pregnancy loss, twin births, stillbirth, neonatal and post-neonatal mortality are proportions (0 to 0.05; 0.03 means 3%).
Recent survey year	Between 2015 and 2030	The allowed range.
Required fields	All	The setup cannot be finished while a field is empty.

Interpretation

- **Are the values from the most recent national survey?** If more recent, validated values exist, enter them.
- **Is every rate a proportion and every coverage a percentage?** A neonatal mortality of 2.5 instead of 0.025, or a coverage of 0.9 instead of 90, distorts every denominator.
- **Do the two survey years cover the period you analyse?** The survey data start year should take in the two most recent surveys.

After the setup

The parameters are used for the expected ratios of the data quality assessment (p. 67) and to compute the denominators in Denominator assessment and selection (p. 95), where the way each parameter enters the calculations is explained.

CHAPTER 7

Data quality assessment

Before coverage is estimated, check whether the health facility data can be trusted: are reports complete, are there implausible values, do related indicators agree, and how good is each year overall?

The health facility data quality assessment checks the monthly data reported by health facilities before coverage is estimated from them: the completeness of reporting by health facilities, extreme outliers, and the internal consistency of the reported numbers, summed up in an annual data quality score.

It is the first step of the analysis after the data are loaded (Analysis setup (p. 44)). What it finds decides what the next step, Data adjustment (p. 75), corrects or removes.

At a glance

- **Answers:** can these facility data be trusted to estimate coverage, and which years and districts are weak?
- **Uses:** the monthly health facility data (reports received and service numbers) and, for the consistency checks, the ANC1 (p. 496), penta1 and penta3 (p. 502) (and OPV1, OPV3 (p. 502)) coverage from the most recent household survey (p. 500).
- **Produces:** indicators for each check, and an annual data quality score.
- **In the apps:** Data quality (p. 306) (**Data quality** in the RMNCAH (p. 503) app, **Data Quality Assessment** in the Vaxx app).

Why

Routinely reported health facility data are an important data source for health indicators at facility and population level. The data are reported by health facilities on events such as immunizations given, or live births attended. As with any data, quality is an issue. Data need to be checked to consider completeness of reporting by health facilities, identify extreme outliers and internal consistency. A standard reporting method for data quality allows assessment of progress over time.

What the assessment checks

Check	The question it answers
Completeness of monthly facility reporting (p. 59)	Did facilities send their monthly reports?
Missing monthly values (p. 62)	Are there months with no number for an indicator?
Extreme outliers (p. 64)	Are some monthly numbers implausibly high or low?
Consistency of annual reporting (p. 67)	Do reported numbers that should be related agree?

Annual data quality score (p. 72)	Summary of performance: how good is each year?

Each indicator is a percentage: the share of reports, values or districts that pass the check, so **higher is better**.

How the checks fit together

Read the checks together, not one at a time. Completeness of reporting tells you whether reports arrived, not whether the numbers in them are right; extreme outliers and consistency look at the numbers themselves. A year can have complete reporting and still fail the consistency checks. The overall score sums up all of them, but always look at which check pulls a weak year down.

Doing it in the app

1. **Open the data quality pages**
   They are under **Data quality** in the RMNCAH app and **Data Quality Assessment** in the Vaxx app: see Data quality (p. 306) for each page and its filters.
2. **Review each check**
   Work through Reporting Rate (p. 309), Data Missingness (p. 315), Outlier Detection (p. 312) and Internal Consistency (p. 317). Note the years and districts that look weak.
3. **Look at the overall score**
   The Overall Data Quality Score (p. 320) is computed for you, one column per year.
4. **Continue to data preparation**
   Correct the problems and, where needed, remove unreliable years in Data preparation (p. 322).

Points to consider for the interpretation

- Is there a data quality pattern by year for which there is an explanation? (Include the explanation.)
- Are there certain regions or other subnational units that are particularly problematic?
- Are there certain reporting forms or services (e.g., antenatal care, labour and delivery, immunization) that are problematic?
- Is there good consistency between reported numbers of ANC1 and penta1?

Common pitfalls

- A low score is a reason to investigate, not always to discard a year. Compare with local events (a strike, a new reporting system) before using Remove Years (p. 326). A year with many outliers or low completeness is a candidate for **Remove Years** or careful adjustment (see Data adjustment (p. 93)).
- District results are noisier than regional ones because they rest on smaller numbers.

7.1 Completeness of monthly facility reporting

The percentage of expected monthly facility reports that were received, nationally and district by district.

Every health facility is expected to send a reporting form (p. 503) each month for each service: antenatal care (p. 496), delivery, immunization and outpatient care (p. 502). Completeness of reporting is the percentage of those expected reports that were received: the reporting rate (p. 503).

At a glance

- **Measures:** the reporting rate for the whole country, and the share of districts that report well.
- **Good result:** reporting rates over 90%, with few districts below the district threshold (p. 498).
- **In the app:** the **Reporting Rate** (p. 309) page of the data quality pages.

Why

If facilities do not report, the services they provided are missing from the totals and coverage (p. 503) is underestimated. Changes in reporting completeness over time affect trend analysis: a coverage trend can come from more or fewer facilities reporting rather than from more or fewer services.

What

Indicator	Numerator	Denominator	Interpretation
% of expected monthly facility reports (national)	Number of monthly facility reports received	Total number of facility reports expected (12 × number of facilities with the service)	Reporting rates over 90% are good; changes in reporting completeness over time affect trend analysis
% of districts with completeness of facility reporting ≥ 90%	Number of districts with at least 90% monthly reporting completeness in a year	Total number of districts	Identifies districts with low reporting rates, in one year or several

Both are averaged over the reporting forms of the analysis: antenatal care, delivery, immunization and outpatient care for RMNCAH (p. 503); antenatal care, delivery and immunization for vaccines.

Formula

$$\text{National reporting rate} = \frac{R}{E} \times 100$$

$$\text{Districts reporting well} = \frac{D_{90}}{D} \times 100$$

where

- R is the number of monthly facility reports received in the year;
- E is the number of reports expected: 12 × the number of facilities with the service;
- D_{90} is the number of districts with a reporting rate of at least 90% in the year;
- D is the total number of districts.

Settings

Setting	Default	What it means
District threshold	90%	A district reports well when its reporting rate is at least this. You can change it on the **Reporting Rate** page (quick picks: 70%, 80%, 90%, 95%).

Interpretation

- **Is the national reporting rate over 90%?** If not, the totals miss a noticeable share of services; the adjustment for incomplete reporting will matter, and trends need care.
- **Do few districts report well while the national rate is good?** A high national rate can hide districts with low reporting. Look at the district and regional views.
- **Does completeness change between years?** Then a change in coverage between those years may come from reporting, not from services.
- **Are some services worse than others?** Compare the reporting forms.
- Completeness shows that reports arrived, not that the numbers in them are right. Read it with extreme outliers (p. 64) and consistency of annual reporting (p. 67).

Example: reporting rate by area and year

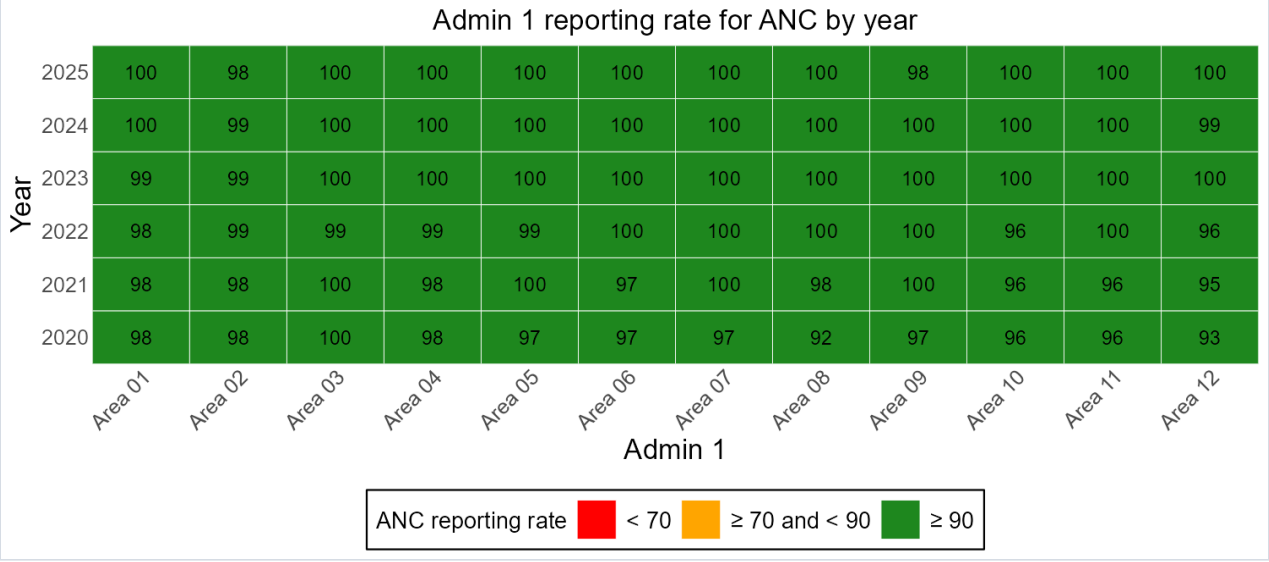


Figure 7.1 — *Reporting rate for antenatal care by Admin 1 area and year*

Reading the chart: each column is an Admin 1 (p. 496) area and each row a year; the number is the reporting rate for antenatal care. Green means 90% or more.

What it shows: every area reported over 90% in every year. The lowest rates were in the first year (92% and 93% in two areas); from the fourth year nearly every area reported 99% or 100%. Reporting for antenatal care improved and is not a concern.

Example: districts with low reporting, by service

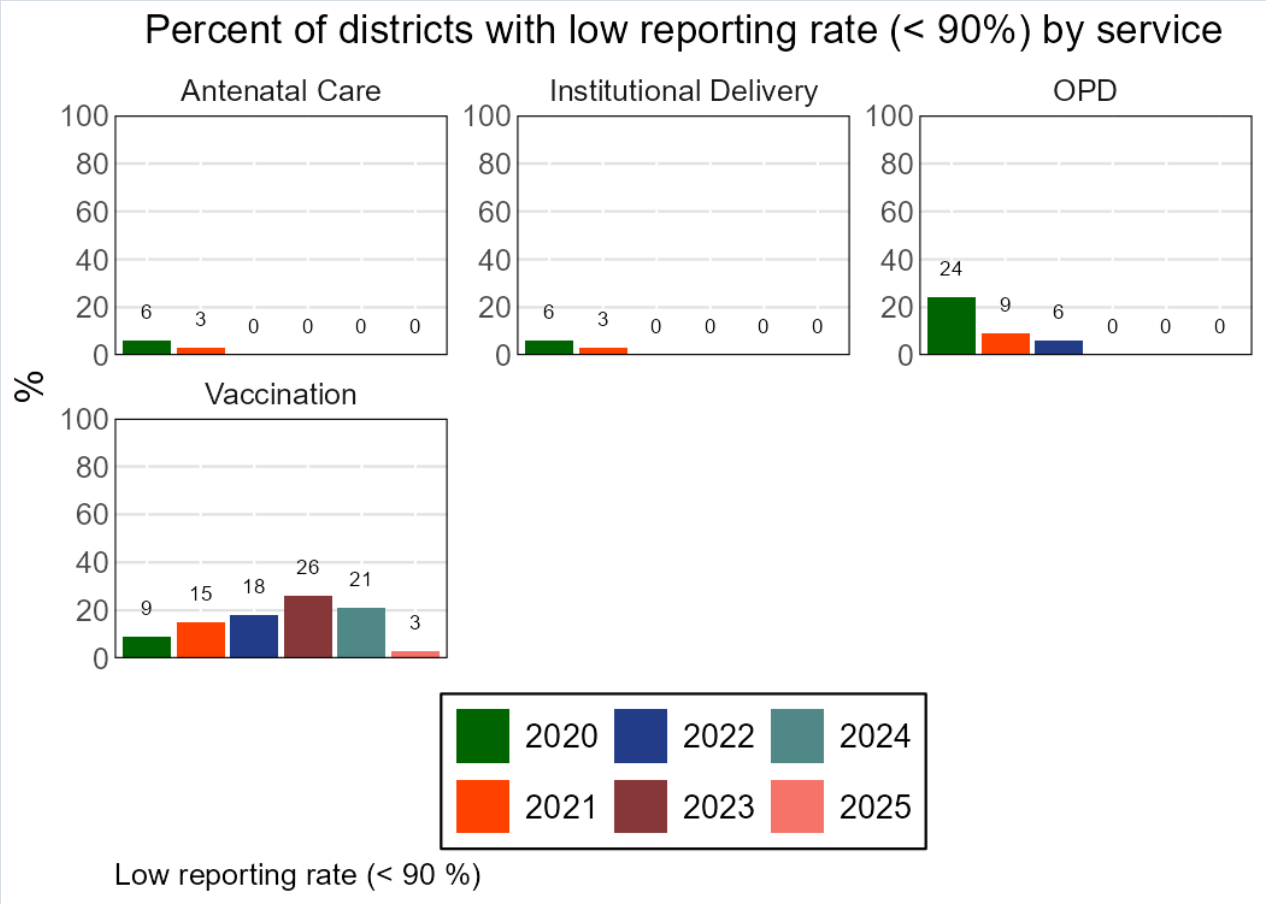


Figure 7.2 — *Percentage of districts with a reporting rate below 90%, by service and year*

Reading the chart: one panel per service; each bar is a year, and its height the percentage of districts that reported below 90%. Lower is better.

What it shows: for antenatal care and delivery, 6% of districts reported below 90% in the first year, 3% in the second and none after. Outpatient reporting started worse (24% of districts) and was complete from the fourth year. Vaccination is the exception: the share of districts reporting below 90% rose from 9% to 26% over four years before falling to 3%. Immunization reporting in those years needs a closer look, district by district, before vaccination coverage is trusted.

After the assessment

In Data adjustment (p. 75), the reported numbers are adjusted for incomplete reporting: facilities that did not report are assumed to have provided services at a lower volume than those that did, set by the k-factor (p. 500) (0.25 by default). A month where a district's reporting rate is below 75% is treated separately. See Configuring the -factor (p. 81) and Critical data threshold rules (p. 84).

7.2 Missing monthly values

The percentage of districts with no missing monthly values in a year, and where the gaps are.

A missing value (p. 502) is a month with no number for an indicator in a district. This check finds those gaps and shows where they are. It differs from completeness of reporting (p. 59), which counts the reports received: here we look for months that have no value at all.

> **At a glance**
>
> - **Measures:** the share of districts with no missing months in the year.
> - **Good result:** as close to 100% as possible.
> - **In the app:** the **Data Missingness** **(p. 315)** page of the data quality pages.

Why

A missing month counts as zero when the year is added up, so the annual total, and the coverage (p. 503) computed from it, is too low. Gaps that appear in some years and not others distort trends.

What

Indicator	Numerator	Denominator	Interpretation
% of districts with no missing monthly values in a year for any of the 4 forms	Number of districts with no missing values for any of the 4 reporting forms (p. 503) in a year	Total number of districts	Shows how the missing values are spread across districts

For vaccines it is averaged over the tracer vaccines (p. 504). It is one of the indicators of the annual data quality score (p. 72).

Formula

$$\text{Districts with no missing values} = \frac{D_c}{D} \times 100$$

where

- D_c is the number of districts with no missing monthly value in the year, for any of the reporting forms;
- D is the total number of districts.

Interpretation

- **Are the gaps in a few indicators?** An indicator missing almost everywhere may not be collected in the country, or only from a certain date; check before using it.
- **Are the gaps in a few areas?** Look at those districts before comparing them with others.
- **Are the gaps in one year?** Look for a cause, such as a change of reporting system, before comparing that year with the others.

Example: missing months by indicator and area

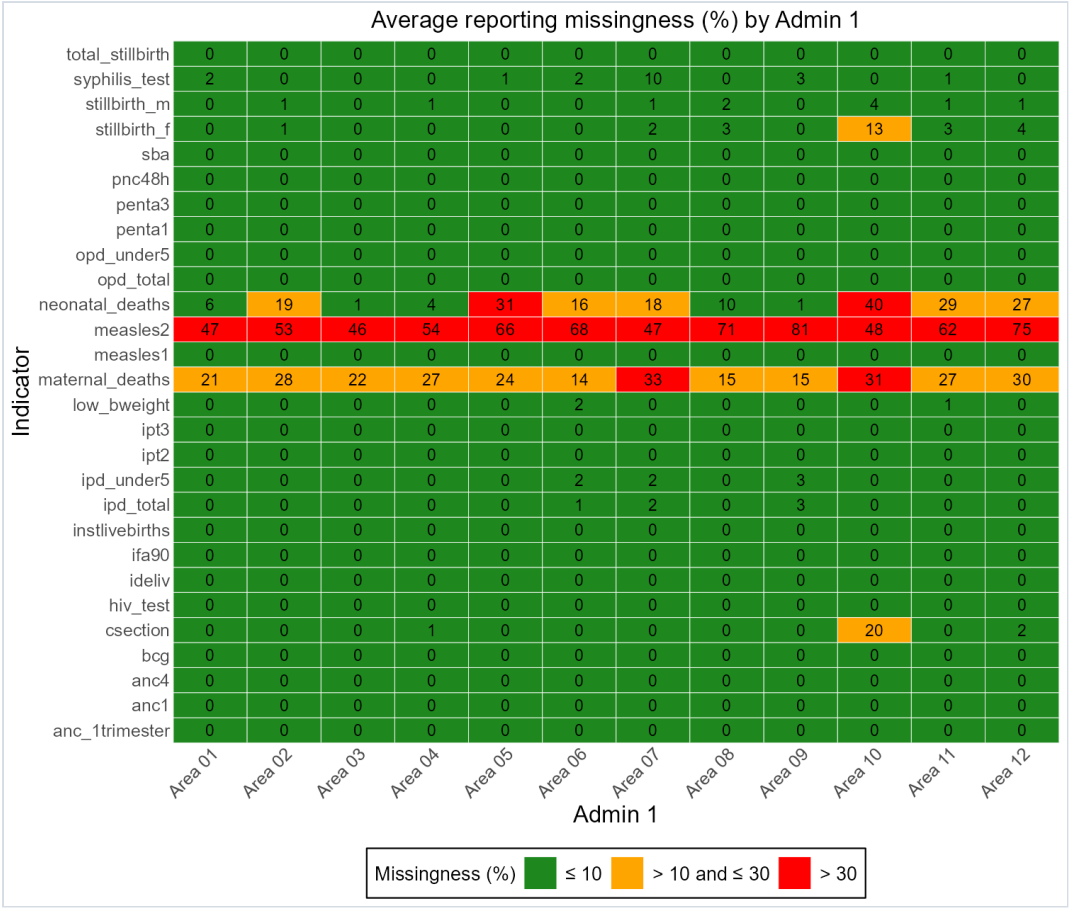


Figure 7.3 — *Missing reports by indicator and Admin 1 area, as a percentage of months*

Reading the chart: each row is an indicator and each column an Admin 1 (p. 496) area; the number is the percentage of months with no value. Green is 10% or less, orange more than 10%, red more than 30%.

What it shows: most indicators have no missing months anywhere. Three stand out: the second dose of measles vaccine (**measles2**) is missing for 46% to 81% of the months in every area, which suggests it is not reported everywhere or was introduced during the period; **maternal deaths** (14% to 33%) and **neonatal deaths** (up to 40%) have many gaps, so mortality figures need care. One area also has local gaps in caesarean sections (20%) and female stillbirths (p. 504) (13%): that area deserves a look before its figures are compared with the others.

After the assessment

In Data adjustment (p. 75), each missing monthly value is replaced with the district's median (p. 501) monthly value for the year, so a gap no longer counts as zero. See Replacing outliers and missing values (p. 87).

7.3 Extreme outliers

The percentage of monthly values that are not extreme outliers, and of districts with no extreme outliers in the year.

An extreme outlier (p. 499) is a monthly number far from the usual values of the same district, such as an accidental extra zero. Most are data errors; some are real events, such as a vaccination campaign.

> **At a glance**
>
> - **Measures:** the share of monthly values, and of districts, without extreme outliers.
> - **Good result:** at least 99% of monthly values, and 90% of districts, without extreme outliers.
> - **In the app:** the **Outlier Detection** **(p. 312)** page of the data quality pages.

Why

Extreme outliers, such as an accidental extra zero in a number, can have a large impact, especially on subnational numbers: a single wrong month can change a district's annual total, and its coverage (p. 503), considerably.

What

A monthly value is an extreme outlier when it is more than 5 times the median absolute deviation (MAD) (p. 501) away from the median (p. 501) monthly value of the same district. The median and MAD are computed from all years except the most recent one, so the latest year is judged against the earlier years.

Indicator	Numerator	Denominator	Interpretation
% of monthly values that are not extreme outliers (national)	Number of monthly values that are not extreme outliers in a specific year	Total number of monthly values	At least 99% of monthly data expected not to be an extreme outlier; consider reasons
% of districts with no extreme outliers in a year	Number of districts with no extreme outliers in a specific year	Total number of districts	At least 90% of districts should have no extreme outliers at all; consider reasons

Formula

A monthly value x is an extreme outlier when

$$|x - m\,| > 5 \times MAD$$

$$\text{Values without outliers} = \frac{V_{ok}}{V} \times 100$$

$$\text{Districts without outliers} = \frac{D_{ok}}{D} \times 100$$

where

- m is the median monthly value of the district, over all years except the most recent;
- MAD is the median absolute deviation of the same values;
- V_{ok} is the number of monthly values in the year that are not extreme outliers, and V the number of monthly values;

- D_{ok} is the number of districts with no extreme outlier in the year, and D the total number of districts.

Worked example: a district's monthly values have a median of 100 and a MAD of 10. Values from 100 − 5 × 10 = **50** to 100 + 5 × 10 = **150** are normal; a month with 1,000 is an extreme outlier.

Settings

Setting	Default	What it means
Outlier bound	5 × MAD	How far from the median a monthly value must be to count as an extreme outlier.
Baseline years	All years except the latest	The years used for the median and MAD, so the latest year cannot hide its own outliers.

Interpretation

- **Are at least 99% of monthly values, and 90% of districts, free of extreme outliers?** If not, find which districts and indicators the outliers come from.
- **Is there a real-world explanation?** Outliers are found by a statistical rule. A campaign, a stock-out or an outbreak can produce a real peak; consider the reasons before treating a value as an error.
- **Is it a single month?** A peak that falls back the next month is typical of a typing error, or of a campaign. Look at the trend around it.
- **Are many outliers in one year?** That suggests a reporting problem in that year.

Example: monthly values without outliers, by area and year

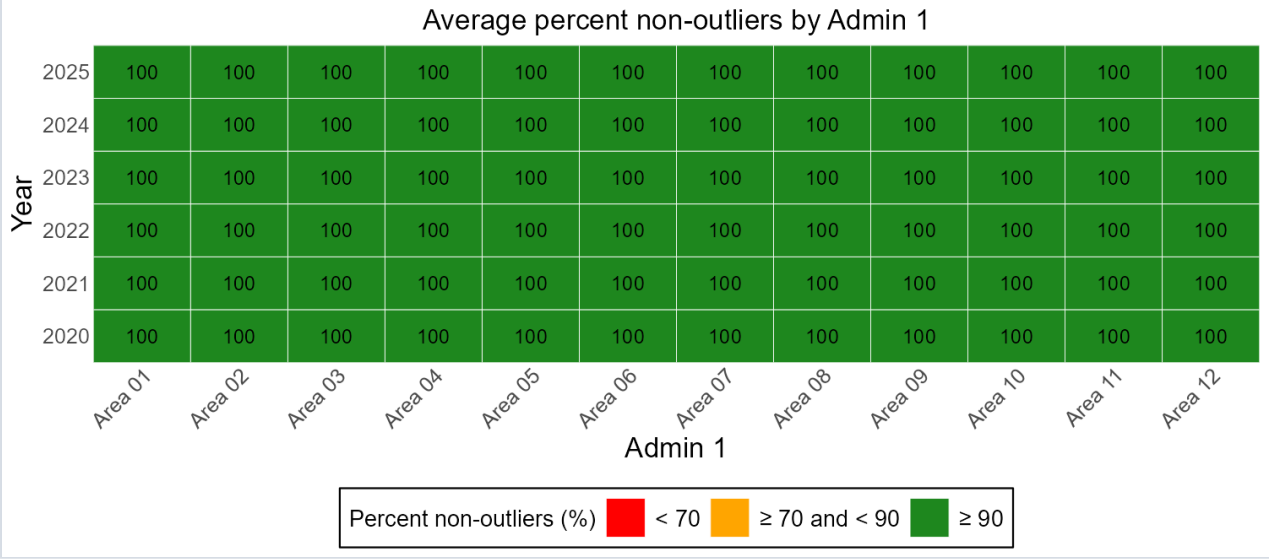


Figure 7.4 — *Percentage of monthly values that are not extreme outliers, by Admin 1 area and year*

Reading the chart: each column is an Admin 1 (p. 496) area and each row a year; the number is the percentage of monthly values that are not extreme outliers. Green is 90% or more.

What it shows: every area has 100% of its monthly values free of extreme outliers in every year: no outlier problem at this level for this indicator. A few outliers in single districts can still disappear in a rounded area average, so the app also lists the districts and months concerned.

Example: one district's months against the band of normal values

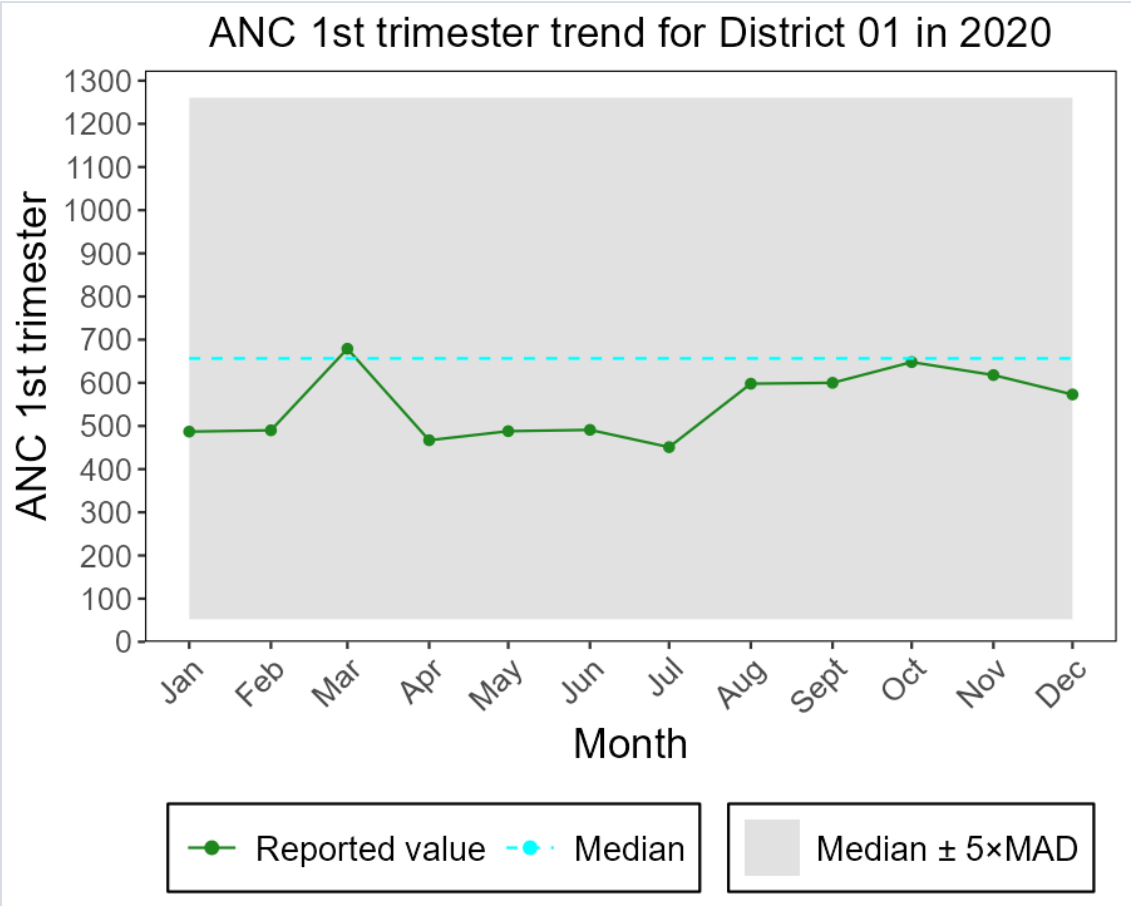


Figure 7.5 — *Monthly values for one district, with the band of normal values (median ± 5 MAD)*

Reading the chart: the green line is the reported value each month; the grey band is the range of normal values (median ± 5 MAD) and the dashed line the median. A month outside the band would be an extreme outlier.

What it shows: the band runs from about 50 to about 1,260 and the median is about 650. Every month lies inside the band, including the peak in March (about 680), so none is an extreme outlier.

After the assessment

In Data adjustment (p. 75), extreme monthly outliers are corrected and given the district's median value for the same year. See Replacing outliers and missing values (p. 87).

7.4 Consistency of annual reporting

Whether reported numbers that should be related agree: ANC1 with penta1, penta1 with penta3, and OPV1 with OPV3, nationally and by district.

Some reported numbers should be related. Most women who make a first antenatal visit (ANC1 (p. 496)) give birth to a baby who later receives the first dose of pentavalent vaccine (penta1 (p. 502)); most children who receive penta1 go on to receive penta3. This check compares their annual totals with the ratio expected between them.

At a glance

- **Compares:** ANC1 with penta1, penta1 with penta3, and for vaccines OPV1 with OPV3 (p. 502).
- **Good result:** each national ratio within ± 0.05 of its expected ratio (p. 499); most districts with ratios between 1.0 and 1.5.
- **In the app:** the **Internal Consistency** (p. 317) page of the data quality pages (**Ratio Plots** and **Consistency Checks**).

Why

There is often an inconsistency between antenatal and immunization data, even though the two should be consistent. Examining the association between ANC1 and penta1 is particularly informative: a number can be complete and free of outliers and still be wrong, and only a comparison with a related number shows it.

What

Indicator	Numerator	Denominator	Interpretation
ANC1 to penta1 ratio in the reported data (national)	Number of ANC1 reported	Number of penta1 reported	National ratio within an expected range (see Expected ratios)
Penta1 to penta3 ratio in the reported data (national)	Number of penta1 reported	Number of penta3 reported	National ratio within an expected range, based on the survey results
OPV1 to OPV3 ratio in the reported data (national; vaccines)	Number of OPV1 reported	Number of OPV3 reported	National ratio within an expected range, based on the survey results
% of districts with the ratio between 1.0 and 1.5 (one for each ratio above)	Number of districts with ratios within the expected range	Total number of districts	For districts there is more variation in the ratio: a wider range is considered

Expected ratios

ANC1 to penta1. ANC1 takes place at about 20 weeks of pregnancy and penta1 at 6-8 weeks after birth. In between, the method assumes a pregnancy loss (p. 503) of 3%, a stillbirth (p. 504) rate of 2%, a twinning rate (p. 504) of 1.5% and neonatal mortality (p. 502) before penta1 of 3%. For every 100 women at ANC1 there are then

$$100 \times \left(1 - 0.03 - 0.02 + 0.015 - 0.03\right) = 93.5 \text{ infants at penta1}$$

so the ratio is 100 ÷ 93.5 = 1.07 when ANC1 and penta1 coverage are the same. The actual coverage of both, from the most recent household survey (p. 500), is then taken into account:

$$E_{\text{ANC1/penta1}} = 1.07 \times \frac{C_{ANC1}}{C_{penta1}}$$

Worked example: survey coverage of 90% for ANC1 and 95% for penta1 gives 1.07 × 0.90 ÷ 0.95 = **1.01**; reported ratios from **0.96 to 1.06** are acceptable.

Penta1 to penta3 (and **OPV1 to OPV3**). The doses are given at 6 and 14 weeks of age; few children die in between, so the main factor is the drop-out (p. 498) between the doses. The expected ratio comes from the latest survey:

$$E_{\text{penta1/penta3}} = \frac{C_{penta1}}{C_{penta3}}$$

Worked example: survey coverage of 95% for penta1 and 85% for penta3 gives 0.95 ÷ 0.85 = **1.12**; reported ratios from **1.07 to 1.17** are acceptable.

Formula

$$\text{Ratio} = \frac{N_1}{N_2}$$

A national ratio is acceptable when it is within 0.05 of its expected ratio E:

$$|\text{Ratio} - E\,| \leq 0.05$$

For districts, the share whose ratio is in the wider range:

$$\text{Districts in range} = \frac{\text{number of districts with } 1.0 \leq \text{Ratio} \leq 1.5}{D} \times 100$$

where

- N_1 and N_2 are the two reported annual numbers (for example ANC1 and penta1);
- C is a coverage rate from the most recent survey;
- D is the total number of districts.

Settings

Setting	Default	What it means
National tolerance	± 0.05	How far a national ratio may be from its expected ratio.
District range	1.0 to 1.5	The range a district's ratio must be in.
Mortality factor	1.07	The ANC1 to penta1 ratio expected when both have the same coverage (above).

If no survey coverage is entered, the expected ratios assume coverage of 98% for ANC1, 97% for penta1 and OPV1, 89% for penta3 and 78% for OPV3; enter the country's survey values in Analysis setup (p. 44).

Interpretation

- **Is a national ratio outside its range?** Flag it, and discuss possible explanations before adjusting anything.

- **Which districts are far from the diagonal** in the consistency plots? For instance, districts where ANC1 is much higher than penta1.
- **Is the survey value right?** The expected ratios come from the survey, so a wrong survey value makes good data look inconsistent.
- Not every deviation is a data problem: an unusual penta1 to penta3 ratio can be real drop-out, and campaigns or service disruptions can explain differences.

Example: national ratios against their expected values

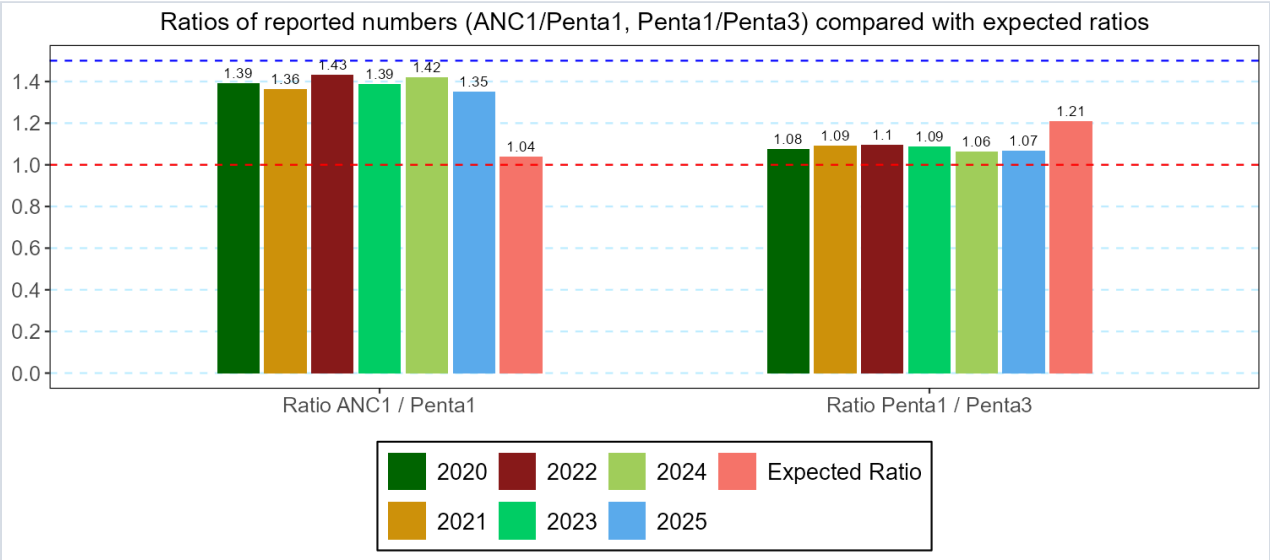


Figure 7.6 — *ANC1 to penta1 and penta1 to penta3 ratios by year, compared with the expected ratios*

Reading the chart: each bar is a year's reported ratio; the last bar of each group (pink) is the expected ratio. The dashed lines mark 1.0 and 1.5, the range used for districts.

What it shows: the ANC1 to penta1 ratio is between **1.35 and 1.43** every year, against an expected **1.04**: far outside the acceptable 0.99 to 1.09. Many more ANC1 visits are reported than penta1 doses: either ANC1 is over-reported (for example, later antenatal visits counted as first visits) or penta1 is under-reported. This needs to be discussed before antenatal or vaccination coverage is used. The penta1 to penta3 ratio is between **1.06 and 1.10**, below its expected **1.21** (acceptable 1.16 to 1.26): the reported drop-out is smaller than the survey suggests.

Example: ANC1 and penta1, district by district

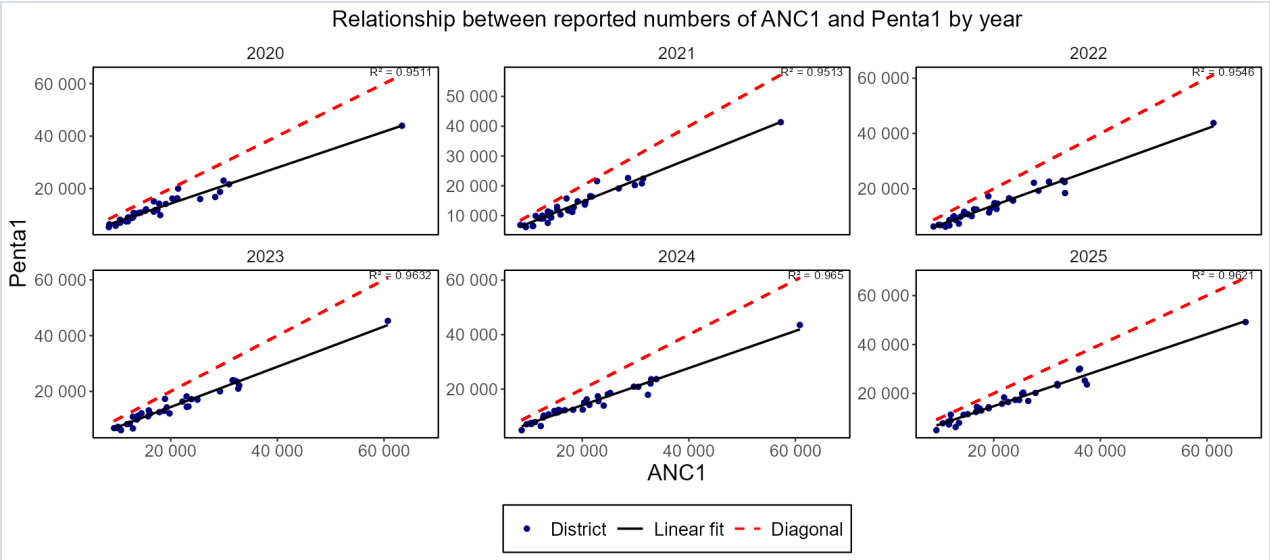


Figure 7.7 — *Penta1 against ANC1 for each district, by year*

Reading the chart: each dot is a district, with its ANC1 (across) and penta1 (up). On the red dashed diagonal the two would be equal; the black line follows the districts.

What it shows: the districts lie below the diagonal, and further below it the larger the district: the gap between ANC1 and penta1 is found in almost every district, not in one or two. The districts follow the line closely every year (R^2 about 0.95 to 0.97), so the pattern is systematic.

Example: penta1 and penta3, district by district

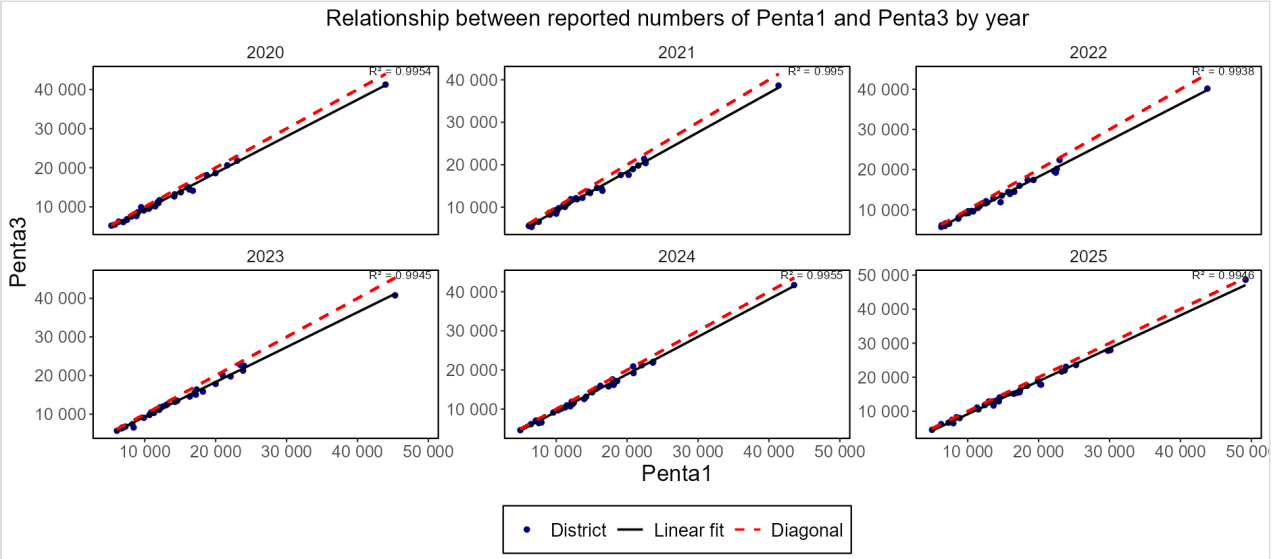


Figure 7.8 — *Penta3 against penta1 for each district, by year*

Reading the chart: as above, with penta1 across and penta3 up.

What it shows: the districts lie almost on the diagonal, very closely every year (R^2 about 0.99): penta1 and penta3 are consistent in every district, with little drop-out.

After the assessment

Consistency is not corrected automatically. Its results feed the annual data quality score (p. 72) and help decide, in Data adjustment (p. 75), whether a year or a district can be used.

7.5 Annual data quality score

The summary of performance: one score per year, the mean of the completeness, missing value, extreme outlier and district consistency indicators.

The annual data quality score is the summary of performance: one number per year that brings the data quality indicators together, so years can be compared at a glance.

> **At a glance**
>
> - **Made from:** the completeness, missing value, extreme outlier and district consistency indicators.
> - **Good result:** a high score, similar across years; a year well below the others needs a closer look.
> - **In the app:** the **Overall Data Quality Score** (p. 320) page, computed for you.

Why

A standard reporting method for data quality allows assessment of progress over time. The score gives one view of all the checks for each year; the individual indicators then show why a year scores as it does.

What

The score is the mean of these indicators for the year. They are all percentages where higher is better, so the score is a percentage too:

- % of expected monthly facility reports (national) and % of districts with completeness of facility reporting ≥ 90% (completeness (p. 59));
- % of districts with no missing monthly values (missing values (p. 62));
- % of monthly values that are not extreme outliers and % of districts with no extreme outliers (extreme outliers (p. 64));
- % of districts with ANC1 to penta1, and with penta1 to penta3, in range, plus OPV1 to OPV3 for vaccines (consistency (p. 67)).

The national ratios themselves are shown in the table but are not in the score.

Formula

$$\text{Score} = \frac{1}{n}\sum_{i=1}^{n} I_i$$

where

- $I_1, \ldots, I_n$ are the indicators listed above, for the year;
- n is their number: 7, or 8 for vaccines (with OPV1 to OPV3).

Settings

The score has no setting of its own, but it depends on the district threshold (p. 498) of completeness (90% by default): changing it changes the score.

Interpretation

- **Is there a data quality pattern by year** for which there is an explanation? Find which indicator pulls a weak year down.
- **Is a low score a reason to drop the year?** Not necessarily: investigate first, and look for a cause such as a strike or a new reporting system.
- **Has the district threshold been changed?** Agree it with your team before comparing years or countries.

Example: the score table

	Data Quality Metrics	2020	2021	2022	2023	2024	2025
1. Completeness of monthly facility reporting (mean of ANC, delivery, immunization)							
1a	% of expected monthly facility reports (national)	96	97	98	98	98	100
1b	% of districts with completeness of facility reporting >= 90	89	93	94	93	95	99
1c	% of districts with no missing values for the 4 forms	90	91	91	91	91	97
2. Extreme outliers (mean of ANC, delivery, immunization)							
2a	% of monthly values that are not extreme outliers (national)	100	100	100	100	99	100
2b	% of districts with no extreme outliers in the year	92	96	94	93	91	87
3. Consistency of annual reporting							
3a	Ratio anc1/penta1	1.39	1.36	1.43	1.39	1.42	1.35
3b	Ratio penta1/penta3	1.08	1.09	1.10	1.09	1.06	1.07
3c	% district with anc1/penta1 in expected ranged	71	79	74	82	82	82
3d	% district with penta1/penta3 in expected ranged	97	100	100	100	97	100
4	**Annual data quality score**	**91**	**94**	**93**	**94**	**93**	**95**

Figure 7.9 — *The overall data quality score table, by year*

Reading the table: one column per year. The rows are numbered as in the app:

Row	Indicator
1a	% of expected monthly facility reports (national)
1b	% of districts with completeness of facility reporting ≥ 90%
1c	% of districts with no missing values for the 4 forms
2a	% of monthly values that are not extreme outliers (national)
2b	% of districts with no extreme outliers in the year
3a, 3b	National ratios: ANC1 to penta1, penta1 to penta3 (not in the score)
3c, 3d	% of districts with ANC1 to penta1, and penta1 to penta3, in range
4	Annual data quality score

Green cells meet the expected level; yellow cells fall short.

What it shows: the score is high and rises from **91** in the first year to **95** in the last. The weakest row every year is the share of districts with ANC1 to penta1 in range (71% to 82%); the consistency check (p. 67) shows why. In the first year, completeness was just short of the mark (89% of districts reporting well); in the last year, the share of districts with no extreme outliers dropped to 87%, so that year's outliers deserve a look. The first year's score is (96 + 89 + 90 + 100 + 92 + 71 + 97) ÷ 7 = 90.7, shown as 91.

After the assessment

The score helps decide, in Data adjustment (p. 75), which years to keep: a year with a very low score is a candidate for Remove years (p. 93).

CHAPTER 8

Data adjustment

How reported facility numbers are corrected for incomplete reporting, low reporting rates, extreme outliers and missing values, and how unreliable years are removed, to obtain a clean data set for analysis.

The health facility data quality adjustments turn the reported monthly numbers into a clean data set for analysis: they adjust for incomplete reporting, correct extreme outliers and fill missing values, and, where a year cannot be relied on, remove it. The adjusted numbers are the numerators of every coverage statistic computed afterwards.

It follows the Data quality assessment (p. 57), which finds the problems, and comes before Denominator assessment and selection (p. 95), which uses the adjusted numbers.

At a glance

- **Answers:** how should the reported numbers be corrected so that coverage is not underestimated, and which years should be left out?
- **Uses:** the monthly health facility data with their reporting rates (p. 503), the extreme outliers and missing values found in the data quality assessment, and the adjustment factor k chosen by the analyst.
- **Produces:** the adjusted numbers (N_{adj}), saved as the master adjusted dataset (p. 501) (`.dta`), used in Denominator assessment and selection (p. 95) and Coverage (p. 115).
- **In the apps:** Data preparation (p. 322) (**Remove Years**, **Data Adjustment** and **Data Adjustment Changes**).

Why

The completeness of reporting may affect the analysis, especially if completeness is low or varies between years. Extreme outliers, such as an accidental extra zero in a number, can have a large impact, especially on subnational numbers. Following the assessments, several steps are necessary to obtain a clean data set for analysis. This implies adjusting for incomplete reporting and correcting for extreme outliers.

If reporting completeness is not considered, all non-reporting facilities are assumed to have provided zero services, which is not likely to be true, and coverage is underestimated.

Summary of adjustments

Summary of adjustments made to the raw health facility data in preparation of a clean data set:

Problem	Action	Adjustment

Low reporting rates	If a district's monthly reporting rate for a service is below 75% (default), or missing	The rate is replaced by the median of the district's reporting rates between 75% and 100% for that service; the month's number is then adjusted with k
Incomplete reporting by districts, variable over time, affecting trend assessment	If reporting rates are ≥ 75% and ≤ 100% (default), an assumption is made about the volume of services provided by the non-reporting facilities	The adjustment factor k is used to adjust for incomplete reporting (state the value used, and whether it is the same for all reporting forms)
Extreme outliers can greatly affect coverage trend assessments	If a monthly value is more than 5 times the Median absolute deviation (MAD) (p. 501) from the district's median monthly value	Extreme monthly outliers are corrected and given the median of the district's other (non-outlier) months in the same year
Missing values	If there is a missing value	The district's median monthly value for the year is imputed for the month with the missing value

How the adjustments fit together

The adjustments are applied one after the other, district by district and month by month, in this order:

1. **Removed years** are taken out of the data first, so nothing else is computed from them.
2. **Completeness of reporting:** each month's number is scaled up with k, using the district's reporting rate for that service; a rate below 75% (or missing) is first replaced (the low reporting rate rule).
3. **Extreme outliers** are then looked for in the scaled-up numbers and replaced.
4. **Missing values** are filled last.

Each step works on the result of the one before. The Data Adjustment Changes (p. 331) chart shows how much each step changed each year.

In RMNCAH (p. 503) analyses, mortality indicators (maternal deaths, stillbirths and neonatal deaths in facilities) are not adjusted for reporting or outliers, because deaths are few and volatile (see Mortality (p. 381)). Inpatient numbers are corrected for extreme outliers only.

The effect of the adjustments

Once the data are adjusted, compare the number of events before (raw) and after adjustment, year by year, for the main indicators: institutional deliveries and penta1 (p. 502) (for vaccines, the main vaccination indicators: BCG, penta1, penta3 and measles 1). The effect is summed up nationally by the relative change (p. 503) the adjustment made to the reported number.

$$\text{Relative change} = \left| \frac{N_{\text{rep}} - N_{\text{adj}}}{N_{\text{rep}}} \right| \times 100$$

where

- N_{rep} is the number reported in the year (raw, unadjusted);
- N_{adj} is the number after adjustment.

Worked example: 100,000 penta1 vaccinations were reported and 102,000 remain after adjustment: |100,000 – 102,000| ÷ 100,000 × 100 = **2%**.

How to read it:

- A small gap between before and after is normal and means reporting was already nearly complete.
- A wide gap or a large jump in one year points to a year with very low reporting, or to an outlier that needs a second look in Outlier Detection (p. 312).
- The effect is largest for services with the lowest reporting rates. Compare with the Reporting Rate (p. 309) page.

Example: penta1 before and after adjustment

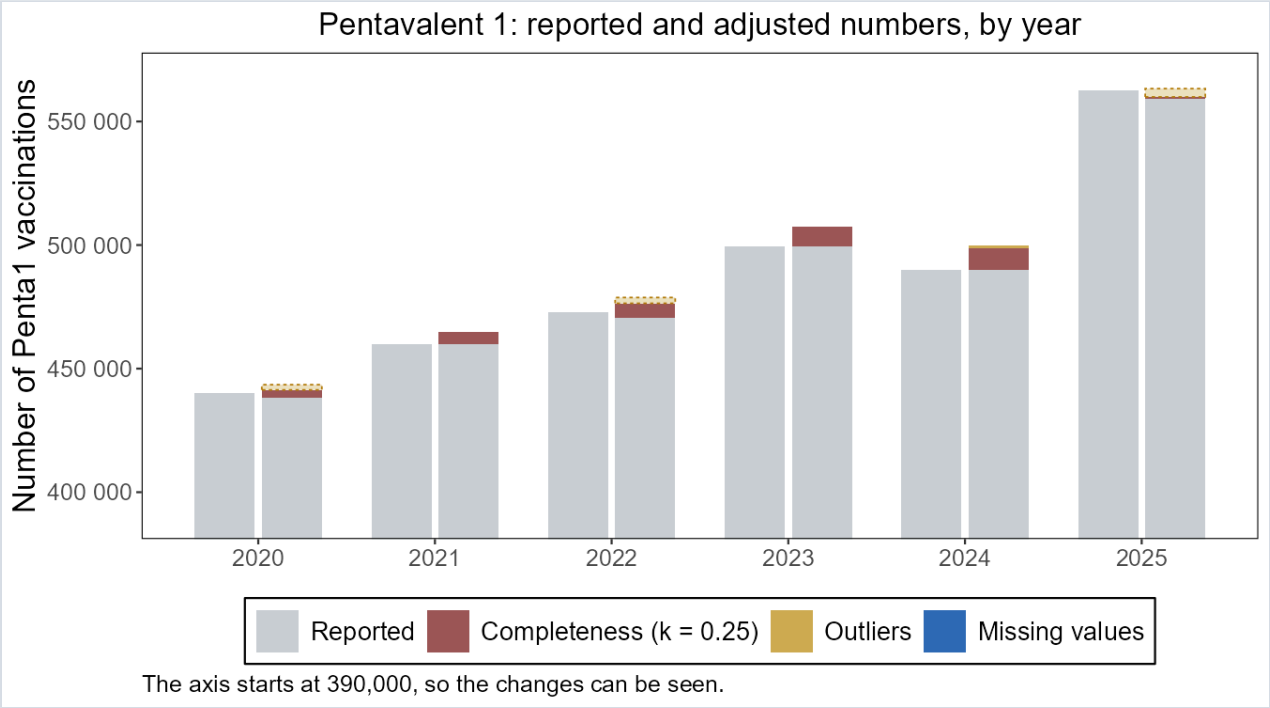


Figure 8.1 — *Penta1 vaccinations reported and adjusted, by year, with the part added for completeness and the part removed as outliers*

Reading the chart: for each year, the left bar (grey) is the number of penta1 vaccinations reported and the right bar the number after adjustment. On the right bar, the red part is what the adjustment for completeness of reporting added (k = 0.25 for every service); a dashed outline marks what the outlier correction removed, and a gold part what it added. Blue would be missing values filled. The axis starts at 390,000 so that the differences can be seen.

What it shows: the adjustment changed penta1 very little. The completeness adjustment added between about 900 (+0.15%, the last year) and 8,800 (+1.8%, the fifth year) vaccinations a year; the outlier correction removed about 2,100, 2,300 and 3,400 in the first, third and last years and added about 1,100 in the fifth. In all, the adjusted number differs from the reported one by 0.3% to 2.0% a year (about 1% on average), the largest change being in the fifth year (+2.0%). Reporting of immunization was nearly complete, so the adjusted trend is the same as the reported one.

Reporting the adjustment

The interpretation should include the selected adjustment factor (factor *k*) that was used to adjust for incomplete reporting (if necessary, by service). If the default factor is used, then report this and explain what this means for the reader.

Report the percentage change that the adjustment made in the reported numbers of institutional deliveries and penta1 (for vaccines, all selected vaccine indicators) over the period, as the average of the years. You may wish to:

- highlight the year with the greatest impact of the adjustment, if there is one;
- interpret whether the impact of the adjustment on coverage rates is large or small;
- make the same description and interpretation for penta1 vaccinations.

Doing it in the app

1. **Remove unreliable years**
   On the **Remove Years** page, select the years to exclude and confirm (see Remove Years (p. 326)). Decide from the data quality pages first: Removing years (p. 93).
2. **Choose the k-factors and adjust the data**
   On the **Data Adjustment** page, choose the *k* value that fits the country context for each service and press **Adjust Data** (see Data Adjustment (p. 324)). How to choose: Choosing k (p. 81).
3. **Review the changes**
   Compare the numbers before and after adjustment on the Data Adjustment Changes (p. 331) page, and report them as described in Reporting the adjustment.

Points to consider for the interpretation

- Which adjustment factor *k* was used, for which services, and why?
- In which year did the adjustment have the greatest impact, and why (low reporting, an outlier)?
- Is the impact of the adjustment on coverage large or small?
- Were any years removed, and for what real-world reason?

Common pitfalls

- Removing a year you should keep silently shortens every later trend.
- Using a high *k* to hide poor reporting hides the problem instead of fixing it. Document your choice.
- Months or districts with reporting below 75% are not left as they are (see Low reporting rates (p. 84)), but adjustment cannot recover what very incomplete reporting missed, so consider removing such years.
- Changing the *k*-factors or the removed years changes every adjusted number: the adjustment must be run again and the denominators re-checked (see Data preparation > Common pitfalls (p. 323) in

the app).

8.1 Adjusting for completeness of reporting

How the reported numbers are scaled up for the facilities that did not report, with the adjustment factor k.

When some facilities do not send their monthly report, the services they provided are missing from the reported numbers. The adjustment for completeness of reporting adds an estimate of those services, using the reporting rate (p. 503) and an assumption about how much the non-reporting facilities provided: the adjustment factor k (p. 500).

At a glance

- **Adjusts:** every monthly number of a district for a service, for the reports that were not received.
- **Uses:** the district's reporting rate for that service in the month, and k (0.25 recommended).
- **In the app:** the **K-Factors (Adjustment Factors)** card of the **Data Adjustment** (p. 327) page.

Why

If we do not consider reporting completeness, that means we assume all non-reporting facilities provided zero services, which is not likely to be true. Coverage computed from the reported numbers is then too low, and if completeness varies between years, the coverage trend reflects reporting rather than services.

What

Adjustments depend on how many services (e.g., pregnancy care, vaccinations) were provided at non-reporting facilities compared to those that reported. The adjustment factor k is the ratio of the volume of services provided by non-reporting facilities to the volume of services provided by reporting facilities. It is used to scale up the reported numbers for incomplete reporting.

Problem	Action	Adjustment
Incomplete reporting by districts, variable over time, affecting trend assessment	If reporting rates are ≥ 75% and ≤ 100%, an assumption is made about the volume of services provided by the non-reporting facilities	The adjustment factor k is used to adjust for incomplete reporting: default value 0.25 (state it if a different value is used, and whether it is the same for all reporting forms or different between forms)

The adjustment is made for each district, month and service, with the district's reporting rate for the reporting form (p. 503) of that service: the antenatal care rate for antenatal care indicators, the delivery rate for delivery indicators (deliveries, live births, caesarean sections, postnatal care…), the immunization rate for vaccinations and, in RMNCAH (p. 503) analyses, the outpatient rate for outpatient (p. 502) visits. A month whose rate is below 75%, or missing, is handled first by the low reporting rate rule (p. 84).

Formula

$$N_{\text{adj}} = N_{\text{rep}} + N_{\text{rep}} \times \left(\frac{1}{C} - 1\right) \times k$$

where

- N_{adj} is the number of events adjusted for incomplete reporting (the adjusted numerator (p. 496));
- N_{rep} is the number of events reported;
- C is the reporting completeness: the proportion of expected facility reports received (0.90 for 90%);
- k is the ratio of the volume of services in non-reporting facilities to that in reporting facilities.

$N_{\text{rep}} \times \left(1 / C - 1\right)$ is the number the non-reporting facilities would have reported if they had provided services at the same volume as the reporting ones; k says what share of that volume they are assumed to have provided.

Worked example: a district reported 900 antenatal care visits in a month, with a reporting rate of 90% (C = 0.90). With k = 0.25: 900 + 900 × (1 ÷ 0.90 − 1) × 0.25 = 900 + 900 × 0.111 × 0.25 = **925**. With k = 1 (non-reporting facilities as busy as reporting ones) it would be 1,000; with k = 0 it stays 900.

Settings

Setting	Default	What it means
Adjustment factor k	0.25 (the recommended value)	The volume of services of non-reporting facilities, as a share of that of reporting facilities.
Starting k in the app	0	A new dataset starts with k = 0 for every service: nothing is adjusted for completeness until you choose k on the **Data Adjustment** page.
Values offered	0, 0.25, 0.5, 0.75, 1	The choices for each service.
Services	One k per service	Four in RMNCAH analyses (**ANC**, **Delivery**, **Vaccines** and **OPD**), three in vaccine analyses (**ANC**, **Delivery** and **Vaccines**). See Choose the K-Factors (p. 327).

Choosing k

As a default value, we use k = 0.25, which means the non-reporting facilities provided services but only at a volume which was a quarter of the reporting health facilities. The factor k can be different for different services.

k	What it assumes about non-reporting facilities
0.00	No services provided.
0.25	Some services provided, but at a volume much lower than reporting facilities (the recommended value).
0.50	Services provided at half the rate of reporting facilities.
0.75	Services provided at nearly the same rate as reporting facilities.
1.00	Services provided at the same rate as reporting facilities.

- **Is there evidence that non-reporting facilities provide a different share of services?** If not, keep the default of 0.25.

- **Do facilities that report poorly provide a large share of a service?** For instance, if private facility reporting is poor but they are in the national system and they provide a considerable number of deliveries, *k* may be greater than 0.25, or even as high as 1.0, for deliveries.
- **Are the non-reporting facilities mostly closed or inactive?** Then a lower *k* fits better.
- Report the *k* used for each service, and what it means (see Reporting the adjustment (p. 78)).

Interpretation

- **How large is the adjustment?** It is small where reporting is nearly complete: at 95% reporting and *k* = 0.25 a number goes up by about 1.3%. It is largest where reporting is lowest.
- **Does it change the trend?** If reporting improved over the years, the adjustment adds more to the early years than to the later ones, and the adjusted trend rises less steeply than the reported one.
- **Does one year change much more than the others?** Look at the reporting rates of that year on the Reporting Rate (p. 309) page.

Example: what the completeness adjustment added to institutional live births

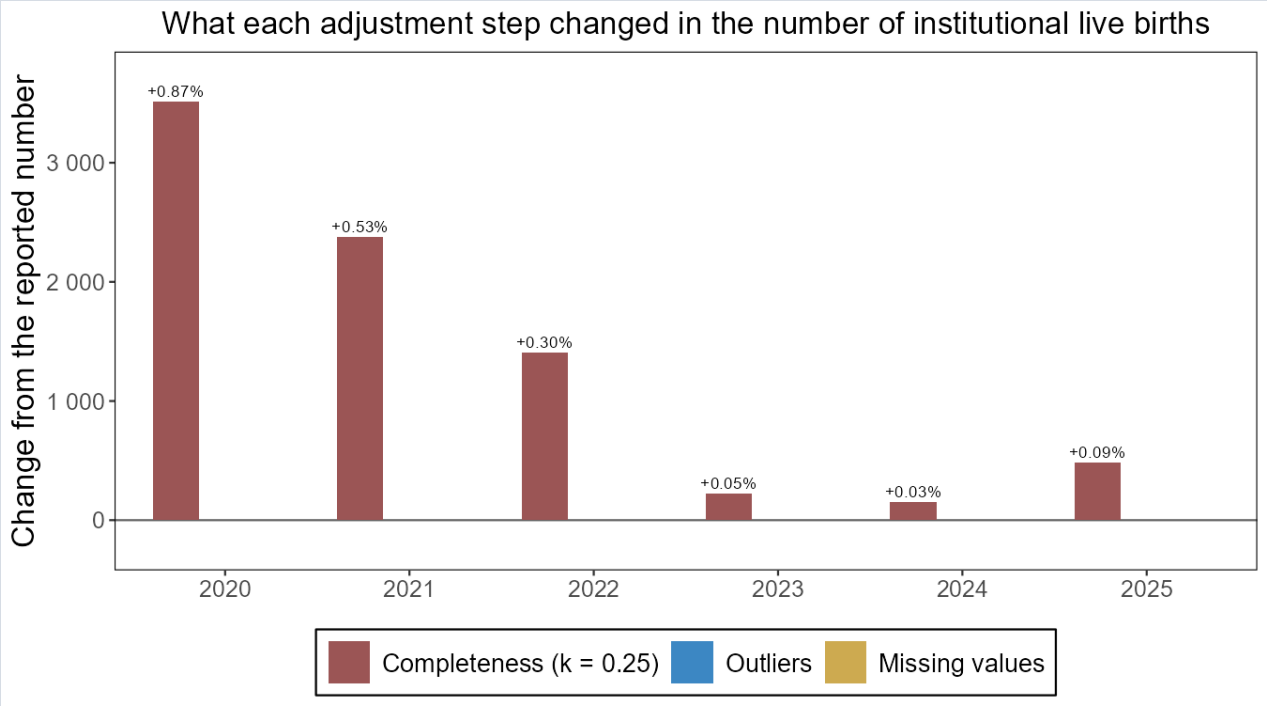


Figure 8.2 — *Change in the number of institutional live births made by each adjustment step, by year*

Reading the chart: each bar is what one adjustment step changed in the national number of institutional live births in a year, with *k* = 0.25 for every service; the label is the change as a percentage of the reported number. Red is the adjustment for completeness of reporting; blue (outliers) and gold (missing values) would show the other steps.

What it shows: only the completeness adjustment changed institutional live births. It added about 3,500 live births (+0.87%) in the first year, 2,400 (+0.53%) in the second and 1,400 (+0.30%) in the third, and less than 0.1% from the fourth year, as reporting of deliveries rose from about 96% to almost

100%. The adjusted trend therefore rises a little less steeply than the reported one.

After the adjustment

The adjusted numbers go, with the corrections for extreme outliers (p. 87) and missing values (p. 90), into the master adjusted dataset, used as the numerators of coverage in Denominator assessment and selection (p. 95) and Coverage (p. 115). Their overall effect is reported as described in The effect of the adjustments (p. 76).

8.2 Low reporting rates

What is done with a district's month whose reporting rate is below 75%, or missing, before the adjustment for completeness.

A low reporting rate is a month in which a district received fewer than 75% of the reports expected for a service. Such a rate is too uncertain to scale the reported number up with: it is replaced by the district's usual reporting rate before the adjustment for completeness of reporting (p. 80).

At a glance

- **Adjusts:** the reporting rate (p. 503) used for a district's month when it is below 75%, or missing.
- **Replaced by:** the median (p. 501) of the district's reporting rates between 75% and 100% for that service.
- **In the app:** applied automatically by **Adjust Data** on the **Data Adjustment** (p. 327) page; the effect shows on **Data Adjustment Changes** (p. 331).

Why

If the facility reporting rate is below 75%, it becomes more difficult to impute district data: the adjustment would rest on a small share of the facilities, and a few very incomplete months could produce impossible estimates. With reporting that low, further analysis to determine coverage with the facility data is not considered sufficiently reliable.

What

Problem	Action	Adjustment
Low reporting rates: identifying low rates that were adjusted	If a district's monthly reporting rate for a service is below 75% (default), or missing	The rate is replaced by the median of the district's reporting rates between 75% and 100% for that service. The month's reported number is kept, and adjusted with the factor k as usual.

The median is taken over all the district's months, in all years, whose reporting rate for that service is between 75% and 100%. If the district has no such month, there is no median, and the month's number is kept as reported.

Formula

$$C = \begin{cases} C_m & \text{if } 75\,\% \le C_m \le 100\% \\ \tilde{C}_{75} & \text{if } C_m < 75\,\% \text{ or missing} \end{cases}$$

where

- C is the reporting rate used to adjust the month's number (in the completeness formula (p. 80));
- C_m is the district's reporting rate for the service in that month;
- $\tilde{C}_{75}$ is the median of the district's reporting rates for that service that are between 75% and 100%.

Worked example: a district reported 600 antenatal care visits in a month in which only 60% of the expected reports arrived. Its other months usually report between 90% and 100%, with a median of 96%. The rate used is 96%, not 60%. With $k = 0.25$ the number becomes $600 + 600 \times (1 \div 0.96 - 1) \times 0.25 = \mathbf{606}$. With the 60% rate it would have become $600 + 600 \times (1 \div 0.60 - 1) \times 0.25 = 700$, an increase resting on a minority of the facilities.

Settings

Setting	Default	What it means
Reporting rate cutoff	75%	A monthly reporting rate below this (or missing) is not used as it is.
Rates used for the replacement	75% to 100%	The district's rates, between these limits inclusive, from which the replacement median is taken.

Interpretation

- **Are there district-months below 75%?** The Reporting Rate (p. 309) page and the districts with low reporting rates show where they are.
- **Did the rule add little to those months?** That is expected: the replacement rate is usually high, so a month with very low reporting is adjusted as if it were a normal month, and its number stays low. Coverage of that district and year is then probably underestimated.
- **Are many months, or whole years, below 75% in a district?** Then the facility data of that district are not considered sufficiently reliable for coverage. Note it in the interpretation and consider removing the year (p. 93) if the problem is national.

Example: a district with low outpatient reporting

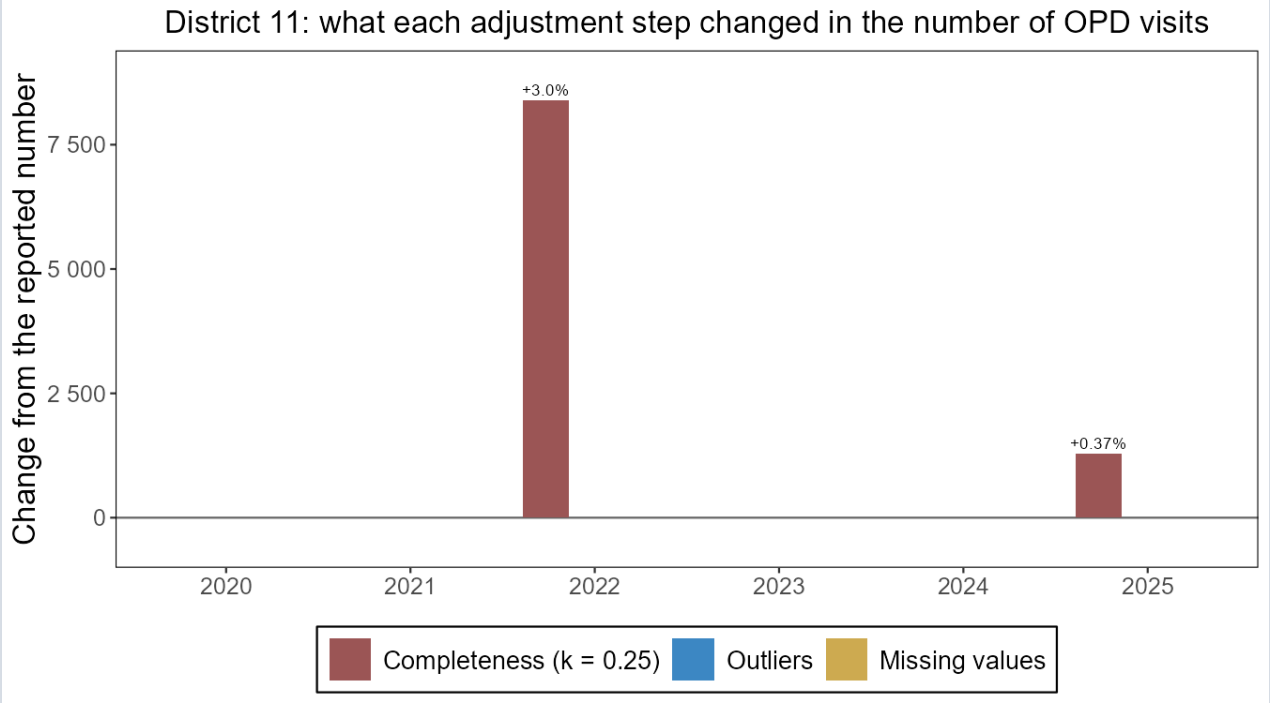


Figure 8.3 — *Change in the number of OPD visits of one district made by each adjustment step, by year*

Reading the chart: each bar is what one adjustment step changed in the number of outpatient (OPD) (p. 502) visits of one district (District 11) in a year, with $k = 0.25$; the label is the change as a percentage of the reported number. Red is the adjustment for completeness of reporting.

What it shows: in the first two years the district's outpatient reporting rate was below 75% in every month (about 68% to 74%). Those rates were replaced by the median of its months at 75% to 100%, which is 100%, so nothing was added in those years. Had the low rates been used, the first year would have gone up by about 11% (some 24,000 visits). In the third year, with reporting around 88% to 90%, the adjustment added 3.0% (about 8,400 visits); in the fourth and fifth years reporting was complete and nothing was added, and in the last year (about 98%) 0.37%. The low numbers of the first two years remain: outpatient coverage of this district in those years is not reliable.

After the adjustment

The month's number, adjusted with the replaced rate, then goes through the extreme outlier (p. 87) and missing value (p. 90) corrections, and into the master adjusted dataset used for Denominator assessment and selection (p. 95) and Coverage (p. 115).

8.3 Extreme outliers

How extreme monthly outliers are corrected: replaced with the median of the district's other months in the same year.

An extreme outlier (p. 499) is a monthly number far from the usual values of the same district, such as an accidental extra zero. The data quality assessment (p. 64) finds and counts them; the data adjustment corrects them, so that a single wrong month does not distort a district's annual total.

> **At a glance**
>
> - **Adjusts:** each monthly value more than 5 × MAD from the district's median monthly value.
> - **Replaced by:** the median (p. 501) of the district's other (non-outlier) months in the same year.
> - **In the app:** applied automatically by **Adjust Data** on the **Data Adjustment** (p. 327) page; the effect shows on **Data Adjustment Changes** (p. 331).

Why

Extreme outliers, such as an accidental extra zero in a number, can have a large impact, especially on subnational numbers, and can greatly affect coverage trend assessments: one month ten times too high can nearly double a district's annual total.

What

Extreme outliers, as defined in the data quality assessment (p. 64), are corrected by imputing the median monthly value of the same year.

Problem	Action	Adjustment
Extreme outliers can greatly affect coverage trend assessments	If a monthly value is greater or smaller than 5 times the Median absolute deviation (MAD) (p. 501) from the district's monthly median value, an adjustment is made	Extreme monthly outliers are corrected and given the median of the district's other (non-outlier) months in the same year

- The median and MAD that define an outlier are computed from all years except the most recent one, so the latest year is judged against the earlier years.
- Outliers are looked for after the adjustment for completeness of reporting (p. 80), in the adjusted numbers.
- All the adjusted service indicators are corrected, and inpatient numbers too. Mortality indicators are not.

Formula

A monthly value x is an extreme outlier when

$$|x - m| > 5 \times MAD$$

and it is then replaced:

$$x \to \tilde{x}_y$$

where

- m is the district's median monthly value, over all years except the most recent;
- MAD is the median absolute deviation of the same values;
- $\tilde{x}_y$ is the median of the district's months in the same year y that are not extreme outliers.

Worked example: a district usually gives about 500 penta1 vaccinations a month (median 500, MAD 40, so the normal range is 500 ± 5 × 40, from 300 to 700). One month was entered as 5,000. It is an extreme outlier and is replaced by the median of the district's other months that year, say 505. The annual total drops from about 10,500 to about 6,000.

Settings

Setting	Default	What it means
Outlier bound	5 × MAD	How far from the median a monthly value must be to be corrected.
Baseline years	All years except the latest	The years used for the median and MAD.
Replacement	Median of the non-outlier months, same district and year	The value given to an extreme outlier.

Interpretation

- **Did the outlier correction change a year's total by more than the completeness adjustment?** Then that year had large outliers: look at them on the Outlier Detection (p. 312) page.
- **Is there a real-world explanation?** Outliers are identified statistically. A vaccination campaign, an outbreak or a stock-out can produce a real peak, and the correction then removes real services; consider the reasons before accepting the correction.
- **Are the outliers mostly in the latest year?** The latest year is compared with the earlier ones, so a real, steady increase can be flagged. Check the trend.

Example: outliers in measles 1 vaccinations

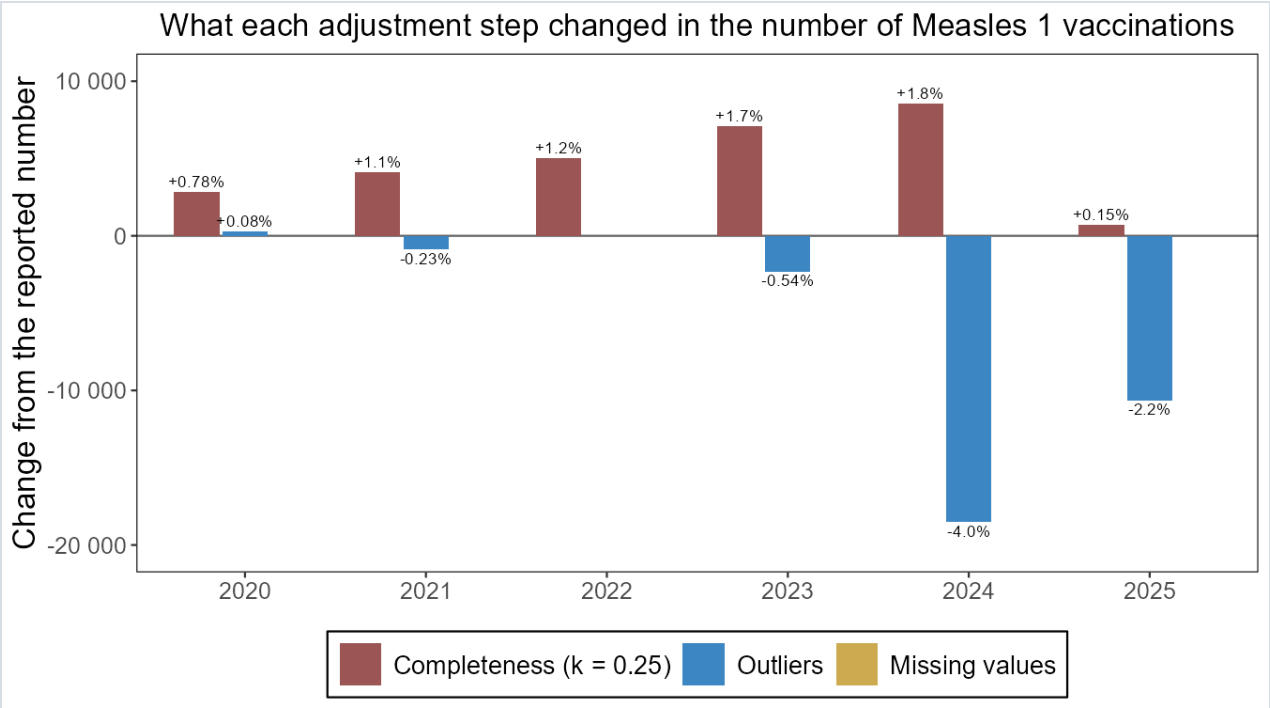


Figure 8.4 — *Change in the number of measles 1 vaccinations made by each adjustment step, by year*

Reading the chart: each bar is what one adjustment step changed in the national number of measles 1 vaccinations in a year; the label is the change as a percentage of the reported number. Red is the adjustment for completeness of reporting (k = 0.25), blue the outlier correction (below zero when outliers were lowered).

What it shows: the completeness adjustment added between 0.15% and 1.8% a year. The outlier correction was small in the first four years (between +0.08% and −0.54%) but removed about 18,500 vaccinations (−4.0%) in the fifth year and 10,600 (−2.2%) in the last, more than the completeness adjustment added: the adjusted numbers of those two years are about 2% lower than the reported ones. Before accepting this, check on the Outlier Detection page which districts and months were corrected, and whether a measles campaign or another real event explains the peaks.

After the adjustment

The corrected numbers then go through the missing value (p. 90) correction and into the master adjusted dataset, used for Denominator assessment and selection (p. 95) and Coverage (p. 115).

8.4 Missing values

How a month with no number is filled: with the district's median monthly value for the year.

A missing value (p. 502) is a month with no number for an indicator in a district. The data quality assessment (p. 62) shows where they are; the data adjustment fills them by imputation (p. 500), so that the month does not count as zero in the annual total.

At a glance

- **Adjusts:** each month with no number for an indicator in a district.
- **Filled with:** the district's median (p. 501) monthly value for the year.
- **In the app:** applied automatically by **Adjust Data** on the **Data Adjustment** **(p. 327)** page; the effect shows on **Data Adjustment Changes** **(p. 331)**.

Why

A missing month counts as zero when the year is added up, so the annual total, and the coverage (p. 503) computed from it, is too low. Filling it with a typical value for the district and year keeps the annual total comparable with the other years and districts.

What

Problem	Action	Adjustment
Missing values	If there is a missing value, data are imputed	The district's median monthly value for the year is imputed for the month with the missing value

- The median is taken over the district's months of the same year that have a value, after the corrections for completeness and extreme outliers.
- If a district has no value at all for the indicator in a year, there is no median and the months stay missing.
- All the adjusted service indicators are filled; inpatient and mortality numbers are not.

Formula

$$x_{\text{missing}} \rightarrow \tilde{x}_y$$

where

- x_{missing} is a month of the district with no value for the indicator;
- $\tilde{x}_y$ is the median of the district's monthly values for the indicator in the same year y.

Worked example: a district reported caesarean sections in 11 months of a year, with a median of 41, and nothing in March. March is given 41, and the annual total goes up by 41.

Settings

The rule has no setting: missing months are always filled with the district's median for the year.

Interpretation

- **Did filling missing values change a year's total noticeably?** Then many months were missing in that year: check on the Data Missingness (p. 315) page which districts and indicators they come from.
- **Are there districts with a whole year missing for an indicator?** They are not filled, so the district's total for that year stays zero. Find out whether the service was not provided, not recorded, or not yet introduced.
- **Is the indicator newly introduced?** Months before its introduction are missing for a real reason; filling within a year cannot create them, but a partial first year will be filled from the months that have values.

Example: missing values in caesarean sections

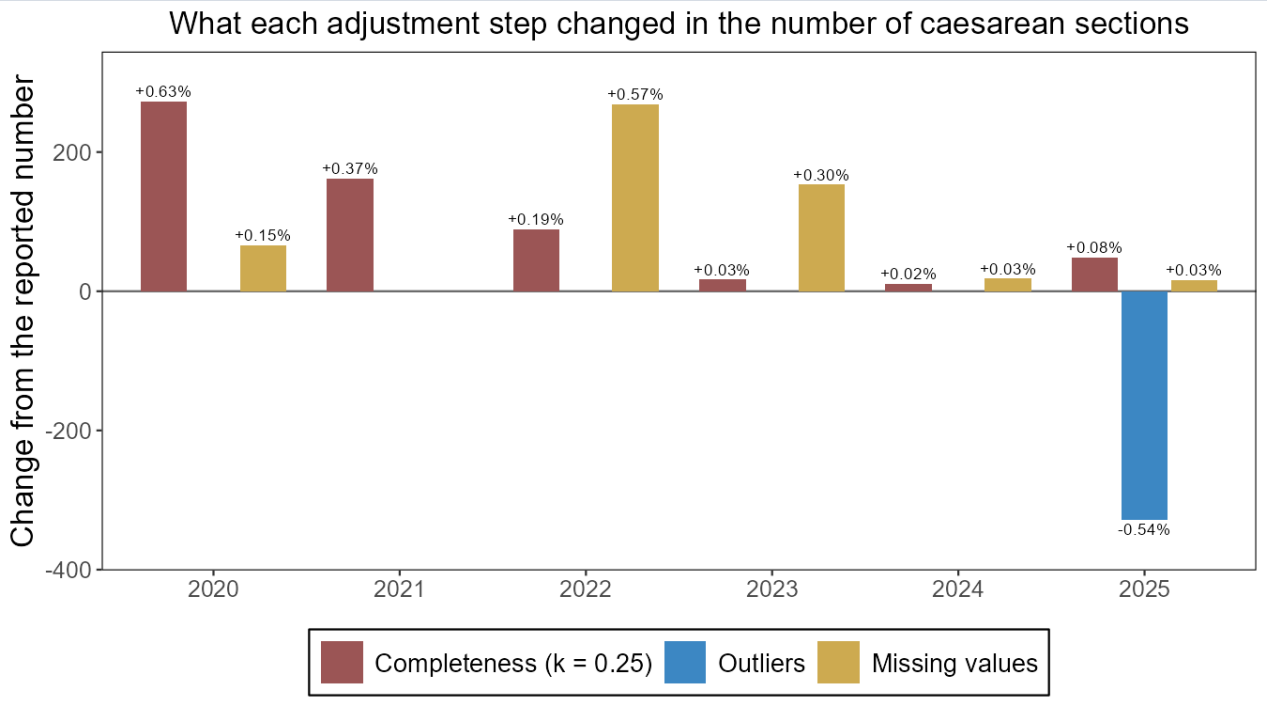


Figure 8.5 — *Change in the number of caesarean sections made by each adjustment step, by year*

Reading the chart: each bar is what one adjustment step changed in the national number of caesarean sections in a year; the label is the change as a percentage of the reported number. Red is the adjustment for completeness of reporting (k = 0.25), blue the outlier correction and gold the missing values filled.

What it shows: filling missing values added about 270 caesarean sections (+0.57%) in the third year and 150 (+0.30%) in the fourth, and less than 0.2% in the other years. It added nothing in the second year: the district with most gaps had no caesarean section number in any month of that year, so there was no median to fill them with, and that district's total for the year stays zero. That district needs a look on the Data Missingness page before caesarean section rates are compared between districts.

After the adjustment

Filling missing values is the last step: the numbers then form the master adjusted dataset, used for Denominator assessment and selection (p. 95) and Coverage (p. 115). The overall effect of all the adjustments is reported as described in The effect of the adjustments (p. 76).

8.5 Removing years

When a whole year of facility data is too unreliable to use, and how it is left out of the analysis.

Some years cannot be corrected: a year with no reporting data, or one whose data the national technical team has found unfit for analysis. Removing a year leaves it out of the data adjustment and of every later step, so trends and coverage rest only on the years that can be relied on.

> **At a glance**
>
> - **Removes:** whole calendar years from the analysis.
> - **When:** only for years with no data, or clearly unusable data and a real-world reason.
> - **In the app:** the **Remove Years** (p. 326) page.

Why

Removing years is only needed by countries that want to exclude one or more of the years because of data issues. A year with almost no reporting, or a known system-wide failure, cannot be repaired by the adjustments: the adjustment for completeness (p. 80) and the low reporting rate rule (p. 84) assume that most facilities reported. Its numbers would distort coverage trend assessments.

What

Problem	Action
A year contains no reporting data	Remove the year
A year's data have been flagged unfit for analysis by the national technical team, because of severe, systemic quality issues	Remove the year

Removed years are taken out of the data before any adjustment, so they are used in no other calculation. A removed year can be brought back at any time; the adjustment must then be run again.

When to remove a year

Remove a year only when the data quality (p. 57) checks show that it is clearly erratic or unusable, and you have a real-world reason to believe it. If you have no reason, keep the year.

Removing a year is recommended when:

- its data quality falls well below the acceptable thresholds (see the Annual data quality score (p. 72));
- the country team has sufficient, verified contextual information that makes that year's data completely unreliable or inaccurate (for example a major system-wide roll-out failure, a prolonged network blackout, or a database corruption).

Interpretation

- **Is the year really unusable, or only weaker than the others?** A low score is a reason to investigate, not always to discard a year. Compare with local events before removing it.

- **Which later results does the removal affect?** Every trend, coverage estimate and comparison with surveys loses that year. Removing a year you should keep silently shortens every later trend.
- **Is the removal reported?** State which years were removed and why in the interpretation of the results.

After the adjustment

The remaining years are adjusted for completeness of reporting (p. 80), extreme outliers (p. 87) and missing values (p. 90). If the removed years change after the adjustment, the adjustment must be run again (see Remove Years (p. 326) in the app).

CHAPTER 9

Denominator assessment and selection

How to assess the DHIS2 population projections, build denominators from ANC1 and Penta1, and select the denominator that gives the most plausible coverage, compared with the survey.

Service coverage (p. 503) is the population who received a service divided by the population who need it (the target population (p. 504)). The numerator comes from the health facility data; the denominator has to be estimated, for the country and for each region and district. This step assesses the population projections in DHIS2 (p. 498) and the other ways of estimating the denominator, and selects the one that gives the most plausible coverage.

It comes after Data adjustment (p. 75), which corrects the numerators, and before Coverage (p. 115), which uses the denominator selected here for every coverage indicator.

At a glance

- **Answers:** are the DHIS2 population projections plausible, and which denominator gives the most plausible coverage, compared with the survey, at national and subnational level?
- **Uses:** the adjusted numbers from Data adjustment (p. 75), the UN (p. 504) and DHIS2 population projections, and the survey values (ANC1 and Penta1 coverage, stillbirth rate, neonatal and post-neonatal mortality) from Analysis setup (p. 54).
- **Produces:** the selected denominator: one for a vaccine analysis, two for an RMNCAH (p. 503) analysis (a maternal denominator (p. 501) and an immunization denominator).
- **In the apps: Denominators (p. 333)** (Denominator Assessment (p. 335) and Denominator Selection (p. 338)).

Why

Service coverage is defined as the population who received the service divided by the population who need the services (also referred to as target population). The numerators of coverage statistics (e.g., number of live births in health facilities) are derived from the health facility data and need to be adjusted as shown in the previous section. The denominator of the coverage statistics (e.g., number of live births in the population) needs to be estimated for national and subnational levels (regions/provinces and districts).

The objective of the health facility denominator analysis is twofold. First, we assess the quality of the population projections in DHIS2 by comparison with the UN projections and the internal consistency. Then, we assess the performance of multiple denominator options for the computation of population-based service coverage indicators from the health facility data. This should lead to a final decision on the denominators that are used for the analyses of population-based coverage indicators based on health facility data.

Because the numerator comes from facility data and the denominator has to be estimated, the choice of denominator matters as much as the data quality. For the scientific basis of facility-based coverage, see Coverage analysis: rationale (p. 115).

How the parts fit together

The first part tells you whether the DHIS2 projections can be trusted as a denominator, and which DHIS2 population (total population, live births, total births or under-1) is the most consistent with the UN projections; that population carries the population-growth options from one year to the next. The second part explains the options themselves. The third compares them all with the survey and ends with the choice. A DHIS2 projection that fails the checks of the first part will usually also give implausible coverage in the third.

Doing it in the app

① Assess the population projections

On **Denominator Assessment** (**Population Trend** in the Vaxx app), compare the UN and DHIS2 projections on the **Population**, **Births** and **Under 1** tabs, and set **Select Best Population** (see Denominator Assessment (p. 335)).

② Compare the denominator options

On **Denominator Selection**, compare coverage under each option with the survey: nationally, by region and over time (see Denominator Selection (p. 338)).

③ Choose the denominator

Set the denominator(s) the rest of the analysis uses, and write down why (see Choose the denominator (p. 339)).

Points to consider for the interpretation

- Are the DHIS2 projections robust: is the total population projection consistent over time with regular population growth, is the live birth projection consistent over time (a regular trend), are the projected total population and live births close to the UN projections, and are they consistent with the UN estimates of the crude birth rate and crude death rate?
- Which denominator methods performed best at the national level for the two indicators (institutional live births and penta3)?
- Which denominator performed best at the subnational level for the two indicators?
- What selection is made for the indicators in the coverage analyses?

Common pitfalls

- Comparing against the survey when the survey year is far from the analysis years. Weigh that in your judgement.

- Choosing a denominator per chart instead of once for the whole analysis. The denominator you select changes every coverage, inequality, target, mortality and health system result.
- Choosing the denominator that gives the highest coverage. The best denominator is the one closest to the survey, with plausible trends.
- Missing survey coverage values in the national parameters: the survey comparison then has nothing to compare with.

9.1 Quality of the DHIS2 population projections

Whether the population projections in DHIS2 are consistent over time and close to the UN projections, at national level.

Most countries keep population projections in DHIS2 (p. 498): the total population, live births, total births and children under 1, by district and year, usually projected from the last census (the population data (p. 502)). They are one of the possible denominators. In the first part of the denominator assessment, their quality is checked at national level, in two ways: their internal consistency over time, and their agreement with the UN population projections (p. 504).

At a glance

- **Measures:** the population growth rate (p. 502) and crude birth rate (p. 498) implied by the DHIS2 projections, over time and against the UN projections, and the crude death rate (p. 498) they imply.
- **Good result:** a regular trend in total population and live births, close to the UN projections, with plausible birth and death rates.
- **In the app:** the **Denominator Assessment** **(p. 335)** page (**Population Trend** in the Vaxx app), card **Population Comparison**.

Why

The DHIS2 projections are used directly as a denominator (the **DHIS2 projections** option), and the population you select here carries the population-growth options from one year to the next (see Facility data derived denominators (p. 106)). The first part assesses the accuracy and consistency of the projected population numbers in DHIS2 by comparing them to external sources. Differences with the UN projections may occur, but large differences suggest that there is a problem with the DHIS2 population projections.

What

The internal consistency of the DHIS2 population growth over time, and the comparison with the UN projections for four indicators, with the values that flag a problem:

Indicator	Numerator	Denominator	Interpretation
Annual population growth rate (DHIS2, over time)	Population in a year minus population in the previous year	Population in the previous year	Expected to be consistent over time: a change of 2 per 1,000 or more between years is inconsistent
Crude birth rate (DHIS2, over time)	Projected live births × 1,000	Total population	Expected to be consistent over time: a change of 2 per 1,000 or more between years is inconsistent
Population size compared with the UN	Absolute difference between the DHIS2 and UN population	UN population	A relative difference greater than 5% should be considered a data quality issue with the DHIS2 projections
Population growth compared with the UN	Absolute difference between the DHIS2 and UN annual growth rates	–	A difference greater than 0.3% is a reason for concern with the data quality

Crude birth rate compared with the UN	Absolute difference between the DHIS2 and UN crude birth rates	–	A difference greater than 5 per 1,000 population: the DHIS2 projection may be problematic
Crude death rate implied by DHIS2	Crude birth rate minus population growth rate (both per 1,000)	–	Compared with the UN; a negative crude death rate, or one below 5 per 1,000 population, means the data are inconsistent

The crude death rate is the difference between the crude birth rate and the population growth if there is no major out- or in-migration. UN projections exist at national level only, so these checks are national.

Formula

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$
$$g_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$$
$$\mathrm{CBR}_t = \frac{B_t}{P_t} \times 1000$$
$$\mathrm{CDR}_t = \mathrm{CBR}_t - 1000 \times g_t$$
$$\Delta P_t = \frac{\left|P_t^D - P_t^U\right|}{P_t^U} \times 100$$

where

- P_t is the total population in year t (P^D in the DHIS2 projections, P^U in the UN projections);
- B_t is the number of live births projected for year t;
- r_t is the annual growth rate used to check the DHIS2 trend, and g_t the annual growth rate compared with the UN;
- CBR_t and CDR_t are the crude birth and death rates, per 1,000 population;
- ΔP_t is the relative difference between the DHIS2 and UN population, in percent.

The growth rates are compared as $\left|g_t^D - g_t^U\right|$ (flag above 0.3%) and the birth rates as $\left|\mathrm{CBR}_t^D - \mathrm{CBR}_t^U\right|$ (flag above 5 per 1,000).

Worked example. A DHIS2 projection of 1,045,000 people in one year and 2,005,000 the next gives $g = \ln(2{,}005 / 1{,}045) = 0.652$: a growth of 65%, against about 2.5% in the UN projections. With 75,700 projected live births, $\mathrm{CBR} = 75.7 / 2{,}005 \times 1000 = 37.8$ per 1,000, and $\mathrm{CDR} = 37.8 - 652 = -614$ per 1,000: a negative death rate, so the projection is inconsistent.

Settings

Setting	Default	What it means
Select Best Population	Live births (DHIS2)	The DHIS2 population that is most consistent with the UN projections: live births, total births, total population or under 1 year. Its growth carries the population-growth options from the reference year to the other years. Set it on the **Denominator Assessment** page.

Interpretation

The DHIS2 projections are considered robust when:

1. the DHIS2 total population projection is consistent over time, with regular population growth;
2. the DHIS2 total live birth projection is consistent over time (a regular trend);
3. the projected numbers of total population and live births are close to the UN population projection;
4. the DHIS2 population projections are consistent with the UN estimates of the crude birth rate and crude death rate.

- **Do the UN and DHIS2 lines follow the same direction, with a stable gap?** Compare the total population, live births and under-1 population on the three tabs. Select the population (in **Select Best Population**) whose trends are consistent across the two sources, with estimates generally close: for example, the under-1 population if the UN estimates and the DHIS2 projections for that group are very close.
- **Is there a large gap between UN and DHIS2?** It suggests an outdated census base or a wrong population entry in the data. DHIS2 projections often drift when they are based on an old census.
- **Does one line break suddenly?** A sudden break usually means a change in how the population was projected, not a real change in births. Check the growth rate and birth rate for that year.
- **Is any check flagged?** Consider reasons, and expect the DHIS2 projections to perform poorly as a denominator in the selection (p. 108).

The app draws the trends; the growth, birth and death rates are computed by the analyst from the projected numbers.

Example: total population, DHIS2 and UN projections

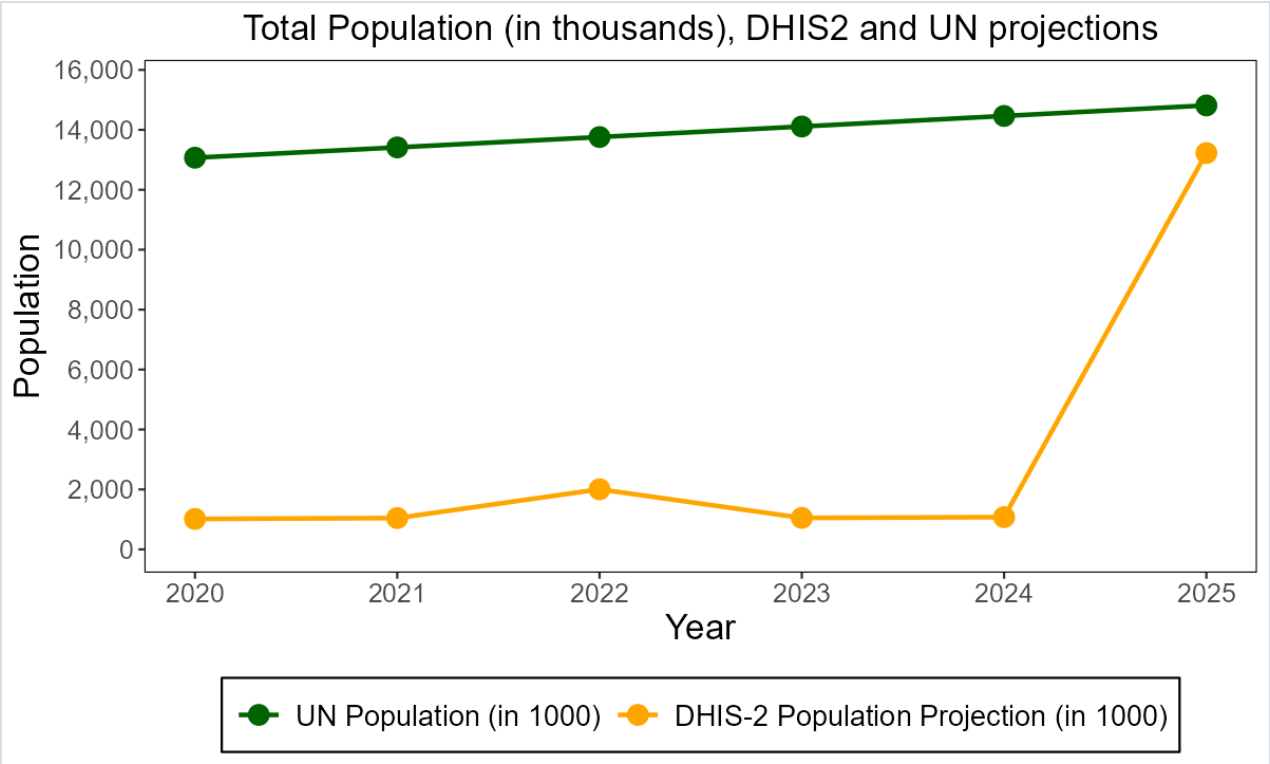


Figure 9.1 — *Total population in thousands, DHIS2 and UN projections, by year*

Reading the chart: each line is one source, by year, in thousands: green for the UN projection, orange for the DHIS2 projection.

What it shows: the UN projection grows regularly, from 13.1 million in the first year to 14.8 million in the sixth (about 2.5% a year). The DHIS2 projection is about 1.0-1.1 million in most years, 92% below the UN, jumps to 2.0 million in the third year, falls back, and jumps to 13.2 million in the last year (11% below the UN). The growth rates it implies (+65%, −65%, +251% in the years with a jump) are far from the UN's 2.5%, and the relative difference is above 5% in every year. The DHIS2 total population fails checks 1 and 3.

Example: live births, DHIS2 and UN projections

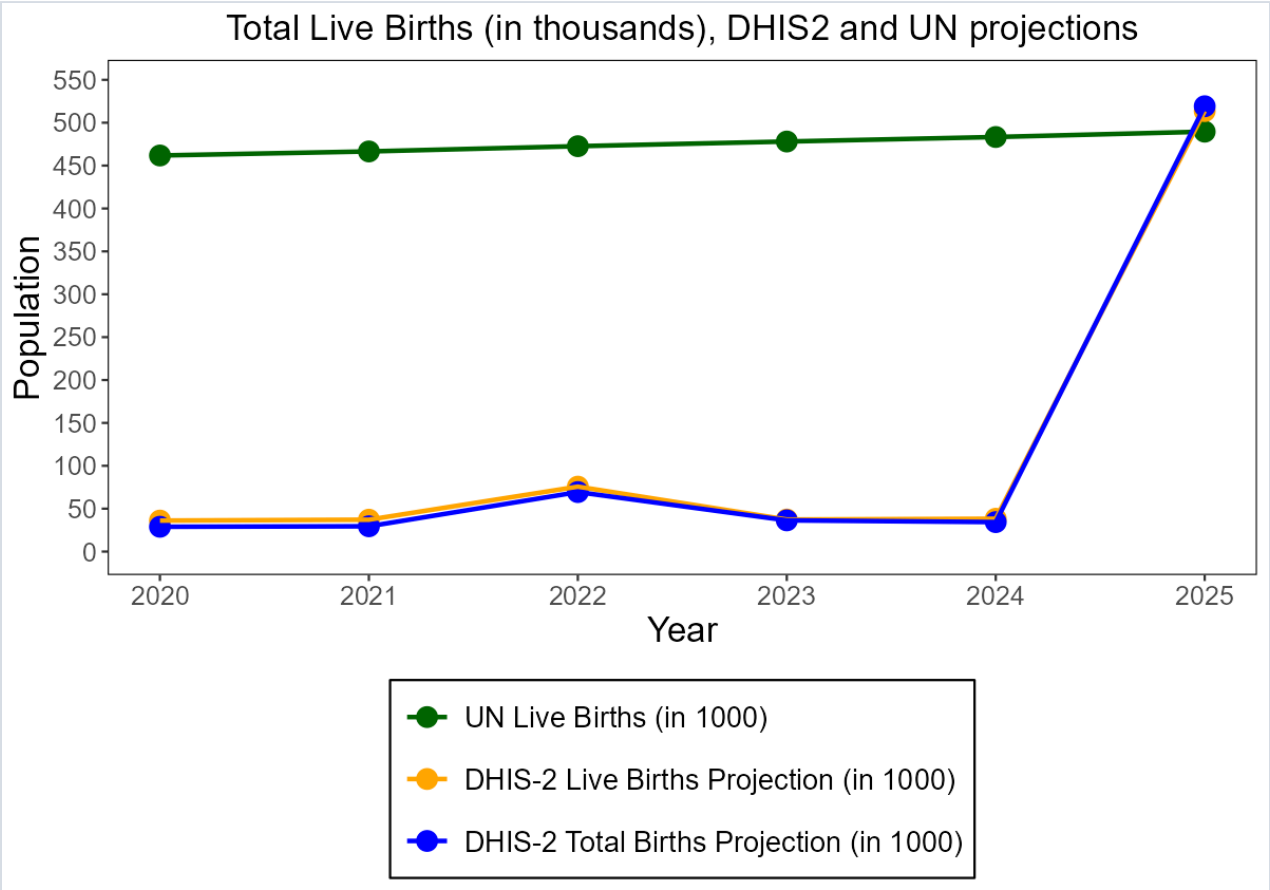


Figure 9.2 — *Total live births in thousands, DHIS2 and UN projections, by year*

Reading the chart: green is the UN projection of live births, orange the DHIS2 projection of live births and blue the DHIS2 projection of total births, in thousands.

What it shows: the UN projects 462,000 to 490,000 live births a year. DHIS2 projects 36,000 to 38,000, with 76,000 in the third year and 514,000 in the last; its total births are even lower than its live births in the first five years, which is not possible. The crude birth rate implied by DHIS2 (35.6 per 1,000 in most years) is close to the UN's (33-35 per 1,000) because births and population are understated together, but it jumps by more than 2 per 1,000 in the third and last years, and differs from the UN by 5.8 per 1,000 in the last year. The implied crude death rate is negative in the years the population jumps. None of the DHIS2 populations is consistent with the UN projections: the DHIS2 projections should not be used as a denominator, and the population-growth options built on them need care.

After the assessment

The result of this part feeds the selection of the best performing denominator (p. 108), where the DHIS2 projections are compared with the facility data derived denominators (p. 103) against the survey. The population selected in **Select Best Population** is used by the population-growth options.

9.2 Facility data derived denominators

The denominator options, and how target populations are derived from the reported numbers of ANC1 visits and Penta1 doses with the survey coverage.

The denominator of a coverage indicator can come from population projections (UN or DHIS2), or it can be derived from the health facility data themselves. A facility data derived denominator (p. 499) starts from a service that nearly everyone uses, ANC1 (p. 496) or Penta1 (p. 502), and adds the share of the population that did not use it, according to the survey (p. 500). This page describes the six denominator options and how the derived ones are computed.

At a glance

- **Options:** UN projections, DHIS2 projections, ANC1-derived, Penta1-derived, ANC1 population growth, Penta1 population growth.
- **Uses:** the adjusted numbers of ANC1 visits and Penta1 doses, the survey coverage of ANC1 and Penta1, and the assumptions on pregnancy loss (p. 503), stillbirths (p. 504), twins (p. 504), neonatal (p. 502) and post-neonatal mortality (p. 503).
- **In the app:** computed for you; listed under **Denominator options (p. 339)** and compared on the **Denominator Selection (p. 338)** page.

Why

The basic idea is that if the coverage of an indicator is high (e.g., over 90%), then the number of events reported by health facilities has to be close to the target population. In other words, the denominators or target population can be derived from the numbers in DHIS2. The best candidate indicators for this approach are ANC1 and DPT/penta1 (BCG also possible in some countries if re-vaccinations are recorded separately). This gives a denominator even where the population projections are poor, and one that is available for every region and district.

Denominator options

Six denominator options are analysed to determine the most accurate baseline:

Option	What it is	Id in the data
UN projections	The UN estimates of live births and population. National only	`un`
DHIS2 projections	The DHIS2 projection of live births; for vaccines, the DHIS2 population found most consistent with the UN projections (live births, total births, total population or under 1)	`dhis2`
ANC1-derived	Derived each year from that year's reported ANC1 visits and the survey's ANC1 coverage	`anc1`
Penta1-derived	Derived each year from that year's reported Penta1 doses and the survey's Penta1 coverage	`penta1`
ANC1 population growth	The ANC1-derived value of the reference year, carried to the other years with annual population growth	`anc1derived`

Penta1 population growth	The Penta1-derived value of the reference year, carried to the other years with annual population growth	`penta1derived`

The ids ending in `derived` are the population-growth (p. 502) options, not the "-derived" ones: see the Data dictionary (p. 483).

> **UN projections are national only.** For the subnational level, no UN projections are available, so below national level only the DHIS2 projections and the ANC1 and Penta1 options are compared. In the apps the UN projections appear in the national comparison; the denominators you can select are DHIS2, ANC1-derived, Penta1-derived, ANC1 population growth and Penta1 population growth.

What

Each coverage indicator uses the population that matches it. For vaccination indicators, the denominator depends on the recommended age of administration:

Indicator	Numerator	Denominator
% of pregnant women with ANC1, ANC4, a first-trimester visit, IPT, iron-folic acid, syphilis or HIV testing	Number of women receiving the service	Number of pregnancies
% of institutional deliveries; % of caesarean sections	Number of institutional deliveries; of caesarean sections	Number of deliveries
% of institutional live births (p. 500); skilled birth attendance, low birth weight, postnatal care within 48 hours	Number of live births in facilities; number receiving the service	Number of live births
% of infants receiving BCG vaccine	N of infants receiving BCG vaccine	N of live births
% of infants receiving penta, OPV, PCV, IPV or rota vaccines	N of infants receiving the vaccine	N of infants eligible for the vaccine: live births minus neonatal deaths
% of infants receiving first dose of measles vaccine	N of infants receiving first dose of measles vaccine	N of infants eligible: live births minus neonatal deaths minus post-neonatal deaths
% of children receiving second dose of measles vaccine	N of children receiving second dose of measles vaccine	N of children eligible: live births minus neonatal deaths minus 2 × post-neonatal deaths

The maternal and newborn health denominators are closely related and can be computed from each other by making assumptions. Starting with pregnancies, the number of live births is closely associated with the number of pregnancies, which are usually identified within the health system at the first antenatal visit, which in most countries is around 4-5 months of pregnancy (according to the surveys). The immunization denominators (p. 500) and measles denominators (p. 501) then follow from live births.

The derived approach requires the following:

- A recent population-based survey is used to obtain an estimate of population-level coverage of ANC1 or penta1. For example, ANC1 coverage is 95% of pregnant women.

- The DHIS2 data on the number of ANC1 and penta1 visits need to be considered complete and accurate (after adjustments and cleaning of the data). For example, 100,000 ANC1 visits were reported.

Formula

The ANC1-derived option starts from the women attending a first antenatal visit and works forward through pregnancy to the eligible infants:

$$N_{preg} = \frac{A}{c_A}$$
$$N_{deliv} = N_{preg} \times \left(1 - p\right)$$
$$N_{births} = \frac{N_{deliv}}{1 - 0.5\,w}$$
$$N_{lb} = N_{births} \times \left(1 - s\right)$$
$$N_{inf} = N_{lb} \times \left(1 - n\right)$$
$$N_{m1} = N_{inf} \times \left(1 - q\right)$$
$$N_{m2} = N_{inf} \times \left(1 - 2q\right)$$

The Penta1-derived option starts from the infants vaccinated with Penta1 and works backward:

$$N_{inf} = \frac{V}{c_V}$$
$$N_{lb} = \frac{N_{inf}}{1 - n}$$
$$N_{births} = \frac{N_{lb}}{1 - s}$$
$$N_{deliv} = N_{births} \times \left(1 - 0.5\,w\right)$$
$$N_{preg} = \frac{N_{deliv}}{1 - p}$$

and the measles targets N_{m1}, N_{m2} from N_{inf} as above.

where

- A is the number of ANC1 visits reported in the year, and c_A the survey ANC1 coverage (as a proportion);
- V is the number of Penta1 doses reported in the year, and c_V the survey Penta1 coverage;
- N_{preg}, N_{deliv}, N_{births}, N_{lb} are the pregnancies, deliveries, births and live births;
- N_{inf} is the infants eligible for Penta (live births minus neonatal deaths); N_{m1}, N_{m2} the measles 1 and measles 2 targets;
- p is the pregnancy loss after the first antenatal visit, w the twinning rate, s the stillbirth rate, n the neonatal mortality and q the post-neonatal mortality (all proportions).

The DHIS2 and UN projections of live births are converted to the other populations with the same assumptions (pregnancies $= N_{lb} \times \left(1 - 0.5\,w\right) / \left(\left(1 - s\right)\left(1 - p\right)\right)$, deliveries, births and infants as above). Coverage is then the numerator divided by the matching population, × 100.

Worked example (ANC1). If ANC1 coverage from the survey is 95% and the number of ANC1 visits from DHIS2 for the year is 100,000, the total number of pregnant women is 100,000 / 0.95 = 105,263. Then:

- at 3% abortion, 105,263 × (1 – 0.03) = 102,105 deliveries;

- at a 1.5% twinning rate, 102,105 / (1 – 0.015/2) = 102,877 births;
- at a 2% stillbirth rate, 102,877 × (1 – 0.02) = 100,819 live births;
- at 3% neonatal mortality, 100,819 × (1 – 0.03) = 97,795 children eligible for DPT1/penta1.

Worked example (Penta1). The survey coverage is the percent of children 12-23 months who received DPT1/penta1; the facility data are the number of infants who received DPT1/penta1. If survey coverage is 92% and there were 100,000 vaccinations given, the target population is 100,000 / 0.92 = 108,696.

Population growth options

The ANC1 population growth and Penta1 population growth options are not re-derived from each year's ANC1 or Penta1. They take the ANC1- or Penta1-derived value of the **reference year**, the year before the survey year (or the first year of data, if that is later), and carry it to the other years with the growth of the population chosen in **Select Best Population** on the DHIS2 population projections (p. 99) page: the value moves up or down by the change in that population between the reference year and the year in question. In the reference year the two versions are equal.

For regions and districts, the region's value is carried in the same way and shared among its districts in proportion to their part of the chosen population. These options therefore follow the trend of the chosen DHIS2 population, and they are only as good as its consistency over time.

Settings

Setting	Default	What it means
ANC1 survey coverage	From the most recent survey; 98% if missing	The share of pregnant women with at least one antenatal visit, used to derive pregnancies from ANC1.
Penta1 survey coverage	From the most recent survey; 97% if missing	The share of children 12-23 months who received Penta1, used to derive infants from Penta1.
Pregnancy loss	3%	Pregnancy loss between 4 and 7 months (28 weeks of pregnancy), after the first antenatal visit.
Stillbirth rate	From the survey; 2% if missing	Stillbirths, or pregnancy loss between 28 weeks and birth.
Twinning rate	1.5%	The share of births that are twins. With pregnancy loss and stillbirths, it gives the number of live births computed from pregnancies.
Neonatal mortality	From the survey; 2.5% if missing	Deaths in the first month, per live birth: live births minus neonatal deaths are the infants eligible for Penta.
Post-neonatal mortality	From the survey; 2.4% if missing	Deaths between 1 and 11 months, per live birth: used for the measles targets.
Survey year	The year of the most recent survey	Sets the year of the survey comparison and the reference year of the population-growth options.

Country-specific values are preferred where available. The survey values are the national parameters of Analysis setup (p. 54): a new dataset starts with the survey's rates and a twinning rate of 1.5% and a pregnancy loss of 3%, and you can change them there. A regional value missing from the survey is filled with the national one.

Interpretation

- **Is the survey recent, and is its year close to the analysis years?** The derived denominators rest on the survey coverage of ANC1 and Penta1; an old survey carries old coverage into every year.
- **Are ANC1 and Penta1 reporting complete and consistent?** The derived denominators assume that most of the population attends the service and that facilities report completely. They are weakest where reporting is poor; if the results look odd, go back to Data quality (p. 57), in particular the consistency of ANC1 and Penta1 (p. 67).
- **Is the coverage of the starting service near 100%?** A denominator derived from Penta1 makes Penta1 coverage equal to the survey's Penta1 coverage by construction, near 100% where the survey coverage is high (and ANC1 likewise for an ANC1-derived denominator). Judge it on other indicators: Penta3, Measles 1 and institutional live births.
- **Do the population-growth options jump between years?** They follow the chosen DHIS2 population; a break in that population becomes a break in coverage (see the example (p. 114)).
- **Subnational denominators can be less reliable than national ones.** Read them together with district data quality.

After the assessment

All the options are compared with the survey in the selection of the best performing denominator (p. 108).

9.3 Selection of the best performing denominator

Which denominator option gives the coverage of institutional live births and penta3 closest to the survey, nationally and by region, with plausible trends.

The second part of the denominator assessment compares the denominator options with each other. Each option gives a different facility-based coverage (p. 499) for the same numerator; the survey (p. 500) coverage for the same year serves as the reference. The option that agrees best with the survey, nationally and by region, with plausible trends over time, is selected for the coverage analyses.

At a glance

- **Measures:** the absolute difference between facility-based and survey coverage of institutional live births (p. 500) and penta3 (p. 502), for each denominator option, nationally and by region.
- **Good result:** coverage close to the survey, not above 100%, with smooth, plausible trends.
- **In the app:** the **Denominator Selection** (p. 338) page, cards **National Coverage**, **Subnational Derived Coverage** and **Population Trend**; the choice is set in Choose the denominator (p. 339).

Why

The numerator of a coverage indicator comes from the facility data; how plausible the coverage is depends on the denominator. The selection of the best performing denominator method is based on a comparison of the performance of the DHIS2 projection and facility-data derived methods for two indicators: institutional live births and penta3. The gold standard is the population coverage rates from a recent survey, for a year as close as possible.

What

First, we compare the results for different denominators at the national level. For the national level, we evaluate the projection methods (DHIS2 estimate and UN estimate) and the facility data derived methods (ANC1-derived and penta1-derived, and their population-growth versions: see Facility data derived denominators (p. 103)). For the subnational level, no UN projections are available, so the UN method is left out.

Indicator	Numerator	Denominator	Interpretation
Coverage of institutional live births, for each denominator option (national and by region, survey year)	Number of institutional live births (adjusted)	Number of live births estimated with the option	Compared with the survey coverage of institutional live births
Coverage of penta3, for each denominator option (national and by region, survey year)	Number of penta3 doses (adjusted)	Number of infants eligible for penta, estimated with the option	Compared with the survey coverage of penta3
National gap	Absolute difference between survey and facility-based coverage	–	Smaller differences are considered better estimates and therefore better denominators

Median subnational gap	Median, over the regions, of the absolute difference between survey and facility-based coverage	–	The best methods have the smallest gaps with the survey results

How many denominators you choose depends on the analysis. A vaccine analysis selects one method for the denominator of all the immunization coverage indicators (the Vaxx app's **Vaccination denominator**). An RMNCAH (p. 503) analysis selects two: a maternal denominator (p. 501) for the antenatal care and delivery indicators and an immunization denominator for the immunization indicators (the RMNCAH app's **Maternal denominator** and **Vaccination denominator**). See RMNCAH and vaccine analyses (p. 258).

Formula

$$C_o = \frac{N}{D_o} \times 100$$
$$G_o = \left| C_o - C_s \right|$$
$$\widetilde{G}_o = \text{median}_r \left| C_{o,r} - C_{s,r} \right|$$

where

- N is the adjusted number from the facility data (institutional live births, or penta3 doses) in the survey year;
- D_o is the matching population estimated with denominator option o, and C_o the facility-based coverage it gives;
- C_s is the survey coverage, at national level;
- G_o is the national gap of option o; $\widetilde{G}_o$ its median gap over the regions r, with the regional facility-based and survey coverage $C_{o,r}$ and $C_{s,r}$.

The best performing option has the smallest national gap and the smallest median subnational gap. (This can also be expressed as a number of standard errors from the survey value, but this requires the standard errors from the surveys; the results will be the same.)

Worked example. The survey gives a penta3 coverage of 66.5%. With the Penta1-derived denominator the facility data give 73.5%, a gap of 7.0 points; with the ANC1-derived denominator, 56.0%, a gap of 10.5 points. The Penta1-derived denominator performs better for penta3 nationally.

Settings

Setting	Default	What it means
Vaccination denominator	Penta1-derived	The denominator used for every immunization coverage indicator. Set it on the **Denominator Selection** page.
Maternal denominator (RMNCAH)	ANC1-derived	The denominator used for the antenatal care and delivery indicators. Set it on the **Denominator Selection** page.
Admin Level	National	The level of the coverage trend chart: national, Admin 1 or Admin 2 (with a region chooser).

The survey year and survey values come from the national parameters (see Facility data derived denominators (p. 106)).

Interpretation

Examine the consistency between the survey estimates and the coverage computed from the DHIS2 data with each denominator method, by comparing three sets of graphs at national and subnational levels. Note down the best options at each step.

Comparison of national estimates

Which options give national coverage close to the survey, in the survey year? Compare the national estimate of each indicator between the routine data and the survey: the bars are the routine estimates, one for each denominator, and the horizontal line is the survey estimate. A denominator that gives coverage close to the survey, without exceeding 100 percent, is a good candidate. Note down the best denominator options (there may be more than one viable option).

Subnational comparison in the survey year

Which option agrees best with the survey across regions? In the second set of figures, one for each denominator option, the horizontal bars are the coverage estimates from the routine data for each region and the "diamond" is the survey estimate. Identify the best denominator option based on the closeness of the routine coverage to the survey coverage. Also look for estimates that are implausible: coverage above 100%, or far above it. Regions that do not match the survey regions are left out of the comparison. Note down the best denominator option.

Trends in national estimates

Are the trends plausible? Examine the trend for each denominator option, looking for consistency and overall plausibility in the historical trends: smooth trends, without breaks that the services cannot explain.

Final selection and synthesis

Combine your notes from the three comparisons above to select the best performing denominator (one or two, depending on the analysis). The results on the national gap and the median subnational gap should both be taken into account to make that choice.

Summarize your choice in a short paragraph, documenting the justification for each denominator you selected.

- The best denominator is not the one that gives the highest coverage, but the one that best agrees with (is closest to) the survey and stays stable and consistent over time.
- Only the options you can select count: the UN projections are shown for comparison but cannot be selected, and they do not exist below national level.

> A denominator derived from Penta1 makes Penta1 coverage equal to the survey's Penta1 coverage by construction, near 100% when the survey coverage is high. Judge it on Penta3 and Measles 1, and against the survey.

Example: national coverage of institutional live births, by denominator

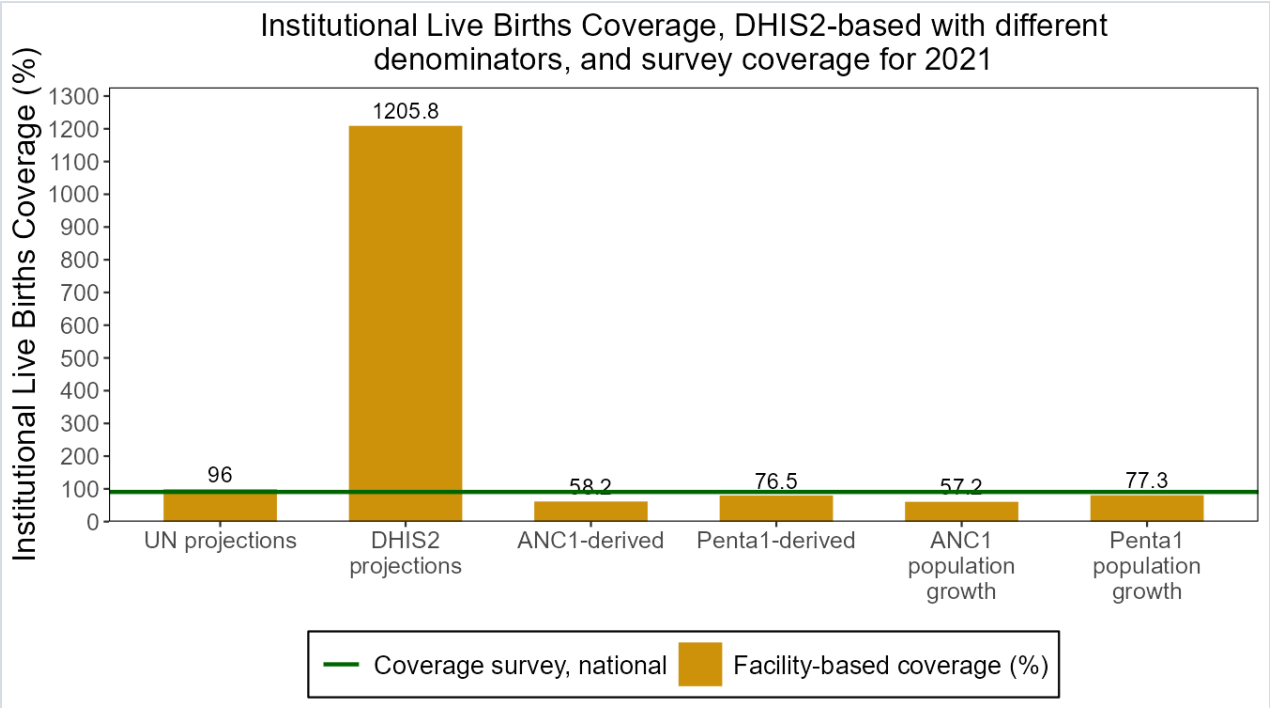


Figure 9.3 — *Institutional live births coverage in the survey year with each denominator, and the national survey coverage*

Reading the chart: each bar is the facility-based coverage of institutional live births in the survey year with one denominator option, with its value on top; the green line is the national survey coverage (90.6%).

What it shows: the UN projections give 96.0%, the closest to the survey (a gap of 5.4 points), but they cannot be selected. The DHIS2 projections give 1,206%, which is impossible: the DHIS2 live births are far too low (see the DHIS2 projections example (p. 102)). Among the options you can select, Penta1-derived (76.5%, a gap of 14.1 points) and Penta1 population growth (77.3%, 13.3 points) are closest; the ANC1 options give about 58%, more than 30 points below the survey.

Example: national coverage of penta3, by denominator

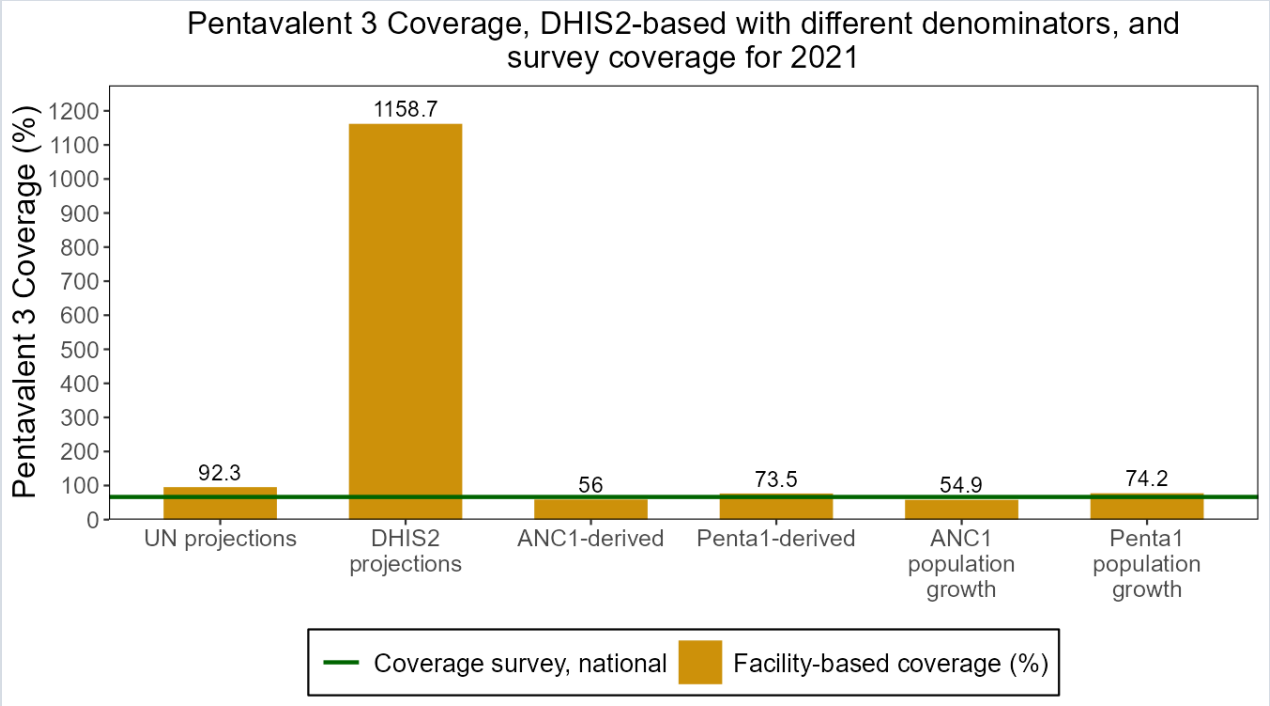


Figure 9.4 — *Penta3 coverage in the survey year with each denominator, and the national survey coverage*

Reading the chart: as above, for penta3; the green line is the survey coverage of penta3 (66.5%).

What it shows: Penta1-derived gives 73.5% (a gap of 7.0 points) and Penta1 population growth 74.2% (7.7 points); the ANC1 options give 54.9-56.0% (10.5-11.6 points below); the UN projections 92.3% (25.8 points above) and DHIS2 1,159%. For penta3 the Penta1 options perform best nationally.

Example: penta3 coverage by region and denominator

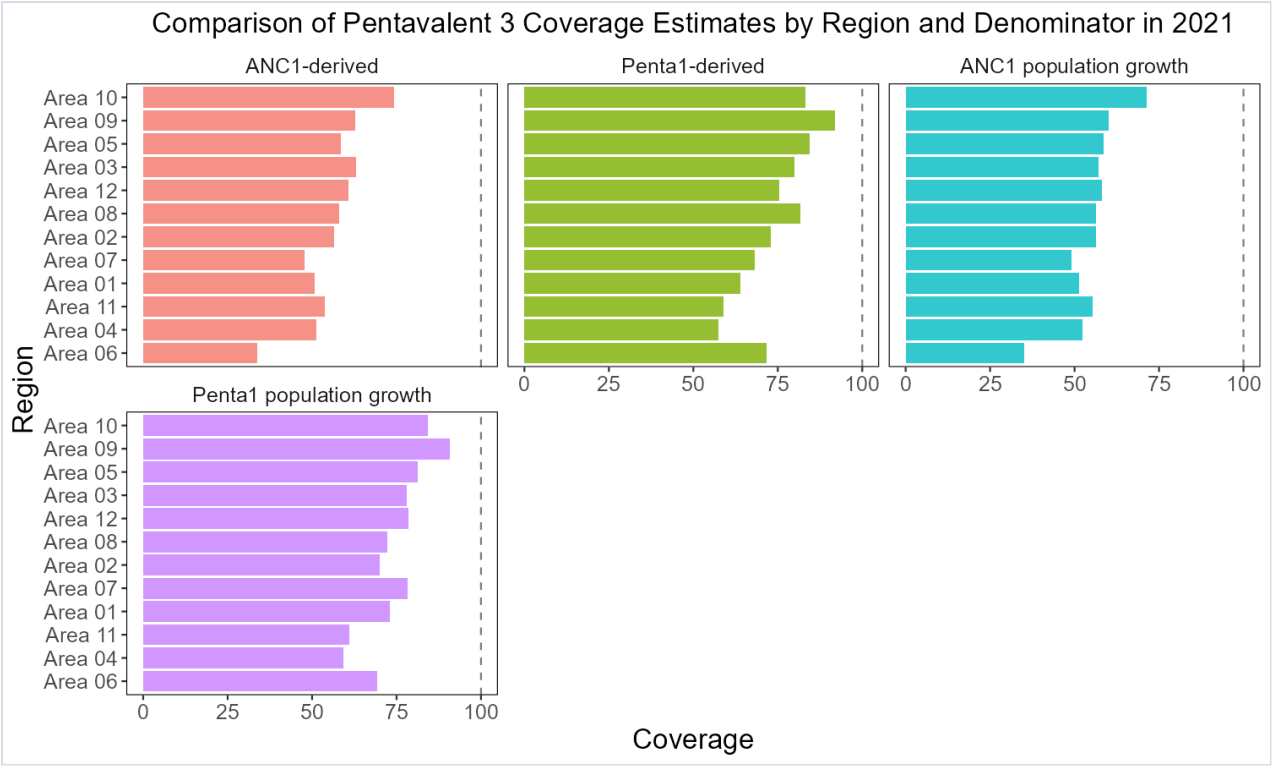


Figure 9.5 — *Penta3 coverage by Admin 1 area in the survey year, one panel per denominator*

Reading the chart: one panel per denominator option; each bar is the penta3 coverage of an Admin 1 (p. 496) area in the survey year, and the dashed line marks 100%. The DHIS2 option is left out of this chart: its coverage is above 800% in every area and would flatten the other panels. The example dataset has no regional survey values, so the survey diamonds are missing.

What it shows: no area is above 100% with any of the four options. Penta1-derived coverage ranges from 57% (Area 04) to 92% (Area 09), ANC1-derived from 34% (Area 06) to 74% (Area 10). The two disagree most in Area 06 (72% with Penta1, 34% with ANC1): consider reasons, such as incomplete ANC1 reporting in that area. With regional survey values, the median gap of each option would decide; here the comparison can only confirm that the derived options are plausible by region.

Example: national coverage over time

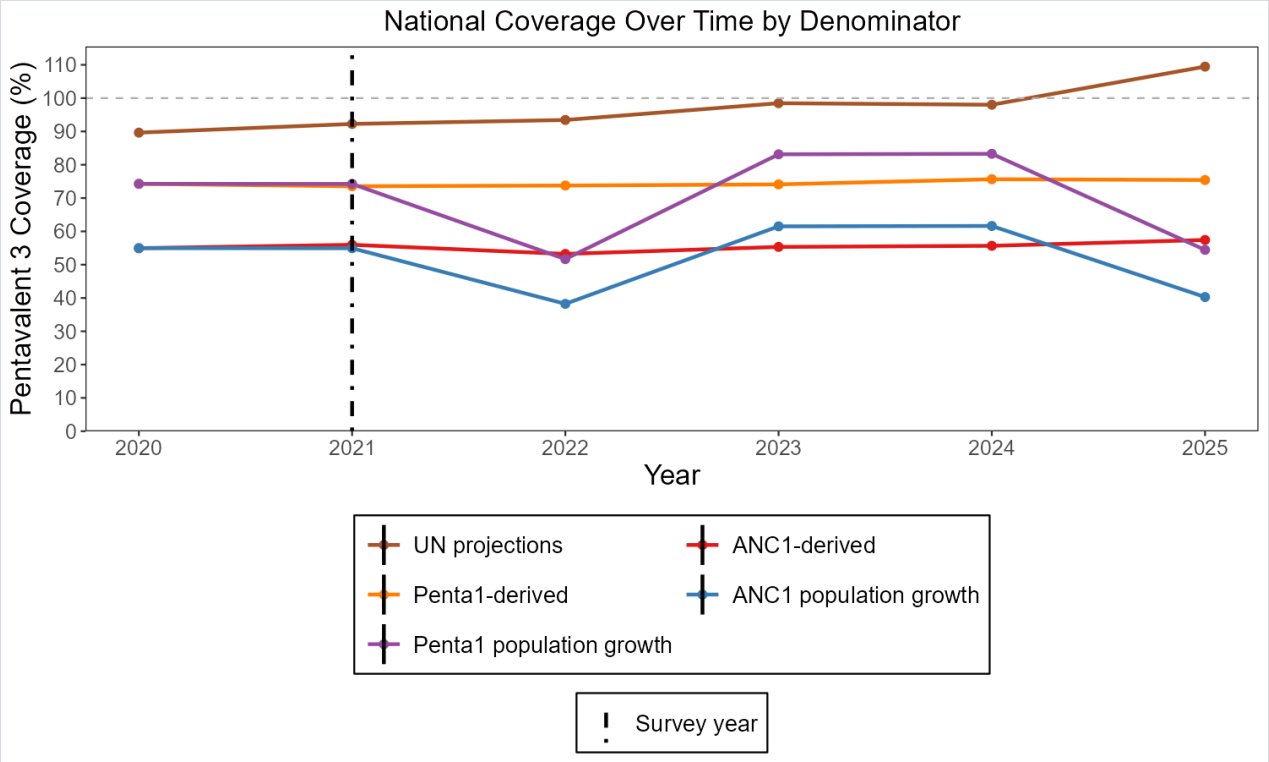


Figure 9.6 — *National penta3 coverage over time with each denominator, and the survey year*

Reading the chart: each line is the national penta3 coverage with one denominator option, by year; the dashed grey line marks 100% and the dot-dash vertical line the survey year. The DHIS2 option is left out: above 580% in the first five years, 104% in the last.

What it shows: Penta1-derived coverage is stable, between 73.5% and 75.7% every year, and ANC1-derived between 53% and 57%. The UN projections give a rising coverage that passes 100% in the last year (109%). The population-growth options start equal to the derived ones in the first year (the reference year) but then fall to 52% and 38% in the third year, rise to 83% and 62%, and fall again in the last year: they follow the breaks of the DHIS2 live births, not changes in services. Together with the national comparisons, this points to the **Penta1-derived** denominator, for the immunization indicators and, among the options that can be selected, for the maternal indicators too.

After the assessment

The selected denominator(s) are used for every coverage indicator in Coverage (p. 115), and from there in the equity, target, mortality and health system analyses. The denominator label appears on every analysis page; if you change the denominator later, review the pages that depend on it and regenerate any exported figures.

CHAPTER 10

National coverage

How annual national coverage is estimated from adjusted facility numerators and selected denominators, for antenatal care, delivery care, immunization and family planning, then compared with surveys, WUENIC and targets.

Coverage of interventions answers the core question of the analysis: *what proportion of the population needing a health service actually received it?* Annual national coverage is estimated for antenatal care, delivery care, immunization and family planning by dividing the adjusted numbers of services reported by health facilities by the best-performing population denominators. The coverage estimates are then compared with recent household survey (p. 500) results, with WUENIC (p. 505) for immunization, and with national and global targets.

It comes after the facility numbers have been checked and adjusted (Data adjustment (p. 75)) and the denominators chosen (Denominator assessment and selection (p. 95)). Its results feed the Bayesian modeling approach (p. 143) and the Subnational analysis (p. 165).

At a glance

- **Answers:** what proportion of the population needing a health service actually received it, year by year at national level, and do surveys (and WUENIC for immunization) confirm the facility-based estimates?
- **Uses:** the adjusted numerators from Data adjustment (p. 75) and the denominators chosen in Denominator assessment and selection (p. 95); household survey estimates, WUENIC for immunization and FPET (p. 499) for family planning.
- **Produces:** annual national coverage estimates and trends for each indicator, compared with surveys and targets; the maternal and newborn and the child continuum of care (p. 497) profiles; heatmaps by Admin 1 (p. 496) area.
- **In the apps: National Coverage (p. 345)** (both apps); **Continuum of Care (p. 376)** and **FPET FP Projection (p. 379)** (RMNCAH app).

Why

Coverage of interventions is a critical and direct output of health systems. Regular tracking of coverage at national and subnational levels has become the mainstay of monitoring progress in national health plans and international initiatives. Reproductive, maternal, newborn, child and adolescent health indicators (in particular vaccination indicators) with targets are the most common indicators of national health plans and global monitoring. Both health facility data and household surveys can provide coverage statistics, and an integrated analytical approach is desirable.

Many coverage indicators can be estimated in both surveys and from health facility data. Both are critical pieces of information and need to be considered in conjunction with each other. The facility data can be used to generate annual estimates of coverage; these coverage results should be compared and interpreted alongside the results from recent surveys.

What the coverage analysis covers

Part	The question it answers
Antenatal care (p. 118)	Did pregnant women make contact with services, often enough and early enough, and receive the care they need?
Delivery care (p. 122)	Were births attended in health facilities, by caesarean section where needed, and followed by postnatal care?
Immunization (p. 126)	Did infants receive their vaccines, and how many dropped out or received none?
Family planning (p. 132)	How many women use modern contraception, and how much of the demand is satisfied?
Comparison with surveys and targets (p. 134)	Do surveys and WUENIC confirm the facility trend, and is the country on track for its targets?
Continuum of care (p. 137)	Where along the care pathway are women and children lost, and in which areas?

Each coverage indicator is the number of people who received the service (the numerator, from the adjusted facility data) divided by the number who need it (the denominator, or target population (p. 503)). The parts differ in their numerators and denominators: pregnant women for antenatal care, live births for delivery care, surviving infants for immunization.

RMNCAH (p. 503) analyses estimate coverage for antenatal care, delivery (including caesarean section and low birth weight), postnatal care, a small set of vaccines and family planning (the FPET projection), and add the maternal and child continuum of care. For vaccines, the analysis estimates coverage for ANC1, institutional deliveries and live births and the routine vaccines (BCG, IPV, measles, OPV, penta, PCV, rotavirus), with dropout, zero-dose and under-vaccinated children. See RMNCAH and vaccine analyses (p. 258).

How the parts fit together

Read the indicator parts with the comparison part. Coverage estimation must account for operational realities: some indicators (institutional deliveries, immunizations) reach implausibly high levels (over 100%) because of data quality issues or duplicate recording, and then need to be reframed as an average number of services per person rather than a percentage. Family planning is the exception: its coverage comes from a specialized demographic model (FPET) that combines several survey data sources and can incorporate facility data if their quality permits, not from the facility data of the analysis.

The analysis produces trend lines, comparisons with global targets (for example 90% institutional delivery by 2025) and side-by-side facility and survey estimates to judge how credible the facility estimates are. The continuum of care then puts the indicators side by side, from pregnancy to delivery

and from birth to the infant vaccines.

Doing it in the app

① Open National Coverage

In the **National Analysis** group, open **National Coverage**. It needs adjusted data and uses the denominators chosen on Denominator Selection (in the Vaxx app, the vaccination denominator chosen in Denominators).

② Pick an indicator tab

One indicator per tab; **Custom** (**Custom Check** in the Vaxx app) covers any other indicator. In the RMNCAH app, the **FP Projection** tab shows the FPET estimates (FPET FP Projection (p. 379)).

③ Read the chart

The chart plots by year the DHIS2 estimate, the WUENIC estimate and the survey estimate with its 95% CI; the caption names the denominator the DHIS2 values use. Screen by screen: National Coverage (p. 345). How to read it: Comparison with surveys and targets (p. 135).

④ Follow the continuum of care

In the RMNCAH app, **National Analysis**, **Continuum of Care** shows the maternal and child continuum profiles and the heatmap by Admin 1 area: Continuum of Care (p. 376).

Points to consider for the interpretation

- Is facility-based coverage increasing, declining or stagnating?
- Are the last two survey estimates consistent with the facility trend?
- For immunization, do facility data, survey data and WUENIC tell a similar story?
- Is the country moving toward the relevant national or global target?
- Where along the continuum of care are women and children lost, and is this concentrated in a few areas or widespread?

Common pitfalls

- Reading coverage above 100% as strong programme performance. It is a data quality signal.
- Forgetting which denominator a result used. The chart caption names it; record it with the result.
- Zero-dose and under-vaccinated children (vaccine analyses) depend on the denominator even more than dose coverage does.
- Reading the FP Projection as a facility indicator. It is model output.
- Interpreting the continuum heatmap without first checking Data quality (p. 57) for the area.

10.1 Antenatal care

The percentage of pregnant women who made contact with antenatal care (a first visit, four or more visits, a first visit in the first trimester) and who received its main interventions.

Antenatal care (p. 496) is the care a woman receives from health services during pregnancy. Its coverage is the percentage of pregnant women in the population who made contact with antenatal care (a first visit (p. 496), four or more visits (p. 496), a first visit in the first trimester (p. 499)) or who received one of its interventions, such as malaria prevention or iron-folic acid tablets.

At a glance

- **Measures:** ANC1, ANC4 and first-trimester ANC coverage, and the coverage of the contents of antenatal care.
- **Good result:** ANC4 coverage moving toward the global target of 90%, with no implausible values over 100%.
- **In the app: National Coverage (p. 345)**, tab **ANC 4 Visits** (other ANC indicators under **Custom**).

Why

Most countries have at least one antenatal care indicator with a target in the national plan. The global ENAP/EPMM (p. 498) coverage targets for 2025 are: globally, at least 90% of pregnant women with 4 or more antenatal care visits, and 90% of countries with at least 70% coverage.

ANC1 is often considered an indicator of basic access to health services. It is high in most countries, and in many instances the numbers of ANC1 visits in the routine health facility data can provide a better denominator for the antenatal care and delivery indicators than population projections (see Denominator assessment and selection (p. 95)).

What

There are several antenatal care indicators that capture:

- **Contact with health services during pregnancy:** ANC first visit, ANC 4 or more visits, ANC first visit in the first trimester.
- **Contents of the services provided:** intermittent preventive therapy (p. 500) (IPT2 or IPT3) against malaria, HIV testing, syphilis testing and iron-folic acid (IFA) supplementation (p. 500) (at least 90 tablets given to the pregnant woman). Some countries will not have policies for all of these diagnostic or therapeutic interventions during pregnancy (e.g., no IPT if there is no malaria risk).

Indicator	Numerator	Denominator (facility data)	Survey denominator	Interpretation
ANC1 coverage	Women with a first antenatal visit	Estimated total pregnant women in the population	Women aged 15-49 years with a live birth in the last 2 years	Basic access to health services; high in most countries

ANC4 coverage	Women with 4 or more antenatal visits	Estimated total pregnant women	Women aged 15-49 years with a live birth in the last 2 years	Continuity of antenatal care; compare with the 90% global target
First-trimester ANC coverage	Women with a first visit in the first trimester	Estimated total pregnant women	Women aged 15-49 years with a live birth in the last 2 years	Early entry into pregnancy care
IPT2, IPT3 coverage	Women receiving 2 or 3 doses of IPT	Estimated total pregnant women	Women aged 15-49 years with a live birth in the last 2 years	Malaria prevention in pregnancy, where there is malaria risk
IFA, HIV and syphilis testing coverage	Women receiving the intervention	Estimated total pregnant women	Women aged 15-49 years with a live birth in the last 2 years	Content of antenatal care

Formula

$$C = \frac{N_{adj}}{P} \times 100$$

where

- C is the coverage of the indicator in the year, in percent;
- N_{adj} is the adjusted number (p. 496) of women who received the service in the year (for example ANC4 visits), from Data adjustment (p. 75);
- P is the estimated number of pregnant women in the year, from the denominator chosen in Denominator assessment and selection (p. 95).

With the ANC1-derived denominator, the number of pregnant women comes from the facility ANC1 numbers and the survey ANC1 coverage:

$$P = \frac{N_{ANC1}}{S_{ANC1}}$$

where N_{ANC1} is the number of ANC1 visits reported in the year and S_{ANC1} the ANC1 coverage in the most recent survey (as a proportion).

Worked example: survey ANC1 coverage is 95% and there were 100,000 ANC1 visits in the year, so there were 100,000 ÷ 0.95 = **105,263** pregnant women. If 52,000 women made 4 or more visits, ANC4 coverage is 52,000 ÷ 105,263 × 100 = **49.4%**. ANC1 coverage with this denominator is 95%, the survey value, by construction.

Settings

Setting	Default	What it means
Maternal denominator	The one chosen in Denominator Selection	The estimate of pregnant women (and of deliveries and live births) used for the antenatal care and delivery indicators. The chart caption names it.
ANC4 coverage target	70%	An area reaches the ANC4 target when its coverage is at least this, on the **Coverage Target** page.

Interpretation

- **Is ANC1 high, and ANC4 much lower?** A large gap between the first and the fourth visit points to weak continuity of antenatal care; see the Continuum of care (p. 137).
- **Is ANC4 moving toward the target?** Compare the trend with the 90% global target and the national target.
- **Is a coverage implausibly high, say over 125%?** This may be because the data quality of the numerator is poor, the denominator is wrong, or the intervention is given and recorded more than once during pregnancy. An example is IFA supplementation. In that case, the computation of coverage is not useful and it is better to express it differently. For instance, if coverage is 200%, compute the average number of courses of 90 IFA tablets that a pregnant woman received (in this case 2.0 per pregnant woman in the population).
- **Does the survey agree?** Compare with the survey estimates as described in Comparison with surveys and targets (p. 134).

Example: national ANC4 coverage

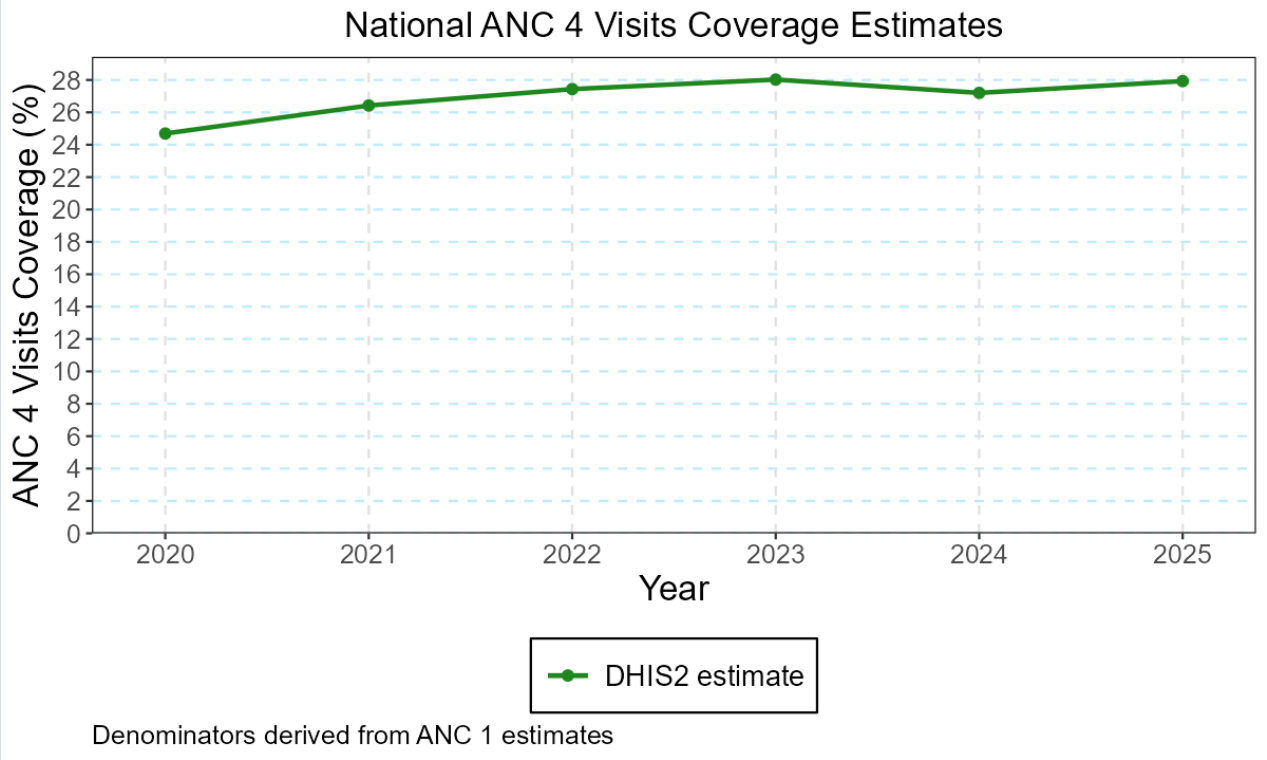


Figure 10.1 — *National ANC4 coverage estimates by year, from the facility data*

Reading the chart: each point is a year's ANC4 coverage from the adjusted facility data (the DHIS2 estimate), with the denominator named in the caption: here the pregnant women derived from ANC1. In a country dataset, the survey estimates (with their 95% confidence interval) are drawn on the same chart; the example dataset has no survey file, so only the facility line is shown.

What it shows: ANC4 coverage rose from **24.7%** in the first year to **28.0%** in the fourth and stayed near **28%** after (27.9% in the last year). In the same dataset ANC1 coverage is **78%** (see the continuum of care example (p. 139)): most women make a first visit but only about a third of them make four. ANC4 is far from the 70% and 90% targets. The survey value for ANC4 is twice as high (see the comparison with the survey (p. 136)), so incomplete recording of fourth visits should be checked before concluding that coverage is this low.

Next steps

ANC4 and first-trimester ANC are modelled in the Bayesian modeling approach (p. 143), which combines them with the survey estimates. Their distribution between areas and districts, and the share of areas that reach the ANC4 target, are in the Subnational analysis (p. 165).

10.2 Delivery care

The percentage of births that take place in health facilities or with a skilled attendant, the caesarean section rate, postnatal care within 48 hours and the percentage of babies born with a low birth weight.

Delivery care covers the care around birth: whether the birth took place in a health facility (institutional live births (p. 500) and deliveries) or with a skilled attendant (p. 504), whether a caesarean section (p. 497) was done when needed, and whether mother and baby received postnatal care (p. 503). The percentage of babies born with a low birth weight (p. 501) is read with them.

> **At a glance**
>
> - **Measures:** institutional live births and deliveries, skilled birth attendance, caesarean section, postnatal care within 48 hours and low birth weight.
> - **Good result:** institutional births moving toward the targets; caesarean sections around 10-15% of deliveries.
> - **In the app: National Coverage (p. 345)**, tabs **Institutional Live Births** and **Low Birth Weight** (other indicators under **Custom**).

Why

All countries have at least one delivery care indicator with a target in the national plan. The global Every Newborn Action Plan (ENAP) and Ending Preventable Maternal Mortality (EPMM) coverage target for 2025 is at least 90% global average coverage, and 90% of countries with at least 80% coverage of skilled birth attendance (SBA). For postnatal care (PNC) within 2 days, the global coverage target is 80%, and 90% of countries with at least 60% coverage.

Caesarean section is a life-saving intervention and an important indicator. Low birthweight is a critical indicator of neonatal health.

What

Institutional (live) birth coverage and SBA are closely related, as almost all deliveries with a skilled attendant occur in health facilities. From the analytical perspective, institutional birth coverage is preferred because it is a more objective measure and avoids issues with varying national definitions of what constitutes a skilled birth. SBA is often preferred from the healthcare perspective because it includes an element of quality; obviously, an institutional delivery without skilled attendance is not desirable, and in some countries home SBA may be part of the healthcare delivery strategy. Either indicator works well to capture delivery care.

Postnatal care within 48 hours is provided to the mother and newborn. Country systems vary in how they record the type of PNC, and household surveys are known to have data quality issues regarding PNC for the mother or the baby. Some countries use multiple definitions for the timing of the PNC visit (e.g., within the first week).

Low birthweight is most meaningful if the distinction between prematurity and small-for-gestational-age is made, but most facility reporting systems and surveys lack such granular data. All babies are supposed to be weighed immediately after birth, and the percentage of newborns weighing less than 2,500 grams is typically reported in DHIS2 (p. 498).

Indicator	Numerator	Denominator (facility data)	Survey denominator	Interpretation
Institutional live births	Live births in health facilities	Estimated live births	Live births in the last 2 years	Preferred measure of delivery care; compare with the targets
Institutional deliveries	Deliveries in health facilities	Estimated deliveries	Live births in the last 2 years	As above, counting deliveries rather than live births
Skilled birth attendance (SBA)	Births attended by skilled health personnel	Estimated live births	Live births in the last 2 years	Includes an element of quality; national definitions vary
Caesarean section rate	Caesarean sections	Estimated deliveries	Live births in the last 2 years	Population need is about 10-15% of deliveries
Postnatal care within 48 hours	Mothers and newborns with a postnatal contact within 48 hours	Estimated live births	Live births in the last 2 years	Care after birth; recording varies between countries
Low birth weight	Live births weighing less than 2,500 grams	Estimated live births	Live births in the last 2 years	Neonatal health; for sub-Saharan Africa about 13.9% (2020)

For vaccines, the analysis includes institutional deliveries and institutional live births.

Formula

$$C = \frac{N_{adj}}{P} \times 100$$

where

- C is the coverage (or, for caesarean sections and low birth weight, the percentage) in the year;
- N_{adj} is the adjusted number (p. 496) of events in the year (live births in facilities, caesarean sections, babies under 2,500 grams…);
- P is the estimated number of live births (or of deliveries, for institutional deliveries and caesarean sections) in the year, from the denominator chosen in Denominator assessment and selection (p. 95).

Worked example: with 105,263 pregnant women and a pregnancy loss (p. 503) of 3%, there are about 102,105 deliveries. If 4,600 of them were by caesarean section, the caesarean section rate is 4,600 ÷ 102,105 × 100 = **4.5%**: below 10%, a sign of unmet need for surgical delivery care.

Settings

Institutional live births coverage target	80%	An area reaches the institutional live births target when its coverage is at least this, on the **Coverage Target** page.

Interpretation

- **Are institutional births moving toward the targets?** Compare the trend with the 80% and 90% targets above.
- **Is the caesarean section rate in the range of 10-15%?** A general rule of thumb put forward by the World Health Organization (WHO) is that the population-level need for caesarean section is in the range of 10-15% of all deliveries.
  - **Below 10%:** indicates an unmet need for surgical delivery care.
  - **Over 10-15%:** implies a likely overuse of caesarean section. It may also imply a combination of unmet need among certain population groups (e.g., the poorest or rural women) alongside overuse in other subgroups (such as urban and wealthier women).
- **Is low birth weight plausible?** As general guidance for data interpretation, low birthweight prevalence in sub-Saharan Africa was estimated at 13.9% (95% credibility interval: 12.4-15.7%) for 2020. A much lower percentage suggests that babies are not all weighed or that low weights are not all recorded.
- **Is postnatal care much lower than institutional births?** That suggests missed opportunities before discharge or weak follow-up after birth; see the Continuum of care (p. 137).
- **Is a coverage over 100%?** Treat it as a data quality signal (see Comparison with surveys and targets (p. 135)).

Example: national institutional live births coverage

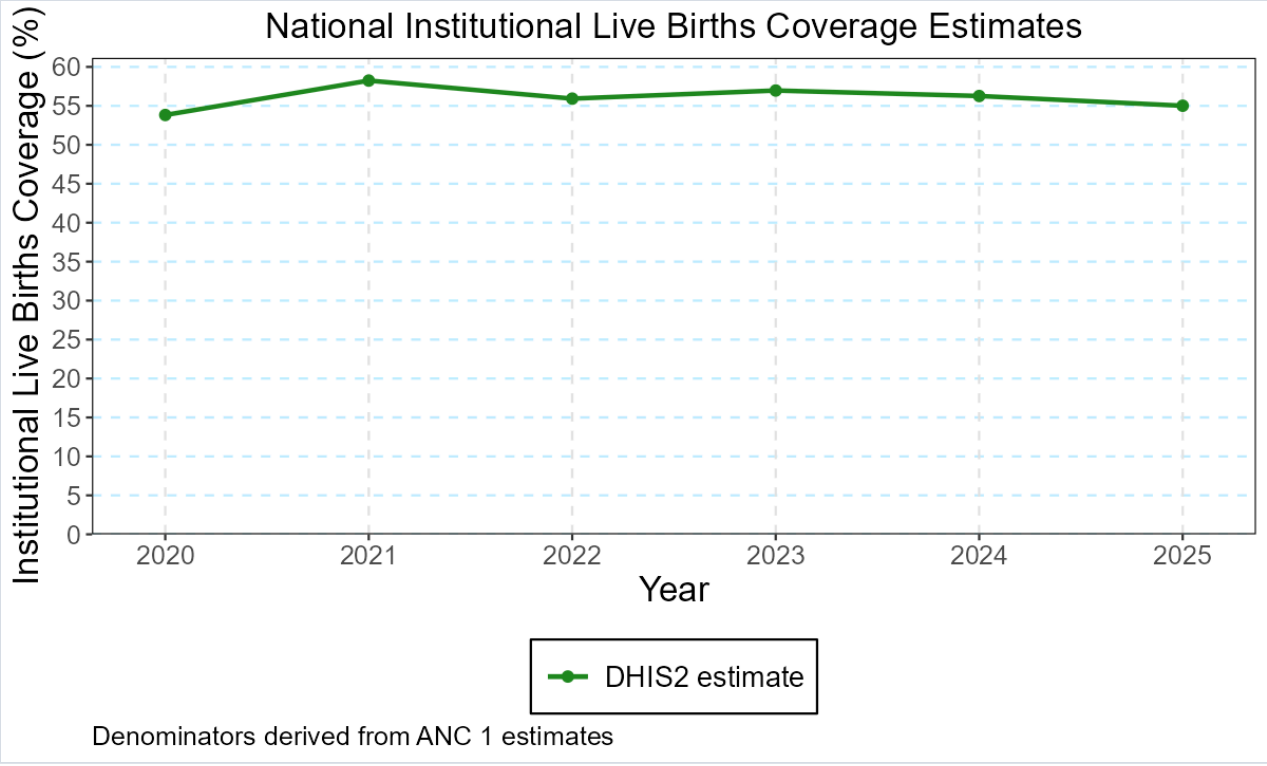


Figure 10.2 — *National institutional live births coverage estimates by year, from the facility data*

Reading the chart: each point is a year's institutional live births coverage from the adjusted facility data (the DHIS2 estimate), using the live births derived from ANC1 (caption). The example dataset has no survey file, so the survey estimates are not drawn.

What it shows: coverage was **53.8%** in the first year, peaked at **58.2%** in the second and stayed between **55% and 57%** after (55.0% in the last year). It is flat and far from the 80% target: about 45% of live births are not recorded as taking place in a health facility. The survey value is much higher (see the comparison with the survey (p. 136)), so check that private facilities and all public facilities report their deliveries.

Example: national low birth weight

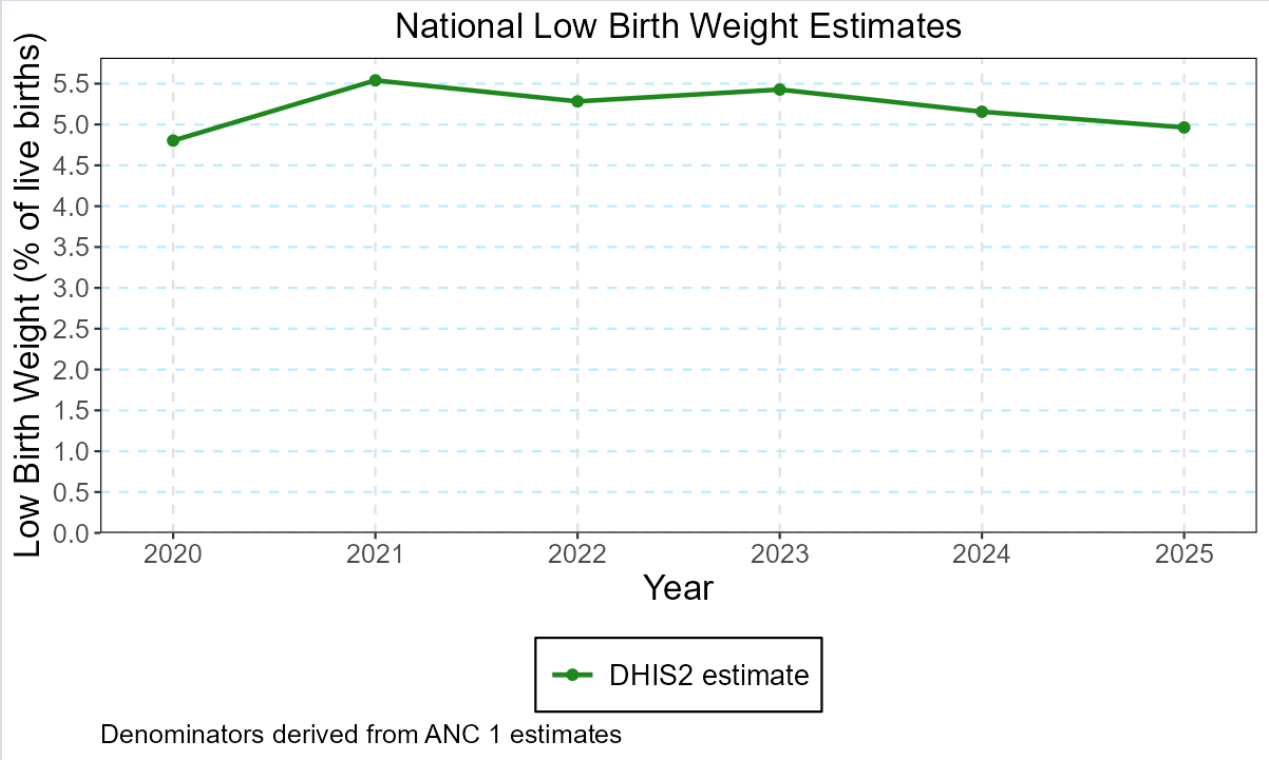


Figure 10.3 — *National percentage of live births with a low birth weight, by year*

Reading the chart: each point is the percentage of live births weighing less than 2,500 grams in a year, from the facility data.

What it shows: between **4.8% and 5.5%** of live births have a low birth weight every year (5.0% in the last year). This is well below the regional estimate of 13.9%: either many babies are not weighed, or low weights are not recorded and reported. The percentage is not plausible as a measure of low birth weight in the population; use it to follow recording, not prevalence.

Next steps

Institutional deliveries are modelled in the Bayesian modeling approach (p. 143). The share of areas and districts that reach the institutional live births target is in the Subnational analysis (p. 165).

10.3 Immunization

The percentage of eligible children who received each routine vaccine (BCG, penta, measles and the other infant vaccines), with dropout between doses and the zero-dose and under-vaccinated children.

Immunization coverage is the percentage of eligible children in the population who received a vaccine dose: BCG (p. 496) at birth, the first and third doses of pentavalent vaccine (penta1 and penta3 (p. 502)), the first and second doses of measles vaccine (p. 501) and the other routine infant vaccines. From these come the dropout (p. 498) between doses and the zero-dose and under-vaccinated (p. 505) children.

At a glance

- **Measures:** coverage of each routine vaccine, dropout between doses, zero-dose and under-vaccinated children.
- **Good result:** at least 90% coverage for the essential vaccines; dropout below 10%.
- **In the app: National Coverage (p. 345)**, tabs **Pentavalent 3** and **Measles 1** (RMNCAH app), headline indicators and **Custom Check** (Vaxx app).

Why

Immunization remains one of the most powerful interventions to enhance child survival and development. Immunization coverage indicators are included in virtually every country health sector monitoring plan. A general target is at least 90% coverage for essential vaccines given in childhood and adolescence.

What

For the national coverage analyses, the focus is on BCG, the first and third doses of penta/DTP (penta1 or DTP1 and penta3 or DTP3), and the first and second doses of measles vaccine, often given in combination with rubella vaccine. BCG and penta/DTP vaccinations are recommended to be given at birth (BCG), 6 weeks (penta1) and 14 weeks (penta3). For facility data, the number of vaccinations given to infants is used, and the denominator is the number of eligible children in the population. Survey data generally provide vaccination coverage among children 12-23 months (and may also include the age at which the vaccination was given, mostly before the first birthday).

The first dose of measles vaccine is generally recommended at age 9 months. For the facility data, recording and reporting usually separate measles doses given to children under 1 year from those given to children 1 year and older, though the quality of recording and reporting by age group may vary (there may be a tendency to record measles vaccinations given after 12 months as given to infants). Here the denominator is the children who survived the first year of life. The second dose of measles uses children aged 24-35 months as the denominator.

Indicator	Numerator	Denominator (facility data)	Survey denominator	Interpretation
BCG	BCG doses given to infants	Estimated live births	Children 12-23 months	Contact with immunization services at birth

Penta1, penta3	First and third doses of pentavalent vaccine given to infants	Surviving infants (beyond the neonatal period)	Children 12-23 months	Access to (penta1) and completion of (penta3) the infant series
Measles 1	First dose of measles vaccine given to infants	Surviving infants (beyond the postneonatal period)	Children 12-23 months	Follow-up after the early infant vaccines
Measles 2	Second dose of measles vaccine	Children aged 24-35 months	Children 24-35 months	Follow-up in the second year of life
OPV1, OPV3, PCV1, PCV3, IPV1, IPV2, rotavirus 1 and 2	Doses given to infants	Surviving infants (beyond the neonatal period)	Children 12-23 months	Coverage of the other routine vaccines (vaccine analyses)
Dropout (penta1 to penta3, penta1 to measles 1, penta3 to measles 1, measles 1 to measles 2)	Children who received the first vaccine but not the second	Children who received the first vaccine		Children lost between two contacts; target below 10%
Zero-dose children	Infants who lack the first dose of DTP-containing vaccine (penta1)	Surviving infants		Children never reached by immunization services
Under-vaccinated children	Infants missing the third dose of DTP-containing vaccine (penta3)	Surviving infants		Children who did not complete the infant series

Dropout, zero-dose and under-vaccinated children are computed for vaccines (in the Vaxx app). The guidebook also lists true zero-dose (infants who have not received any routine vaccine), under-immunized (infants missing at least one routine vaccine) and fully immunized children (infants receiving all four basic vaccines: BCG, polio, DTP and MCV); these come from surveys.

Formula

$$C = \frac{N_{adj}}{P} \times 100$$

with the population P that matches the vaccine:

$$P_{\text{penta}} = LB - ND$$
$$P_{\text{measles 1}} = LB - ND - PND$$
$$P_{\text{measles 2}} = LB - ND - 2 \times PND$$
$$\text{Dropout} = \frac{N_1 - N_2}{N_1} \times 100$$
$$\text{Zero-dose} = 100 - C_{\text{penta1}} \qquad \text{Under-vaccinated} = 100 - C_{\text{penta3}}$$

where

- N_{adj} is the adjusted number (p. 496) of doses given in the year;
- LB is the estimated number of live births in the year (the denominator for BCG), from the denominator chosen in Denominator assessment and selection (p. 95);
- ND is the number of neonatal deaths (p. 502) and PND the number of postneonatal deaths (p. 503) among them (immunization denominator (p. 500) and measles denominators (p. 501));

- N_1 and N_2 are the children who received the first and the second vaccine of the pair (for example penta1 and penta3);
- C_{penta1} and C_{penta3} are the penta1 and penta3 coverage.

Worked example: with 100,000 live births, a neonatal mortality of 2.5% and a postneonatal mortality of 2.4%, there are 100,000 − 2,500 = **97,500** surviving infants for penta, 97,500 − 2,400 = **95,100** for measles 1 and 97,500 − 4,800 = **92,700** for measles 2. If 85,000 children received penta3, penta3 coverage is 85,000 ÷ 97,500 × 100 = **87.2%**. If 92,000 received penta1, the penta1 to penta3 dropout is (92,000 − 85,000) ÷ 92,000 × 100 = **7.6%**.

Settings

Setting	Default	What it means
Vaccination denominator	The one chosen in Denominator Selection (Denominators in the Vaxx app)	The estimate of live births and surviving infants used for the vaccines. The chart caption names it.
Vaccine coverage target (national)	90%	At national level a vaccine (BCG, penta3, measles 1) reaches the target when its coverage is at least this.
Dropout rate target	10%	Dropout (penta1 to penta3, penta3 to measles 1) meets the target when it is below this, on the **Coverage Target** page (vaccine analyses).

Below national level the vaccine target is 80%: see the Subnational analysis (p. 165).

Interpretation

- **Do the essential vaccines reach 90%?** Compare BCG, penta3 and measles 1 with the target and the national plan.
- **Is there a gap between penta3 and measles 1?** It may suggest dropout after the early infant vaccine series: weak follow-up, missed opportunities, caregiver mobility or reporting gaps.
- **Is dropout below 10%?** Higher dropout means children are reached once but not brought back.
- **Which denominator do zero-dose and under-vaccinated use?** They depend on the denominator even more than dose coverage does. With the penta1-derived denominator, penta1 coverage equals the survey penta1 coverage every year by construction, so zero-dose children do not change from year to year.
- **Do facility data, survey and WUENIC agree?** Compare the three sources as described in Comparison with surveys and targets (p. 134).

Example: national penta3 coverage

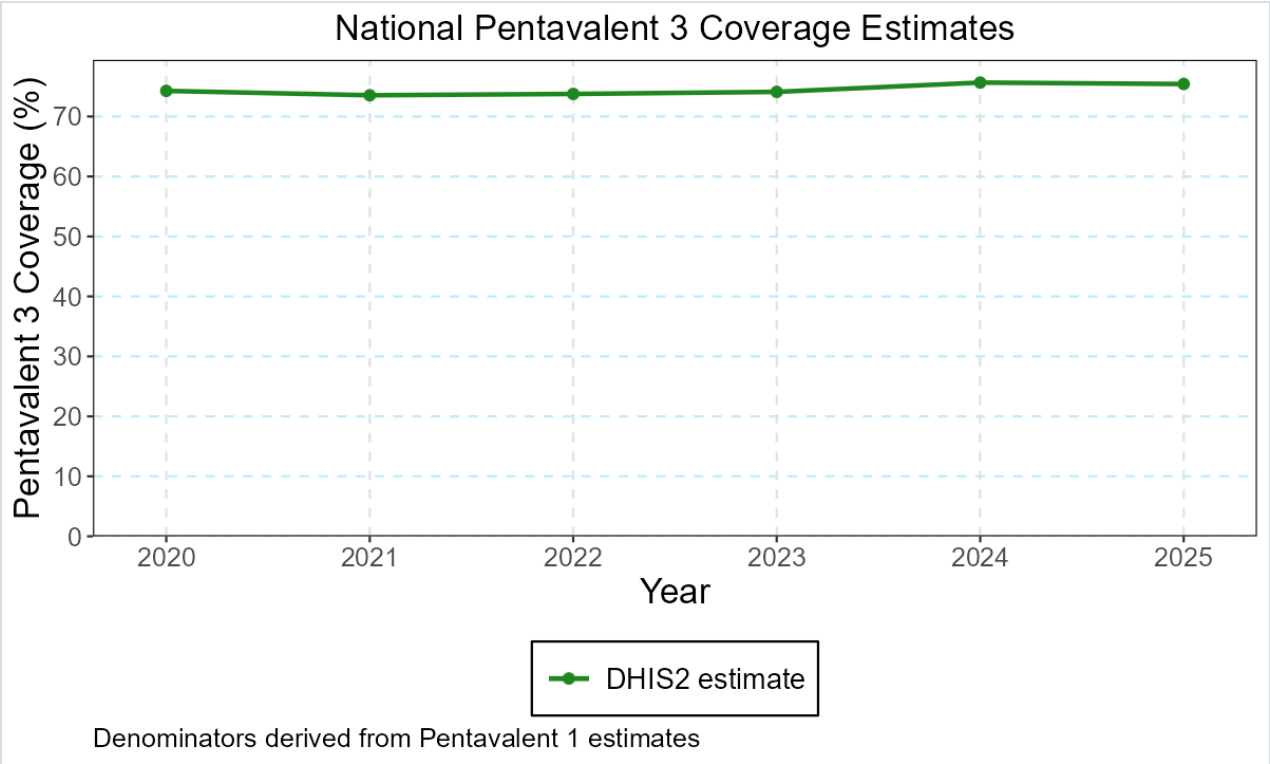


Figure 10.4 — *National penta3 coverage estimates by year, from the facility data*

Reading the chart: each point is a year's penta3 coverage from the adjusted facility data (the DHIS2 estimate), using the surviving infants derived from penta1 (caption). The example dataset has no survey or WUENIC file, so only the facility line is shown.

What it shows: penta3 coverage was stable between **73.5% and 75.7%** (75.4% in the last year): below the 90% target throughout, with no progress.

Example: national measles 1 coverage

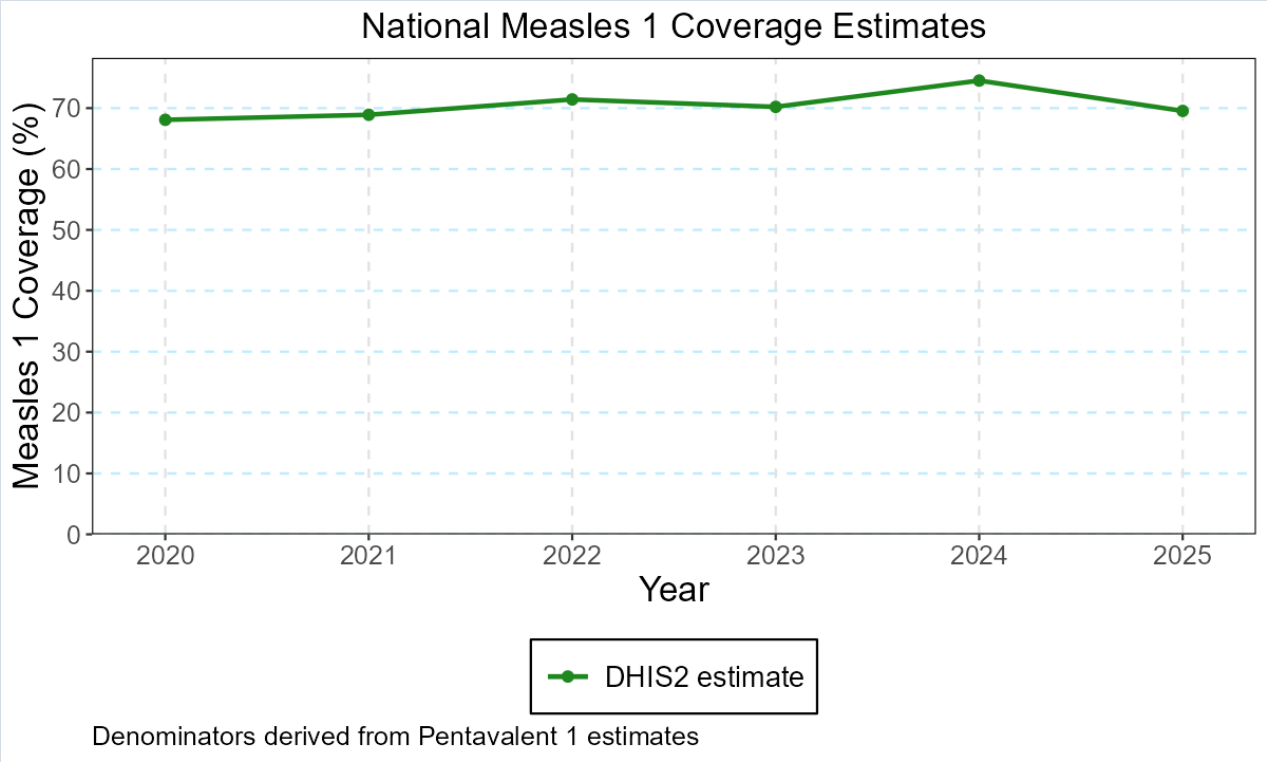


Figure 10.5 — *National measles 1 coverage estimates by year, from the facility data*

Reading the chart: as above, for the first dose of measles vaccine.

What it shows: measles 1 coverage rose from **68.1%** to **74.5%** in the fifth year, then fell back to **69.5%** in the last year. It is below penta3 every year, by 1 to 6 points (about 6 in the first and the last year): some children who complete the penta series do not come back for measles vaccination at 9 months.

Example: zero-dose, under-vaccinated and dropout

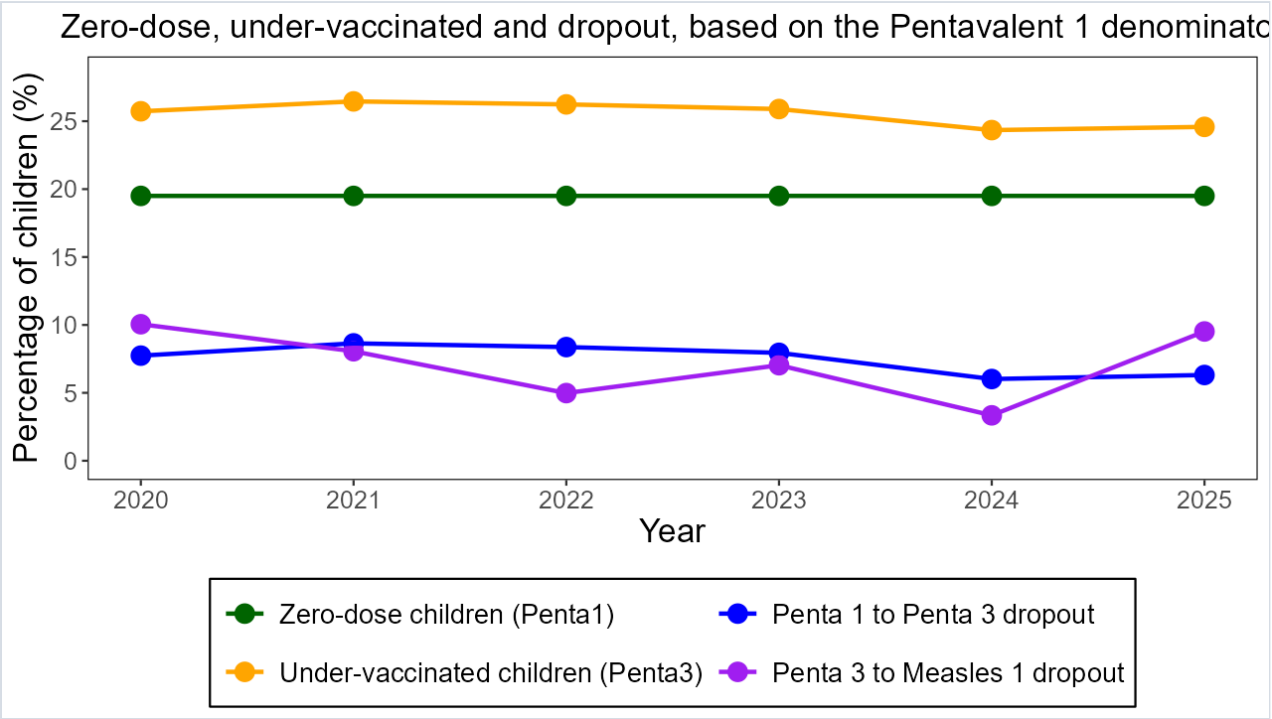


Figure 10.6 — *Zero-dose and under-vaccinated children and dropout rates by year, with the penta1-derived denominator*

Reading the chart: each line is an indicator of the vaccine analysis, by year: the percentage of infants without penta1 (zero-dose) and without penta3 (under-vaccinated), and the dropout from penta1 to penta3 and from penta3 to measles 1. Lower is better.

What it shows: zero-dose children are **19.5%** every year: with the penta1-derived denominator, penta1 coverage is the survey value (80.5%) every year, so this line only repeats the survey. Under-vaccinated children are **24% to 27%**. Penta1 to penta3 dropout is **6% to 9%**, below the 10% target. Penta3 to measles 1 dropout varies more, from **3.4%** in the fifth year to **10.0%** in the first and **9.5%** in the last year: at the limit of the target.

Next steps

Penta3 and measles 1 are modelled in the Bayesian modeling approach (p. 143). The share of areas and districts that reach the vaccine and dropout targets, and the inequality between them, are in the Subnational analysis (p. 165).

10.4 Family planning

Modern contraceptive use, unmet need for modern contraceptives and demand satisfied with modern methods, estimated with the Family Planning Estimation Tool (FPET).

Family planning coverage is measured by three related indicators: the use of modern contraceptives (modern contraceptive prevalence (p. 502)), the unmet need (p. 505) for modern contraceptives, and the demand satisfied (p. 498) with modern methods. Unlike the other parts, they are not computed from the facility data of the analysis: they come from the Family Planning Estimation Tool (FPET) (p. 499).

At a glance

- **Measures:** modern contraceptive prevalence and demand satisfied with a modern method, among currently married women aged 15-49.
- **Good result:** demand satisfied rising over time.
- **In the app: FPET FP Projection (p. 379)**: the **FP Projection** tab of National Coverage and of Sub-National Coverage (RMNCAH app).

Why

Family planning coverage estimates are derived from a collaboration between Countdown to 2030 and Track20. Track20 developed FPET, an advanced model that brings together the available data on contraceptive use, so that one consistent trend is available every year even though surveys are few.

What

FPET focuses on three indicators: modern contraceptive use, unmet need for modern contraceptives, and demand satisfied with modern methods. FPET uses statistical modeling that incorporates all available survey data and can integrate facility data if their quality is sufficient. It allows various types of survey data to be combined, fitting a trend line that draws strength from multiple data points and minimizes the risks associated with directly comparing different surveys.

In the apps, the FPET estimates are built-in reference data for the country; they do not come from your facility data. The chart shows two indicators: **Prevalence of Modern Methods (mCPR)** and **Demand Satisfied with a Modern Method**.

Indicator	Numerator	Denominator	Interpretation
Modern contraceptive prevalence (mCPR)	Women using a modern method of contraception	Currently married women aged 15-49	Level of modern contraceptive use
Unmet need for modern contraceptives	Women who want to stop or delay childbearing but use no modern method	Currently married women aged 15-49	Demand not covered by services
Demand satisfied with a modern method	Women using a modern method	Women using a modern method plus women with an unmet need	The coverage indicator of family planning

Formula

The three indicators are closely related:

$$DS = \frac{U}{UN + U} \times 100$$

where

- DS is the demand satisfied with modern methods, in percent;
- U is the use of modern contraceptives (modern contraceptive prevalence);
- UN is the unmet need for modern contraceptives.

Worked example: if 30% of women use a modern method and 20% have an unmet need, demand satisfied is 30 ÷ (20 + 30) × 100 = **60%**.

Interpretation

- **Is demand satisfied rising?** It is the coverage indicator: the share of women who want to avoid pregnancy and use a modern method.
- **How to read the lines and bands:** lines are median estimates and shaded bands are 95% credible intervals.
- **Is there a survey behind each year?** No: FPET fits a trend through all available surveys, so a smooth curve does not mean a survey exists for each year.
- **Is it a facility indicator?** No. The FP Projection is model output; do not compare it with the facility-based coverage of the other parts as if it came from the same data.

The example dataset has no FPET data, so no chart is shown here; the chart is described on FPET FP Projection (p. 379).

Next steps

The FP Projection also appears on the **Sub-National Coverage** page for each area; see the Subnational analysis (p. 165).

10.5 Comparison with surveys and targets

Whether household surveys and, for immunization, WUENIC confirm the facility-based coverage trend, and whether coverage is moving toward national and global targets.

Facility-based coverage is compared with the coverage measured by recent household surveys (p. 500) and, for immunization, with WUENIC (p. 505), the WHO and UNICEF estimates of national immunization coverage. Where they agree, the facility trend is credible; where they differ, the comparison points to a numerator or denominator problem. The same chart shows how far coverage is from its targets.

At a glance

- **Compares:** the facility (DHIS2) estimate with the survey estimates and their 95% confidence interval (p. 497), and with WUENIC for vaccines.
- **Good result:** facility and survey estimates close, the facility line passing through the survey intervals.
- **In the app: National Coverage (p. 345)**, every indicator tab.

Why

Many coverage indicators can be estimated in both surveys and from health facility data. Both are critical pieces of information and need to be considered in conjunction with each other: the facility data give annual estimates, the surveys a population-based benchmark every few years. Comparing them is how the analysis validates its findings and identifies systematic bias.

What

For most indicators, the surveys provide coverage estimates for the national level, with 95% confidence intervals. For most coverage indicators the survey data refer to a period before the survey: e.g., institutional birth coverage for the live births in the two years preceding the survey. This means that the midpoint of the coverage estimate lies one year before the survey.

WHO and UNICEF work with countries to produce annual estimates of immunization coverage based on all data sources. These national estimates, WUENIC, are published and available for 2020-2024. Their time trends are included in the outputs to compare the 2019-2024 annual facility data estimates with the survey results.

The national trend chart shows annual facility-based coverage, usually from 2020-2025 or the latest available period. Survey estimates are overlaid using at least the last two household surveys, with confidence intervals where available. For immunization indicators, WUENIC estimates may also be shown as an external reference.

Comparison	Facility value	Reference	Interpretation
Facility and survey coverage	DHIS2 estimate for the years of the survey	Survey estimate and its 95% confidence interval	Close: stronger confidence; a persistent gap: a numerator or denominator problem

Facility coverage and WUENIC (vaccines)	DHIS2 estimate by year	WUENIC estimate by year	Lines that track each other support the chosen denominator
Coverage and targets	Latest DHIS2 estimate and trend	National and global targets (below)	Is the country moving toward the target?

Targets. A general target is at least 90% coverage for the essential vaccines given in childhood and adolescence. The global ENAP and EPMM targets for 2025 are at least 90% of pregnant women with 4 or more antenatal care visits (and 90% of countries with at least 70% coverage), at least 90% skilled birth attendance (90% of countries with at least 80%) and 80% postnatal care within 2 days (90% of countries with at least 60%). National plans set their own targets, which the comparison should use as well.

Formula

$$G = C_F - C_S$$

where

- G is the gap between the two sources, in percentage points;
- C_F is the facility-based coverage for the year of the survey (or the year of its midpoint);
- C_S is the survey coverage.

A positive gap means the facility data give higher coverage than the survey.

Settings

Setting	Default	What it means
Survey year	The year of the latest survey, entered in Analysis setup	The year whose survey values are compared with the facility estimate, and used for the derived denominators.
High performer threshold (other indicators)	80%	An area counts as a high performer for an indicator without its own target when its coverage is at least this.

The targets of each indicator are on its part page.

Interpretation

- **Is facility-based coverage increasing, declining or stagnating?** Are the last two survey estimates consistent with the facility trend? For immunization, do facility data, survey data and WUENIC tell a similar story? Is the country moving toward the relevant national or global target?
- **Are facility and survey estimates close?** Then confidence in the trend is stronger: DHIS2 lines that track the survey and WUENIC points support the chosen denominator.
- **Are facility estimates much higher than survey estimates?** Check for denominator issues, over-reporting, or incomplete private-sector adjustment.
- **Are facility estimates much lower than survey estimates?** Check for incomplete facility reporting, missing outreach services, missing private-sector data, or extraction problems.
- **Is the gap persistent?** A persistent gap points to a denominator or numerator problem; go back to Data quality (p. 57) and Denominators (p. 95).

- **Is the survey interval wide?** Wide survey confidence intervals mean the survey point should be interpreted cautiously. Survey points are few and have confidence intervals: do not over-read small differences.
- **Is coverage above 100%?** Treat it as a data quality signal, not as strong programme performance.

Example: facility and survey coverage in the survey year

The example dataset has no survey file for the chart, but it holds the survey values entered for its latest survey (the second year of data). The facility coverage of that year, with the denominators derived from ANC1 (maternal indicators) and penta1 (vaccines):

Indicator	Facility	Survey	Gap (points)
ANC1	78.1%	78.1%	0
ANC4	26.4%	52.6%	−26.2
Institutional live births	58.2%	90.6%	−32.4
Caesarean section	5.7%	5.8%	−0.1
Low birth weight	5.5%	8.0%	−2.5
BCG	84.7%	86.4%	−1.7
Penta1	80.5%	80.5%	0
Penta3	73.5%	66.5%	+7.0
Measles 1	68.9%	53.7%	+15.2

Reading the table: each row compares the facility-based coverage of the survey year with the survey value; the gap is facility minus survey.

What it shows: ANC1 and penta1 agree exactly because the derived denominators are built from them: they cannot be used to judge agreement. The other indicators split in two. The facility data give much lower coverage than the survey for ANC4 (−26 points) and institutional live births (−32 points): check whether fourth visits and deliveries, particularly in private facilities, are reported. They give higher coverage than the survey for penta3 (+7) and measles 1 (+15): check for over-reporting of doses, or doses given to older children recorded as given to infants. BCG and caesarean sections agree within 2 points. Before concluding, compare the survey intervals: a single survey point has its own uncertainty.

After the comparison

What the comparison finds decides how the coverage results are used: a persistent gap sends the analysis back to Data adjustment (p. 75) and Denominator assessment and selection (p. 95). The Bayesian modeling approach (p. 143) combines the survey levels with the facility trend into one estimate per year, and the Subnational analysis (p. 165) measures the share of areas and districts that reach the targets.

10.6 Continuum of care

Coverage along the care pathway, from pregnancy to delivery and the postnatal period, and from birth to the infant vaccines, nationally and by Admin 1 area.

The continuum of care (p. 497) puts the coverage indicators of the other parts side by side, in the order a woman or a child meets them: first-trimester antenatal care, ANC1, ANC4, institutional live births and postnatal care for the maternal and newborn continuum; BCG, penta1, penta3 and measles 1 for the child continuum. Each step that is lower than the one before shows where women and children are lost along the pathway. The same steps are shown for each Admin 1 (p. 496) area in a heatmap.

At a glance

- **Shows:** the latest facility, survey and (for vaccines) WUENIC values for each step of the pathway, nationally; the latest facility values by Admin 1 area.
- **Good result:** high coverage at every step, with small losses between steps, in every area.
- **In the app: Continuum of Care (p. 376)** (RMNCAH app), tabs **Maternal Continuum** and **Child Continuum**.

Why

Coverage of a single indicator does not show where along the pathway women and children are lost. The maternal and newborn continuum follows the pathway from pregnancy to delivery and the immediate postnatal period; the child continuum focuses on child health contacts after birth, especially immunization, and helps assess whether children remain in contact with services after birth and have access to essential immunization services.

What

Continuum	Steps (indicators)	Interpretation
Maternal and newborn	First-trimester ANC, ANC1, ANC4, institutional live births, postnatal care within 48 hours	Where women and newborns are lost between pregnancy, delivery and the postnatal period
Child	BCG, penta1 (DTP1), penta3 (DTP3), measles 1 (MCV1)	Whether children stay in contact with immunization services after birth

Typical maternal indicators are first-trimester ANC, ANC4+, institutional live births and postnatal care within 48 hours. Priority child indicators are DTP3/penta3 and MCV1/measles 1; BCG, DTP1, MCV2, PCV3, rotavirus, full immunization, zero-dose, ORS and pneumonia care-seeking can be added where available.

Each step is the coverage of its part, with the same numerator and denominator: see Antenatal care (p. 118), Delivery care (p. 122) and Immunization (p. 126). The national profile shows the latest facility value, the latest survey value and, where relevant, the WUENIC value of each step. The heatmap shows the latest facility value for each area and step.

Interpretation

Along the maternal pathway:

- **Is first-trimester ANC low?** That suggests delayed entry into pregnancy care.
- **Is there a gap between first-trimester ANC and ANC4+?** That suggests weak continuity of antenatal care.
- **Is ANC high but institutional live births lower?** That may indicate barriers to delivery care, such as distance, transport, cost, referral gaps, or perceived quality of care.
- **Are institutional live births high but postnatal care low?** That suggests missed opportunities before discharge or weak follow-up after birth.
- **Do facility and survey estimates agree?** If they are consistent, the interpretation is stronger. If they differ, check indicator definitions, reporting completeness, denominator quality, private-sector reporting, and survey timing.

Along the child pathway:

- **Is there a gap between DTP3/penta3 and MCV1?** It may suggest dropout after the early infant vaccine series. Low MCV1 compared with DTP3 may point to weak follow-up, missed opportunities, caregiver mobility, or reporting gaps.
- **Do facility data, survey estimates and WUENIC agree?** If all three sources align, confidence in the immunization coverage pattern is stronger. If facility data are higher than both survey and WUENIC, check denominators, reported doses, and possible over-reporting. If facility data are lower than survey or WUENIC, check completeness of reporting, outreach data capture, private-sector reporting, and extraction issues.

In the heatmap by area: red is lower coverage (weaker performance), yellow and orange intermediate coverage, green higher coverage (stronger performance). Read across a row to see how one area performs across the indicators, and down a column to see how one indicator varies across areas.

- **A row with many red cells** suggests broad service delivery, access, or reporting challenges in that area.
- **A column with many red cells** suggests that the indicator is a national bottleneck.
- **A mixed row** suggests that the area performs well for some services but poorly for others.
- **Green areas** may provide useful lessons, but they should still be checked for data quality.
- **Red areas** should trigger deeper review, but low values may reflect reporting or denominator problems rather than true low coverage, so check quality before acting.

Best and worst performing areas

Best and worst areas should not be classified using latest coverage alone. Interpretation should consider the latest coverage level, the change over time, reporting completeness, denominator quality, and implausible values above 100%. A useful classification is:

- **High and improving:** strong performance; potential learning area.

- **High but declining:** early warning area.
- **Low but improving:** progress area needing continued support.
- **Low and declining:** priority area for review and corrective action.

Areas with coverage above 100% should be flagged and excluded from best performance rankings unless the data are reviewed. Countries should examine whether poor performance is concentrated in a few areas or widespread across many areas.

Example: maternal and newborn continuum profile

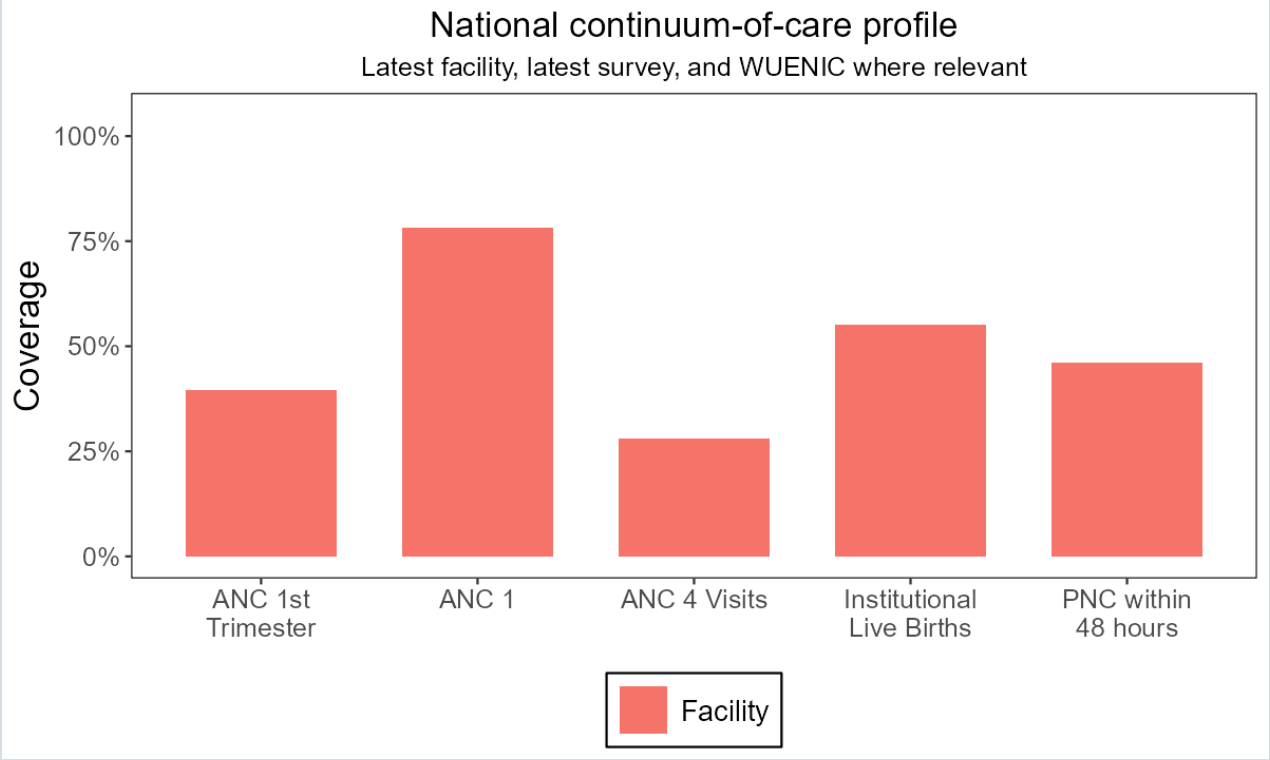


Figure 10.7 — *National maternal and newborn continuum-of-care profile, latest facility values*

Reading the chart: each bar is a step of the maternal pathway, with its latest facility value (the last year of data). In a country dataset, the latest survey value is drawn next to it; the example dataset has no survey file.

What it shows: **78%** of pregnant women make a first antenatal visit, but only **40%** make it in the first trimester and **28%** make four or more visits. **55%** of live births take place in a health facility and **46%** of mothers and newborns receive postnatal care within 48 hours. The largest loss is between ANC1 and ANC4: most women are reached once but antenatal care is not continued. Entry into care is also late.

Example: child continuum profile

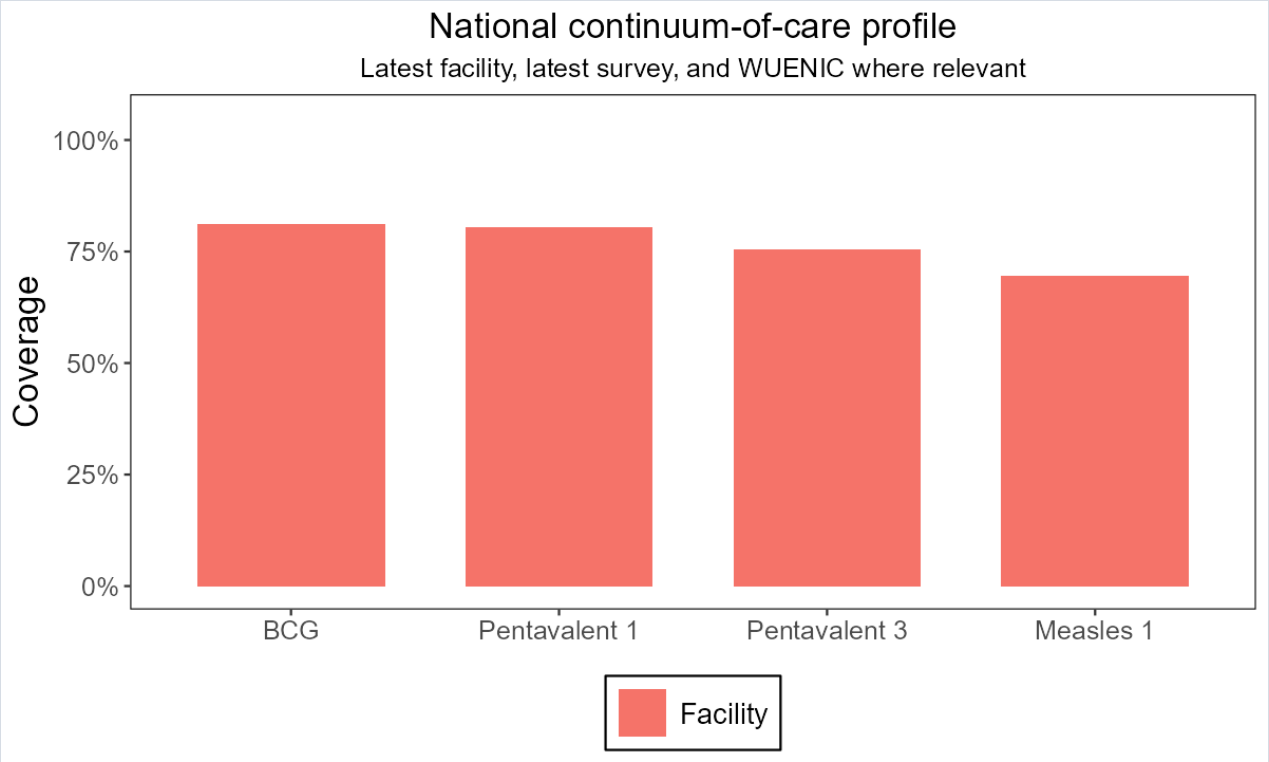


Figure 10.8 — *National child continuum-of-care profile, latest facility values*

Reading the chart: as above, for the infant vaccines.

What it shows: coverage falls gradually along the pathway: **81%** for BCG, **80%** for penta1, **75%** for penta3 and **70%** for measles 1. The largest loss is between penta3 and measles 1 (about 6 points): children who complete the penta series are not all brought back at 9 months.

Example: maternal continuum by area

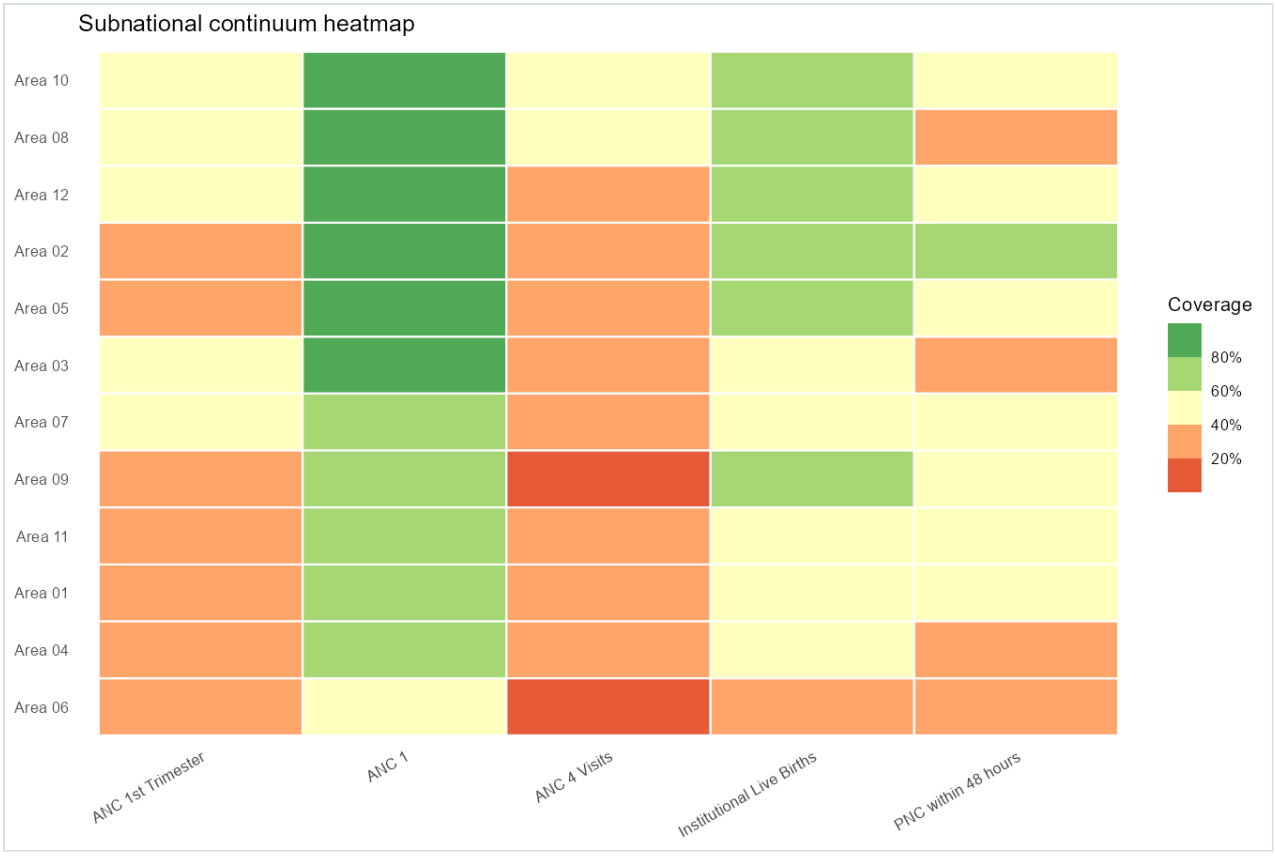


Figure 10.9 — *Heatmap of the maternal continuum steps by Admin 1 area, latest year*

Reading the chart: each row is an Admin 1 area and each column a step of the maternal pathway; the colour shows the area's latest facility coverage, from red (below 20%) through orange (20-40%) and yellow (40-60%) to light green (60-80%) and green (80% or more). The areas are sorted by their overall level.

What it shows: the ANC4 column is orange or red in every area except two (yellow, 41% and 44%): low ANC4 is a national bottleneck, not a problem of a few areas. ANC1 is green in six areas (81-97%). Area 06 has the lowest coverage at every step (ANC1 48%, ANC4 10%, institutional live births 31%): it is the priority area, after checking its reporting. Area 09 has the second lowest ANC4 (18%) although 76% of its women make a first visit.

Example: child continuum by area

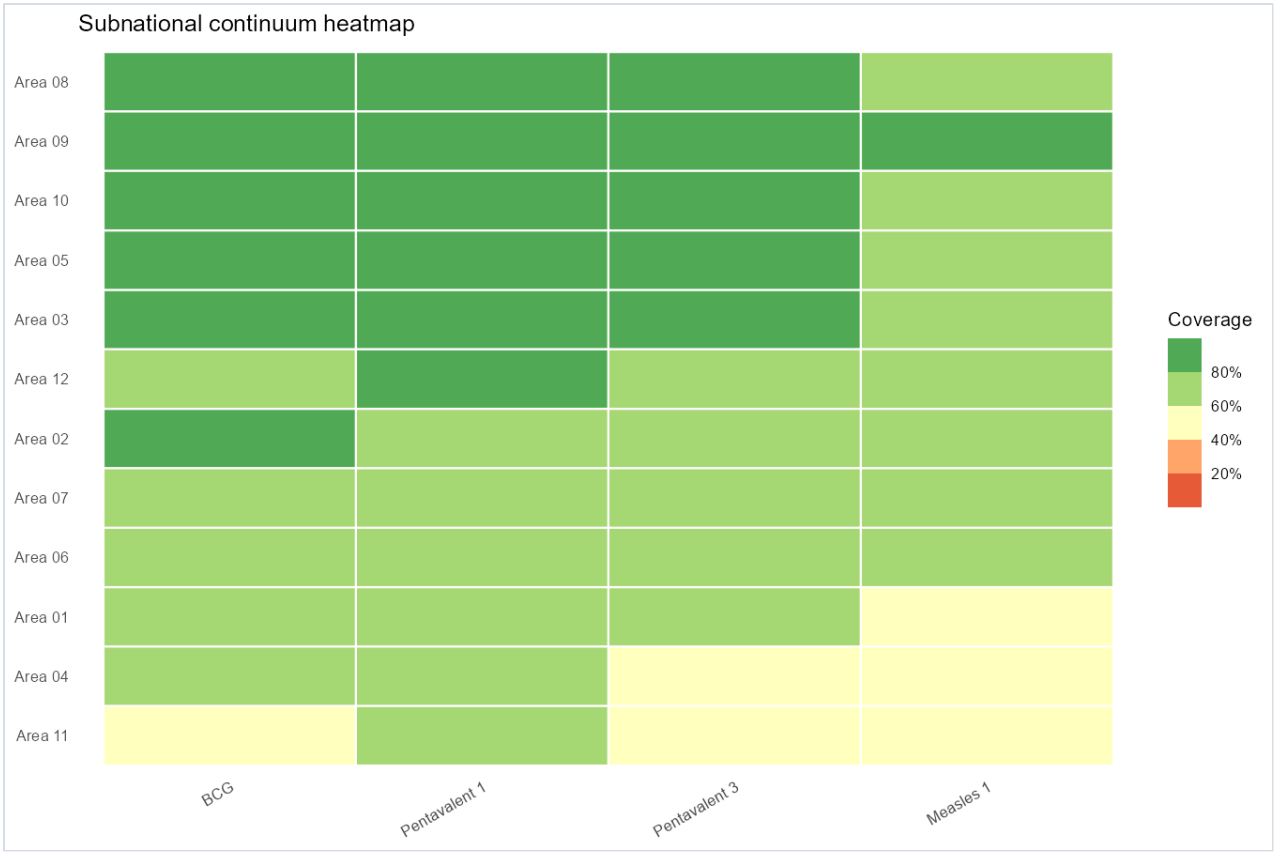


Figure 10.10 — *Heatmap of the child continuum steps by Admin 1 area, latest year*

Reading the chart: as above, for the infant vaccines.

What it shows: no cell is below 40%: vaccination reaches every area better than antenatal care. Five areas are green (80% or more) for BCG, penta1 and penta3; measles 1 is the weakest column, below 60% in Areas 01, 04 and 11. Area 08 is green, but its BCG coverage is **111%**: a value over 100% that must be flagged and reviewed before Area 08 is counted as a best performer. Areas 04 and 11 have the lowest coverage (penta3 59%, measles 1 53-56%).

Next steps

The continuum by district, the inequality between areas and the share of areas that reach the targets are in the Subnational analysis (p. 165). Check the data quality of an area (Data quality (p. 57)) before acting on its continuum.

CHAPTER 11

Bayesian modeling approach

A Bayesian model combining household surveys with facility data trends to estimate coverage every year, with uncertainty and short-term forecasts, nationally and by region.

A statistical model estimates coverage at the national level and for subnational regions up to the most recent year with data, and makes short-term forecasts for the years ahead. It combines the coverage levels measured by household surveys (p. 500) with the year-to-year trend of the routine facility data, and gives each year a best estimate and an uncertainty interval (p. 504). It is a Bayesian model (p. 496): it starts from what is known about how coverage usually changes and updates it with the country's data.

It comes after National coverage (p. 115), whose facility-based coverage it uses, and before the Equity assessment (p. 153).

At a glance

- **Answers:** what is the best estimate of coverage in each year, nationally and for subnational regions, up to the most recent year with data and for the years just ahead, and how uncertain is it?
- **Uses:** household survey data (their coverage levels, weighted by their uncertainty), the routine facility coverage from National coverage (p. 115) (only its year-to-year trend), and the overall data quality score (p. 502) from the Data quality assessment (p. 57).
- **Produces:** for ANC4, ANC in the first trimester, institutional deliveries, measles 1 and penta3, a best estimate for each year with a 95% uncertainty interval, nationally and by region, and short-term forecasts.
- **In the apps: Bayesian Analysis (p. 373)** (RMNCAH app only).

Why

Reliable information on health coverage indicators is essential for monitoring progress, informing programs, and guiding policy decisions. However, relying on survey and routine data alone is often challenging. Routine data from facilities may suffer from underreporting or inconsistent data quality, while survey data may be uncertain or not recent.

To address these challenges, statistical models can be used to produce estimates and forecasts over time, combining multiple data sources and accounting for their limitations. This results in estimates of coverage every year (between data points and past data) with an indication of how certain or uncertain we are about each year. As a result, models enable more complete, accurate, and timely assessments of health coverage, especially in data-limited settings.

The model is the tool to use when surveys are old or few and facility data are imperfect, and a single defensible trend is needed, for example for a synthesis report.

How the parts fit together

The two data sources play different roles. Surveys set the level of coverage, each point counting as much as it is reliable; routine data set how fast coverage changes between and after surveys, as much as their quality allows. The model joins them into many possible paths of coverage over time (trajectories (p. 504)) and summarises them in the estimate and its interval. Read the estimates together with the data points drawn on the same chart: where the estimate departs from a survey point or from the routine trend, the survey and routine data pages explain why.

Doing it in the app

① Open Bayesian Analysis

In the RMNCAH app, the **Bayesian Analysis** group of the sidebar has two entries: **National Analysis** and **Sub-National Analysis** (regions). The Vaxx app does not have these pages.

② Pick an indicator tab

Each tab is an indicator the model is defined for: **ANC 4 Visits**, **ANC 1st Trimester**, **Institutional Deliveries**, **Measles 1** and **Pentavalent 3** (see Bayesian Analysis (p. 373)).

③ Read the chart

The central line is the best estimate and the band is the uncertainty range. How to read it: Estimates and short-term forecasts (p. 146).

Points to consider for the interpretation

- Does the estimate follow the survey points, and where it does not, is the survey point very uncertain?
- After the most recent survey, does the routine trend show faster or slower change than the survey-based trend?
- How wide is the uncertainty band, and in which years?
- Is the routine data quality score low, so that the estimate leans on the surveys?

Common pitfalls

- Using Bayesian estimates without stating that they are modelled. Label them as such in reports.
- Reading the model as a replacement for good data: with few or old surveys and poor routine data, the band is wide.
- Expecting the estimate to match the facility-based coverage of National coverage (p. 115): the model uses the routine data's trend, not their level.

11.1 Estimates and short-term forecasts

How the model produces a best estimate of coverage and an uncertainty interval for every year, nationally and by region, with forecasts for the years just ahead.

The model looks at the available data and identifies different ways coverage could have changed over time. Each of these possible paths is a trajectory (p. 504). For each year, the best estimate is the middle value across all trajectories, and the uncertainty interval (p. 504) the range they cover once the most extreme are left out. Years after the last data point are short-term forecasts.

At a glance

- **Produces:** a best estimate and a 95% uncertainty interval for each year, nationally and for each region, with forecasts for the years ahead.
- **Good result:** a narrow band that passes close to the reliable survey points.
- **In the app: Bayesian Analysis** (p. 373), **National Analysis** and **Sub-National Analysis** (RMNCAH app).

Why

Surveys measure coverage only every few years, and the most recent one may be several years old. Estimates are needed for every year, including the years between surveys and after the last one, with an honest indication of how certain each of them is.

What

Informed by global analyses, the model captures that over time coverage usually rises from low to high levels, following an S-shaped trend, with periods of faster or slower progress along the way depending on what the data indicate. It uses survey and routine data to choose the trajectories that match what has been observed. The model also draws on trends seen in other areas with similar coverage levels.

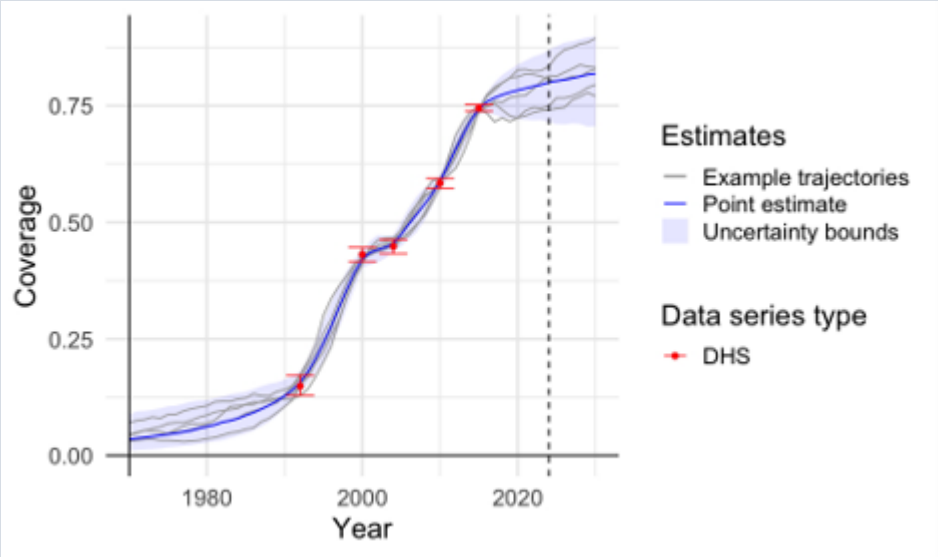


Figure 11.1 — *Possible coverage trajectories over time, with the point estimate and its uncertainty bounds, based on survey data*

The figure illustrates the method: the grey lines are example trajectories that agree with the survey points (red, with their uncertainty); the blue line is the point estimate (the middle of all trajectories each year) and the shaded band the uncertainty bounds. After the last survey (dashed line) the trajectories spread out: the forecast is less certain.

Estimates are produced for:

Indicator	What it measures
ANC4	The proportion of women who receive the recommended minimum of at least four antenatal care visits during pregnancy
ANC in the first trimester	The proportion of women whose first antenatal visit is in the first trimester
Institutional deliveries	The proportion of births that take place in a health facility
Measles 1	The proportion of children who receive the first dose of measles vaccine
Vaccination coverage (penta3)	The proportion of children who receive the recommended three doses of pentavalent vaccine

Formula

For each year t:

$$\hat{C}_t = \text{median of } C_t^{(1)}, \ \ldots, C_t^{(K)}$$
$$95\%\ \text{interval} = \left[q_{2.5\%},\ q_{97.5\%} \right]$$

where

- $\hat{C}_t$ is the best estimate of coverage in year t;
- $C_t^{(k)}$ is the coverage of trajectory k in year t, among the K trajectories drawn by the model;
- $q_{2.5\%}$ and $q_{97.5\%}$ are the values below which 2.5% and 97.5% of the trajectories lie in that year.

Settings

Setting	Default	What it means
Denominator	The one chosen in Denominator Selection	The routine coverage whose trend the model uses is computed with this denominator.
Year of the subnational comparison	2025	The **Sub-National Analysis** chart compares the regions' estimates for this year.

The model has no other settings in the app.

Interpretation

- **Reading the national chart:** survey data and routine data are drawn as dots (grey for routine data, red and blue for the surveys, with vertical bars showing how uncertain each survey point is); the line is the model's point estimate and the shaded area its 95% uncertainty interval. The title is *Bayesian Coverage Model Estimates* followed by the indicator, and the source note reads *Survey Data & DHIS2 Routine Data*.

- **Reading the subnational chart:** dots are the regions' point estimates for 2025 and horizontal lines their 95% uncertainty intervals.
- **Is the band narrow or wide?** A narrow band means the data agree; a wide band means few or conflicting data points. The band widens after the most recent survey and into the forecast years.
- **Does the estimate follow an S-shaped rise?** That is typical of coverage; a flat or falling estimate comes from the data.
- **Is the routine data quality score low?** Then the estimate leans on the surveys; see How routine data are used (p. 150).
- Do not read the model as a replacement for good data.

The example dataset used for the charts of this guide has no survey data, so the model cannot be fitted on it and no output of the app is shown here.

After the estimates

Label the estimates as modelled when you use them in reports. The survey and facility estimates they combine are compared directly in Comparison with surveys and targets (p. 134); the inequality between groups and areas is in the Equity assessment (p. 153).

11.2 How survey data are used

How the household survey points set the level of the estimated coverage, each weighted by its uncertainty, nationally and for each region.

The model takes its coverage levels from the household surveys (p. 500): if a survey indicates coverage at 60% for a particular year, the model uses that value directly. Each survey point counts according to how certain it is: a precise point pulls the estimate close to it; a very uncertain one has less influence.

> **At a glance**
>
> - **Uses:** the national (and, for regions, the regional) survey estimates of the indicator with their standard errors.
> - **Good result:** several recent survey points with narrow intervals; the estimate passes through them.
> - **In the app:** the survey points drawn on the **Bayesian Analysis** (p. 373) charts.

Why

Survey data help us understand coverage, but they come with some uncertainty. This happens because surveys only include a sample of the population, not everyone, and there can be errors during data collection.

What

To best use survey data, the model tries to strike the right balance: smoothing out random ups and downs that are probably not real, while still capturing important patterns in the data. To do this, the model looks at how uncertain each data point is:

Survey point	Effect on the estimate
High uncertainty (a wide interval)	Less impact on estimated coverage: the model gives it less weight and is more likely to smooth over it
High reliability (a narrow interval)	The estimates align closer to that value

For subnational regions, the model uses survey data from the specific region when available.

Interpretation

- **Does the estimate follow the survey points?** It usually does, unless a data point is very uncertain and does not match the overall trend.
- **Does the estimate miss a survey point?** Look at the point's vertical bar: if it is wide, the model has smoothed over it. In the example with four regions (p. 151), the most recent survey point in Region B is lower than expected and very uncertain, so the model's recent estimates do not exactly align with it.
- **How wide is the survey interval?** The survey's 95% confidence interval shows how much the point can count: compare the widths of the points on the chart before reading a change between

two surveys.

After the estimates

The years after the most recent survey are informed by the routine data: see How routine data are used (p. 150).

11.3 How routine data are used

How the facility data inform the estimated trend through their year-to-year change, not their level, and how their data quality score limits their influence.

The coverage estimates from routine facility data can be biased, so the model treats them differently from survey data. It does not rely on the coverage levels observed in routine data. Instead, it uses their **trend over time**: how coverage changes from year to year. How much that trend counts depends on the quality of the routine data, measured by the overall data quality score (p. 502).

At a glance

- **Uses:** the facility-based coverage of National coverage (p. 115) (only its year-to-year change) and the annual data quality score of the Data quality assessment (p. 57).
- **Good result:** a high data quality score, so that the routine trend informs the years after the most recent survey.
- **In the app:** the routine data points drawn on the **Bayesian Analysis** (p. 373) charts.

Why

Routine data can be biased due to issues such as incomplete reporting and data collection inconsistencies, so their coverage levels are not reliable. But they are available every year, including the years after the most recent survey, and the change they show from one year to the next is informative.

What

For surveys, the model uses the actual coverage levels. For routine data, it uses the change in coverage from year to year. If the routine data show faster or slower changes than expected, the model adjusts the estimates, but only as much as the data can be trusted.

Accounting for data quality. The model accounts for routine data quality by adjusting the uncertainty of the trend estimates. Observations with lower data quality are treated as more uncertain and given less weight in determining true rates of change. Specifically, uncertainty in the estimated rate of change increases as the Countdown mean data quality score (the annual data quality score) decreases. Consequently, lower-quality data exert less influence on the modelled rate of change.

Routine data	Used by the model	Effect
Coverage level	No	Biased levels do not move the estimate up or down
Year-to-year change	Yes	Informs how fast coverage changes, especially after the most recent survey
Data quality score of the year	Yes	A lower score makes the routine change more uncertain, so it counts less

Formula

$$\Delta_t = R_t - R_{t-1}$$

where

- Δ_t is the change in routine coverage between year $t-1$ and year t, in percentage points;
- R_t is the facility-based coverage in year t, with the denominator chosen in Denominator assessment and selection (p. 95).

Worked example: if routine data show coverage increasing from 50% in one year to 53% in the next, the model interprets this as a 3 percentage point increase, whatever the survey level is.

Interpretation

- **Does the routine trend change the estimate after the most recent survey?** Routine data help inform how quickly coverage is changing over time, particularly after the most recent survey, when no new survey data are available.
- **Is the routine level far from the survey level?** That does not affect the estimate: only the change counts. A gap between the levels is examined in Comparison with surveys and targets (p. 134).
- **Is the data quality score low?** Then the routine trend counts less and the estimate leans on the surveys; the band stays wide after the last survey.

Example: data and Bayesian estimates in four regions

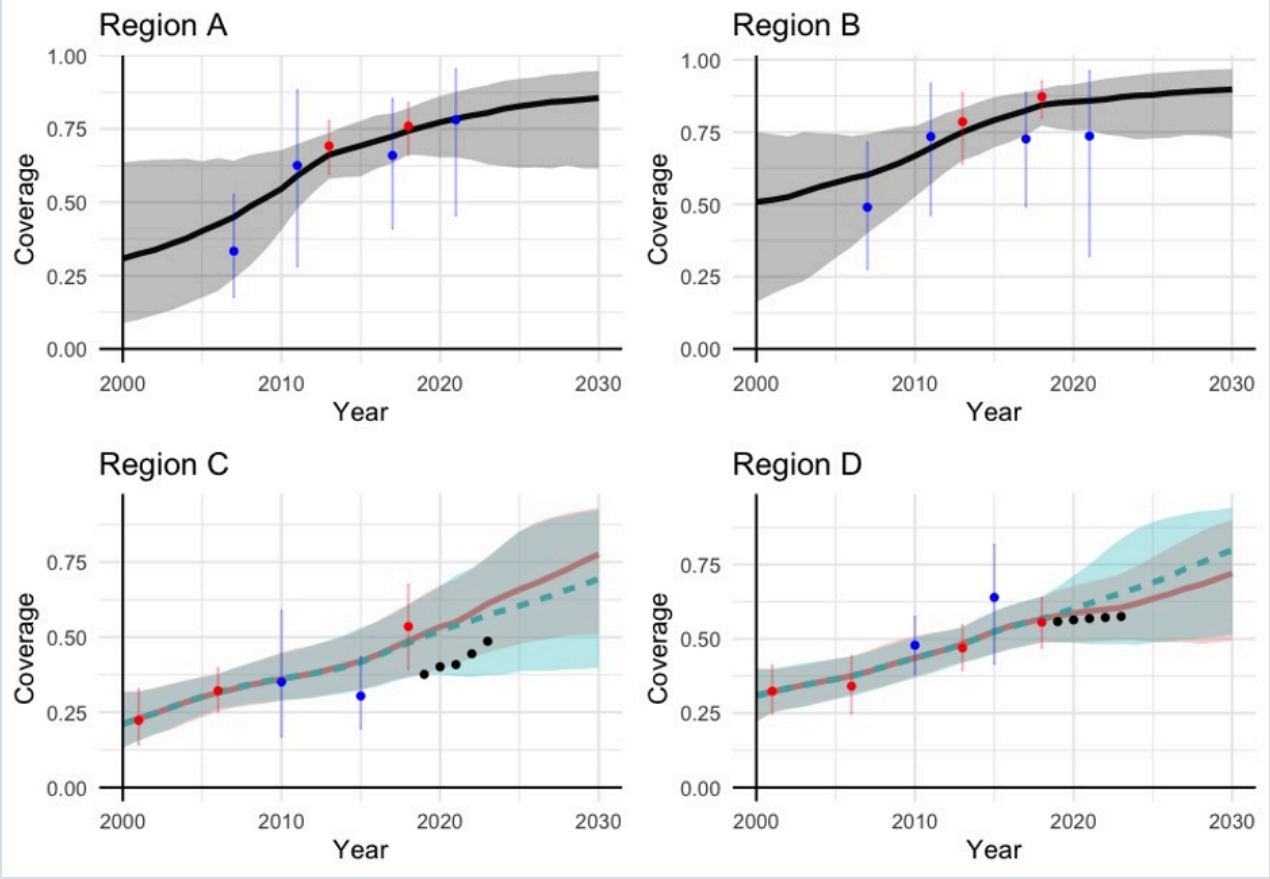


Figure 11.2 — *Survey data, routine data and Bayesian estimates for a coverage indicator in four regions*

Reading the chart: an illustration of the model's behaviour (not an output of the example dataset), for a coverage indicator in four regions. Survey data from DHS and MICS are shown in red and blue, with vertical bars indicating how uncertain each survey point is. Routine data are shown in black. The lines represent the model's point estimates with shaded areas highlighting uncertainty: in Regions C and D, red for the estimates that include the routine data and dashed green for the estimates based on the surveys only.

What it shows:

- The model's estimates usually follow the survey data, unless a data point is very uncertain and does not match the overall trend. In **Region B**, the most recent MICS survey point is lower than expected and very uncertain, so the model's recent estimates do not exactly align with it.
- **Region C:** the routine data show an increase in coverage after the last survey. As a result, the estimates that include routine data show a faster rate of change during that period compared to the estimates based only on survey data.
- **Region D:** routine data suggest that coverage has not increased recently, and the estimates that include routine data reflect that information: they rise more slowly than the survey-only estimates.
- In Region C the routine levels are well below the last survey point, yet the estimates stay at the survey level: only the routine change is used.

After the estimates

Label the estimates as modelled when you use them in reports. To raise the influence of the routine data, improve their quality: see Data quality assessment (p. 57) and Data adjustment (p. 75).

CHAPTER 12

Equity assessment (equiplots)

How equiplots of household survey data show coverage gaps by wealth quintile, maternal education and area of residence, and whether they narrow or widen.

National coverage averages mask disparities that decide whether health systems are truly leaving no one behind. The equity assessment uses household survey (p. 500) data to show the coverage of population subgroups, by household wealth, the mother's education and area of residence (urban or rural), and whether the gaps between them have narrowed over time.

It follows the national coverage analysis (National coverage (p. 115) and the Bayesian modeling approach (p. 143)). The next step, the Subnational analysis (p. 165), looks at inequality between regions and districts with the health facility data.

At a glance

- **Answers:** are the coverage gaps between population subgroups narrowing or widening, and which pattern of inequality do the wealth quintiles show?
- **Uses:** coverage by wealth quintile (p. 505), maternal education and area of residence from the household surveys loaded with the data.
- **Produces:** an equiplot (p. 499) for each indicator and dimension and, for wealth, the pattern of inequality: top (p. 504), linear (p. 501) or bottom inequality (p. 497).
- **In the apps: Equity Assessment (p. 355)** (in the Vaxx app, the **Survey Data** page under **National Inequality**).

Why

Universal health coverage implies ensuring equitable coverage of quality interventions among all populations. Monitoring is critical to draw attention to those population subgroups who are left behind. Equity analysis has multiple dimensions, including inequalities by individual or household characteristics, such as socio-economic (such as household wealth and education), demographic (such as maternal age, parity, sex of the child), and inequalities by place of residence (such as urban/rural or region/district). Some national plans have equity-specific targets for selected indicators.

The analyses consider the equity issue from two angles. The first uses the health facility data by region and by district: the Subnational analysis (p. 165). The second, this assessment, uses survey data to assess inequalities by wealth, the mother's education and area of residence. Tracking them over time shows whether equity interventions work and whether the health system is moving towards universal coverage for all population groups, not just on average.

What the assessment covers

Part	The question it answers
Inequality by wealth quintile (p. 156)	How far are the poorest behind the richest, and which pattern of inequality is it?
Inequality by maternal education (p. 160)	How far are the children of mothers with no education behind those of mothers with secondary or higher education?
Inequality by area of residence (p. 163)	How far is coverage in rural areas behind urban areas?

Each part is read on an equiplot of one indicator (for example Penta3 coverage), with one line per survey.

How the parts fit together

All three parts are read the same way. In an equiplot each line is a survey year and each dot the coverage of one group; the distance between the dots is the absolute inequality between the groups. Focus on the latest data point, the change since 2010 and whether the gap is closing, and look at the level too: a gap can close because the better-off groups lose coverage, which is not progress.

The wealth quintiles are the only dimension where the type of inequality can be assessed, because all five groups are the same size. Education and residence groups differ in size, so for them read the gap and its trend. Read the three together: the same children are often poor, rural and born to mothers with little education, so a gap in one dimension usually shows in the others.

Doing it in the app

1. **Open Equity Assessment**
   In the RMNCAH app it is **National Analysis** > **Equity Assessment**; in the Vaxx app, the **Survey Data** page under **National Inequality**. See Equity Assessment (p. 355) for the page.
2. **Choose the Equity Type**
   Wealth Quintile, **Maternal Education** or **Area** (rural, urban).
3. **Choose the indicator**
   Pick the indicator, from the tabs listed in Choose the indicator (p. 355).
4. **Read the equiplot**
   Read it as described on each part's page.

The equiplots use the survey files and the survey year loaded with the data (see Equity Assessment (p. 355)).

Points to consider for the interpretation

The interpretation should focus on whether inequalities have reduced over time and to what extent global targets for coverage have been met. Consider your audience(s) and what key messages and insights you want to get across. The following questions should guide and be answered by the interpretation:

- What is the level of inequality in the most recent data point?
- How have inequalities changed over time?
- Is there any inequality pattern that can be observed?
- What will be the best approaches to reduce inequalities?

Common pitfalls

- Equity charts come from survey files and can be several years old. Compare them with routine trends, not with a single year of DHIS2 data.
- There is no numeric threshold for what counts as a large gap: judge it against the level of coverage, the target and its change over time.
- For dropout, lower is better: the disadvantaged group has the higher value.
- In the Vaxx app, an equiplot drawn from the built-in reference data is not the country's own survey: check that the survey files were loaded.

12.1 Inequality by wealth quintile

Coverage in each household wealth quintile, the gap between the poorest and the richest, and the pattern of inequality.

In a household survey (p. 500), households are ranked by their wealth and split into five groups of the same size, the wealth quintiles (p. 505), from the poorest fifth (Q1) to the richest fifth (Q5). The equiplot (p. 499) shows the coverage of one indicator in each quintile, survey by survey: how far the poorest are behind the richest, and which pattern of inequality the quintiles show.

> **At a glance**
>
> - **Measures:** the coverage of an indicator in each wealth quintile, and the gap between the poorest and the richest quintile, in each survey.
> - **Good result:** a small gap between the poorest and the richest that narrows over time, with every quintile near the target.
> - **In the app: Equity Assessment (p. 355)**, Equity Type **Wealth Quintile**.

Why

Household surveys provide critical information on inequalities. For wealth, household wealth quintiles are used. Equiplots are used to assess whether the country has made progress since 2010 in reducing the poor-rich gap.

For wealth quintiles, it can be assessed what the type of inequality is, as all categories are of the same size. Each pattern requires a different strategy of health programming and targeting, so knowing the pattern tells a programme whom to reach.

What

Indicator	Numerator	Denominator	Interpretation
Coverage in a wealth quintile	Children (or women, or births) of the quintile who received the intervention	All children (or women, or births) of the quintile who need it	Compare the quintiles in the same survey, and the same quintile between surveys
Poor-rich gap	Coverage in the richest quintile minus coverage in the poorest quintile		The absolute inequality: has the gap been reduced since 2010?

The pattern of inequality is read from the spacing of the five quintiles:

Pattern	What the equiplot shows	Meaning
Top inequality (mass deprivation) (p. 504)	The richest are well ahead of all other wealth quintiles	Only the rich escape
Linear inequality (p. 501)	The coverage differences are equidistant	Increasing household wealth is linearly associated with higher coverage
Bottom inequality (marginalization) (p. 497)	The poorest are left behind	Marginalization of the poorest

The coverage figures are the survey's own estimates for each quintile, taken from the survey files loaded with the data.

Formula

$$G = c_5 - c_1$$

where

- G is the poor-rich gap, in percentage points;
- c_5 is the coverage in the richest quintile (Q5);
- c_1 is the coverage in the poorest quintile (Q1).

For example, Penta3 coverage of 82.7% in the richest quintile and 51.5% in the poorest gives a gap of $82.7 - 51.5 = 31.2$ percentage points.

Interpretation

- **What is the level of inequality in the most recent survey?** Read the gap between the poorest and the richest dot on the latest line. Compare every quintile with the target (for vaccines, 90%): a small gap far below the target is not success.
- **How has the gap changed over time?** Compare the lines. Has it narrowed since 2010? Check why: a gap that narrows because the poorest catch up is progress; one that narrows because the richest lose coverage is not.
- **Is there a pattern?** Only the richest ahead is top inequality; evenly spaced quintiles are linear inequality; the poorest far behind is bottom inequality. Each pattern calls for a different programme response: mass approaches that reach the whole population for top inequality, pro-poor targeting along the whole gradient for linear inequality, and action targeted at the poorest (such as education initiatives) for bottom inequality.
- **What will be the best approaches to reduce inequalities?** Use the pattern and its trend to say whom the programme must reach.

Example: Penta3 coverage by wealth quintile

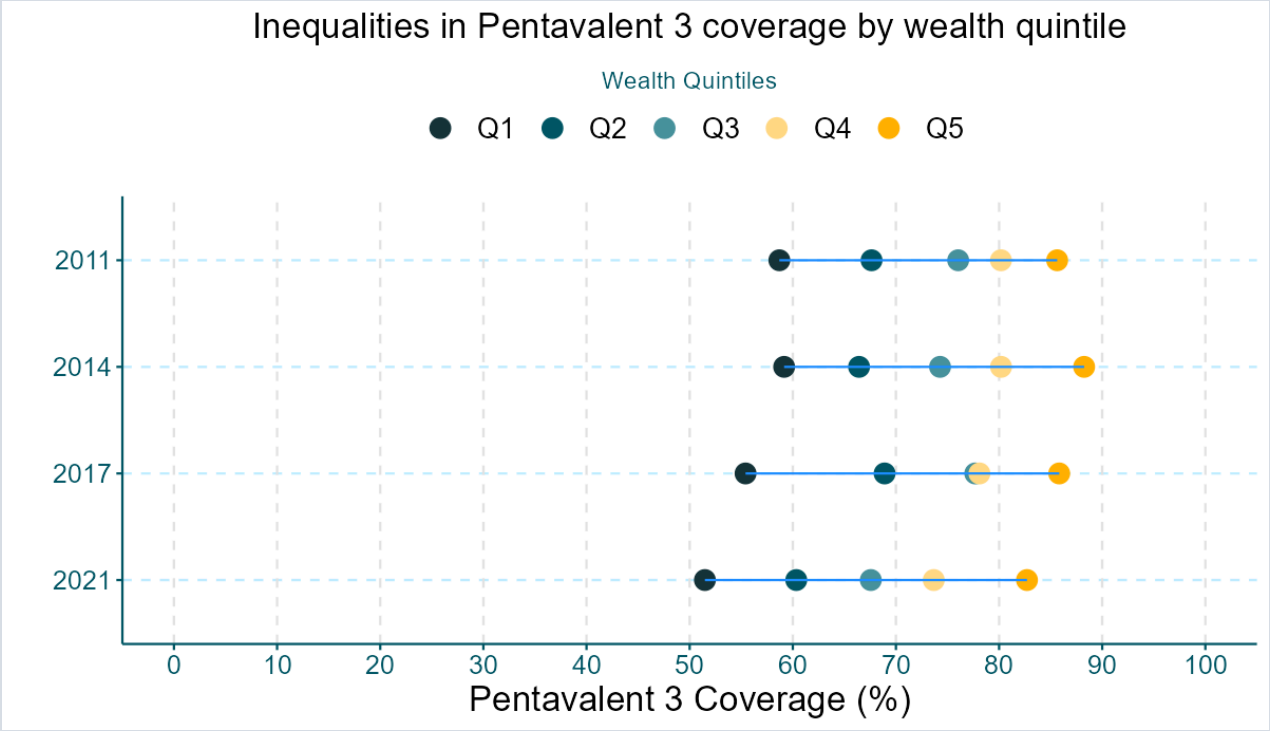


Figure 12.1 — *Equiplot of Pentavalent 3 coverage by wealth quintile in four surveys*

Reading the chart: each row is a survey (2011, 2014, 2017 and 2021, the latest at the bottom); each dot is a wealth quintile, from the poorest (Q1, darkest) to the richest (Q5, orange), placed at its Penta3 coverage on the horizontal axis. The blue line joins the poorest and the richest: its length is the poor-rich gap.

What it shows: in the latest survey coverage rose step by step with wealth, from 51.5% in the poorest quintile to 60.3%, 67.6%, 73.7% and 82.7% in the richest: steps of 6 to 9 points, a linear inequality. The poor-rich gap widened a little from survey to survey: 26.9 points in 2011, 29.0 in 2014, 30.4 in 2017 and 31.2 in 2021. Between the last two surveys coverage fell in every quintile, most in the second and middle ones (from 68.9% to 60.3% and from 77.7% to 67.6%), and no quintile reached 90%. The gap has not been reduced since 2011: raising coverage along the whole wealth gradient, with priority to the poorer quintiles, is needed.

Example: institutional live births by wealth quintile

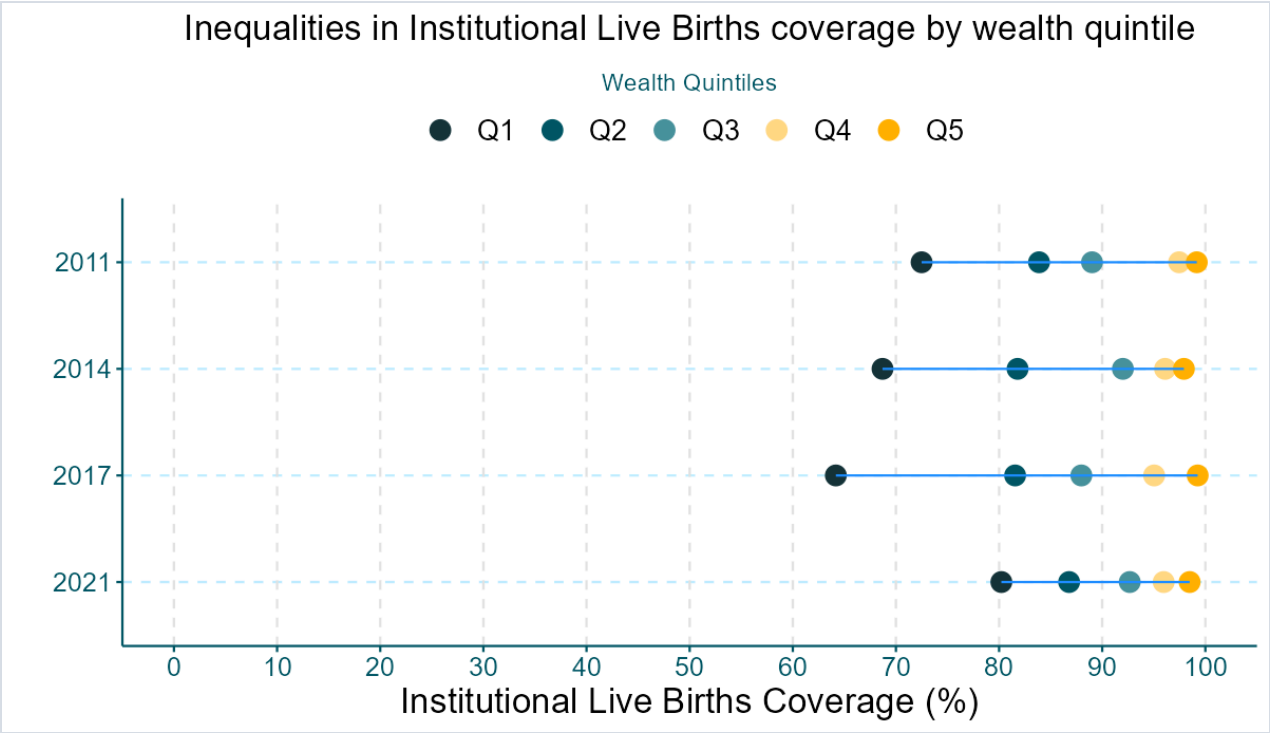


Figure 12.2 — *Equiplot of institutional live births coverage by wealth quintile in four surveys*

Reading the chart: as above, for the coverage of institutional live births.

What it shows: the richest quintiles are close to 100% (Q4 and Q5 at 95% to 99% in every survey) while the poorest quintile is far behind the rest: a bottom inequality. In 2017 the poorest were at 64.2%, 17.4 points behind the next quintile and 35.1 points behind the richest. In the latest survey the poorest rose to 80.2% and the gap fell to 18.3 points, the smallest of the four surveys, because the poorest caught up while the richest stayed at 98.5%. This is real progress; action targeted at the poorest households should continue until they reach the others.

Using the results

The equiplots and their interpretation go into the national inequality report (**Generate Report** on the RMNCAH page) and the chartbook (see Dissemination (p. 245)). Read them with the maternal education (p. 160) and area of residence (p. 163) equiplots of the same indicator, then look at inequality between regions and districts in the Subnational analysis (p. 165).

12.2 Inequality by maternal education

Coverage by the mother's level of education, and the gap between mothers with no education and those with secondary or higher education.

Household surveys (p. 500) record the education of the mother. The equiplot (p. 499) shows the coverage of one indicator in three groups, no education, primary, and secondary or higher, survey by survey: how far the children (or births) of mothers with no education are behind those of mothers with more education.

At a glance

- **Measures:** the coverage of an indicator by the mother's level of education, and the gap between no education and secondary or higher education, in each survey.
- **Good result:** a small gap that narrows over time, with every group near the target.
- **In the app: Equity Assessment** (p. 355), Equity Type **Maternal Education**.

Why

Household surveys provide critical information on inequalities. For education, the focus is on the education of the mother. Equiplots are used to assess whether the country has made progress since 2010 in reducing the gap between women with no education or low education and women with higher education.

What

Indicator	Numerator	Denominator	Interpretation
Coverage by maternal education	Children (or women, or births) of the education group who received the intervention	All children (or women, or births) of the education group who need it	Compare the groups in the same survey, and the same group between surveys
Education gap	Coverage for mothers with secondary or higher education minus coverage for mothers with no education		The absolute inequality: has the gap been reduced since 2010?

The three groups are not the same size, so the type of inequality is assessed on the wealth quintiles (p. 156). The equiplot can still show a pattern, for example the most educated far ahead of the rest.

Formula

$$G = c_S - c_N$$

where

- G is the education gap, in percentage points;
- c_S is the coverage for mothers with secondary or higher education;
- c_N is the coverage for mothers with no education.

For example, Penta3 coverage of 81.3% for mothers with secondary or higher education and 57.9% for mothers with no education gives a gap of $81.3 - 57.9 = 23.4$ percentage points.

Interpretation

- **What is the level of inequality in the most recent survey?** Read the gap between no education and secondary or higher on the latest line, and how far each group is from the target.
- **How has the gap changed over time?** Has it narrowed since 2010, and is it because the least educated catch up?
- **Where does primary education sit?** Close to secondary or higher means mothers with no education are left behind; close to no education means only the most educated are ahead.
- **What will be the best approaches to reduce inequalities?** Groups with no education often need different outreach (information, community services) from the rest.

Example: Penta3 coverage by maternal education

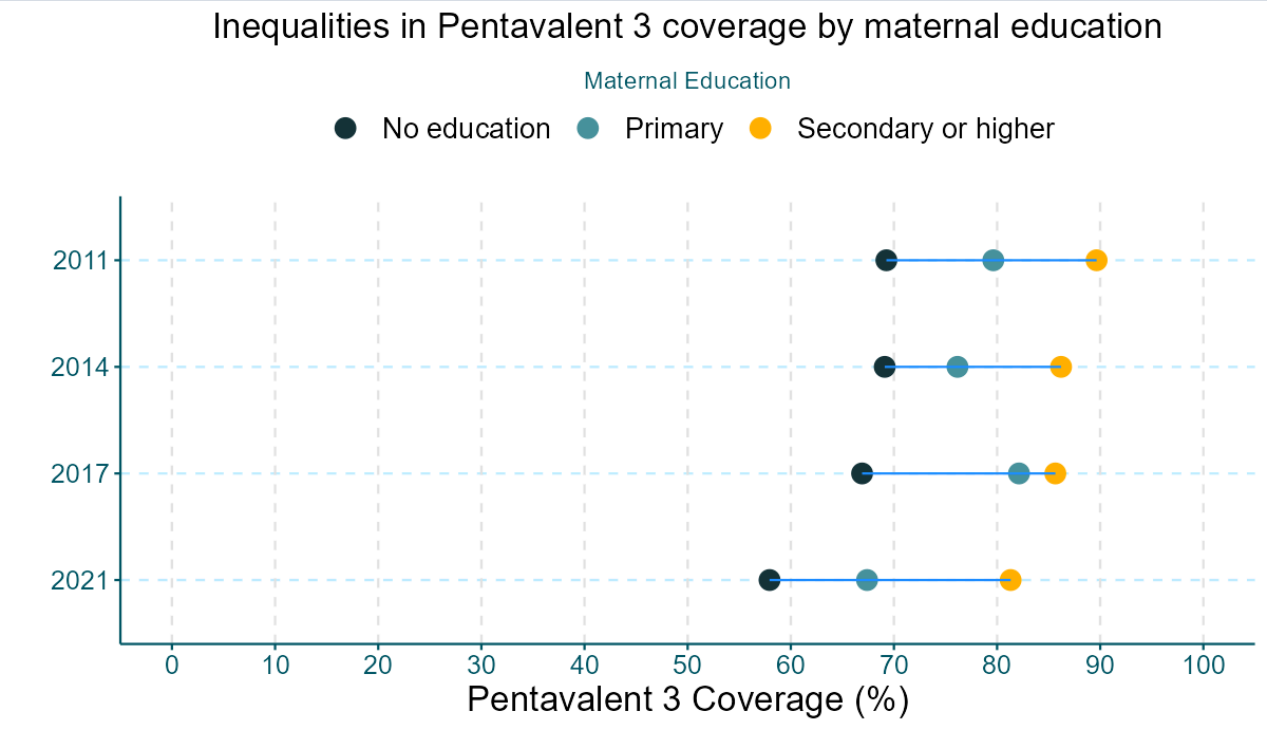


Figure 12.3 — *Equiplot of Pentavalent 3 coverage by maternal education in four surveys*

Reading the chart: each row is a survey (2011, 2014, 2017 and 2021, the latest at the bottom); each dot is an education group (no education, primary, secondary or higher) placed at its Penta3 coverage on the horizontal axis. The blue line spans the gap between the groups.

What it shows: the gap between mothers with no education and those with secondary or higher education was 20.3 points in 2011, 17.1 in 2014 and 18.8 in 2017, and widened to 23.4 points in the latest survey (57.9% against 81.3%). Coverage fell in all three groups between the last two surveys, by 9.0 points with no education, 14.7 with primary education and only 4.4 with secondary or higher. In 2017 children of mothers with primary education were close to those with secondary or higher (82.1% and 85.7%); in 2021 the most educated were far ahead of both other groups. The education gap has not been reduced: the fall since 2017 hit the less educated hardest.

Using the results

The equiplots and their interpretation go into the national inequality report (**Generate Report** on the RMNCAH page) and the chartbook (see Dissemination (p. 245)). Read them with the wealth quintile (p. 156) and area of residence (p. 163) equiplots of the same indicator, then look at inequality between regions and districts in the Subnational analysis (p. 165).

12.3 Inequality by area of residence

Coverage in urban and rural areas, and the gap between them.

Household surveys (p. 500) record whether a household lives in an urban or a rural area. The equiplot (p. 499) shows the coverage of one indicator in urban and in rural areas, survey by survey: how far rural coverage is behind urban coverage.

> **At a glance**
>
> - **Measures:** the coverage of an indicator in urban and rural areas, and the gap between them, in each survey.
> - **Good result:** a small urban-rural gap that narrows over time, with both near the target.
> - **In the app: Equity Assessment** (p. 355), Equity Type **Area**.

Why

Inequalities by place of residence are one of the dimensions of equity analysis. For area of residence, the focus is on urban and rural areas: services are often closer, and easier to reach, in towns. Differences between regions and districts are assessed with the health facility data in the Subnational analysis (p. 165).

What

Indicator	Numerator	Denominator	Interpretation
Coverage by area of residence	Children (or women, or births) of urban (or rural) households who received the intervention	All children (or women, or births) of urban (or rural) households who need it	Compare urban and rural in the same survey, and each between surveys
Urban-rural gap	Urban coverage minus rural coverage		The absolute inequality: has the gap been reduced since 2010?

Formula

$$G = c_U - c_R$$

where

- G is the urban-rural gap, in percentage points;
- c_U is the coverage in urban areas;
- c_R is the coverage in rural areas.

For example, Penta3 coverage of 75.2% in urban areas and 60.3% in rural areas gives a gap of $75.2 - 60.3 = 14.9$ percentage points.

Interpretation

- **What is the level of inequality in the most recent survey?** Read the gap between the two dots on the latest line, and how far each is from the target.

- **How has the gap changed over time?** Has it narrowed since 2010? Check whether rural coverage rose or urban coverage fell.
- **What will be the best approaches to reduce inequalities?** A large rural gap points to access (distance, outreach services); look at which regions and districts are behind in the Subnational analysis (p. 165).

Example: Penta3 coverage by area of residence

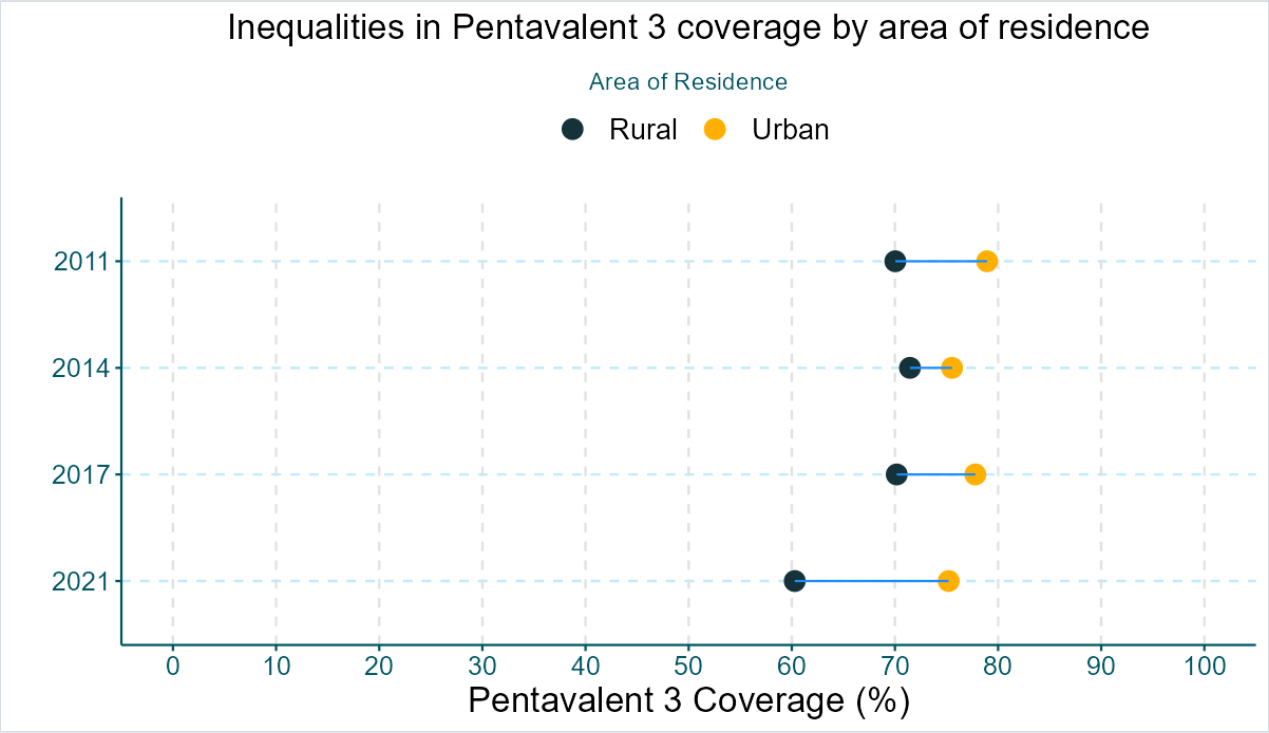


Figure 12.4 — *Equiplot of Pentavalent 3 coverage by area of residence in four surveys*

Reading the chart: each row is a survey (2011, 2014, 2017 and 2021, the latest at the bottom); the dark dot is rural and the orange dot urban coverage of Penta3. The blue line between them is the urban-rural gap.

What it shows: the gap was 8.9 points in 2011 (70.0% rural, 78.9% urban), narrowed to 4.0 points in 2014 and was 7.6 points in 2017. In the latest survey it widened to 14.9 points: rural coverage fell by 9.9 points, to 60.3%, while urban coverage fell by 2.6 points, to 75.2%. Neither reaches 90%. The fall since 2017 is mostly a rural one: rural services, and the regions where they are weakest, need attention.

Using the results

The equiplots and their interpretation go into the national inequality report (**Generate Report** on the RMNCAH page) and the chartbook (see Dissemination (p. 245)). Read them with the wealth quintile (p. 156) and maternal education (p. 160) equiplots of the same indicator, then look at inequality between regions and districts in the Subnational analysis (p. 165).

CHAPTER 13

Subnational analysis

How coverage varies across regions and districts: subnational coverage estimates and the regional one-pager, inequality between units (MADM), and the share of units meeting global coverage targets.

The subnational analysis uses the health facility data by region (Admin 1 (p. 496)) and by district (Admin 2 (p. 496)). It assesses the extent to which the country programmes have succeeded in reducing the inequalities between regions and between districts, and the extent to which districts have reached a coverage target.

It repeats the national analysis (National coverage (p. 115)) for each region and district, with the same Data adjustment (p. 75) and the best-performing denominator chosen in Denominator selection (p. 95). It follows the Equity assessment (p. 153), which looks at inequalities between population groups with survey data, and comes before Mortality (p. 179).

At a glance

- **Answers:** how does coverage vary across regions and districts? Have inequalities between subnational units reduced? What percentage of regions and districts meet each target?
- **Uses:** the adjusted health facility data by region and district, and the denominators chosen for the national analysis.
- **Produces:** coverage by region and district, the regional subnational statistical summary (one-pager), the MADM (p. 501) for each indicator and year, and the percentage of units meeting each target (p. 502); used in the reports (Reports (p. 415)).
- **In the apps: Sub-national analysis (p. 359)**; the **National Inequality (p. 352)** and **Coverage Target (p. 349)** pages of the National Analysis group.

Why

Universal health coverage implies ensuring equitable coverage of quality interventions among all populations, and inequalities by place of residence, between regions and districts, are one of its dimensions. National averages hide the regions and districts that are being left behind, and regions and districts are where programme action is usually targeted. Monitoring subnational data helps identify geographical areas where coverage is low, signalling potential inequities in health service access and prompting targeted interventions. This is particularly important for drawing attention to populations who are left behind and ensuring resources are directed where they are most needed.

Reducing geographic inequality is essential for equitable health systems and achieving the SDGs. Subnational inequalities reveal inconsistencies in service delivery and highlight systemic barriers to health care access. Tracking inequality indicators helps assess whether health systems are becoming

more equitable and whether targeted interventions are working.

How the parts fit together

Start with the coverage of each region and district: the inequality and target results are only as good as these estimates. A unit with an implausible coverage (far above 100%, or jumping between years) moves the MADM and can be counted as meeting a target it has not met, so find such units first and check their data quality and denominator.

Then read the spread between units with the MADM, and the share of units meeting each target. The two measure different things: the MADM says whether units are drawing closer to each other, the targets whether they are reaching high coverage. Units can all draw closer while staying below the target, or more units can reach the target while the weakest fall further behind.

Doing it in the app

1. **Look at the regions against the country**
   In the **National Analysis** group, **National Inequality** (in the Vaxx app, **Routine Data** under **National Inequality**) shows the coverage of every Admin 1 region by year against the national coverage, with the MADM. **Coverage Target** shows the percentage of regions or districts meeting each target. See Coverage, equity and targets (p. 349).

2. **Open Sub-National Coverage**
   In the **Sub-National Analysis** group, choose the area and the indicator: see Sub-National Coverage (p. 362).

3. **Read the inequality between units**
   Sub-National Inequality plots the regional and district results by year, with the MADM as the summary measure to assess whether inequalities have reduced. See Sub-National Inequality (p. 365).

4. **Read the coverage targets**
   Coverage Target gives the percentage of districts in a region meeting each target, by year, and lists the districts meeting it. See Coverage Target (p. 368).

5. **Produce the one-pagers**
   On the Reports (p. 415) page, the **Sub-national one-pager** gives one region on one slide; duplicate it for each region.

Points to consider for the interpretation

The interpretation should focus on whether inequalities have reduced over time and to what extent global targets for coverage have been met.

- Have the inequalities between regions, and between districts, been reduced (is the MADM falling)?
- What percentage of regions and districts have reached each target, and is it rising?

- Are there indicators with implausible results (e.g., far too high)? They should be excluded or carefully interpreted.

Common pitfalls

- Reading a single district year in isolation. Look at the trend and the neighbouring years.
- Small districts give noisy percentages. Read several years, not one.
- Comparing regions that have very different populations without checking numerators.
- Comparing Admin 1 and Admin 2 results directly: units differ in size and reliability.
- Forgetting that the denominator selected for the whole country also drives regional results. If one region looks impossible (above 100%), test another denominator before flagging the region.
- Ranking areas on the latest year only. Combine level, change over time, reporting completeness and denominator quality, as in National coverage > Best and worst subnational performance (p. 138).
- Using a map whose legend shows only a few percentage points between colours: small differences then look large (see Subnational coverage (p. 170)).

13.1 Subnational coverage

Coverage estimates for each region (Admin 1) and district (Admin 2), and the regional subnational statistical summary (one-pager).

Subnational coverage is the coverage (p. 503) of an intervention (for example ANC4, institutional live births, Penta3) estimated for each region (Admin 1 (p. 496)) and district (Admin 2 (p. 496)) from the health facility data, in the same way as the national coverage. Each region is then summarised, with its districts, in a one-page statistical summary.

At a glance

- **Measures:** the coverage of each indicator in each region and district, by year.
- **Good result:** plausible coverage (not above 100%, no jumps between years) that is close to the survey estimates for the region and rising towards the target.
- **In the app: Sub-National Coverage (p. 362)**; the one-pager on the **Reports (p. 415)** page.

Why

Subnational analysis of health intervention coverage is essential for advancing universal health coverage, which aims to ensure equitable access to quality health services for all populations. National averages often mask critical disparities across regions and districts, and the underperforming districts that pull the national figure down are where resources are most needed. The one-pager communicates the coverage situation of each region to decision-makers in an accessible, local format.

What

Indicator	Numerator	Denominator	Interpretation
Coverage in a region or district	Number of services in the unit (e.g., institutional live births, Penta3 doses), adjusted for incomplete reporting and outliers	Target population of the unit (e.g., live births, surviving infants), from the best-performing denominator	Compare with the national coverage, with the survey estimate for the region, and between years

The regional and district analyses use the best performing denominator for the coverage computations, as decided from the denominator analysis (p. 95). There will be more "noise" in the subnational data, with improbably high or low coverage rates, compared to the national analyses, and more so in the district analyses than in the regional (Admin 1) analyses. District analyses are affected by small numbers (more fluctuations, which may be random or due to data quality issues) and by the health service utilization patterns of women and children. For instance, a municipal district may get more deliveries than an adjacent rural district because of the location of the hospitals in the municipal district. Small numbers, patient movement between districts and reporting gaps all produce implausibly high or low values.

Subnational statistical summary (one-pager)

The aim is to produce a one-pager for each Admin 1 unit (generally region, province or county) that contains the regional information as well as a summary of the districts within the region: the regional subnational statistical summary. It includes:

- **Summary of key demographic information for the region and the districts:** a table with the expected number of births in the latest year according to the DHIS2 projections and according to the preferred denominators derived from the health facility data (based on ANC1 for live births, and on Penta1 for immunization indicators).
- **Line graphs with the trend in coverage of institutional deliveries and Penta3 vaccination:** the best estimates for the recent years based on the health facility data, as well as the estimates from the most recent surveys for the same indicators (with confidence intervals if possible).
- **Bar graphs that compare the latest year's coverage in the region with all other regions:** this puts the region into the lowest, middle or upper third of regions in terms of coverage, for both institutional deliveries and Penta3 vaccination.
- **A table that summarizes the coverage of institutional deliveries and Penta3 vaccination in the latest year by district:** a localized breakdown of performance across the region.

Other indicators can be used as prioritized by the country (e.g., ANC4, Measles 1). The coverage situation according to the health facility statistics for the latest year can also be summarised at regional level, and shown on a map, with the composite coverage index (p. 497): an average of seven mother and child health indicators, with equal weight given to all indicators.

Formula

$$C = \frac{N}{P} \times 100$$

$$\text{CCI} = \frac{c_{\text{ANC4}} + c_{\text{ILB}} + c_{\text{SBA}} + c_{\text{IPT2}} + c_{\text{PNC}} + c_{\text{Penta3}} + c_{\text{MCV1}}}{7}$$

where

- C is the coverage of the region or district, in percent;
- N is the adjusted number of services in the unit in the year;
- P is the target population of the unit in the year, from the chosen denominator;
- CCI is the composite coverage index of the unit;
- c_{ANC4}, c_{ILB}, c_{SBA}, c_{IPT2}, c_{PNC}, c_{Penta3} and c_{MCV1} are the coverage of four or more antenatal care visits, institutional live births, skilled attendance at birth, the second dose of intermittent preventive treatment of malaria in pregnancy, postnatal care, the third dose of pentavalent vaccine and the first dose of measles vaccine.

For example, a district with 4,500 institutional live births and 6,000 expected live births has a coverage of $4{,}500 \div 6{,}000 \times 100 = 75\%$.

Interpretation

- **Is the survey estimate for the region far from the DHIS2 line?** A survey point far from the DHIS2 line in one region often signals a denominator or reporting problem there. Check the region's reporting rate (Data quality (p. 57)) before drawing conclusions.
- **Is a district's coverage over 100%, or does it flip between years?** District denominators are the least reliable. Such a district is more likely to have a population or reporting problem than a real change: cross-check it against the district's data quality, and test another denominator before flagging it.
- **How does the region compare with the national coverage and with the other regions?** Is it in the lowest, middle or upper third, and is its gap with the country closing?
- **Before using a map, check its legend.** If the legend shows only a few percentage points between the colour gradings, the map should not be used: it can mislead the reader to assume geographical inequalities are large when they are, in fact, small. Maps help when there are many units and large differences between them.

Example: Penta3 coverage in one area

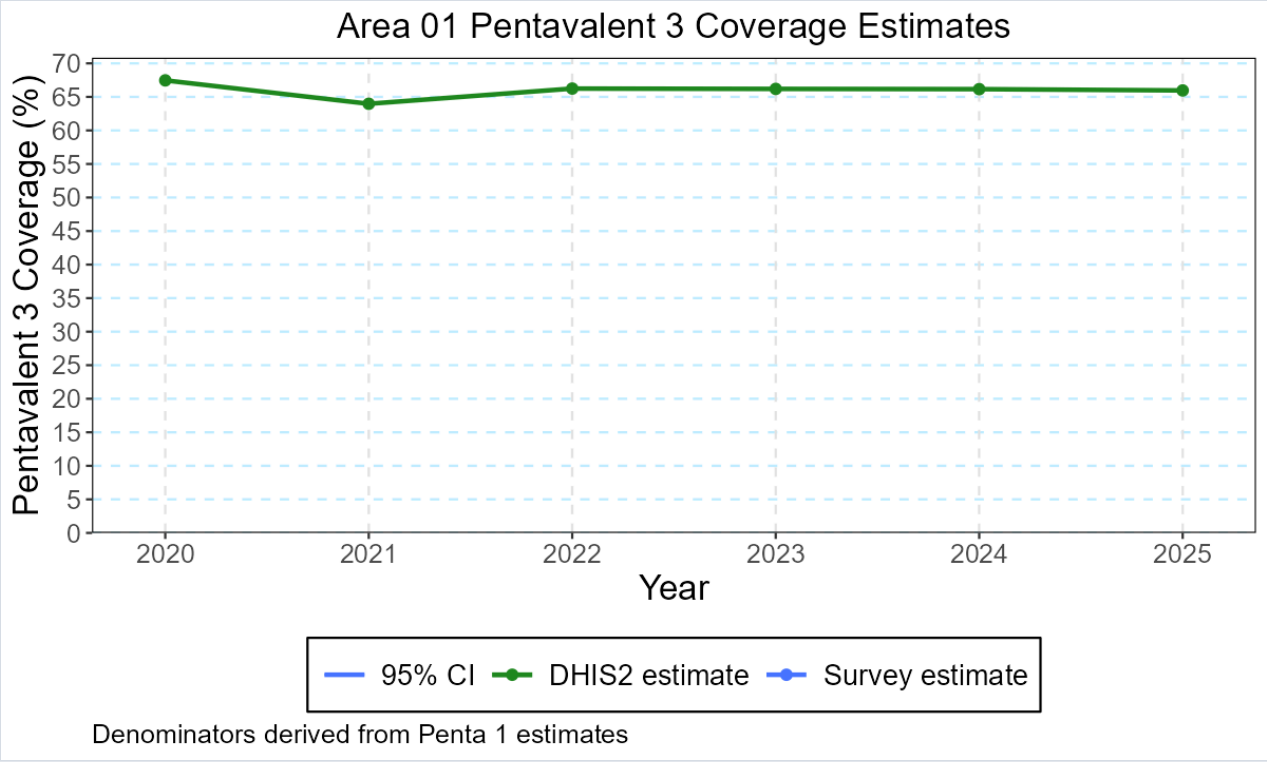


Figure 13.1 — *Penta3 coverage estimates for Area 01 by year*

Reading the chart: the green line is the coverage of Penta3 in one Admin 1 area estimated from the health facility data (DHIS2 estimate), by year; the caption names the denominator used (derived from Penta1). Survey estimates, with their 95% confidence interval, appear as a second line when the survey files have regional results; this example has none, so only the DHIS2 line is drawn.

What it shows: Penta3 coverage in the area was 67.5% in 2020, fell to 64.0% in 2021 and stayed at about 66% from 2022 to 2025. The national coverage was 74% to 76% over the same years (Subnational inequality (p. 174)), so the area stays 8 to 10 points below the country, with no sign of catching up. Before concluding that access is weaker there, check the area's reporting rate and compare with the survey estimate for the area when one is available.

Using the results

The coverage of each region and district is what the Subnational inequality (p. 172) and Global coverage targets (p. 175) are computed from. The one-pager is the **Sub-national one-pager** on the Reports (p. 415) page: one region on one slide, with its coverage against the other regions and over time, the data quality score table and a coverage table by district. Share one per region with the regional teams (see Dissemination (p. 245)).

13.2 Subnational inequality (MADM)

How far the coverage of regions or districts is from the national coverage, summarised each year by the MADM.

Subnational inequality is the spread of coverage between regions (Admin 1 (p. 496)) or districts (Admin 2 (p. 496)). The chart shows the coverage of every unit, year by year, against the national coverage, and summarises the spread each year in one number, the MADM (p. 501): how far, on average, the units are from the national coverage.

At a glance

- **Measures:** the coverage of each region or district against the national (or regional) coverage, and the MADM, by year.
- **Good result:** units close to the national coverage, and an MADM that falls over the years.
- **In the app: Sub-National Inequality (p. 365)**; for the regions of the whole country, **National Inequality (p. 352)**.

Why

The focus is on assessing the extent to which the country programmes have succeeded in reducing the inequalities between regions and between districts. The MADM gives an indication of whether the country has been successful in reducing inequalities between subnational units. A falling MADM shows that the weaker units are drawing closer to the national level; a stable or rising one that the national average still hides units that are left behind.

What

Indicator	Numerator	Denominator	Interpretation
MADM (median absolute distance from the median)	Sum of the absolute differences between each unit's coverage and the national coverage	Number of units	The smaller it is, the closer the units are to each other; a decrease over the years means that inequalities between subnational units are being reduced

The units are the Admin 1 regions or the districts of the country, compared with the national coverage. When you choose one region, its districts are shown against the region's coverage. The inequalities are computed indicator by indicator, with the same denominator as the coverage.

Formula

$$\text{MADM} = \frac{1}{n}\sum_{i=1}^{n} |c_i - c_N|$$

where

- c_i is the coverage of unit i (a region or a district) in the year;
- c_N is the national coverage in the year (the region's coverage when the units are the districts of one region);
- n is the number of units;

- the MADM is in percentage points.

For example, three districts with coverage of 60%, 75% and 90% in a country with 75% coverage are 15, 0 and 15 points away from it: MADM $= \left(15 + 0 + 15\right) \div 3 = 10$ points.

Interpretation

In the chart each dot is a subnational unit and the cross is the national (or regional) coverage; the MADM of each year is written at the top.

- **Is the spread of dots wide?** A wide spread means high inequality: the national average hides weak areas.
- **Is the spread, and the MADM, narrowing over the years?** Then inequality is falling. If the MADM is stable or rising, inequalities between units have not been reduced.
- **Is a unit always low, or low only in one year?** A district that is always low is different from one that is low only in a bad year: the second case may be a data problem.
- **Could the differences come from the data?** Facility inequality between regions can come from denominator or reporting differences, not real differences in access. Check units above 100% or far from the rest in Data quality (p. 57) before reporting them.

Example: institutional live births by Admin 1 area

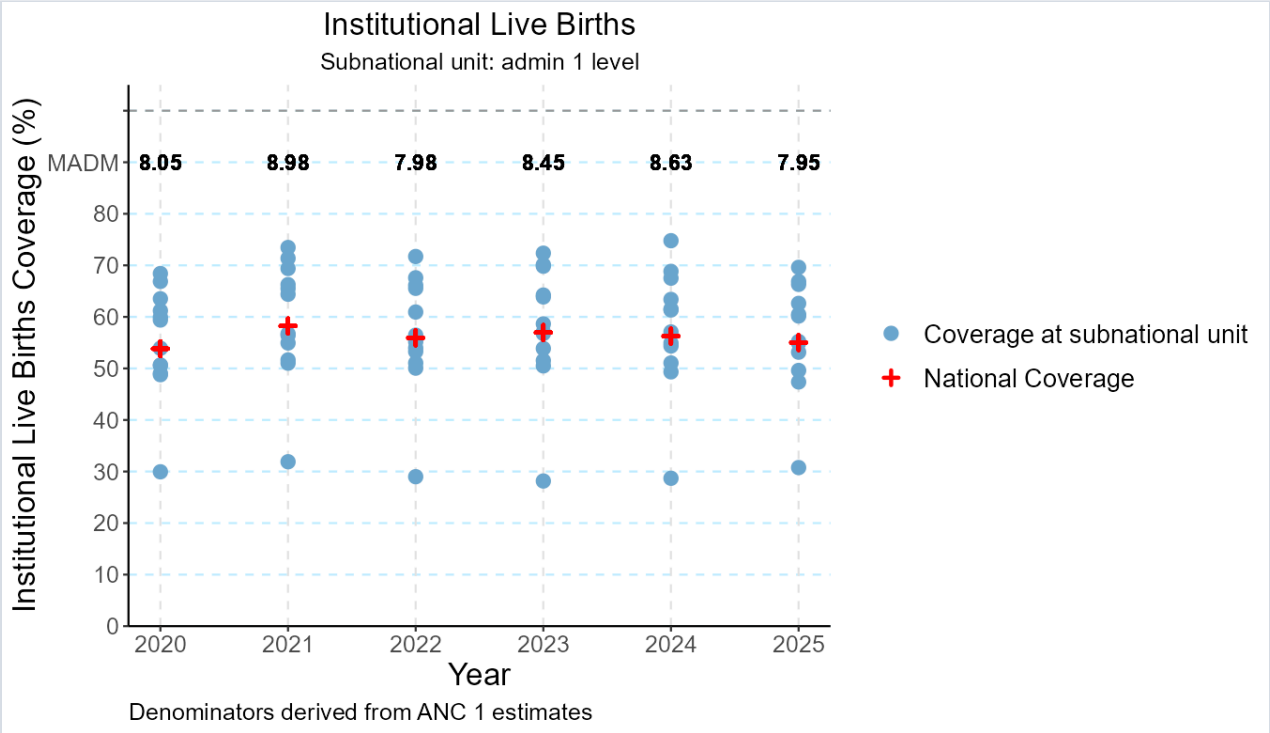


Figure 13.2 — *Institutional live births coverage in each Admin 1 area against the national coverage, with the MADM, by year*

Reading the chart: each blue dot is the coverage of institutional live births in one of the twelve Admin 1 areas, each red cross the national coverage, by year. The number on the MADM line is the MADM of the year, in percentage points. The caption names the denominator (derived from ANC1); the dashed line marks 100%.

What it shows: national coverage was between 54% and 58% every year. The areas ranged from 29.9% to 68.4% in 2020 and from 30.8% to 69.6% in 2025, and the MADM stayed between 7.95 and 8.98 points (8.05 in 2020, 7.95 in 2025): inequality between areas has not been reduced. One area is far below all others every year (28% to 32%): a lasting gap, not a bad year, so look at that area's access, reporting and denominator. The national level itself is far below the latest survey's 91% for institutional live births: check the denominator before reading these levels as real.

Example: Penta3 coverage by district

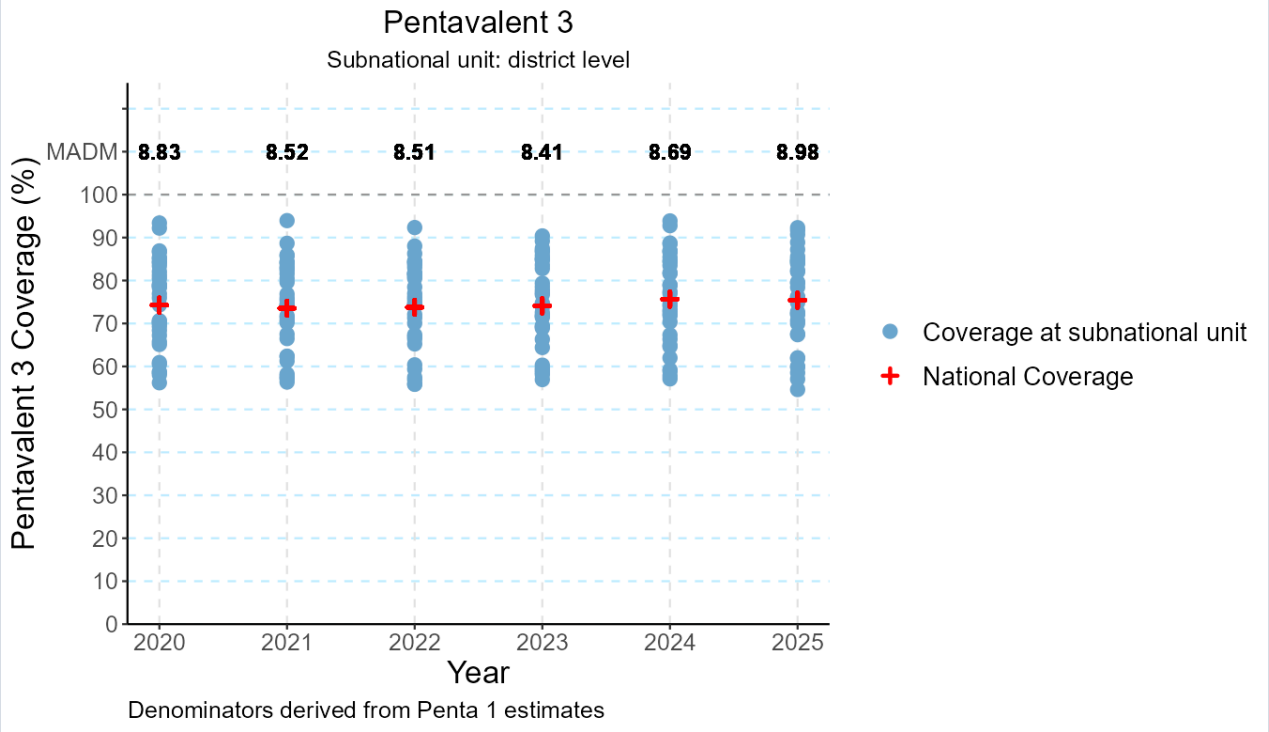


Figure 13.3 — *Penta3 coverage in each district against the national coverage, with the MADM, by year*

Reading the chart: as above, with one dot per district (34 districts) for Penta3 coverage, with the denominator derived from Penta1.

What it shows: national coverage was 74% to 76%. Districts ranged from about 55% to 94% every year and none was above 100%. The MADM was 8.83 points in 2020, 8.41 in 2023 and 8.98 in 2025: on average a district is 8 to 9 points from the national coverage, and this did not fall. The districts at the bottom (55% to 60%) are the ones to look at first.

Using the results

Read the MADM with the share of units that reach the targets (Global coverage targets (p. 175)): units can draw closer to each other while all staying below the target. The charts and their interpretation go into the reports and the chartbook (see Reports (p. 415) and Dissemination (p. 245)).

13.3 Global coverage targets

The percentage of regions or districts whose coverage reaches a global target, by year.

Progress towards international and national targets can be measured by computing the percentage of regions (Admin 1 (p. 496)) and districts (Admin 2 (p. 496)) that have achieved these targets: the percent of subnational units with coverage above a target (p. 502).

At a glance

- **Measures:** the percentage of regions or districts that meet the target of each indicator, by year.
- **Good result:** 100%: every region and district meets the target.
- **In the app: Coverage Target (p. 368)** in the Sub-National Analysis group (districts of one region) and, for the whole country, **Coverage Target (p. 349)** in the National Analysis group.

Why

This indicator provides information on the extent to which a country has been successful in reaching high to universal coverage at the subnational level. The goal is for all regions and districts to have met the target. Higher percentages mean less (geographic) inequality.

What

Three global targets for coverage indicators are of particular interest:

Indicator	Target	Source
Institutional delivery	80% of districts with coverage of at least 80%	EPMM / ENAP (p. 498)
Routine immunization	90% coverage for all routine vaccines	IA2030 (p. 500)
DTP3 (Penta3)	Geographic equity: coverage in the 20% of districts with the lowest coverage	IA2030
DTP3 (Penta3)	At least 80% coverage in every district or equivalent administrative unit	

Indicator	Numerator	Denominator	Interpretation
% of subnational units meeting the target	Number of regions or districts whose coverage reaches the target in the year	Total number of regions or districts	The goal is 100%; higher percentages mean less geographic inequality

A unit meets the target when its coverage is at least as high as the target coverage; for dropout, when its dropout (p. 498) rate is below the target. A target such as "Penta3 at least 80% in every district" is met when the percentage reaches 100%. The geographic equity target (Penta3 coverage in the 20% of districts with the lowest coverage) is not one of the app's outputs: compute it from the district coverage (**Download Data** on Sub-National Coverage (p. 362)).

Formula

$$U = \frac{M}{T} \times 100$$

where

- U is the percentage of units meeting the target in the year;
- M is the number of regions or districts whose coverage reaches the target (for dropout, is below it);
- T is the total number of regions or districts.

For example, if 12 of 34 districts have Penta3 coverage of at least 80%, $U = 12 \div 34 \times 100 = 35.3\%$.

Settings

Setting	Default	What it means
Vaccine target, national	90%	At national level BCG, Penta3 and Measles 1 meet the target when their coverage is at least this.
Vaccine target, regions and districts	80%	Below national level BCG, Penta3 and Measles 1 meet the target when their coverage is at least this.
ANC4 target	70%	A unit meets the ANC4 target when its coverage is at least this (RMNCAH).
Institutional live births target	80%	A unit meets the institutional live births target when its coverage is at least this (RMNCAH).
Dropout target	10%	A unit meets the dropout target (Penta1 to Penta3, Penta3 to Measles 1) when its dropout rate is below this (vaccines).
Other indicators	80%	The table of units meeting the target lists, for an indicator without its own target, the units with coverage of at least this.

Interpretation

- **What percentage of units meet the target, and is it rising?** The goal is 100%. Read the chart title for the rule applied (for dropout, lower is better).
- **Is a rising percentage real?** A rising percentage of units meeting the target is good news only if coverage does not exceed plausible levels. Check units far above 100% for denominator problems.
- **Are some results implausible?** Indicators with implausible results (e.g., far too high) should be excluded or carefully interpreted.
- **Which units do not meet the target?** Use the table to find them, and compare with Subnational inequality (p. 172): are they the same units every year?

Example: districts meeting the vaccine target

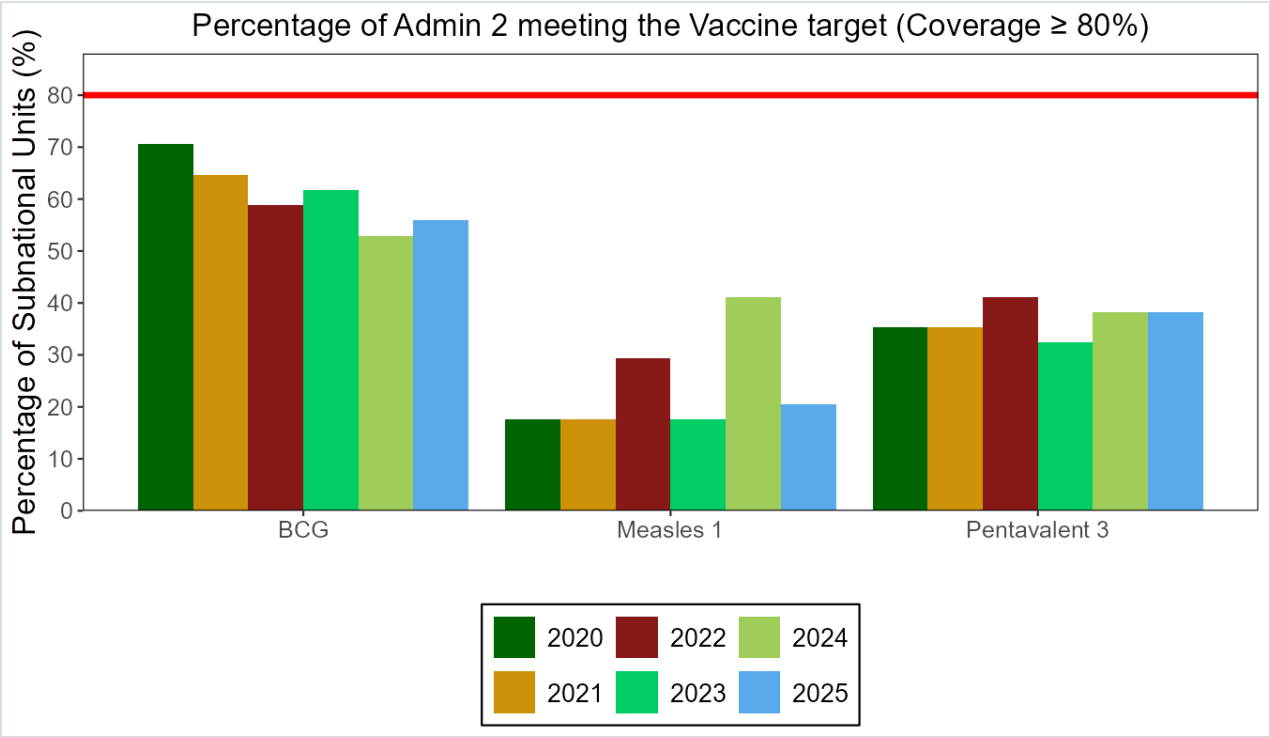


Figure 13.4 — *Percentage of districts with BCG, Measles 1 and Penta3 coverage of at least 80%, by year*

Reading the chart: one group of bars per vaccine, one bar per year; the height of a bar is the percentage of the 34 districts whose coverage is at least 80%, the target for regions and districts. The red line marks 80% of units.

What it shows: BCG is the closest to the goal, but the share of districts reaching 80% fell from 70.6% in 2020 (24 districts) to 52.9% in 2024 and 55.9% in 2025. For Penta3 it stayed between 32.4% and 41.2% (11 to 14 districts), 38.2% in 2025: the target of at least 80% Penta3 coverage in every district is far from met. Measles 1 varied between 17.6% and 41.2% with no trend (20.6% in 2025). Most districts remain below the target for every vaccine, and fewer reach it for BCG than five years earlier.

Example: districts meeting the institutional live births target

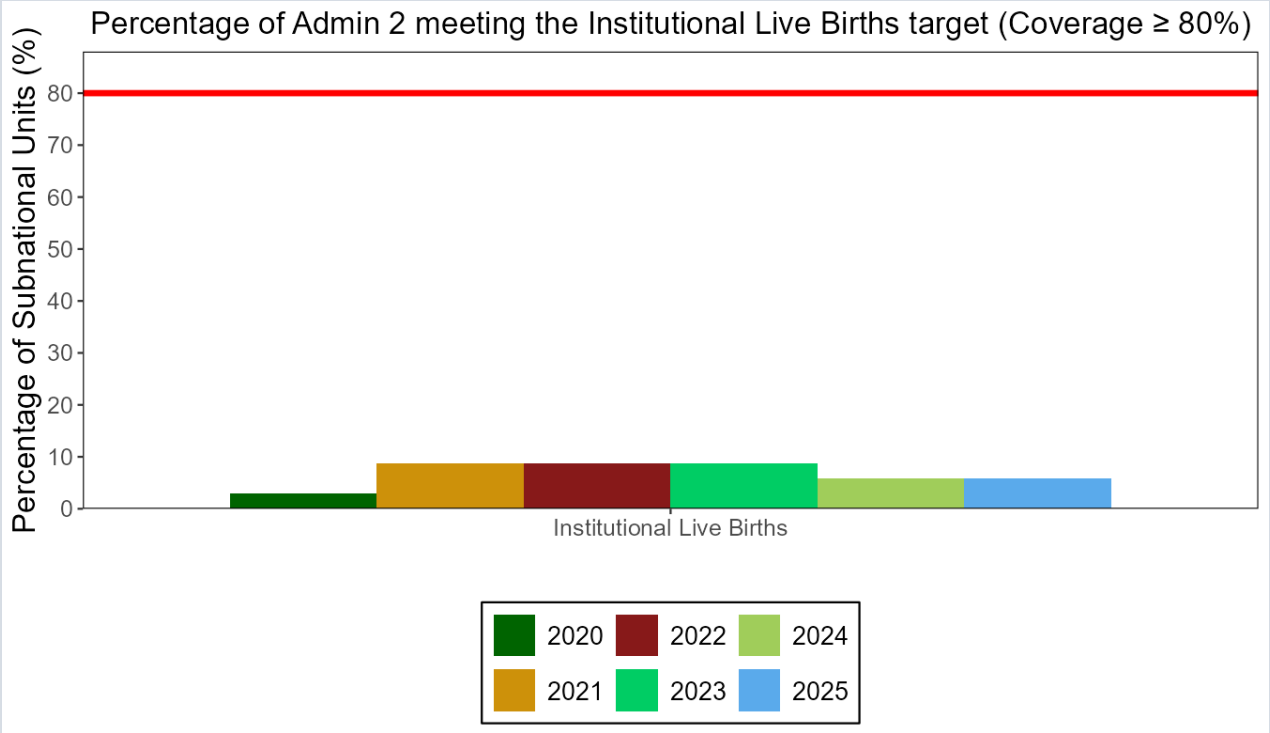


Figure 13.5 — *Percentage of districts with institutional live births coverage of at least 80%, by year*

Reading the chart: one bar per year: the percentage of the 34 districts whose coverage of institutional live births is at least 80%. The red line marks 80% of units, the share that the global target (80% of districts with coverage of at least 80%) asks for.

What it shows: only 1 district (2.9%) reached 80% coverage in 2020, 3 (8.8%) from 2021 to 2023 and 2 (5.9%) in 2024 and 2025, far from the 80% of districts the target asks for. With a national facility-based coverage of about 55% (Subnational inequality (p. 173)), check the denominator before reporting this as the true level.

Using the results

The percentages of units meeting each target go into the reports and the chartbook (see Reports (p. 415) and Dissemination (p. 245)), with the MADM (p. 172) for the same indicators. The next step of the analysis is Mortality (p. 179).

CHAPTER 14

Institutional mortality

How to judge whether facility-reported maternal deaths, stillbirths and neonatal deaths are plausible, and how complete facility death reporting is.

Health facilities record the maternal deaths, stillbirths (p. 504) and neonatal deaths that happen in them. This step computes institutional mortality (p. 500) from those reports, checks whether it is plausible, and compares it with population estimates to see how complete facility reporting of deaths is.

It comes after the Subnational analysis (p. 165). It uses the reported numbers as they are, without the Data adjustment (p. 75) applied to other indicators, but its completeness estimate uses the share of live births in health facilities, computed with the denominator selected in Denominators (p. 95).

At a glance

- **Answers:** are the institutional mortality rates plausible? Is under-reporting more likely for stillbirths or for maternal deaths? How complete is the facilities' reporting of deaths?
- **Uses:** the unadjusted deaths and births (or live births) reported by health facilities, and population mortality estimates: the UN estimates (p. 504), or a recent national survey or census.
- **Produces:** the institutional maternal mortality ratio, stillbirth rate and neonatal deaths before discharge by region and year, data quality checks, and the completeness of facility reporting across modelled scenarios; they go into the mortality report (Reports (p. 415)).
- **In the apps:** Mortality (p. 381) (in the RMNCAH (p. 503) app only).

Why

Maternal and perinatal mortality in health facilities are critical indicators of the quality of care. If coverage of births in health facilities is high (e.g., over 75%), the facility-based statistics also become a useful input into the estimation of population levels of mortality.

The main challenge with mortality data from health facilities is **under-reporting of deaths**. Deaths may not be recorded in the maternity register, or not reported. Maternal deaths do not only occur during birth, but also in pregnancy and post-partum, and maternal deaths in other hospital wards are more likely to be missed. Reporting of stillbirths requires well-maintained maternity registers and cross-checks with operation theatre registers. Because maternal deaths and stillbirths are few and fluctuate, and low institutional mortality in less-developed regions is more likely to signal under-reporting than better outcomes, the rates must be checked for plausibility and compared with population estimates before they are used.

How the parts fit together

Start with the levels: the institutional maternal mortality ratio, stillbirth rate and neonatal deaths before discharge, by region and year. A level that is implausibly low is the first sign of under-reporting. The ratio of stillbirths to maternal deaths then shows which of the two is more likely to be under-reported; it can only say that reporting of both is good when the levels are also in the expected range. The completeness of facility death reporting puts a number on the under-reporting by comparing the reported level with the level expected from population estimates. Check for outliers first: one data entry error can move a whole year.

Doing it in the app

① Open Institutional Mortality

It is under **Mortality** in the **Analysis** group of the RMNCAH app; see Institutional Mortality (p. 384). The first card (p. 384) shows the rates (**Maternal Mortality**, **Still Birth Rate**, **Neonatal Deaths**), with regions as dots and the country as a line, by year.

② Check plausibility

The second card, **Institutional Mortality DQA (p. 384)**, shows fresh stillbirths as a percentage of all stillbirths, the stillbirth to maternal death ratio, and neonatal deaths per maternal death, each against its plausible range.

③ Look at the regions on a map

Mortality Mapping (p. 387) maps **Maternal Mortality** and **Still Birth Rate** by region for the selected years.

④ Estimate completeness

The **Institutional MMR** and **Institutional SBR** tabs of Mortality Completeness (p. 389) show the completeness of facility reporting for the lower bound, median and upper bound of the UN estimate, across community-to-institutional ratios. **Generate Report** opens the mortality report.

Points to consider for the interpretation

- Is the level of MMR from the facility data as expected? What can be said about the regional variation? Are there regions with low MMR and is this plausible?
- What percent of regions has very low MMR (< 25) and very low SBR (< 6)? Is this plausible or is under-reporting of deaths likely?
- What is the ratio of stillbirths to maternal deaths? Is it in the expected range? Is this suggestive of under-reporting of maternal deaths relative to stillbirths?
- Is the level of stillbirths from the facility data as expected? Are there regions with low stillbirth rates, and is this plausible?

- What can be said about the estimated level of completeness of reporting of maternal deaths and stillbirths? Which assumptions are most plausible, in terms of the population level (median, lower or upper bound of the UN estimates) and of the community to institutional ratio?
- What is the neonatal mortality (before discharge) nationally? How does it compare with the national estimate of neonatal mortality per 1,000 live births?

Common pitfalls

- Treating institutional mortality as population mortality. Facilities capture only part of deaths.
- Comparing regions with very different shares of institutional births.
- Ignoring completeness. A low completeness means the reported rate is a lower bound.
- Reading a fall in institutional mortality as better care without checking the other parts: a fall in reported maternal deaths together with a rise in the ratio of stillbirths to maternal deaths points to under-reporting.

14.1 Institutional maternal mortality ratio (iMMR)

The number of maternal deaths in health facilities per 100,000 live births in health facilities, by region and year.

The institutional maternal mortality ratio is the number of maternal deaths (p. 501) that health facilities report for every 100,000 live births (p. 501) in those facilities. It is one of the three rates of institutional mortality (p. 500), and it is computed for each region (p. 496) and for the whole country, year by year.

At a glance

- **Measures:** maternal deaths in health facilities per 100,000 facility live births.
- **Good result:** a level consistent with the population maternal mortality, with no region implausibly low (under 25) and no extreme high values.
- **In the app:** the **Maternal Mortality** tab of the first card of **Institutional Mortality** **(p. 384)**, and the **Maternal Mortality** map of **Mortality Mapping** **(p. 387)**.

Why

Maternal mortality in health facilities is a critical indicator of the quality of care. The main challenge with mortality data from health facilities is under-reporting of deaths: maternal deaths do not only occur during birth, but also in pregnancy and post-partum, and deaths in other hospital wards are more likely to be missed. An implausibly low ratio is therefore more often a sign of missing reports than of good care.

What

A maternal death is the death of a woman from any cause related to or aggravated by pregnancy or its management (excluding accidental or incidental causes) during pregnancy and childbirth or within 42 days of termination of pregnancy, irrespective of the duration and site of the pregnancy. The definitions for institutional, community and population maternal mortality are outlined in the matrix below.

Indicator	Numerator	Denominator	Interpretation
Institutional maternal mortality ratio (iMMR)	Number of maternal deaths in health facilities	Number of live births in health facilities / 100,000	Less than 25 per 100,000 live births is implausibly low; low rates in less-developed regions indicate under-reporting
Population maternal mortality (MMR)	Number of maternal deaths in the population	Number of live births in the population / 100,000	Comes from the UN estimates or a recent national survey or census
Community maternal mortality ratio (cMMR)	Number of maternal deaths in the community	Number of live births in the community / 100,000	Deaths outside health facilities; not reported, but follows from the other two

The annual ratios are computed using the unadjusted data on reported deaths and live births in health facilities: they are not adjusted for reporting rates or outliers (why: Outliers in mortality data (p. 201)). The institutional ratio is read against the population maternal mortality ratio (p. 503), which includes

community maternal mortality ratio (p. 497); how the three fit together is on Completeness of facility death reporting (p. 194).

Formula

$$\text{iMMR} = \frac{D_m}{B_l} \times 100{,}000$$

where

- D_m is the number of maternal deaths reported by health facilities in the year;
- B_l is the number of live births reported by health facilities in the year.

Example: 1,000 maternal deaths among 464,020 facility live births give $1{,}000 \div 464{,}020 \times 100{,}000 = 215.5$ maternal deaths per 100,000 live births.

The population (MMR) and community (cMMR) ratios use the same formula with the deaths and live births in the population or in the community.

Settings

Setting	Default	What it means
Very low institutional MMR	25 per 100,000 live births	A region whose ratio is below this value is flagged as very low, a sign of under-reported deaths.

Interpretation

- **Are there extreme high values?** Assess if there are any extreme outliers on the high side that may be due to major data entry errors. Flag them for potential correction or exclusion (see Outliers in mortality data (p. 201)).
- **How many regions have an implausibly low ratio?** It is arbitrarily defined as **less than 25 per 100,000 live births** (25 is roughly two times the average MMR in high-income countries of 12.5).
- **Are the low regions more advanced areas, or less-developed ones?** In more advanced areas lower mortality is expected. Low mortality in less-developed regions strongly indicates a major under-reporting of institutional deaths rather than superior clinical outcomes.
- **How far is the iMMR from the population estimate?** Compare it with the most recent population estimate, from a recent national survey or census, or the UN estimates (p. 504) for maternal mortality, including their uncertainty range. This difference is used to estimate the completeness of facility death reporting (p. 193).
- **Where do the deaths come from?** If data collection relies only on maternity ward logs or MPDSR (p. 502) audits, deaths upon re-admission or outside the labour ward (intensive care, emergency, or general surgical wards, for example after post-abortal sepsis or delayed haemorrhage) are frequently omitted. A facility-wide all-cause mortality registry, where the underlying maternal causes are accurately coded, reduces this under-reporting.
- **What does the map show?** A map of the iMMR by region (in Mortality Mapping (p. 387)) helps to focus on spatial patterns and identify potential under-reporting across regions. High values on a

map are not always high mortality: regions with few births can swing widely, and regions that report well can look worse than regions that under-report.

Example: institutional maternal mortality by region and year

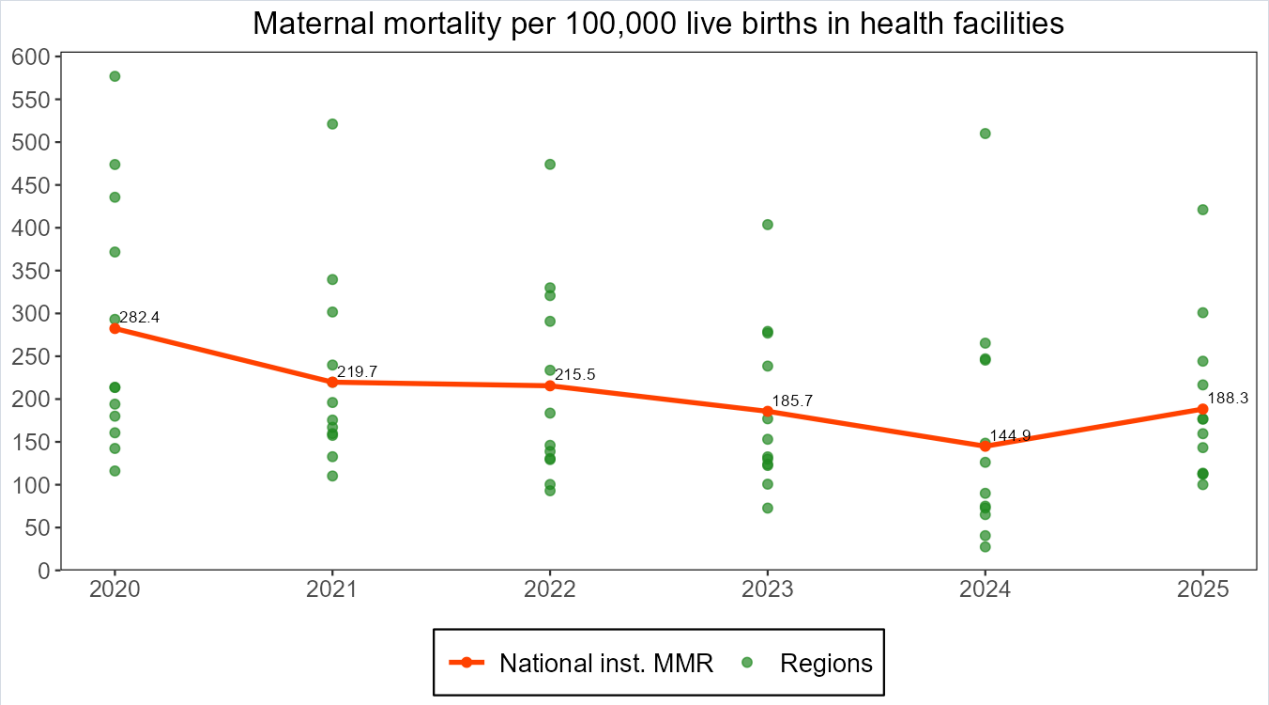


Figure 14.1 — *Maternal mortality per 100,000 live births in health facilities, by region and year*

Reading the chart: each green dot is a region and the orange line the country, with its value for each year. The vertical axis is maternal deaths per 100,000 live births in health facilities.

What it shows: the national ratio fell from 282 in the first year to 145 in the fifth and rose again to 188 in the last. Regions are spread widely, from 116 to 577 in the first year. No region is below 25, but in the fifth year one region is at 27 and several are under 100. The fall in the fifth year comes from fewer reported maternal deaths (694, against 900 the year before) while facility live births stayed level; the ratio of stillbirths to maternal deaths (p. 199) rose the same year, which points to under-reported maternal deaths rather than better care. Check the reports of the low regions for that year.

After the assessment

The reported ratio is compared with the ratio expected from population estimates to estimate the completeness of facility death reporting (p. 193). The ratios and the answers to the points to consider (p. 180) go into the mortality report (Reports (p. 415)).

14.2 Institutional stillbirth rate (iSBR)

The number of stillbirths in health facilities per 1,000 births in health facilities, and the share of fresh stillbirths.

The institutional stillbirth rate is the number of stillbirths (p. 504) that health facilities report for every 1,000 births (live births and stillbirths) in those facilities. Facilities record stillbirths as fresh (p. 499) or macerated (p. 501), and the share of fresh stillbirths is a check on how well they are recorded.

At a glance

- **Measures:** stillbirths in health facilities per 1,000 facility births, and fresh stillbirths as a percentage of all stillbirths.
- **Good result:** no region implausibly low (under 6 per 1,000 births), and fresh stillbirths between 30% and 55% of all stillbirths.
- **In the app:** the **Still Birth Rate** tab of the first card of **Institutional Mortality** (p. 384), the **Fresh Stillbirths** tab of its **Institutional Mortality DQA** (p. 384) card, and the **Still Birth Rate** map of **Mortality Mapping** (p. 387).

Why

Perinatal mortality in health facilities is a critical indicator of the quality of care, and stillbirths are its largest part. Reporting of stillbirths requires well-maintained maternity registers and cross-checks with operation theatre registers; where they are missing, stillbirths go unreported or are recorded in the wrong category.

What

A stillbirth is a baby who dies after 28 weeks of pregnancy (gestation), but before or during birth. Clinical records typically distinguish between:

- **Ante-partum stillbirths (macerated):** deaths occurring before the onset of labour, showing signs of tissue deterioration.
- **Intra-partum stillbirths (fresh):** deaths occurring during labour or delivery.

Indicator	Numerator	Denominator	Interpretation
Institutional stillbirth rate (iSBR)	Number of stillbirths in health facilities	Number of births in health facilities / 1,000	Less than 6 per 1,000 births is implausibly low; low rates in less-developed regions indicate under-reporting
Population stillbirth rate (SBR)	Number of stillbirths in the population	Number of births in the population / 1,000	Comes from the UN estimates or a recent national survey or census
Community stillbirth rate (cSBR)	Number of stillbirths in the community	Number of births in the community / 1,000	Stillbirths outside health facilities
Fresh stillbirths	Number of fresh stillbirths in health facilities	Number of stillbirths in health facilities / 100	Plausible range 30% to 55%; far outside it suggests fresh and macerated stillbirths are misclassified

The annual rates are computed using the unadjusted reported stillbirths and births, not adjusted for reporting rates or outliers (why: Outliers in mortality data (p. 201)).

Formula

$$\text{iSBR} = \frac{S}{B_l + S} \times 1{,}000$$

$$\text{Fresh stillbirths} = \frac{S_f}{S} \times 100$$

where

- S is the number of stillbirths reported by health facilities in the year;
- B_l is the number of live births reported by health facilities in the year, so that $B_l + S$ is all facility births;
- S_f is the number of fresh stillbirths among them.

Example: 9,187 stillbirths and 464,020 live births give $9{,}187 \div 473{,}207 \times 1{,}000 = 19.4$ stillbirths per 1,000 births. If 4,015 of the stillbirths are fresh, fresh stillbirths are $4{,}015 \div 9{,}187 \times 100 = 43.7\%$.

Settings

Setting	Default	What it means
Very low institutional stillbirth rate	6 per 1,000 births	A region whose rate is below this value is flagged as very low, a sign of under-reported stillbirths.

The plausible range for fresh stillbirths (30% to 55%) is fixed.

Interpretation

- **How many regions have an implausibly low iSBR?** It is defined as **less than 6 per 1,000 births** (6 is approximately two times the average SBR in high-income countries).
- **Are the low regions more advanced areas, or less-developed ones?** In more advanced regions mortality is expected to be lower due to better clinical infrastructure; in less-developed regions a low rate is a clear sign of major under-reporting of deaths.
- **How far is the iSBR from the population estimate?** Compare it with the most recent population estimate, from a recent national survey or census, or the UN estimates (p. 504) of stillbirth rates, including their uncertainty range. This difference is used to estimate the completeness of facility stillbirth reporting (p. 196).
- **Are fresh stillbirths between 30% and 55% of all stillbirths?** A share far outside the range suggests that fresh and macerated stillbirths are misclassified. Regions that sit outside the range in every year are the priority for data review.

Example: institutional stillbirth rate by region and year

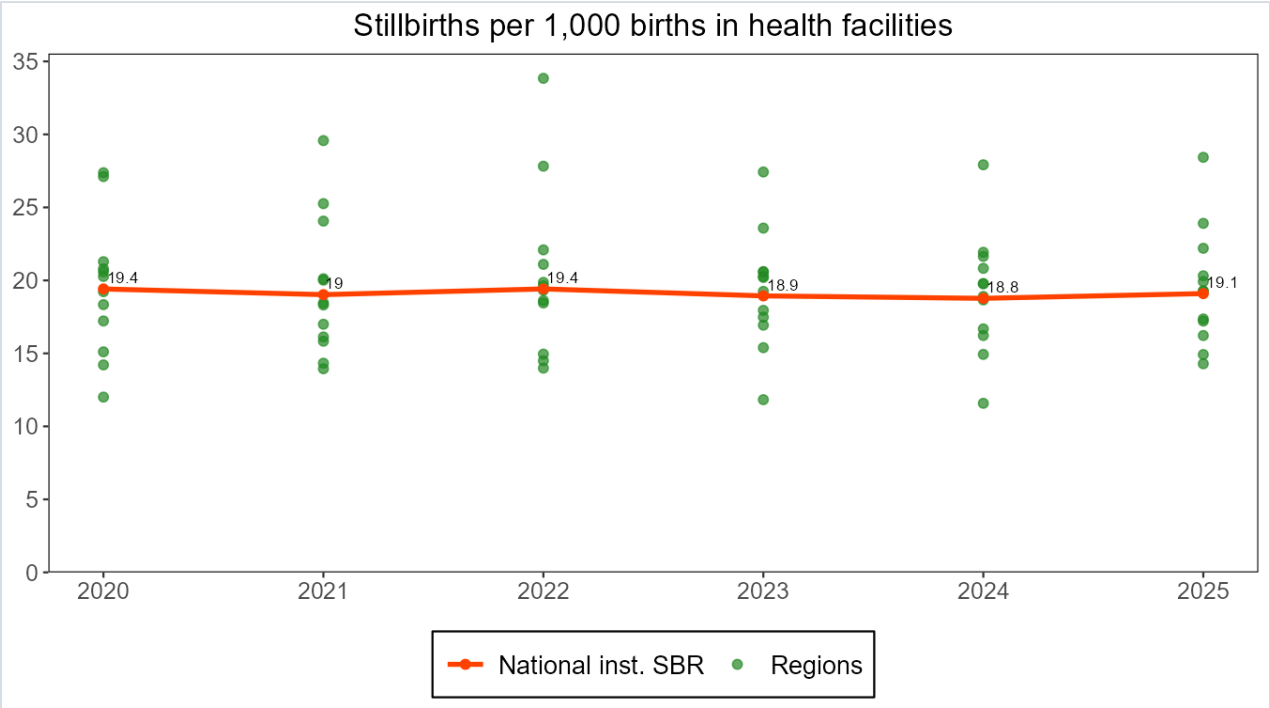


Figure 14.2 — *Stillbirths per 1,000 births in health facilities, by region and year*

Reading the chart: each green dot is a region and the orange line the country, with its value for each year. The vertical axis is stillbirths per 1,000 births in health facilities.

What it shows: the national rate is steady at 19 per 1,000 births in every year (18.8 to 19.4). Regions range from about 12 to 34; none is below 6. The level is plausible; whether it is complete is answered by comparing it with the population estimate (see the stillbirth example (p. 196)).

Example: fresh stillbirths as a share of all stillbirths

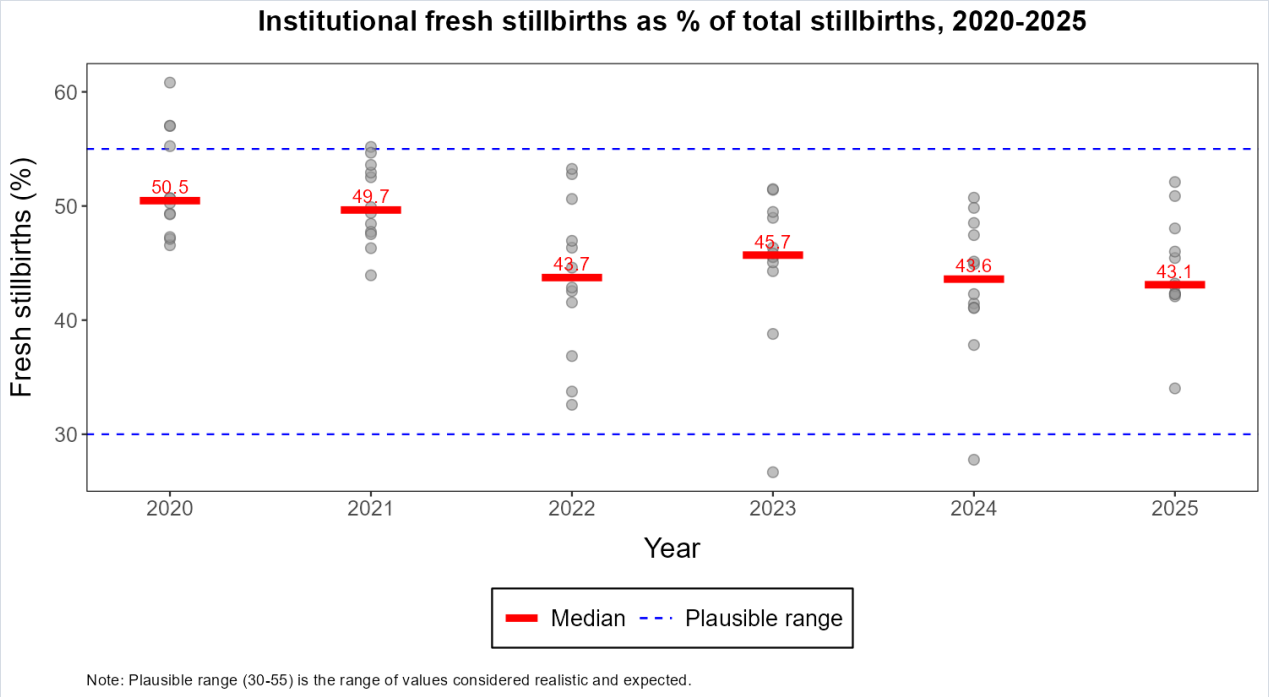


Figure 14.3 — *Institutional fresh stillbirths as a percentage of total stillbirths, by region and year*

Reading the chart: each grey dot is a region; the red bar is the median (p. 501) of the regions, with its value, and the dashed blue lines the plausible range of 30% to 55%.

What it shows: the median fell from 50.5% in the first year to about 43% from the third year, inside the range. In the first year four regions are above 55%; after that at most one region a year is outside the range (just above 55% in the second year, 27% in the fourth and 28% in the fifth). One region is outside the range in three of the six years: review how its facilities classify fresh and macerated stillbirths.

After the assessment

The reported rate is compared with the rate expected from population estimates to estimate the completeness of facility stillbirth reporting (p. 196), and compared with maternal deaths in the ratio of stillbirths to maternal deaths (p. 198).

14.3 Neonatal deaths before discharge

The number of neonatal deaths in health facilities before discharge per 1,000 live births in health facilities, and per maternal death.

Health facilities also report the babies born alive who die before they leave the facility. Divided by the facility live births, they give the institutional neonatal mortality rate, which is mostly referred to as **neonatal deaths before discharge per 1,000 live births**, because it only covers the time a baby stays in the facility.

At a glance

- **Measures:** neonatal deaths before discharge per 1,000 live births in health facilities, and per maternal death.
- **Good result:** at least half of the population neonatal mortality (p. 502), and 5 to 10 neonatal deaths per maternal death.
- **In the app:** the **Neonatal Deaths** tab of the first card of **Institutional Mortality** (p. 384) and the **Neonatal / Maternal Deaths** tab of its **Institutional Mortality DQA** (p. 384) card.

Why

Neonatal deaths in health facilities are graphed like the maternal mortality ratio and the stillbirth rate, but have to be interpreted with additional caution. Almost all babies stay at least 24 hours after delivery in the hospital, but after that many are discharged and the observation time in health facilities is variable. The rate includes day 1, some deaths on day 2, fewer deaths on day 3 and so on. Because they are deaths of live-born babies before discharge, they understate community neonatal mortality. Under-reporting is still the main question: a rate far below the population neonatal mortality means deaths are missing.

What

Indicator	Numerator	Denominator	Interpretation
Neonatal mortality (before discharge)	Number of neonatal deaths before discharge	Number of live births in health facilities / 1,000	Should be at least half of the neonatal mortality in the population
Neonatal deaths per maternal death	Number of neonatal deaths before discharge in health facilities	Number of maternal deaths in health facilities	Plausible range 5 to 10

The rates are computed using the unadjusted reported deaths and live births (why: Outliers in mortality data (p. 201)).

Formula

$$\text{Neonatal deaths before discharge} = \frac{D_n}{B_l} \times 1{,}000$$

$$\text{Neonatal deaths per maternal death} = \frac{D_n}{D_m}$$

where

- D_n is the number of neonatal deaths before discharge reported by health facilities in the year;
- B_l is the number of live births reported by health facilities in the year;
- D_m is the number of maternal deaths reported by health facilities in the year.

Example: 4,504 neonatal deaths among 464,020 live births give $4{,}504 \div 464{,}020 \times 1{,}000 = 9.7$ per 1,000 live births; with 1,000 maternal deaths, that is 4.5 neonatal deaths per maternal death.

A rough guide to assess reporting completeness is that neonatal deaths before discharge should be at least half of neonatal mortality in the population:

$$\text{Neonatal deaths before discharge} \geq 0.5 \times NMR$$

where NMR is the neonatal mortality rate in the population, per 1,000 live births. For instance, if population neonatal mortality is 20 per 1,000 live births, we expect at least 10 neonatal deaths before discharge per 1,000 live births in the health facilities.

Settings

The plausible range of 5 to 10 neonatal deaths per maternal death is fixed.

Interpretation

- **What is the neonatal mortality before discharge nationally, and how does it compare with the national estimate of neonatal mortality?** If it is below half of the population neonatal mortality, the facilities under-report neonatal deaths.
- **Are some regions far below the others?** Regions with very low rates are likely to be under-reporting, unless babies are discharged unusually early there.
- **Is the number of neonatal deaths per maternal death between 5 and 10?** Below 5, neonatal deaths are probably under-reported compared with maternal deaths; above 10, maternal deaths may be the ones missing (compare with the ratio of stillbirths to maternal deaths (p. 198)). Regions that sit outside the range in every year are the priority for data review.

Example: neonatal deaths before discharge by region and year

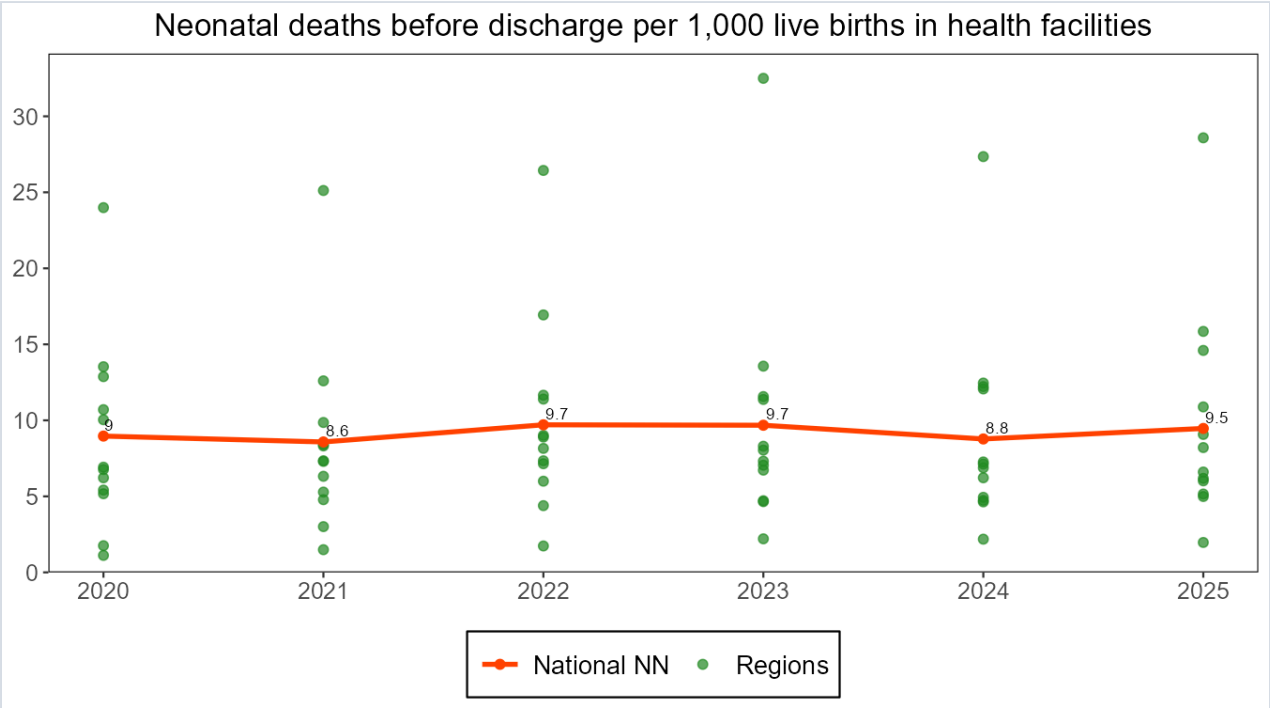


Figure 14.4 — *Neonatal deaths before discharge per 1,000 live births in health facilities, by region and year*

Reading the chart: each green dot is a region and the orange line the country, with its value for each year. The vertical axis is neonatal deaths before discharge per 1,000 live births in health facilities.

What it shows: the national rate is between 8.6 and 9.7 per 1,000 live births in every year. If population neonatal mortality were 20 per 1,000, facilities would be expected to report at least 10: the national rate is just under that, and in every year at least one region reports fewer than 2.5, which is very likely under-reporting. At the other end, the highest region each year reports 24 to 32 per 1,000, far above the others: check whether its numbers include deaths after discharge or entry errors.

Example: neonatal deaths per maternal death

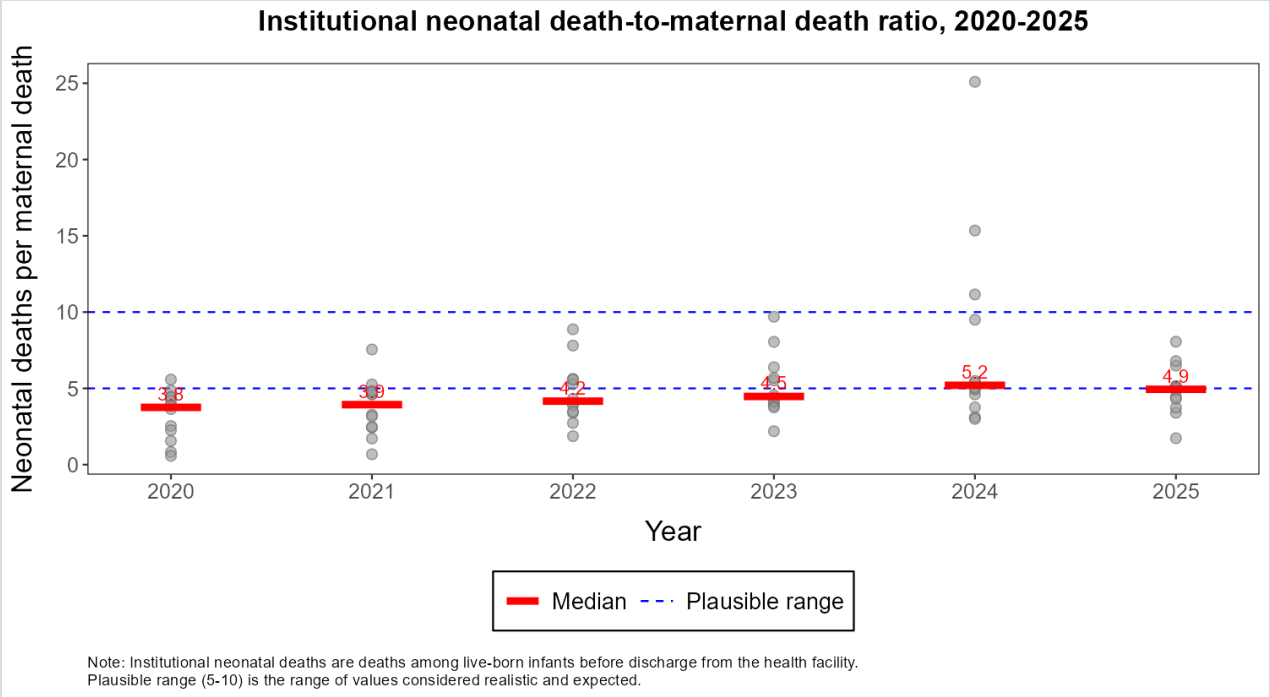


Figure 14.5 — *Institutional neonatal death-to-maternal death ratio, by region and year*

Reading the chart: each grey dot is a region; the red bar is the median (p. 501) of the regions, with its value, and the dashed blue lines the plausible range of 5 to 10.

What it shows: the median is below 5 in four of the six years (3.8 to 4.9) and most regions are below the range: 11 of 12 in the first year, 6 in the last. Neonatal deaths are under-reported compared with maternal deaths. In the fifth year three regions are above 10 (up to 25), the same year in which reported maternal deaths fell: there, maternal deaths are the ones more likely to be missing.

After the assessment

The rate and its comparison with population neonatal mortality go into the mortality report (Reports (p. 415)). Neonatal deaths are not used in the completeness of facility death reporting (p. 193), which is estimated for maternal deaths and stillbirths only.

14.4 Completeness of facility death reporting

The reported institutional maternal mortality ratio (or stillbirth rate) as a percentage of the one expected from population estimates.

How many of the deaths that happen in health facilities are reported? It cannot be counted directly, but it can be estimated: from the population mortality and the share of births in facilities, a model gives the expected institutional maternal mortality ratio (p. 499), and the reported ratio divided by the expected one is the completeness of facility death reporting (p. 497). The same is done for stillbirths.

At a glance

- **Measures:** the reported institutional maternal mortality ratio (and stillbirth rate) as a percentage of the expected one, for several assumptions.
- **Good result:** close to 100% in the scenarios that fit the country; a low completeness means the reported rate is a lower bound.
- **In the app:** the **Institutional MMR** and **Institutional SBR** tabs of **Mortality Completeness** (p. 389).

Why

The population MMR, the community MMR and the institutional MMR must be mathematically consistent. While variations will occur between different populations, it is statistically improbable to observe a community MMR of 1,000 when the institutional MMR is reported at 100. Comparing the reported institutional ratio with the ratio expected from population estimates therefore shows how complete the reporting by health facilities is.

What

An expected MMR in health facilities can be modelled from two sets of parameters:

1. **The maternal mortality ratio in the whole population (M_p):** it accounts for both community and institutional deaths. Teams can use the lower bound, median or upper bound of the global estimates for each country (the UN estimates (p. 504)), or results from a recent national household survey (p. 500).
2. **The ratio of community to institutional maternal mortality (p. 497) (M_c / M_i):** assumptions range from 0.5 (the community MMR is half the institutional MMR) through 1.0 (they are equal) up to 2.0 (the community MMR is twice the institutional MMR). Which ratio fits depends on the local proportion of births occurring in health facilities (P_i): evidence suggests that the community-to-institutional ratio tends to expand as institutional birth rates and overall facility utilization rates increase.

The ratios used (0.5, 1.0, 1.5 and 2.0) are derived from empirical field studies that cross-referenced institutional deaths against total population mortality estimates, accounting for the proportion of all national maternal deaths that occurred in health facilities and the matching percentage of births that took place in health facilities.

Indicator	Numerator	Denominator	Interpretation
Completeness of facility maternal death reporting	Reported institutional MMR	Expected institutional MMR	One value per population MMR (lower bound, best estimate, upper bound) and community-to-institutional ratio; narrow the scenarios with local evidence
Completeness of facility stillbirth reporting	Reported institutional stillbirth rate	Expected institutional stillbirth rate	The same, with all births instead of live births and the population stillbirth rate

Formula

The population MMR is the sum of the institutional and community ratios, weighted by the proportion of live births in each setting:

$$M_p = M_i \times P_i + M_c \times \left(1 - P_i\right)$$

where

- M_p is the maternal mortality ratio in the population;
- M_i is the institutional maternal mortality ratio;
- M_c is the maternal mortality ratio in the community;
- P_i is the proportion of live births occurring in health facilities.

Example: if $M_i = 100$, $M_c = 200$ and 75% of births take place in health facilities ($P_i = 0.75$), the population MMR is $100 \times 0.75 + 200 \times 0.25 = 75 + 50 = 125$. The community-to-institutional ratio is $200 \div 100 = 2.0$.

Since M_i comes from the reported facility data and M_p from independent population estimates, the community MMR can be back-calculated:

$$M_c = \frac{M_p - M_i \times P_i}{1 - P_i}$$

Conversely, from a population MMR and an assumed community-to-institutional ratio $R = M_c \,/\, M_i$, the expected institutional MMR is:

$$E_i = \frac{M_p}{P_i + R \times \left(1 - P_i\right)}$$

and the completeness of facility reporting is the reported institutional MMR divided by the expected one:

$$C = \frac{M_i}{E_i} \times 100$$

where

- E_i is the expected institutional MMR;
- R is the assumed ratio of community to institutional maternal mortality;
- C is the completeness of facility reporting, in percent.

Example: if 75% of women deliver in health facilities ($P_i = 0.75$), the population MMR is 200 and the assumed ratio is 2, the expected institutional MMR is $200 \div \left(0.75 + 2 \times 0.25\right) = 200 \div 1.25 = 160$. If the reported institutional MMR is 100, completeness is $100 \div 160 = 62.5\%$.

In the app:

- the reported institutional MMR M_i is the mean of the national yearly ratios over the years of the analysis;
- the population MMR M_p is the most recent UN estimate, with its lower bound, best estimate and upper bound;
- P_i is the mean share of live births in health facilities over the same years, computed with the denominator selected in Denominators (p. 95).

For stillbirths, the same formulas are used with all births instead of live births, the institutional stillbirth rate for M_i and the population stillbirth rate for M_p.

Settings

Setting	Default	What it means
Completeness ratios compared	0.5, 1, 1.5, 2	The community-to-institutional ratios at which the institutional mortality is compared with the UN estimates.
Population estimates	Built-in UN estimates	The population MMR and stillbirth rate, with their bounds. Use the built-in country data, or upload your own in the Upload Data step of Getting started (p. 295).

Interpretation

Not all modelled scenarios apply equally to every country context. Evaluate them and narrow them down with local evidence:

- **Is the true population MMR thought to be lower or higher than the UN median?** If external validation indicates that it sits below the UN median, pick the scenario of the lower bound (the lower-bound line); if above, the upper bound.
- **How much higher is community mortality than facility mortality?** If country data indicate that community maternal mortality is roughly 1.5 to 2.0 times facility mortality, read completeness only at those ratios: the range narrows to two values on each line.
- **Is completeness low in every plausible scenario?** Then the reported rates are a lower bound of institutional mortality, and the reporting of deaths needs strengthening (registers, all-cause facility mortality, MPDSR (p. 502)).
- **Is completeness above 100%?** Facilities cannot report more deaths than occur: that scenario does not fit, because the population estimate is too low or the assumed ratio too high.

Example: completeness of facility maternal death reporting

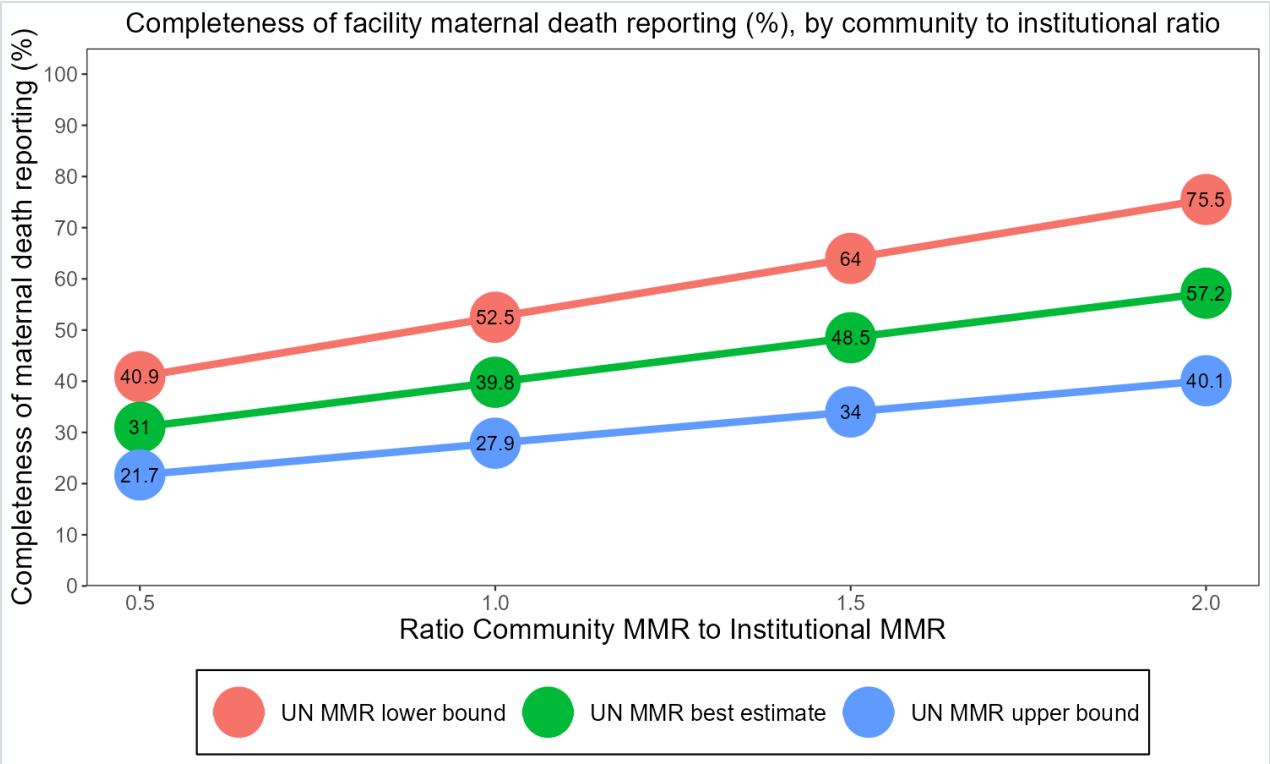


Figure 14.6 — *Completeness of facility maternal death reporting by community to institutional ratio*

Reading the chart: the horizontal axis is the assumed community-to-institutional ratio (0.5 to 2.0); each line is one population MMR: the UN lower bound (red), best estimate (green) and upper bound (blue). Each circle gives the completeness of reporting in percent for that scenario.

What it shows: the reported institutional MMR averages 206 per 100,000 live births over the six years; the most recent UN estimate is 518 (393 to 740), and 56% of live births are in health facilities. At a ratio of 1 with the best estimate, the expected institutional MMR is $518 \div (0.56 + 0.44) = 518$, so completeness is $206 \div 518 = 39.8\%$. Across all scenarios completeness goes from 22% to 76%. If community mortality is 1.5 to 2.0 times facility mortality, completeness is 49% to 57% with the best estimate, and 64% to 76% with the lower bound. Whatever the assumption, a large share of maternal deaths in facilities is not reported.

Example: completeness of facility stillbirth reporting

There is little research on the community to institutional stillbirth ratio (partly because community level stillbirth reporting is more uncertain), but it is likely that the ratios are lower than for maternal mortality, as institutional mortality levels are much higher for stillbirth rates than for MMR. A range of 0.5 to 2.0 may be used. For the UN stillbirth estimates, the lower and upper bounds are 90% uncertainty intervals from the model.

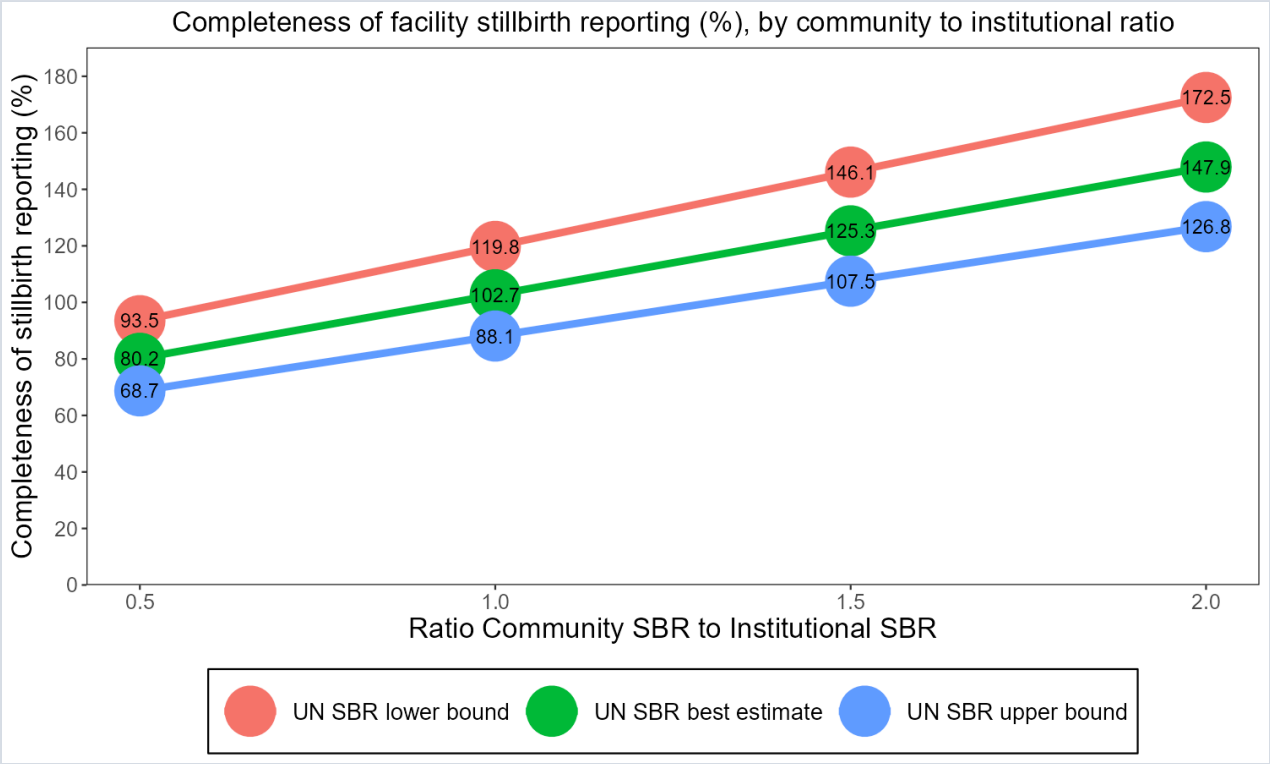


Figure 14.7 — *Completeness of facility stillbirth reporting by community to institutional ratio*

Reading the chart: as above, for stillbirths: each line is one UN stillbirth rate (lower bound, best estimate, upper bound) and each circle the completeness of facility stillbirth reporting for that ratio.

What it shows: the reported institutional stillbirth rate averages 19.1 per 1,000 births, about the same as the UN estimate of 18.6 (16.0 to 21.7). From a ratio of 1 with the best estimate, completeness is over 100% (103% to 148%), which cannot be: those scenarios do not fit. At a ratio of 0.5, which the lower stillbirth ratios make likely, completeness is 69% to 94%. Stillbirth reporting is much more complete than maternal death reporting.

After the assessment

The completeness estimates and the scenarios chosen go into the mortality report (Reports (p. 415)). They tell readers how far the reported institutional rates understate the deaths in facilities; the reported numbers themselves are not corrected.

14.5 Ratio of stillbirths to maternal deaths

The number of reported stillbirths divided by the number of reported maternal deaths, nationally and by region.

For every maternal death, how many stillbirths (p. 504) do health facilities report? Stillbirths and maternal deaths (p. 501) share many causes, so the two move together, and a ratio far from the usual range shows which of the two is more likely to be under-reported.

At a glance

- **Measures:** reported stillbirths per reported maternal death, in a year.
- **Good result:** nationally between 5 and 25; for regions, within the plausible range of 6 to 15 on the chart.
- **In the app:** the **SBR to Maternal Death Ratio** tab of the **Institutional Mortality DQA** (p. 384) card of Institutional Mortality.

Why

We expect maternal mortality and stillbirth to be positively correlated, given the commonalities in causes. Under-reporting rarely affects both in the same way: maternal deaths in other wards are missed, stillbirths are missed where maternity registers are poorly kept. Comparing the two reported numbers is a data quality check that needs no population estimate.

What

Based on a review of the global estimates, historical data and health facility studies, we expect the ratio of stillbirths to maternal deaths to be in the range of **5 to 25** for countries in sub-Saharan Africa.

Indicator	Numerator	Denominator	Interpretation
Ratio of stillbirths to maternal deaths (national)	Number of reported stillbirths in a period (usually a year)	Number of reported maternal deaths in the same period	Outside 5 to 25, raise a "data quality flag"
Stillbirths per maternal death (regions)	Number of stillbirths in health facilities in the region	Number of maternal deaths in health facilities in the region	Plausible range 6 to 15; regions outside it in every year are the priority for data review

The numbers come from DHIS2 or MPDSR (p. 502). The regional chart uses a narrower plausible range (6 to 15) than the range used to flag national data (5 to 25); the question in the mortality report asks whether the ratio is in the range of 5 to 15.

Formula

$$\text{Ratio} = \frac{S}{D_m}$$

where

- S is the number of stillbirths reported by health facilities in the year;
- D_m is the number of maternal deaths reported by health facilities in the year.

Example: 9,187 stillbirths and 1,000 maternal deaths give a ratio of 9,187 ÷ 1,000 = 9.2, inside the 5 to 25 range.

Settings

The ranges (5 to 25 for the national flag, 6 to 15 on the chart) are fixed.

Interpretation

- **Is the ratio lower than 5?** Under-reporting of stillbirths is likely greater than under-reporting of maternal deaths.
- **Is the ratio equal to or greater than 25?** Under-reporting of maternal deaths is likely to be the main issue; under-reporting of stillbirths is less serious than for maternal deaths.
- **Is the ratio between 5 and 25?** Under-reporting of maternal deaths and under-reporting of stillbirths are both possible, or reporting of both is of good quality. Good quality requires that the levels are also in the expected range (see the maternal mortality ratio (p. 183) and the stillbirth rate (p. 186)).
- **Is a very low maternal mortality ratio paired with a low ratio of stillbirths to maternal deaths?** It can point to under-reporting of deaths, not to good care.
- **Do some regions sit outside the range in every year?** They are the priority for data review.

Example: stillbirths per maternal death by region and year

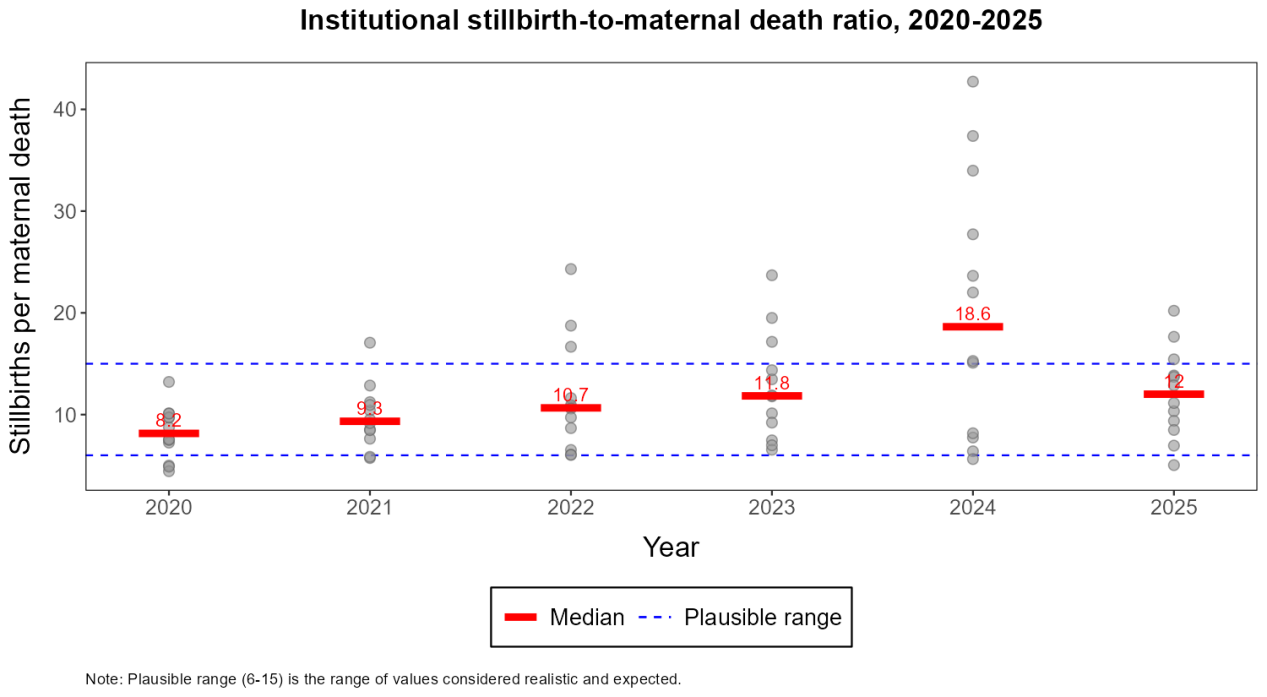


Figure 14.8 — *Institutional stillbirth-to-maternal death ratio, by region and year*

Reading the chart: each grey dot is a region; the red bar is the median (p. 501) of the regions, with its value, and the dashed blue lines the plausible range of 6 to 15.

What it shows: nationally, the ratio rose from 7.0 in the first year to 13.2 in the fifth, inside the 5 to 25 range in every year. The regional median rose the same way, from 8.2 to 18.6 in the fifth year, when it left the range: 9 of the 12 regions are outside it and four are above 25. That year reported maternal deaths fell sharply while stillbirths did not, so under-reporting of maternal deaths is the likely issue, and the fall in the maternal mortality ratio (p. 184) that year is not a real improvement. In the first year three regions are below 5, where stillbirths are more likely to be under-reported. One region is outside the range in five of the six years: review its registers first.

After the assessment

The ratio and its reading go into the mortality report (Reports (p. 415)), next to the levels and the completeness of facility death reporting (p. 193), which puts a number on the under-reporting.

14.6 Outliers in mortality data

Reported numbers of deaths far from the usual values, annual or monthly, that may be data entry errors.

An outlier is a reported number of deaths far from the usual values of the same series, such as a year in which a region reports ten times its usual number of maternal deaths. Mortality data are not corrected automatically, so the outliers that are clearly data entry errors are found and replaced by hand before the rates are read.

> **At a glance**
>
> - **Measures:** annual values more than 3 standard deviations from the long-term annual mean, and monthly values more than 5 MAD from the median.
> - **Good result:** no annual value that clearly indicates a data entry error.
> - **In the app:** maternal deaths, stillbirths (fresh, macerated and total) and neonatal deaths in **Outlier Detection** (p. 312); the yearly values in the first card of **Institutional Mortality** (p. 384).

Why

The annual mortality rates are computed using the unadjusted data on reported deaths and births/live births in health facilities. We do not adjust for reporting rates and outliers (as is done for other interventions, in Data adjustment (p. 75)) because it is difficult to adjust maternal deaths and stillbirths, where the number of deaths is small and fluctuating.

It is however advisable to check the data for any extreme outliers in the annual data (e.g., numbers of deaths that clearly indicate data entry errors) and replace these with the average of the surrounding years.

What

Indicator	Rule	Interpretation
Annual outlier	An annual value deviating by more than 3 standard deviations from the long-term annual mean	Inspect it; if it clearly indicates a data entry error, replace it with the average of the surrounding years
Monthly outlier	A monthly value more than 5 times the median absolute deviation (MAD) from the median	The same rule as for extreme outliers (p. 64) in the data quality assessment; with few deaths a month, many normal months are flagged

Formula

$$\left|x_y - \bar{x}\right| > 3\sigma$$
$$\left|x_m - \tilde{x}\right| > 5 \times MAD$$

where

- x_y is the number of deaths reported in a year, $\bar{x}$ the long-term mean of the annual values and σ their standard deviation (p. 504), taken from the other years so that the year checked does not hide itself;

- x_m is the number of deaths reported in a month, $\tilde{x}$ the median (p. 501) of the monthly values and MAD their median absolute deviation (p. 501) (computed as in the data quality assessment (p. 64)).

Example: a region reports 20, 24, 22, 230, 21 and 23 maternal deaths in six years. The other years have a mean of 22 and a standard deviation of 1.6, so anything beyond $22 \pm 3 \times 1.6$, from 17 to 27, is an outlier. The 230 looks like an extra zero: replace it with the average of the surrounding years, $(22 + 21) \div 2 = 21.5$.

Settings

Setting	Default	What it means
Annual bound	3 standard deviations	How far from the long-term annual mean a yearly number must be to be inspected.
Monthly bound	5 × MAD	How far from the median a monthly number must be to count as an outlier.

Interpretation

- **Is there a year with an extreme high (or low) number of deaths in a region?** Assess whether it is a major data entry error. If it is, replace it with the average of the surrounding years; if not (a real event, a facility added to the reporting), keep it and say so.
- **Are monthly values flagged where the numbers are small?** When a district reports none or one death in most months, the median and the MAD are both 0, and every month with a death is flagged. Judge those districts on their annual totals, not on the monthly flags.
- **Do the outliers fall in a year that already looks odd?** Read them with the maternal mortality ratio (p. 182) and the ratio of stillbirths to maternal deaths (p. 198) for that year.

Example: monthly maternal deaths in a district with few deaths

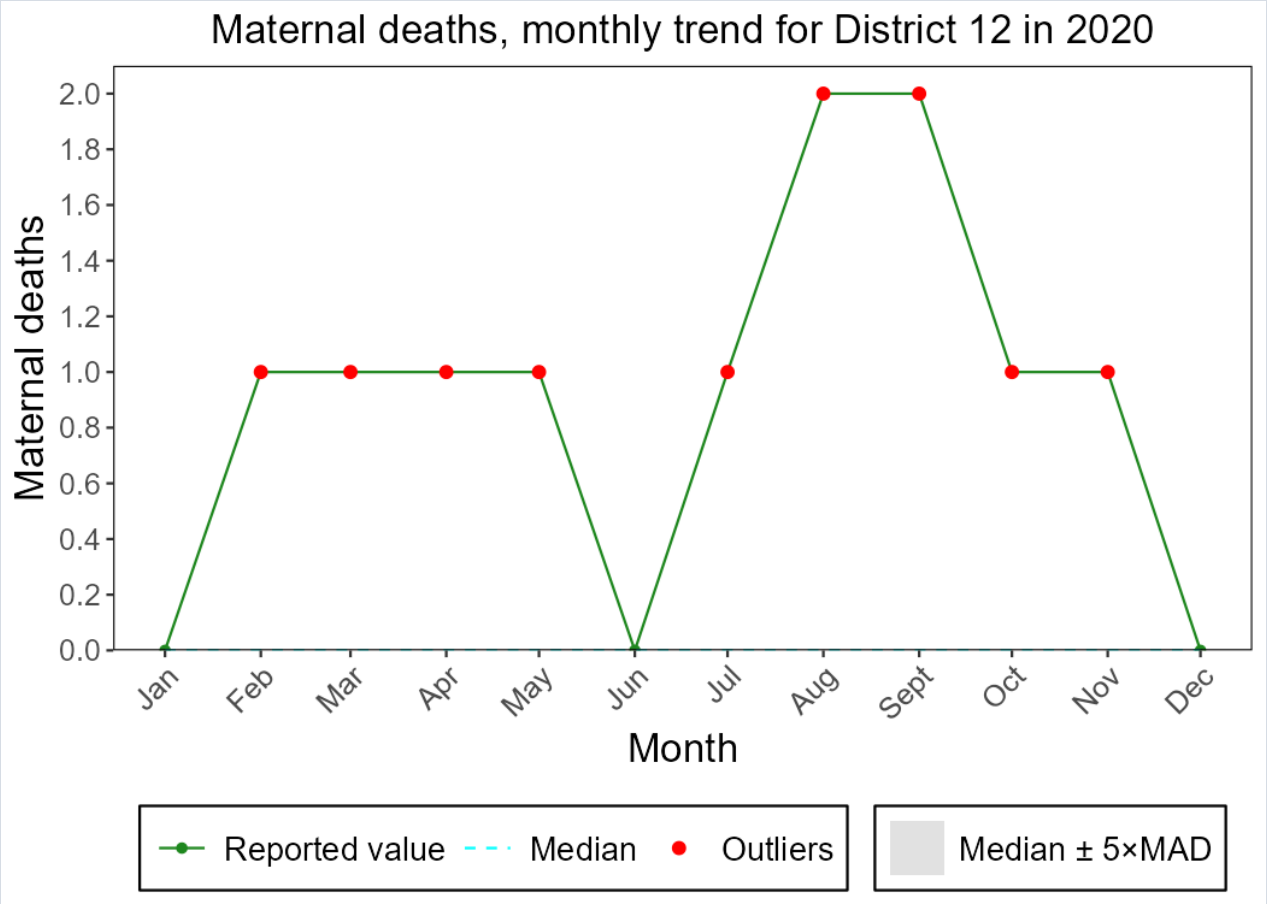


Figure 14.9 — *Monthly maternal deaths for one district in one year, with outliers marked*

Reading the chart: the green line is the number of maternal deaths reported each month by District 12 in the first year; red dots are the months flagged as outliers. The median (dashed line) and the band of median ± 5 × MAD (grey) are both at 0.

What it shows: the district reported 0, 1 or 2 maternal deaths a month. Most of its months have none, so the median and the MAD are 0 and all nine months with a death are flagged. These are not errors: with such small numbers the monthly rule does not work, and the year's total (11 deaths) is what to check against the district's other years. In the example dataset, four or five of the 34 districts have flagged months for maternal deaths in each year, for the same reason.

After the assessment

The replaced annual values are used in the rates of the maternal mortality ratio (p. 182), stillbirth rate (p. 185) and neonatal deaths before discharge (p. 189). No other adjustment is made to mortality data.

CHAPTER 15

Curative health services utilization for sick children

Outpatient visits, admissions and case fatality for children under 5, measuring use of curative services and serving as a proxy for quality of care.

Routine health facility data on outpatient visits and hospital admissions show how far children under 5 reach curative services, and facility case fatality serves as a proxy for the quality of clinical care. This step produces these indicators at national and sub-national level.

It uses the data checked and adjusted in Data adjustment (p. 75), and comes after the Mortality (p. 179) analyses. Its utilization rates are then compared with the inputs of the health system in the next step, Health system performance (p. 225).

At a glance

- **Answers:** how much outpatient and inpatient care do children under 5 receive, and what does facility case fatality suggest about the quality of clinical care?
- **Uses:** OPD (p. 502) visits (new visits and re-visits) and hospital admissions (new admissions and re-admissions), for children under 5 and individuals 5 years and older, adjusted for incomplete reporting and extreme outliers (p. 499) (case fatality uses unadjusted data); deaths in hospital; the DHIS2 (p. 498) projected under-5 population.
- **Produces:** OPD visits per child per year, the under-5 share of OPD visits, admissions per 100 children under 5 per year, the under-5 share of admissions and the pediatric case fatality rate, nationally and by region, and an index of preventive against curative service use by region.
- **In the apps:** Service utilization and health system (p. 391) (**Service Utilization** in the RMNCAH (p. 503) app only).

Why

There is limited routine information regarding curative service utilization, even though diarrhea and pneumonia remain leading causes of child mortality. Service utilization statistics on care-seeking behavior among children with recent illnesses (diarrhea, acute respiratory infection, or fever in the preceding 2 weeks) are traditionally obtained from household surveys (p. 500), which rely on maternal recall.

Routine health facility data on outpatient department (OPD) visits serve as a direct indicator of access to curative services: fewer than one visit per person per year is generally considered an indicator of poor geographic or financial access. Similarly, data on hospital admissions reflect service access, while

institutional facility-based mortality (case fatality) serves as a proxy indicator for the quality of clinical care.

How the parts fit together

As with maternal care, newborn care and immunization, the data quality is assessed first: the OPD and IPD (p. 500) data are checked for completeness of reporting and extreme outliers, then adjusted for incomplete reporting and extreme outliers. The adjusted data are used for the trends and comparisons, nationally and sub-nationally: visits per child and admissions per 100 children measure access, and the under-5 shares of visits and admissions check the consistency of the data. Case fatality is computed from the unadjusted data. The index brings preventive and curative use together for each region, and the utilization rates are compared with health system inputs in Health system performance (p. 225).

Doing it in the app

① Open the Service Utilization pages

They are in the **Service Utilization** group of the **Analysis** section of the RMNCAH app, and need Data adjustment (p. 75) to have been run. See Service utilization and health system (p. 391) for each page and its filters.

② Check the data quality

Open Service Utilization DQA (p. 393) to judge whether the OPD and IPD data are good enough to interpret (Quality of outpatient and inpatient data (p. 207)).

③ Read national and sub-national utilization

National Utilization (p. 395) shows OPD visits, IPD admissions, the under-5 shares, the case fatality rate and deaths, with maps by region; Sub-National Utilization (p. 398) shows the same for a chosen region, and the distribution across Admin 1 (p. 496) regions.

④ Compare preventive and curative use

MCH Preventive vs Curative Index (p. 401) places each Admin 1 region by its use of preventive and curative services.

Points to consider for the interpretation

Outpatient visits

- What can be concluded about the data quality for OPD visits? Is there a high level of consistency in the reported raw numbers between consecutive years?
- What is the number of OPD visits per child per year in each year of the analysis, and is there an upward trend?
- Is the national or sub-national rate lower than 1 visit per year, suggesting low access to services?
- What do the data reveal about OPD visits per child per year across regions or provinces in the latest year? How large is the gap between the highest and the lowest sub-national units?

Inpatient admissions

- Is there internal consistency in the reported raw numbers of admissions and in the admission rates over time?
- What is the number of admissions per 100 children under 5 per year in each year of the analysis?
- Is the rate high or low compared with the benchmarks? What sub-national patterns emerge when comparing admissions per 100 children under 5 across regions or provinces?

Common pitfalls

- Reading utilization trends without checking the data quality first (Service Utilization DQA (p. 207)). If completeness or missingness is poor, treat utilization trends with caution.
- Reading a change in a rate as a change in use without looking at the numbers behind it: a change in the under-5 population used as denominator changes the rate too.
- Reading a year with no deaths reported as a year with no deaths: case fatality can only be computed for the years in which deaths are reported.

15.1 Quality of outpatient and inpatient data

Whether the OPD and IPD data are complete, free of extreme outliers and consistent enough to interpret, year by year.

Before outpatient visits and admissions are interpreted, their data are checked like any other facility data: are the monthly reports complete, are there missing values (p. 502) or extreme outliers (p. 499), and are the numbers for children under 5 consistent with those for all ages? The checks are those of the Data quality assessment (p. 57), applied to the OPD (p. 502) and IPD (p. 500) data, and gathered in one table.

At a glance

- **Measures:** completeness of reporting, missing values and extreme outliers in the OPD and IPD data, and three consistency checks.
- **Good result:** reporting and data quality indicators of 90% or more every year, under-5 shares within their usual range.
- **In the app:** the *Data Quality Metrics* table of **Service Utilization DQA (p. 393)**.

Why

As with maternal care, newborn care and immunization, data quality is systematically assessed first. Adjustments are then applied to correct for incomplete reporting, and extreme outliers are resolved, in Data adjustment (p. 75). This adjusted, clean dataset is then used for the trends and comparisons at national and sub-national level.

Reporting rates from hospitals and facilities with inpatient departments can be more complex to track than routine outpatient data, and extreme monthly outliers are common in inpatient data because of compilation errors: a year can look like a change in use when it is a change in reporting.

What

Indicator	Numerator	Denominator	Interpretation
% of expected monthly facility reports (national)	Number of monthly OPD reports received	Number of monthly OPD reports expected	Reporting rates over 90% are good
% of districts with reporting rates ≥ 90%	Number of districts with at least 90% reporting completeness in the year	Total number of districts	Identifies districts with low reporting
% of non-missing OPD (IPD) monthly values	Number of monthly under-5 OPD (IPD) values that are not missing	Total number of monthly values expected	Shows how much of the year is missing
% of districts with no missing OPD (IPD) values	Number of districts with no missing under-5 OPD (IPD) value in the year	Total number of districts	Shows whether missing values are spread or concentrated
% of OPD monthly values that are not extreme outliers	Number of monthly under-5 OPD values that are not extreme outliers	Total number of monthly values	At least 99% expected; consider reasons

% of districts with no OPD extreme outliers	Number of districts with no extreme outlier in under-5 OPD in the year	Total number of districts	At least 90% expected; consider reasons
Ratio of OPD to IPD, under 5	Number of under-5 OPD visits	Number of under-5 admissions	Usually between 16 and 47
% of OPD visits that are children under 5	Number of under-5 OPD visits	Number of OPD visits, all ages	Usually 15% to 45% (see Outpatient service utilization (p. 210))
% of admissions that are children under 5	Number of under-5 admissions	Number of admissions, all ages	Usually 10% to 40% (see Inpatient service utilization (p. 215))

How each check is computed (reporting rate, missing values, the extreme outlier rule) is explained in Completeness of monthly facility reporting (p. 59), Missing monthly values (p. 62) and Extreme outliers (p. 64).

Settings

Setting	Default	What it means
District threshold	90%	A district reports well when its reporting rate is at least this. It is set on the **Reporting Rate** page of the data quality assessment.
Extreme outlier bound	5 × MAD	A monthly value is an extreme outlier when it is more than 5 median absolute deviations (p. 501) from the district's median (p. 501).

Interpretation

- **Are reporting and the other data quality indicators at 90% or more every year?** If not, note the years and treat their utilization rates with caution.
- **Are IPD values missing in some districts or years?** Missing admissions lower the admission rates and raise the ratio of OPD to IPD; check them before reading the inpatient (p. 215) results.
- **Is the ratio of OPD to IPD for children under 5 between 16 and 47?** A much higher ratio means few admissions for the number of visits: under-reported admissions, or missing inpatient reports.
- **Are the under-5 shares of OPD visits and of admissions in their usual range, and stable between years?** A change of more than 5 percentage points between years flags a data quality problem.

Example: the data quality table, national

	Data Quality Metrics	2020	2021	2022	2023	2024	2025
1. Completeness of reporting							
1a	% expected monthly facility reports (National)	94	97	99	100	100	99
1b	% districts with reporting rates >=90	76	91	94	100	100	100
1c	% non-missing OPD monthly values	100	100	100	100	100	100
1d	% non-missing IPD monthly values	100	100	100	100	100	97
1e	% districts with no missing OPD values	100	100	100	100	100	100
1f	% districts with no missing IPD values	100	100	100	100	100	85
2. Extreme outliers							
2a	% OPD monthly values not extreme outliers	100	100	100	100	100	100
2b	% districts with no OPD extreme outliers	94	100	100	97	100	91
3. Service DQA indicators							
3a	Ratio OPD/IPD under 5	136.88	609.21	58.50	47.48	35.97	48.54
3b	% OPD under 5	30	30	31	30	28	29
3c	% IPD under 5	12	3	26	22	19	16

Reference: ○ 3a 16-47 ○ 3b and 3c 15 - 45%

Figure 15.1 — *Service utilization data quality table by year: reporting, missing values, extreme outliers and consistency checks*

Reading the chart: one column per year. Rows 1a to 1f are completeness of reporting and missing values, 2a and 2b extreme outliers, 3a to 3c the consistency checks (3a the ratio of OPD to IPD for children under 5, 3b and 3c the under-5 shares of OPD visits and admissions). Cells are green at 90% or more, yellow from 70% to 90% and red below 70%. The footnote gives the reference ranges of the consistency checks as the table shows them.

What it shows: reporting is complete (94% to 100% of expected reports), except that only **76%** of districts reached the 90% threshold in the first year. OPD values are never missing and almost never extreme outliers. IPD values are complete until the last year, when **3%** are missing and only **85%** of districts have no missing IPD value. The consistency checks show the real problem: the ratio of OPD to IPD for children under 5 is **137** and **609** in the first two years, far above 16 to 47, and only **12%** and **3%** of admissions are children under 5. Under-5 admissions are under-reported in those two years, so their inpatient rates cannot be compared with later years. The under-5 share of OPD visits is stable at **28% to 31%**.

After the assessment

The OPD and IPD data are adjusted for incomplete reporting and extreme outliers in Data adjustment (p. 75), like the other indicators. For inpatient data, whether to adjust for incomplete reporting is the country team's judgement (see Inpatient service utilization (p. 215)). Years with a weak table are read with caution in the outpatient (p. 210) and inpatient (p. 215) analyses.

15.2 Outpatient service utilization

The mean number of OPD visits per child under 5 per year, and the share of OPD visits contributed by children under five.

Outpatient department (OPD (p. 502)) visits are the consultations of patients who go home the same day. The number of visits per child per year shows how far children under 5 reach curative services, and the share of all visits made by children under 5 checks the consistency of the data. The visits include both new visits and re-visits, and are usually reported for two groups: children under 5 years of age, and individuals 5 years and older.

At a glance

- **Measures:** OPD visits per child under 5 per year (p. 502), and the percent of OPD visits contributed by children under five.
- **Good result:** a stable rate (a change of less than 0.2 visits per child between years), at least 1 visit per child per year, and an under-5 share between 15% and 45%.
- **In the app:** the **OPD Visits** and **Under 5** tabs and the **OPD Map** of **National Utilization** (p. 395), and the **OPD Visits** tab of **Sub-National Utilization** (p. 398).

Why

Routine health facility data on OPD visits serve as a direct indicator of access to curative services: fewer than one visit per person per year is generally considered an indicator of poor geographic or financial access. Care-seeking for sick children is otherwise known only from household surveys (p. 500), every few years and from maternal recall.

What

Indicator	Numerator	Denominator	Interpretation
Mean number of OPD visits per child per year	Adjusted number of under-5 OPD visits (new and re-visits) in the calendar year	Projected population of children under 5 (DHIS2 (p. 498) projections)	Expected to change by less than 0.2 visits per child between years; fewer than 1 visit per child per year indicates that access barriers are likely
Percent of OPD visits contributed by children under five	Number of under-5 OPD visits	Number of OPD visits, all ages	Usually 15% to 45%; in high-fertility settings (a total fertility rate (p. 504) above 4) often over 30%; outside the range, or a change of more than 5 percentage points between years, flags a data quality problem

The rate is expected to remain stable over time. A gradual, steady increase points either to genuine improvements in physical access to OPD services or to an increasing disease burden among children. While there is no rigid global cut-off, a rate of fewer than 1 visit per child per year indicates that access barriers are likely restricting service utilization. OPD statistics are calculated and tracked at national and regional or provincial level. The app also shows the visits per person per year for all ages, with the total population as denominator.

Formula

$$V = \frac{O_5}{P_5}$$

$$S = \frac{O_5}{O} \times 100$$

where

- V is the number of OPD visits per child per year;
- S is the percent of OPD visits contributed by children under five;
- O_5 is the adjusted number of under-5 OPD visits in the year;
- O is the adjusted number of OPD visits in the year, all ages;
- P_5 is the projected population of children under 5.

Worked example: 1,200,000 under-5 OPD visits in a year, out of 4,000,000 visits for all ages, with 800,000 children under 5, give 1,200,000 ÷ 800,000 = **1.5 visits per child per year**, and 1,200,000 ÷ 4,000,000 × 100 = **30%** of visits made by children under 5.

Settings

Setting	Default	What it means
Completeness adjustment factor (k)	0.25	How many services are assumed in the facilities that did not report, when the OPD numbers are adjusted for incomplete reporting. Chosen in Data adjustment (p. 75).

Interpretation

- **Is the data quality for OPD visits good, and are the reported raw numbers consistent between consecutive years?** Check the data quality table (p. 207) first.
- **Does the rate change by less than 0.2 visits per child between years?** If not, look at the numbers of visits and at the under-5 population: a jump in either changes the rate.
- **Is there an upward trend?** A gradual, steady increase points to better access or to more illness among children.
- **Is the national or sub-national rate lower than 1 visit per child per year?** If so, access barriers are likely restricting use.
- **How large is the gap between the highest and the lowest regions** in the latest year?
- **Is the under-5 share between 15% and 45%, and stable?** If not, consider a data quality problem.

Example: OPD visits per person per year, national

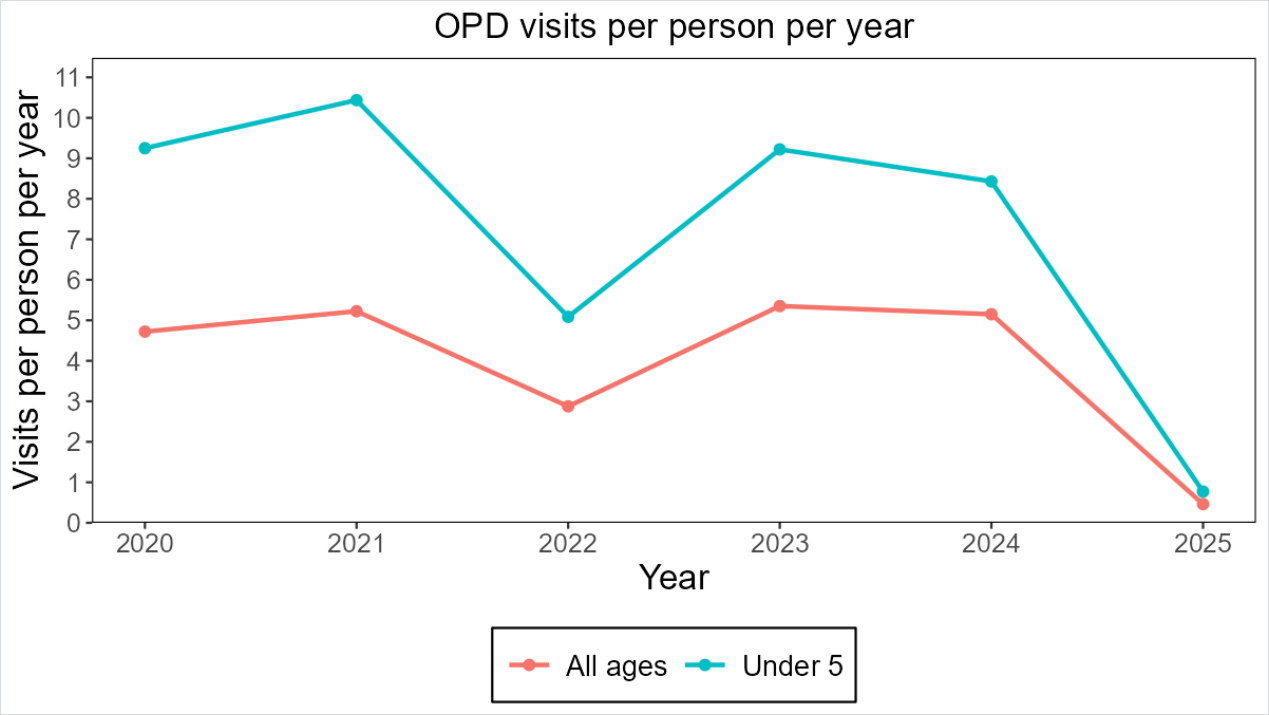


Figure 15.2 — *OPD visits per person per year, under 5 and all ages, by year*

Reading the chart: one line for children under 5 (visits per child) and one for all ages (visits per person), year by year.

What it shows: the under-5 rate swings between **10.4** and **0.8** visits per child, and the all-ages rate between **5.4** and **0.5**: far more than 0.2 between years. The reported visits do not swing: they stay between about 1.4 and 1.8 million for children under 5 every year. The swings come from the population used as denominator, which jumps from about 1 million to 2 million in the third year and to 13 million in the last year. Before reading a trend here, the population data must be checked (see Denominators (p. 95)): only the last year, at **0.8** visits per child, agrees with the regional figures below.

Example: OPD visits per child by region

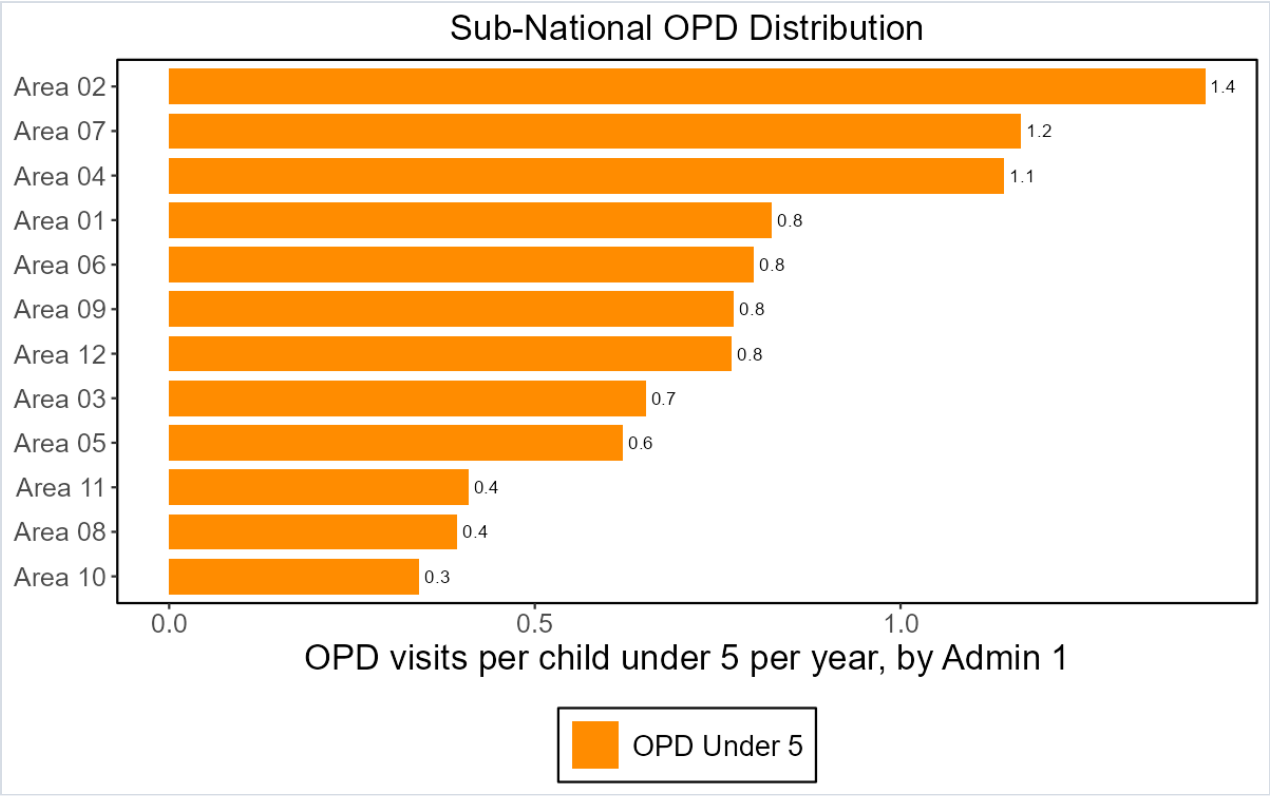


Figure 15.3 — *OPD visits per child under 5 per year by Admin 1 region, latest year*

Reading the chart: one bar per Admin 1 (p. 496) region for the latest year, sorted from the highest rate to the lowest, with the value at the end of each bar.

What it shows: only three regions reach 1 visit per child per year (**1.4**, **1.2** and **1.1**); the other nine are between **0.3** and **0.8**, where access barriers are likely. The gap between the highest and the lowest region is **1.1** visits per child.

Example: under-5 share of OPD visits and admissions

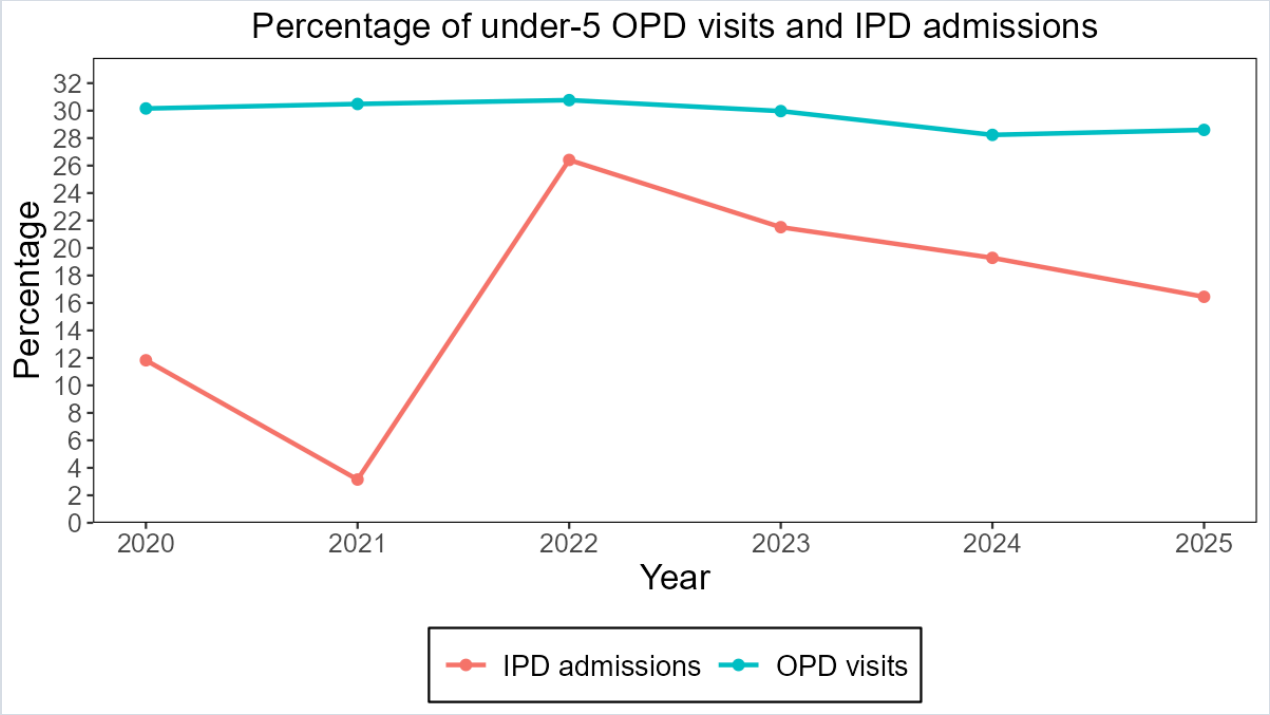


Figure 15.4 — *Percentage of OPD visits and IPD admissions that are children under 5, by year*

Reading the chart: the share of OPD visits (blue) and of admissions (red) made by children under 5, year by year.

What it shows: children under 5 make **28% to 31%** of OPD visits every year, inside the usual 15% to 45%, and the share never changes by more than 5 points: the OPD data are consistent. The share of admissions is discussed in Inpatient service utilization (p. 216).

After the analysis

The OPD visits per child are compared with health workforce density across regions in Health systems outputs by inputs (p. 235), and enter the curative service use index in the MCH preventive vs curative index (p. 222).

15.3 Inpatient service utilization

The number of hospital admissions per 100 children under 5 per year, and the share of admissions contributed by children under five.

Inpatient department (IPD (p. 500)) data count the patients admitted to a hospital or a facility with beds. The number of admissions per 100 children under 5 per year shows access to inpatient care, and the share of all admissions made by children under 5 checks the consistency of the data. The admissions include both new admissions and re-admissions, and are reported for children under 5 and for individuals 5 years and older.

At a glance

- **Measures:** admissions per 100 children under 5 per year (p. 496), and the percent of admissions contributed by children under five.
- **Good result:** at least 2 admissions per 100 children per year, a change of less than 1 between years, and an under-5 share between 10% and 40%.
- **In the app:** the **IPD Admissions** and **Under 5** tabs and the **IPD Map** of **National Utilization** (p. 395), and the **IPD Admissions** tab of **Sub-National Utilization** (p. 398).

Why

Data on hospital admissions reflect access to inpatient services for sick children, as OPD visits reflect access to outpatient services.

Where available, patient discharge (p. 498) data are the preferred input:

$$\text{Discharges} = \text{Admissions} - \text{Deaths}$$

A review of the completeness of reporting and of extreme outliers is used to evaluate the quality of inpatient data (see Quality of outpatient and inpatient data (p. 207)). Because reporting rates from hospitals and facilities with inpatient departments can be more complex to track than routine outpatient data, the decision to adjust for incomplete reporting relies on the judgement of the country team about how reliable the registers are. Extreme monthly outliers are common in inpatient data because of compilation errors, so corrections must be applied cautiously: assess and compare the unadjusted and adjusted results side by side.

What

Indicator	Numerator	Denominator	Interpretation
Number of admissions per 100 children under 5 per year	Number of under-5 admissions in the year × 100	Projected population of children under 5	Fewer than 2 indicates very low access to inpatient services; the median across sub-Saharan African countries for 2018-2022 was 4.5; a change of 1 or more between years is unlikely unless there is an epidemic or outbreak
Percent of admissions contributed by children under five	Number of under-5 admissions	Number of admissions, all ages	Usually 10% to 40%; outside the range, a data quality problem or an exceptional epidemiological event; a change of more than 5 percentage points between years flags a data quality problem

The app also shows the admissions per 100 persons per year for all ages, with the total population as denominator.

Formula

$$A = \frac{I_5}{P_5} \times 100$$
$$S = \frac{I_5}{I} \times 100$$

where

- A is the number of admissions per 100 children under 5 per year;
- S is the percent of admissions contributed by children under five;
- I_5 is the number of under-5 admissions in the year;
- I is the number of admissions in the year, all ages;
- P_5 is the projected population of children under 5.

Worked example: 32,000 under-5 admissions in a year, out of 160,000 admissions for all ages, with 800,000 children under 5, give 32,000 ÷ 800,000 × 100 = **4 admissions per 100 children per year**, and 32,000 ÷ 160,000 × 100 = **20%** of admissions made by children under 5.

Settings

Setting	Default	What it means
Completeness adjustment factor (k)	0.25	How many admissions are assumed in the facilities that did not report, when the IPD numbers are adjusted for incomplete reporting. Chosen in Data adjustment (p. 75), where the country team decides whether inpatient data are adjusted at all.

Interpretation

- **Is there internal consistency in the reported raw numbers of admissions and in the admission rates over time?** A change of 1 or more admissions per 100 children between years is unlikely without an epidemic or outbreak: look for a reporting change first.
- **Is the rate below 2 admissions per 100 children per year?** If so, access to inpatient services is very low. Compare it with the sub-Saharan median of 4.5.
- **What sub-national patterns emerge?** A region with almost no admissions may have missing inpatient reports rather than no admissions: check its data.
- **Is the under-5 share of admissions between 10% and 40%, and stable?** If not, consider a data quality problem or an exceptional epidemiological event. The share for the example is in the chart of under-5 shares on the outpatient page (p. 214): **3% to 26%**, with a jump of 23 points between the second and third years, so under-5 admissions were under-reported in the first two years.
- **Do the adjusted and unadjusted results differ much?** If so, discuss which better reflects the registers before using either.

Example: admissions per 100 persons per year, national

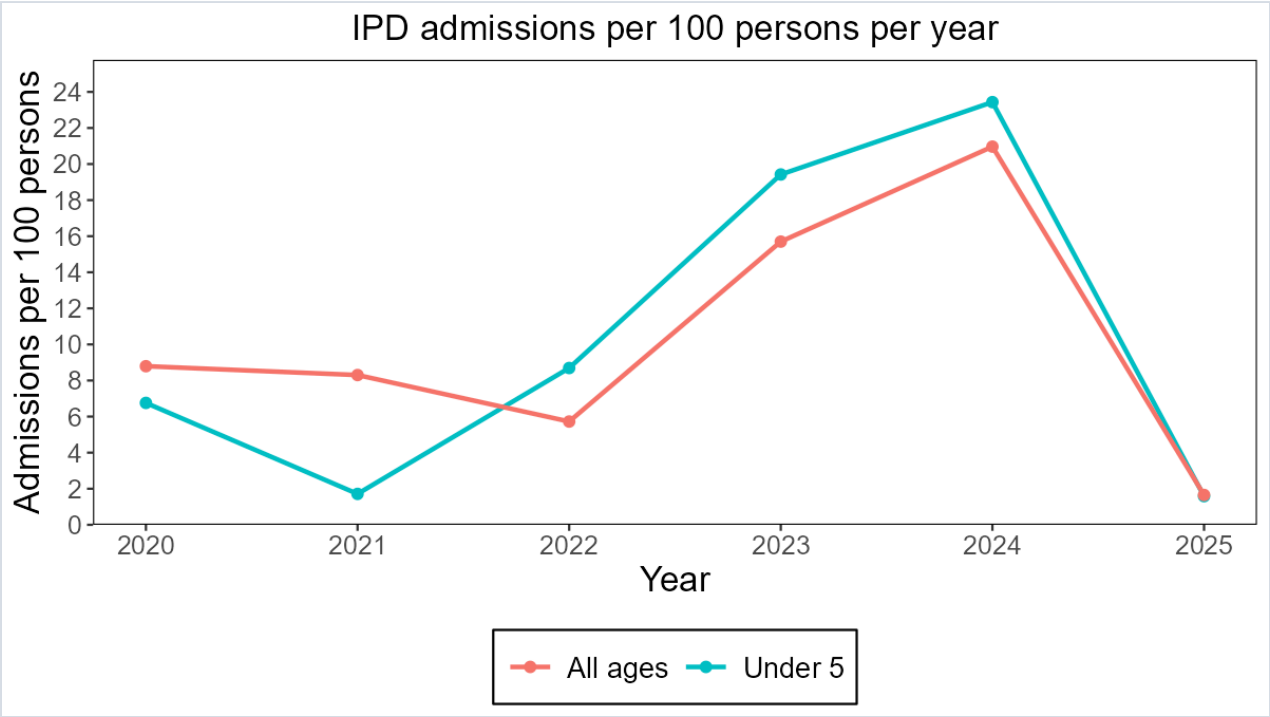


Figure 15.5 — *IPD admissions per 100 persons per year, under 5 and all ages, by year*

Reading the chart: one line for children under 5 (admissions per 100 children) and one for all ages (admissions per 100 persons), year by year.

What it shows: the under-5 rate goes from **6.8** to **1.7**, then up to **23.4**, and down to **1.6** in the last year: changes of far more than 1 per year. Two things are behind it. The reported under-5 admissions jump from about 2,700 in the second year to 30,000 in the third, which is under-reporting in the first years (see the data quality table (p. 209)); and the population used as denominator changes from year to year, as for OPD visits (p. 212). None of these years can be read as a trend in access.

Example: admissions per 100 children by region

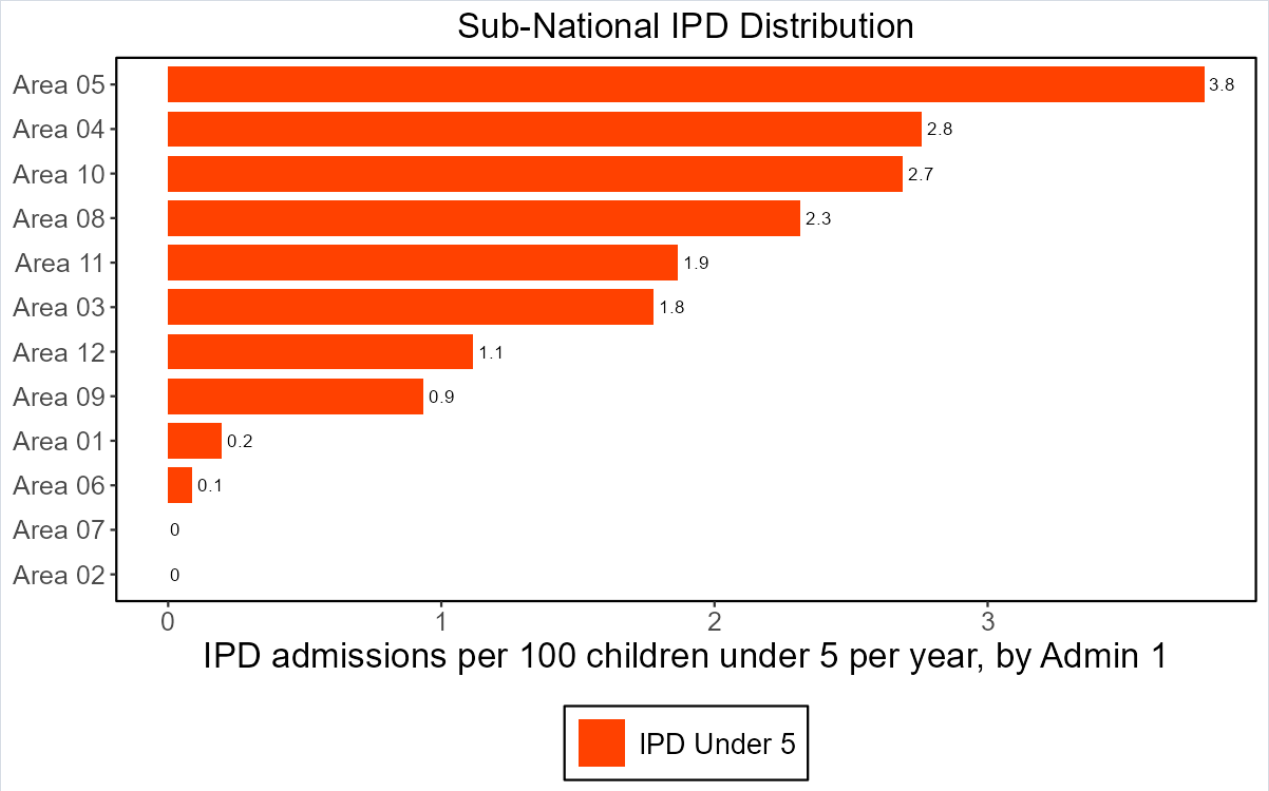


Figure 15.6 — *IPD admissions per 100 children under 5 per year by Admin 1 region, latest year*

Reading the chart: one bar per Admin 1 (p. 496) region for the latest year, sorted from the highest rate to the lowest, with the value at the end of each bar.

What it shows: no region reaches the sub-Saharan median of 4.5; the highest is **3.8** admissions per 100 children. Eight of the twelve regions are below 2, which means very low access. Two regions show **0** and two **0.1** to **0.2**: in the last year 15% of districts have missing IPD values, so these are more likely missing reports than regions with no admissions, and need checking before the regions are compared.

After the analysis

The admissions per 100 children are compared with hospital and bed density across regions in Health systems outputs by inputs (p. 235), enter the curative service use index in the MCH preventive vs curative index (p. 222), and are the denominator of the case fatality rate (p. 219).

15.4 Case fatality rate

The percentage of children admitted to hospital who die there, a proxy for the quality of clinical care.

The pediatric case fatality rate (p. 502) is the share of the children admitted to a hospital or a facility with beds who die there. It serves as a proxy indicator for the quality of clinical care in the wards: the lower the rate, the better the care.

At a glance

- **Measures:** the percentage of under-5 admissions (and of admissions of all ages) that end in death in the facility.
- **Good result:** a low rate, stable or falling over time, computed from years in which deaths are reported.
- **In the app:** the **Case Fatality Rate** and **Deaths** tabs of **National Utilization** (p. 395) and **Sub-National Utilization** (p. 398).

Why

Institutional facility-based mortality (case fatality) serves as a proxy indicator for the quality of clinical care. A lower case fatality rate within health facilities corresponds to a higher standard of acute clinical care.

The case fatality rate is computed from **unadjusted data only**: no adjustment for incomplete reporting and no correction of extreme outliers is applied to the number of deaths in the facility or to the number of admissions.

What

Indicator	Numerator	Denominator	Interpretation
Pediatric case fatality rate	Number of children who die in hospital	Total number of admissions (discharges (p. 498) + deaths)	A lower rate means a higher standard of clinical care
Proportion of deaths that are children under 5	Number of under-5 deaths in the facility	Number of deaths in the facility, all ages	Shows how much of facility mortality concerns young children

The app shows the case fatality rate for children under 5 and for all ages.

Formula

$$F = \frac{D_5}{I_5} \times 100$$

where

- F is the pediatric case fatality rate, in percent;
- D_5 is the number of children under 5 who die in hospital in the year;
- I_5 is the total number of under-5 admissions in the year: discharges + deaths.

Both numbers are unadjusted.

Worked example: 600 children die in hospital out of 20,000 under-5 admissions in a year: 600 ÷ 20,000 × 100 = **3%**.

Interpretation

- **Are deaths reported in every year?** A year with no deaths reported has a case fatality rate of 0 that means nothing: leave it out.
- **Is the rate falling over time?** A falling rate suggests better clinical care, if admissions and deaths are reported in the same way every year.
- **Does a change follow a change in admissions?** Since the denominator is admissions, under-reported admissions raise the rate; check the inpatient data (p. 215) first.
- No threshold is given for what counts as a high case fatality rate: compare regions and years with each other, and with what the country team knows of its hospitals.

Example: case fatality rate, national

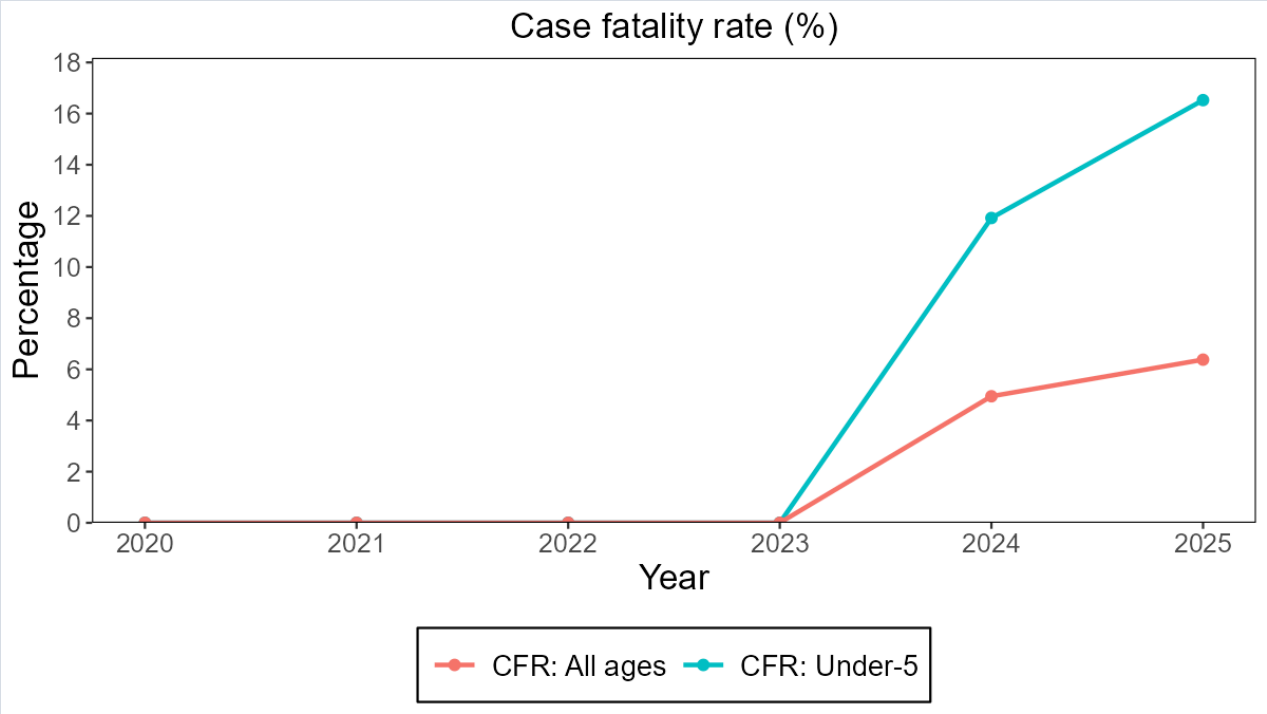


Figure 15.7 — *Case fatality rate (%) for children under 5 and all ages, by year*

Reading the chart: the percentage of admissions that end in death, for children under 5 and for all ages, year by year.

What it shows: the rate is **0** for the first four years: no deaths were reported in those years, so these points cannot be interpreted. In the last two years, **11.9%** and then **16.5%** of the children under 5 admitted died, against **4.9%** and **6.4%** for all ages. Before concluding that care for sick children worsened, check that the deaths and admissions counted come from the same facilities and wards, as the rise is large for two years.

After the analysis

The case fatality rate is reported with the other utilization results, nationally and by region, and read with the admissions (p. 215) it is based on. The quality of clinical care it points to is one of the questions discussed in Health system performance (p. 225).

15.5 MCH preventive vs curative index

Each region placed by its use of preventive maternal and child health services and of curative services for children under 5.

The index places each Admin 1 (p. 496) region by two numbers: its use of preventive maternal and child health (MCH) services, the MCH preventive services index (p. 501), and its use of curative services for children under 5, the curative service use index (p. 498). Read together, they show where people use both kinds of services, only one, or neither.

At a glance

- **Measures:** for each region in the latest year, preventive services and curative services per child under 5.
- **Good result:** regions with high preventive use; the regions with low preventive and low curative use are the priority.
- **In the app:** the scatter plot of **MCH Preventive vs Curative Index** (p. 401).

Why

Preventive services (antenatal care, delivery care, postnatal care, immunization) and curative services (outpatient visits and admissions for sick children) are delivered by the same facilities to the same families. Comparing the two shows whether a region's families reach the health system mainly for prevention, mainly when a child is ill, or rarely at all, which points to different actions.

What

Indicator	What it counts	Per	Interpretation
MCH preventive services index	Antenatal, delivery, postnatal and immunization services, weighted by importance	Child under 5	Higher means more use of preventive services
Curative service use index	OPD visits and IPD admissions of children under 5, an admission counted ten times a visit	Child under 5	Higher means more use of curative services

Median (p. 501) lines divide the regions into four quadrants: low or high preventive, and low or high curative.

Formula

$$M = \frac{A_1 + 3A_4 + N + B + \frac{P_1 + P_3}{2} + P_3 + R_1 + R_2 + 10\,L}{P_5}$$

$$C = \frac{O_5 + 10\,I_5}{P_5}$$

where

- M is the MCH preventive services index and C the curative service use index;
- A_1 and A_4 are the first and fourth antenatal care visits (ANC1 (p. 496) and ANC4);
- N is postnatal care within 48 hours of birth;

- B is BCG (p. 496) vaccination, P_1 and P_3 are penta1 and penta3 (p. 502), R_1 and R_2 the first and second doses of measles vaccine;
- L is institutional deliveries;
- O_5 is the number of under-5 OPD (p. 502) visits and I_5 the number of under-5 IPD (p. 500) admissions;
- P_5 is the projected population of children under 5.

All numbers are the adjusted numbers of the latest year, for each region.

Worked example: a region with 200,000 children under 5, 180,000 under-5 OPD visits and 4,000 under-5 admissions has a curative service use index of (180,000 + 10 × 4,000) ÷ 200,000 = **1.1**.

Interpretation

- **Which regions have low preventive and low curative use?** They are the priority: families there reach the health system rarely, for prevention or for care.
- **Which regions have high curative use with low preventive use?** This may show that people reach facilities only when ill: look at the preventive services there.
- **Which regions have high preventive use with low curative use?** Children are reached for prevention; check whether sick children are brought to the facilities, and whether OPD and IPD are fully reported.
- **Are the regions far from the others real?** A very high or very low index can come from the population used as denominator, or from missing OPD or IPD reports (Quality of outpatient and inpatient data (p. 207)).

Example: preventive against curative index by region

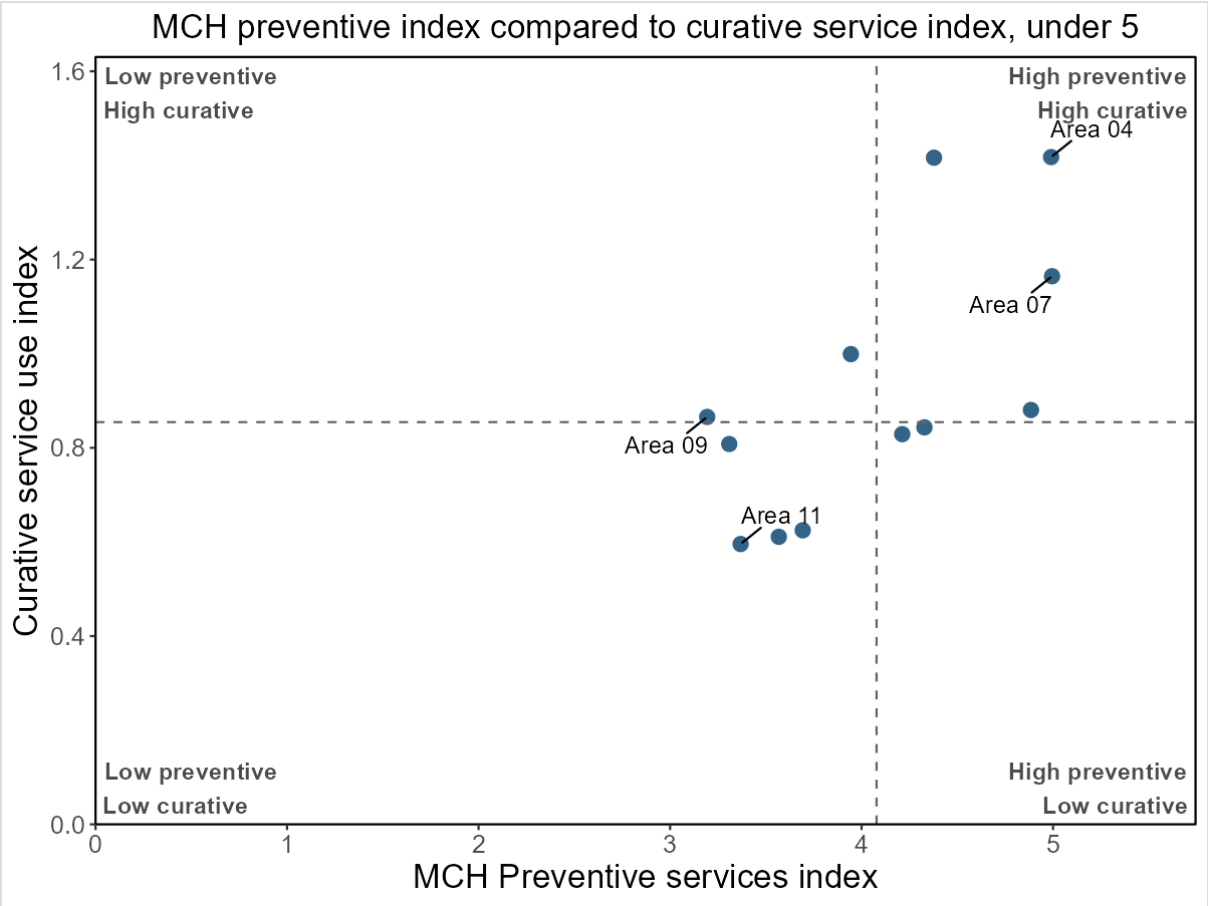


Figure 15.8 — *MCH preventive services index against curative service use index, by Admin 1 region, latest year*

Reading the chart: each dot is a region, with its preventive index across and its curative index up. The dashed lines are the medians of the regions; the corners name the four quadrants. The regions with the highest and lowest value of each index are labelled.

What it shows: the two indices go together: regions with more preventive use also tend to have more curative use. Four regions, among them Area 04 (the highest curative index, about **1.4**) and Area 07 (the highest preventive index, about **5.0**), are high on both. Four regions are low on both, with Area 11 lowest on curative use (about **0.6**): they are the priority. Area 09, the lowest on preventive use (about **3.2**), sits just above the curative median, as does one other region: these may be places where families come mainly when a child is ill.

After the analysis

The two indices are combined into the service index of the PHC Performance charts, and compared with facility and workforce density in Health systems outputs by inputs (p. 235).

CHAPTER 16

Health system performance

Analysis of health system inputs and outputs, such as infrastructure, workforce and financing, to find where inputs fail to translate into coverage.

This step analyses health systems performance: the availability of health services, the health workforce and health financing, compared with the services the system delivers. The analysis aims to identify gaps and challenges in the health system that may affect the delivery of reproductive, maternal, newborn, child, adolescent health and nutrition (RMNCAH-N (p. 503)) services.

It follows Service utilization (p. 204), whose outpatient and inpatient rates it uses as outputs, and comes before Planning for dissemination, decisions and action (p. 245).

At a glance

- **Answers:** where do high health system inputs fail to translate into RMNCAH-N service coverage, and which regions deliver services efficiently?
- **Uses:** health system inputs (p. 499) (infrastructure, workforce and, where available, financing) at national and Admin 1 (p. 496) level; under-5 OPD visits and hospital admissions from Service utilization (p. 204); institutional live birth coverage from a recent household survey (p. 500) or the preferred DHIS2 (p. 498) estimate (National coverage (p. 115)).
- **Produces:** input indicators checked against benchmarks (p. 497), descriptive comparisons of outputs by inputs across Admin 1 units, and the share of private facilities.
- **In the apps:** Service utilization and health system (p. 391) (**Health System Performance** in the RMNCAH app only).

Why

The assessment of the burden of disease, intervention coverage, quality, equity and health system inputs should directly guide policy-making and program targeting. A comprehensive evaluation requires a comparison of health system inputs (such as financing, health workforce density and infrastructure) with the corresponding outputs (p. 499) of the system (service utilization and coverage) at sub-national level (Admin 1).

The objective is to find the sub-national bottlenecks where high system inputs fail to translate into the expected RMNCAH-N service coverage, or, conversely, the high-performing regions that deliver services efficiently.

Analytical scope

The analysis is a direct, descriptive comparison of health system outputs (coverage of tracer interventions) with system inputs (infrastructure, workforce and financing) at sub-national level (Admin 1). More complex methods, such as data envelopment analysis, frontier efficiency modelling, or composite indices controlling for socioeconomic development, are outside its scope.

How the parts fit together

Start with the inputs: check that their data are plausible and see how far each region is from the benchmarks. Then relate the inputs to the outputs region by region: the expected pattern is that more inputs go with more services, and the regions that do not follow it are the ones to discuss. The private sector share adds context, since private facilities are counted in the facility density but may not report their services.

Doing it in the app

1. **Open the Health System Performance pages**
   They are in the **Health System Performance** group of the **Analysis** section of the RMNCAH app, and need Data adjustment (p. 75) to have been run. See Service utilization and health system (p. 391) for each page.

2. **Review the inputs**
   National Health System (p. 403) tabulates the input indicators for the latest year; Sub-National Health System (p. 405) compares each region with the national average and the benchmarks (Health systems inputs (p. 228)).

3. **Relate inputs to service use**
   PHC Performance (p. 407) plots an input against a service index for the Admin 1 regions (Health systems outputs by inputs (p. 235)).

4. **Review the private sector share**
   Public/Private Share (p. 409) shows the ownership of health facilities (Private sector and RMNCAH-N services (p. 242)).

Points to consider for the interpretation

- Are the input data plausible across regions, and consistent with each other (for example, workforce density and bed density)?
- Which regions are far below the benchmarks for facilities, beds and workforce?
- Do regions with more hospitals or beds have more under-5 admissions, and regions with more health workers more under-5 OPD visits and institutional births?
- Which regions have low inputs but high outputs, or high inputs but low outputs, and what explains them, from what is known of the situation in these regions?

Common pitfalls

- Comparing health system densities with different denominators or population years.
- Treating a benchmark as a target for every region. It is a reference, and context matters.
- Interpreting a quadrant position as cause and effect. A comparison of inputs with a service index, as in **PHC Performance**, is descriptive, not proof of efficiency.

16.1 Health systems inputs

The density of health facilities, hospitals, inpatient beds and core health professionals per population, nationally and by region, checked for plausibility and against benchmarks.

Health system inputs (p. 499) are the resources the system works with: health facilities, hospitals and inpatient beds (infrastructure), health workers (workforce), and money (financing). Each is expressed per population, as a density, so that regions of different sizes can be compared with each other and with benchmarks (p. 497).

At a glance

- **Measures:** health facilities and inpatient beds per 10,000 population, hospitals per 100,000, and core health professionals per 10,000, for the latest year.
- **Good result:** plausible values across regions, and densities at or above the benchmarks: 2 facilities, 25 beds and 44.5 health workers per 10,000.
- **In the app:** the table of **National Health System** (p. 403) and the density tabs of **Sub-National Health System** (p. 405).

Why

Inputs are the first half of the comparison of inputs with outputs: before asking whether a region turns its resources into services, one needs to know what resources it has, and whether the figures for them can be trusted. Data on the health workforce in particular are often of poor quality and difficult to obtain.

Data quality of the inputs

First, the assessment focuses on the quality of the data for the health system indicators, at national and sub-national level. For selected indicators, it covers:

- **Global benchmarking:** comparison with global data repositories for selected indicators (national level only).
- **Plausibility:** assessing the plausibility of indicator values across sub-national units, to find major outliers or improbable patterns.

It is also useful to explore the cross-correlations of health system inputs with one another (for example, workforce density against inpatient bed density): this is effective for detecting data inconsistencies across Admin 1 units (provinces, regions or counties). Finally, comparing the health system indicators across administrative levels (for example, Admin 1 totals against the districts within them) helps find structural anomalies, reporting discrepancies or geographical outliers.

Health financing

Health financing indicators at district or Admin 1 level are difficult to obtain and are often limited to budgeted allocations rather than actual expenditures. They also tend to capture government resources only, missing external donor aid, private insurance and out-of-pocket spending.

> Financing data are not loaded by default. If reliable sub-national expenditure data are available to the country team, they should be used to enrich the health system inputs.

What

Indicator	Numerator	Denominator	Interpretation
Core health professionals per 10,000 population (p. 497)	Number of physicians, non-physician clinicians (p. 502), professional nurses and midwives	Total population ÷ 10,000	At least 23 was the 2006 WHO threshold for skilled birth attendance; at least 44.5 is the threshold for universal health coverage (p. 505)
Number of health facilities per 10,000 population (p. 499)	Number of health facilities of all tiers, public and private	Total population ÷ 10,000	Below 2 indicates low access to infrastructure (an indicative reference)
Number of hospitals per 100,000 population (p. 500)	Number of hospitals	Total population ÷ 100,000	Capacity for acute inpatient services; no benchmark
Inpatient bed density per 10,000 population (p. 500)	Number of official inpatient beds in all functioning health facilities	Total population ÷ 10,000	Structural capacity; the benchmark drawn is 25

Core health professionals are physicians (medical doctors); non-physician clinicians (for example clinical officers or health officers with surgical training and several years of certification, but without a full medical degree); professional nurses; and certified midwives. In 2006, WHO set a minimum of **23 core health professionals per 10,000 population** as the threshold needed to achieve the skilled birth attendance that reduces maternal and child mortality. Updated global targets set a higher threshold of at least **44.5 per 10,000** to achieve universal health coverage (p. 505) goals.

Health facilities include all tiers of care: hospitals, comprehensive health centres, and lower-level primary care points such as health posts and dispensaries, in both the private and the public sector. This is a crude indicator, since a small dispensary weighs as much as a tertiary referral hospital. A value of **2 facilities per 10,000 population** is used as an indicative reference: a sub-national unit below 2 has low access to infrastructure.

Hospitals per 100,000 gives insight into the infrastructure available for acute inpatient services, and **inpatient beds per 10,000** is an additional measure of structural capacity.

The national table also shows the share of facilities that are hospitals and the skills mix ratio (p. 504) (nurses and midwives per physician), and the shares of private and NGO facilities (see Private sector and RMNCAH-N services (p. 242)).

Formula

$$W = \frac{H}{T} \times 10{,}000$$
$$F = \frac{N_F}{T} \times 10{,}000$$
$$R = \frac{N_H}{T} \times 100{,}000$$
$$B = N$$

$$\frac{B}{T} \times 10{,}000$$

where

- W, F, R and B are the workforce, health facility, hospital and inpatient bed densities;
- H is the number of core health professionals (physicians + non-physician clinicians + nurses and midwives);
- N_F is the number of health facilities, N_H the number of hospitals and N_B the number of inpatient beds;
- T is the total population.

All are for the latest year of the data. A district with no value for an input is given the median (p. 501) of the districts of its Admin 1 region.

Worked example: a region of 1,500,000 people with 240 health facilities, 6 hospitals, 1,800 inpatient beds and 1,200 core health professionals has 240 ÷ 1,500,000 × 10,000 = **1.6 facilities per 10,000**, 6 ÷ 1,500,000 × 100,000 = **0.4 hospitals per 100,000**, **12 beds per 10,000** and **8 health workers per 10,000**.

Settings

Setting	Default	What it means
Health facility density benchmark	2 per 10,000	Regions at or above it are coloured **Meets benchmark**, below it **Below benchmark**.
Inpatient bed density benchmark	25 per 10,000	The same, for beds.
Workforce density benchmark	44.5 per 10,000	The same, for core health professionals (the UHC threshold).
Hospital density benchmark	None	Hospital density has no benchmark line.

The benchmarks are fixed in the app.

Interpretation

- **How do the national values compare with health system data in the WHO database and in other countries?** A large difference suggests under- or over-reporting.
- **Are the values plausible across regions?** A region far above or below all the others on one input may have a reporting problem; check it before reading it as a real difference.
- **Do the inputs agree with each other?** A region with many health workers but few beds, or the reverse, is worth checking.
- **How far is each region from the benchmarks?** Note the regions below 2 facilities, 25 beds and 44.5 (or 23) health workers per 10,000.
- **How does each region compare with the national average?** The regions furthest below it are where inputs are scarcest.

Example: national health system indicators

2025		
Indicator	**Value**	**Unit**
Health infrastructure		
Health facility density	1.4	per 10,000
Share of facilities that are hospitals	3.1	%
Hospital density	0.4	per 100,000
Inpatient bed density	10.6	per 10,000
Health workforce		
Health workforce density	9.5	per 10,000
Skills mix ratio nurse-midwives per physician	3.1	
Role of private sector		
Share of Private facilities	49.0	%
Share of NGO facilities	0.5	%

Figure 16.1 — *National health system indicators for the latest year: infrastructure, workforce and private sector*

Reading the chart: one row per indicator for the latest year, with its value and unit, grouped into health infrastructure, health workforce and the role of the private sector.

What it shows: the example country has **1.4** health facilities and **10.6** inpatient beds per 10,000, both below their benchmarks (2 and 25), and **0.4** hospitals per 100,000; hospitals are **3.1%** of facilities. The workforce density of **9.5** per 10,000 is far below both the 2006 threshold of 23 and the UHC threshold of 44.5, with about **3** nurses and midwives per physician. Almost half the facilities (**49%**) are private.

Example: health facility density by region

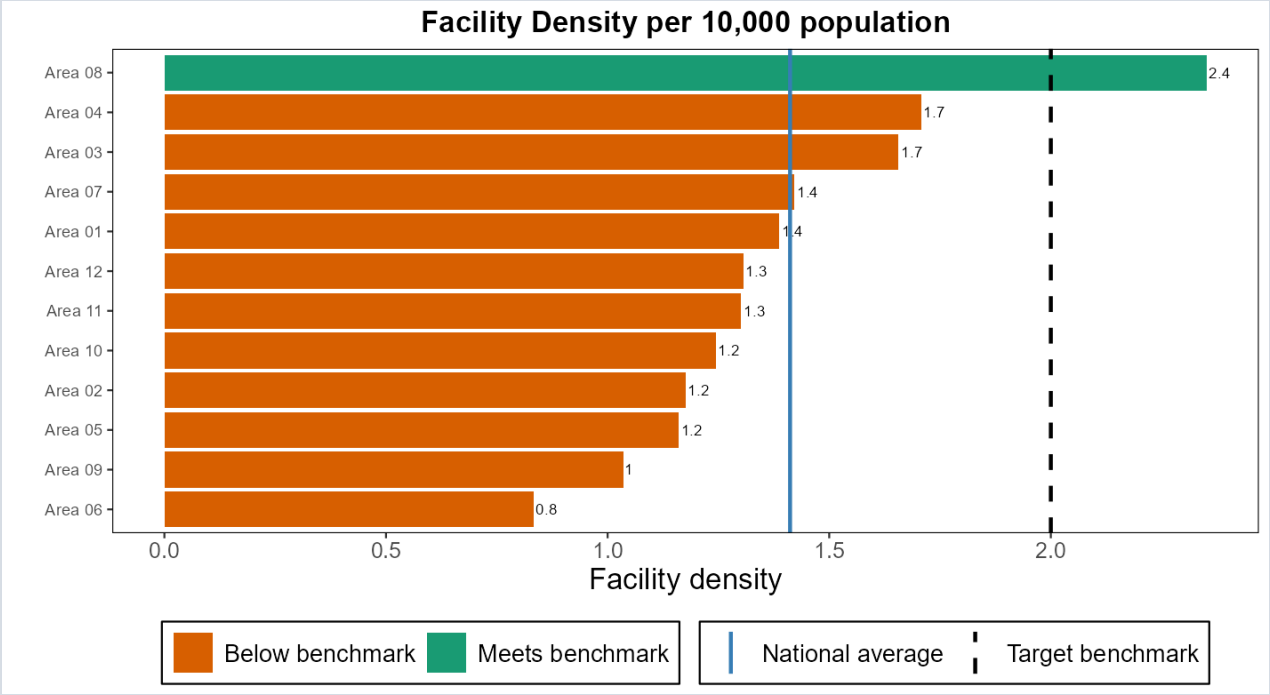


Figure 16.2 — *Health facility density per 10,000 population by Admin 1 region, with national average and benchmark*

Reading the chart: one bar per region, sorted, green when the region meets the benchmark and orange when it is below. The solid blue line is the national average, the dashed black line the benchmark of 2 per 10,000.

What it shows: only Area 08 meets the benchmark, with **2.4** facilities per 10,000; the other regions range from **0.8** to **1.7**, around a national average of 1.4.

Example: hospital density by region

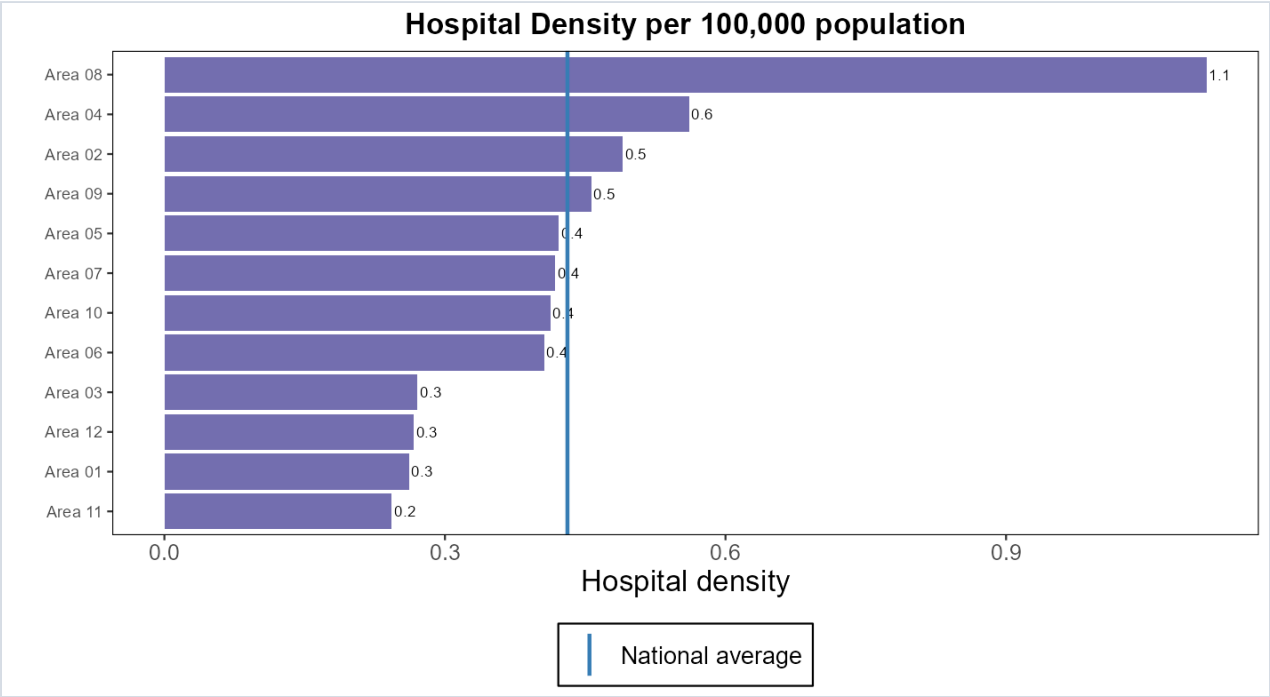


Figure 16.3 — *Hospital density per 100,000 population by Admin 1 region, with national average*

Reading the chart: as above, without a benchmark: all bars have the same colour.

What it shows: Area 08 again stands out, with **1.1** hospitals per 100,000, twice the next region (**0.6**); the others are between **0.2** and **0.5**.

Example: health workforce density by region

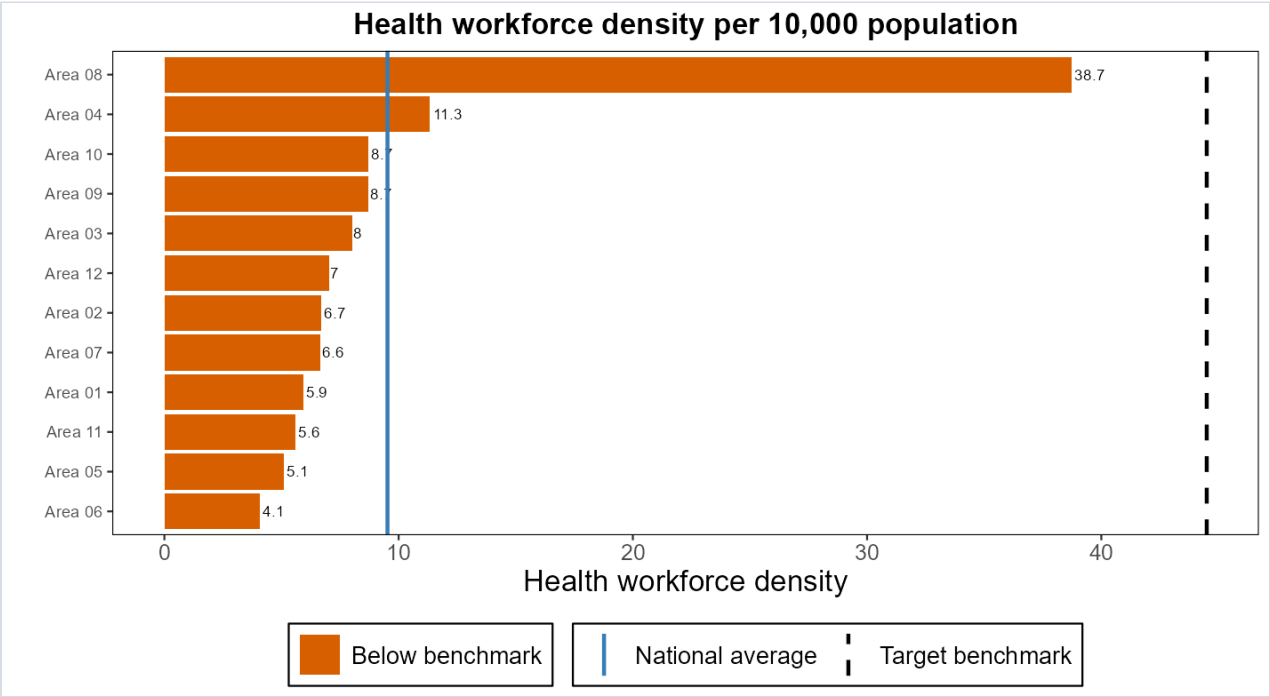


Figure 16.4 — *Health workforce density per 10,000 population by Admin 1 region, with national average and benchmark*

Reading the chart: as for facility density, with the benchmark at 44.5 per 10,000.

What it shows: no region meets the benchmark. Eleven regions have between **4.1** and **11.3** health workers per 10,000; Area 08 has **38.7**, more than three times the next region. Such a gap fails the plausibility check: before reading it as a real concentration of staff (for example in a region with the main referral hospitals), check whether staff are counted where they work, and not counted twice.

Example: inpatient bed density by region

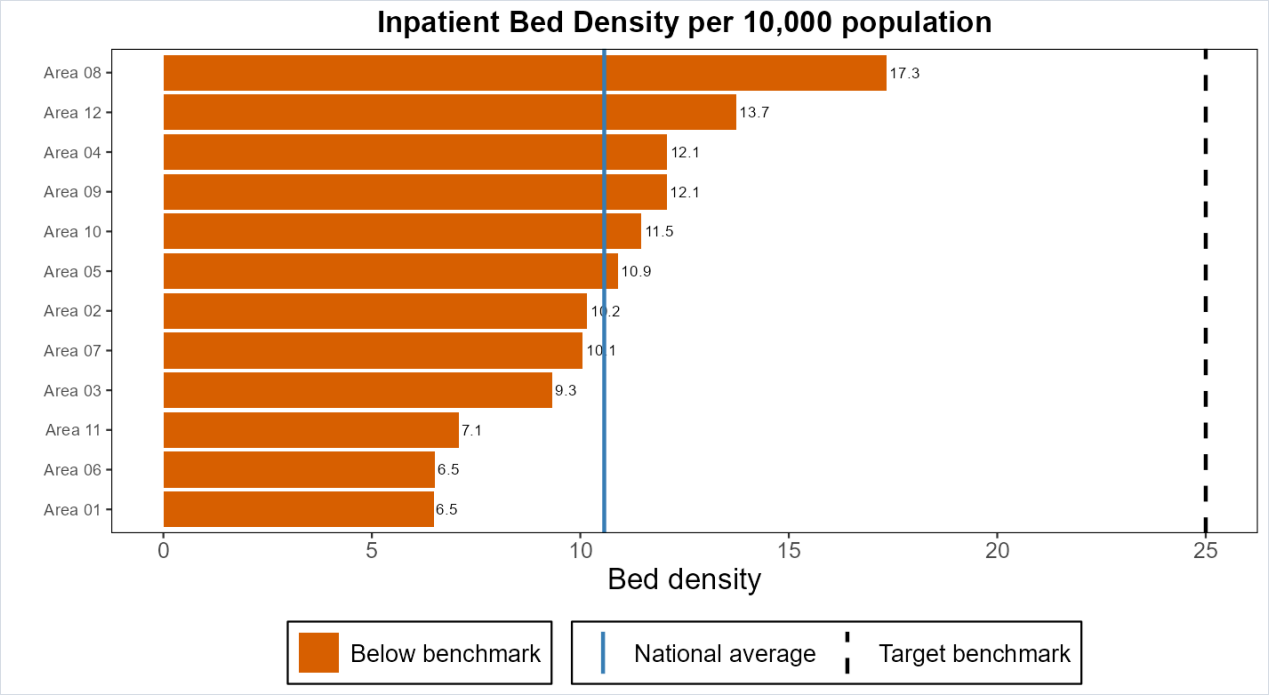


Figure 16.5 — *Inpatient bed density per 10,000 population by Admin 1 region, with national average and benchmark*

Reading the chart: as for facility density, with the benchmark at 25 per 10,000.

What it shows: no region reaches 25 beds per 10,000; the regions range from **6.5** to **17.3**. Area 08 is highest on beds too, which agrees with its facility and hospital densities: the cross-check suggests a better-equipped region, while its workforce figure remains far out of line.

After the analysis

The densities are related to service use and coverage region by region in Health systems outputs by inputs (p. 235).

16.2 Health systems outputs by inputs

Whether regions with more hospitals, beds, facilities or health workers deliver more services, and which regions do not fit the pattern.

The second part of the analysis looks at the association between health system inputs (p. 499) and the matching service utilization or coverage indicators, the outputs (p. 499), across sub-national units (Admin 1 (p. 496)). Each region is a point on a chart with an input on one axis and an output on the other; the regions that do not follow the general pattern are the ones to discuss.

At a glance

- **Compares:** hospital and bed density with under-5 admissions, workforce density with under-5 OPD visits and with institutional live birth coverage, and facility and workforce density with a service index.
- **Good result:** a positive association (more inputs, more services), and an explanation for each region far from it.
- **In the app:** the **Health Facility Density** and **Workforce Density** tabs of **PHC Performance** (p. 407); the other scatter plots of outputs by inputs are in the **Health system performance** standard report of **Reports** (p. 416).

Why

The objective is to find the sub-national bottlenecks where high system inputs fail to translate into the expected RMNCAH-N service coverage, or, conversely, the regions that deliver services efficiently. The comparison is descriptive: it shows where to look, not why.

What

Association	Input	Output	Expected
Inpatient infrastructure and under-5 hospital admissions	Hospitals per 100,000 population (p. 500), or inpatient beds per 10,000 population (p. 500)	Admissions per 100 children under 5 per year (p. 496)	Regions with lower hospital or bed density have lower admission rates, and regions with higher capacity higher rates: a positive slope of the linear fit (p. 500)
Health workforce density and under-5 outpatient utilization	Core health professionals per 10,000 population (p. 497)	OPD visits per child under 5 per year (p. 502)	Regions with lower workforce density have lower OPD use, and regions with higher density higher use: a positive slope
Health workforce density and institutional live birth coverage	Core health professionals per 10,000 population	Institutional live birth coverage, from a recent household survey (p. 500) or the preferred DHIS2 (p. 498) estimate	Regions with lower workforce density have lower institutional delivery coverage: a positive slope
PHC performance	Health facilities per 10,000, or core health professionals per 10,000	The PHC service index (p. 502)	Regions are compared in four quadrants: high or low density, high or low service index

Inpatient infrastructure and under-5 hospital admissions

When reviewing the scatter plot, look for major outliers, such as regions with very low infrastructure density but unusually high admission rates for children. Possible explanations include:

- **Infrastructure under-reporting:** hospitals or beds are under-reported in the system by these regions.
- **High clinical throughput:** the existing hospitals in these regions have exceptionally high admission turnover and bed occupancy.
- **Numerator inflation:** hospital admissions are over-reported or double-counted in these areas.

The narrative should be based on knowledge of the actual situation in these sub-national units.

Health workforce density and under-5 outpatient utilization

If a region has low health workforce density with high OPD use, the country team should consider whether:

- **Workforce under-reporting:** the number of core health professionals is under-reported by these regions.
- **High staff workload:** health workers in these areas have an exceptionally high clinical workload and patient volume.
- **Numerator inflation:** OPD visits are over-reported because of data quality or documentation issues.

The interpretation must reconcile these data with what is known of the affected regions.

Health workforce density and institutional live birth coverage

For delivery coverage, use data from a recent household survey or the preferred DHIS2 facility-based estimate (see National coverage (p. 115)). The considerations for the outliers are the same as above.

PHC performance

The PHC Performance charts relate facility density or workforce density to a service index for each region. Dashed median (p. 501) lines split the chart into four quadrants, for example high density with a low service index. The regions with a service index in the lowest or highest quarter of the regions are labelled.

Formula

$$S = \frac{M + C}{2}$$

where

- S is the PHC service index of a region;
- M is its MCH preventive services index (p. 501) and C its curative service use index (p. 498) (see MCH preventive vs curative index (p. 222)).

On the scatter plots, the line is the linear fit through the regions, and the R^2 (p. 503) under the chart says how closely the regions follow it, from 0 (not at all) to 1 (exactly).

Interpretation

- **Is the slope positive, as expected?** If the regions with more inputs do not have more services, look at the regions that break the pattern.
- **Which regions have low inputs but high outputs?** Consider input under-reporting, high throughput or workload, or an inflated numerator.
- **Which regions have high inputs but low outputs?** Consider an over-reported input, under-reported services, or resources that do not reach the population.
- **Is the association driven by one region?** With twelve or so regions, one region far from the others can make or break the line: check it first, as in Health systems inputs (p. 230).
- **Which quadrant is each region in?** High density with a low service index points to a bottleneck; low density with a high service index to efficient delivery, or to under-counted inputs. A quadrant is a description, not proof of efficiency.

Example: under-5 admissions by hospital density

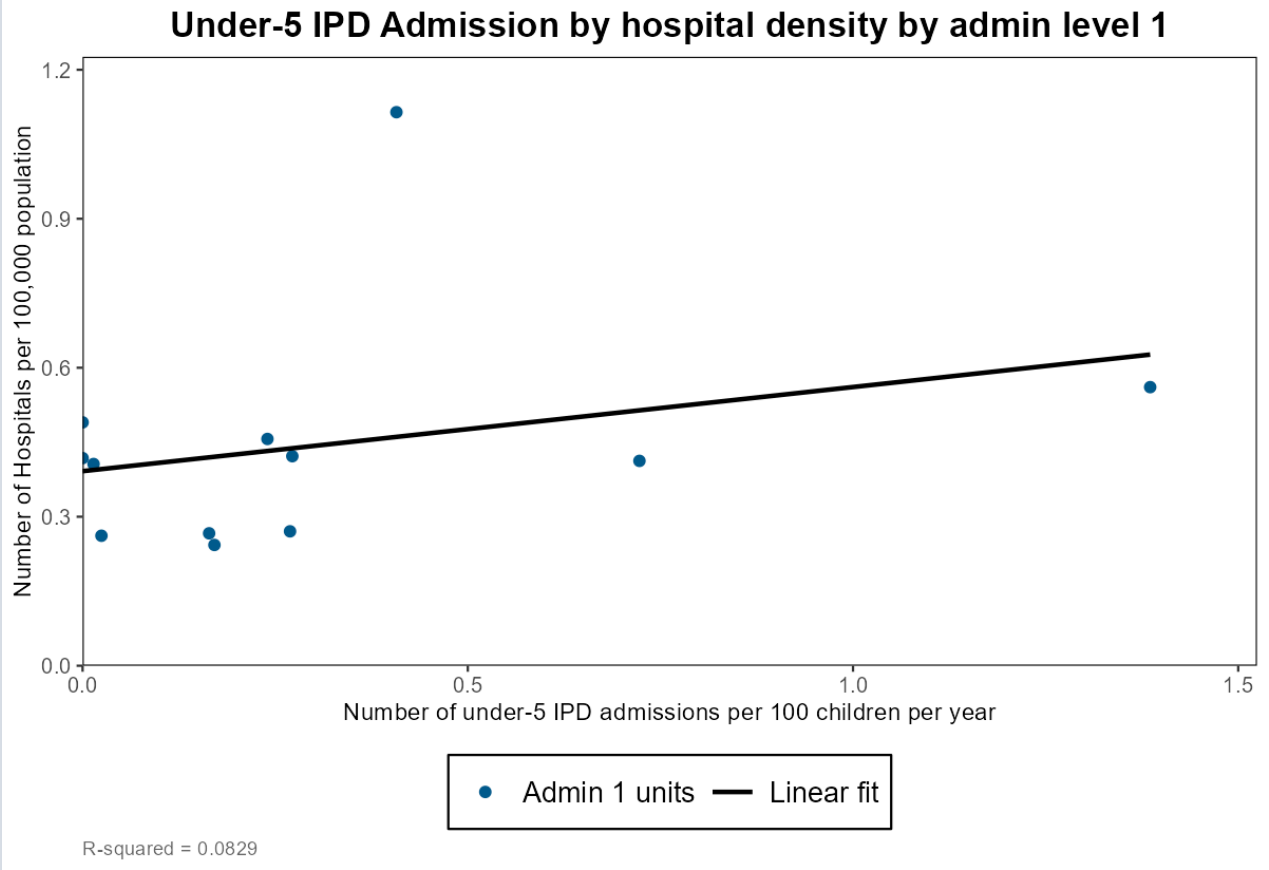


Figure 16.6 — *Hospitals per 100,000 population against under-5 admissions per 100 children, by Admin 1 region, with linear fit*

Reading the chart: each dot is a region, with its under-5 admissions across and its hospital density up. The black line is the linear fit; the R^2 is under the chart.

What it shows: the slope is positive but weak (**$R^2 = 0.08$**). Several regions have almost no admissions whatever their hospital density: they are the regions with missing inpatient reports (see Inpatient service utilization (p. 218)). One region has a hospital density of about **1.1** per 100,000, twice

any other, with only middling admissions: a high input without a matching output, to discuss.

Example: under-5 admissions by bed density

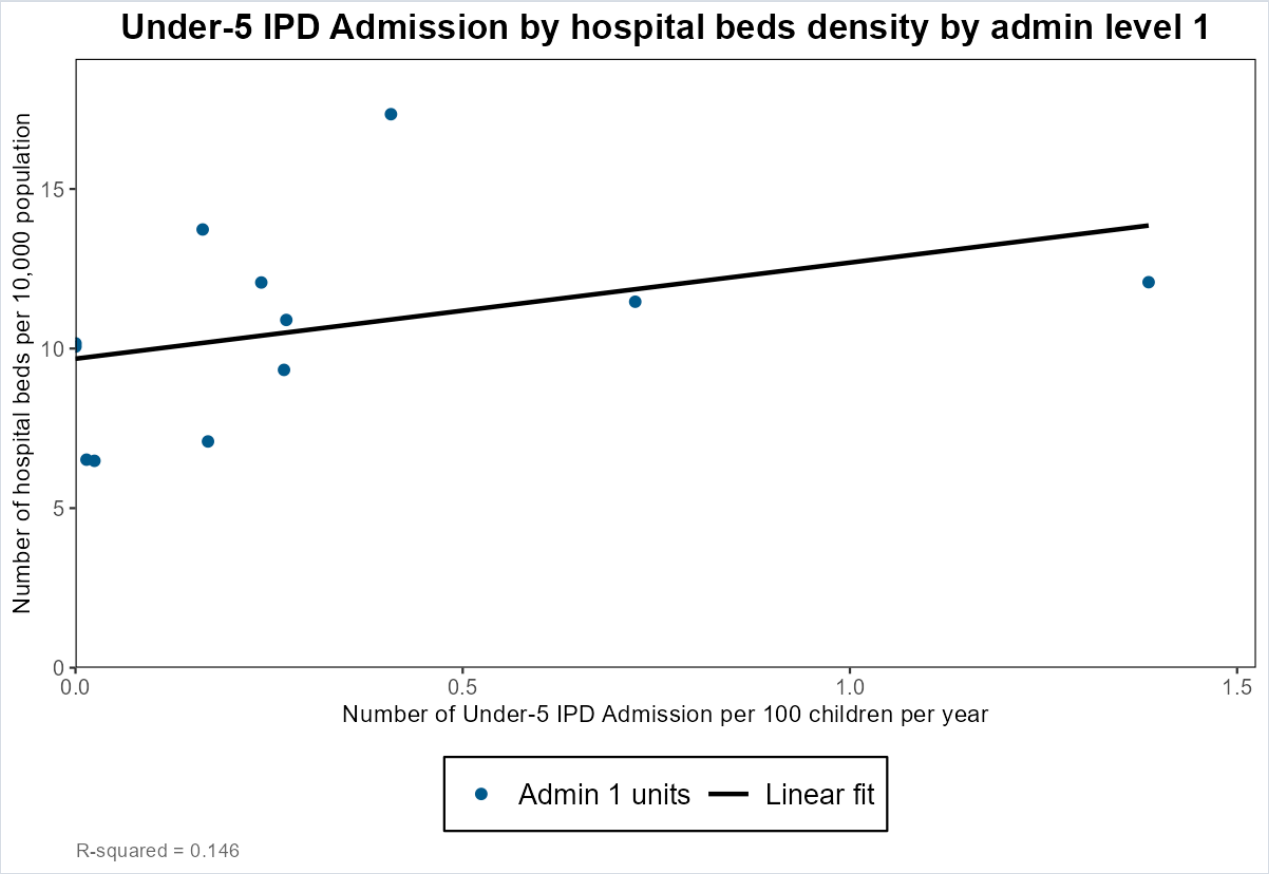


Figure 16.7 — *Inpatient beds per 10,000 population against under-5 admissions per 100 children, by Admin 1 region, with linear fit*

Reading the chart: as above, with inpatient beds per 10,000 up.

What it shows: the association with beds is a little closer (**R^2 = 0.15**), still positive. The region with the most beds (about **17** per 10,000) again has only middling admissions, and the regions with missing inpatient reports sit at zero admissions with 6 to 10 beds per 10,000.

Example: under-5 OPD visits by workforce density

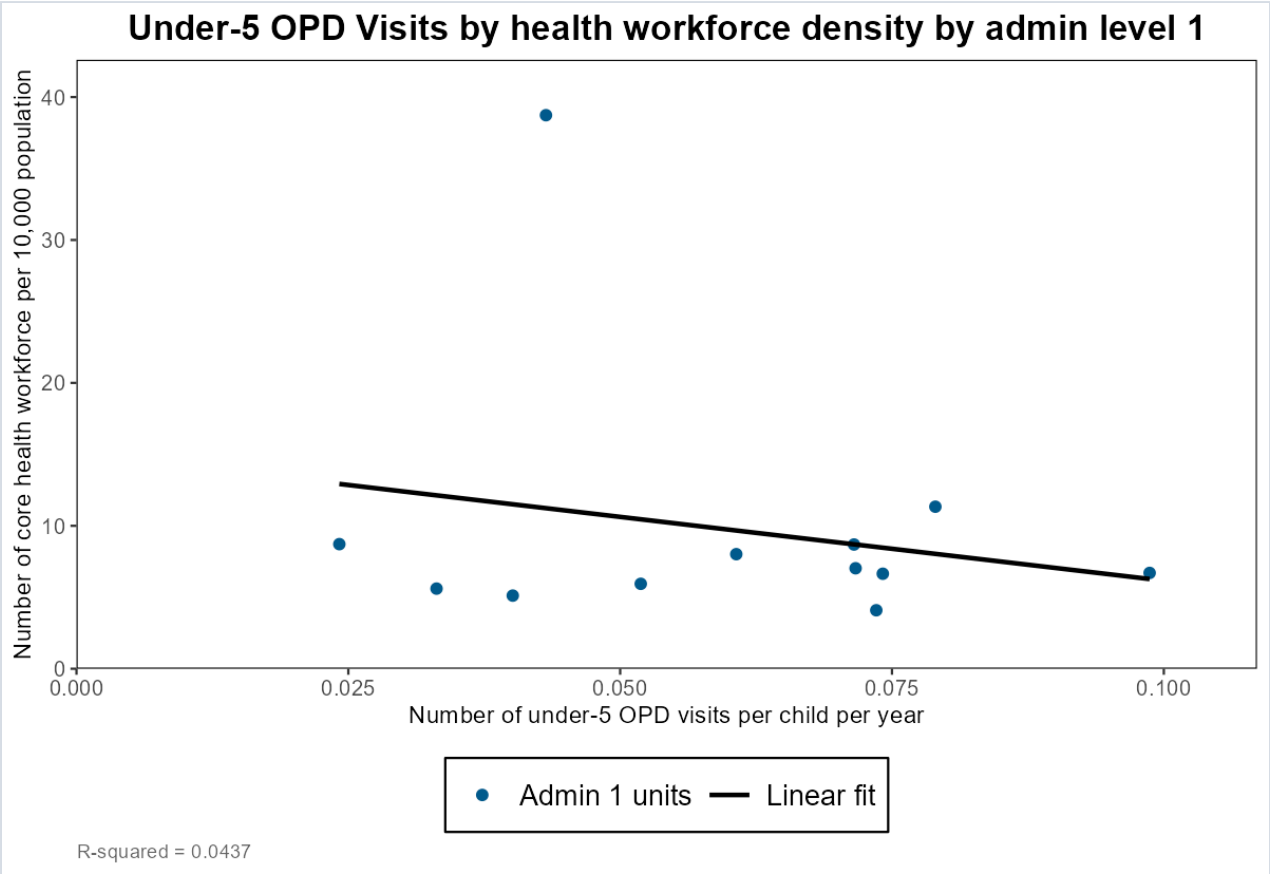


Figure 16.8 — *Core health workforce per 10,000 population against under-5 OPD visits per child, by Admin 1 region, with linear fit*

Reading the chart: as above, with under-5 OPD visits across and core health workforce per 10,000 up.

What it shows: there is no positive association: the line slopes slightly down (**R^2 = 0.04**). It is pulled by the region with **39** health workers per 10,000, whose OPD use is low; the other eleven regions, between 4 and 11 health workers per 10,000, show a wide range of OPD use. That region's workforce figure already failed the plausibility check in Health systems inputs (p. 233): settle it before discussing efficiency.

Example: facility density and the PHC service index

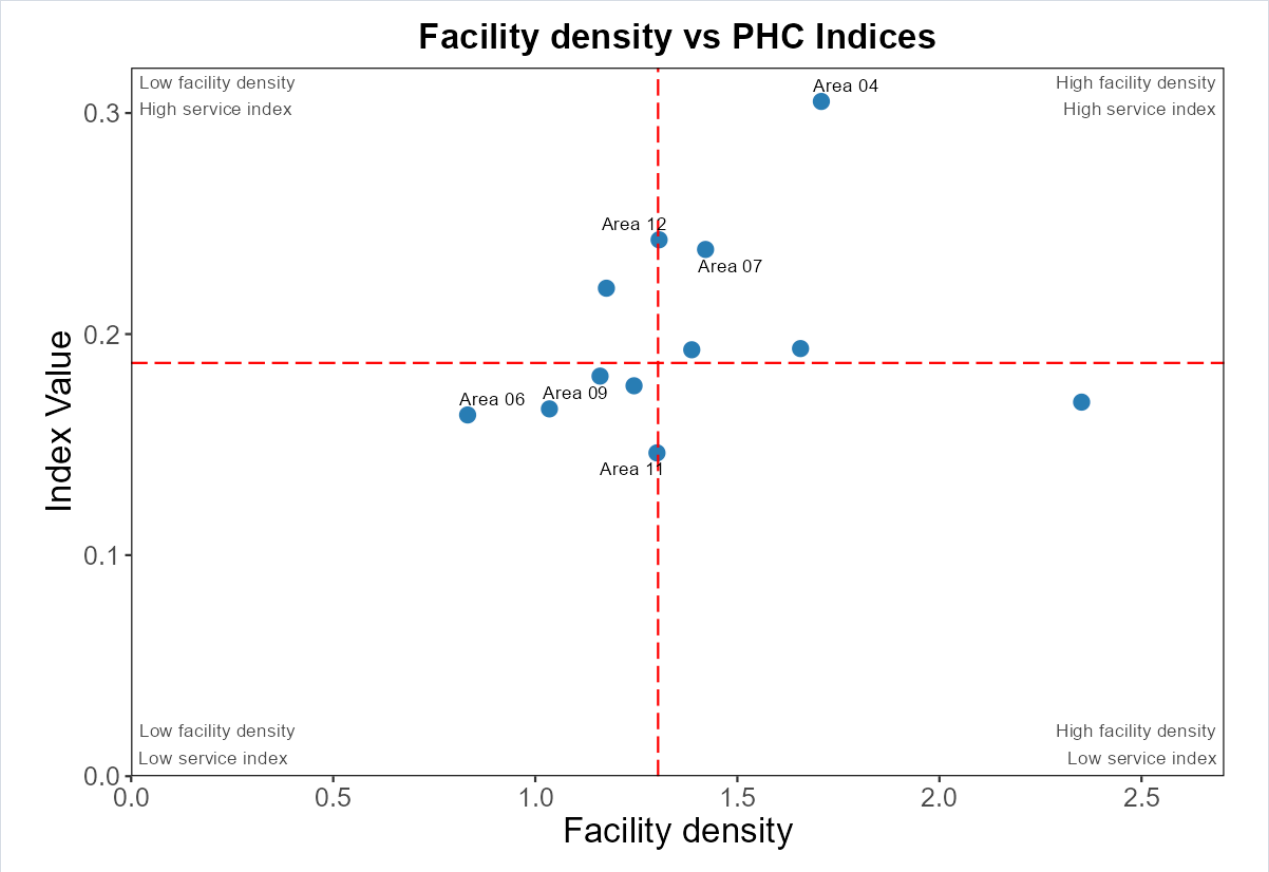


Figure 16.9 — *Health facility density against the PHC service index, by Admin 1 region, with median lines*

Reading the chart: each dot is a region, with its facility density across and its PHC service index up. The red dashed lines are the medians; the corners name the four quadrants. Labelled regions have a service index in the lowest or highest quarter.

What it shows: Area 04 is high on both. Area 07 and Area 12, near the facility median, have high service indices. Area 06 has the fewest facilities (about **0.8** per 10,000) and a low service index, and Area 11 the lowest service index with a median facility density. The region with the highest facility density (about **2.4**) has a service index below the median: high density with a low service index, a possible bottleneck.

Example: workforce density and the PHC service index

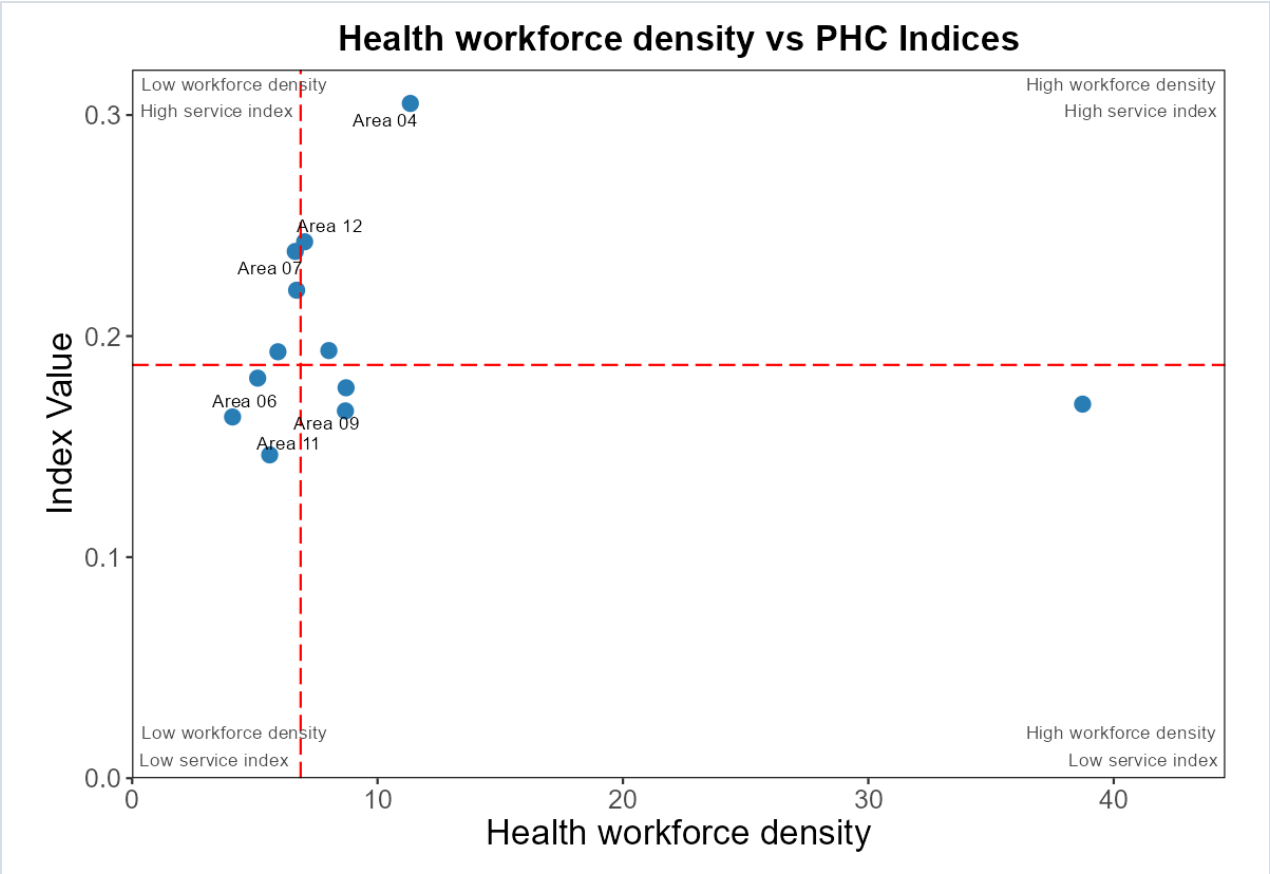


Figure 16.10 — *Health workforce density against the PHC service index, by Admin 1 region, with median lines*

Reading the chart: as above, with workforce density across.

What it shows: most regions sit between 4 and 12 health workers per 10,000. Area 04 is high on both; Area 06 and Area 11 are low on both; Area 07 and Area 12 reach a high service index with a workforce near the median. The region with **39** health workers per 10,000 has a service index below the median, the same region as in the charts above.

The chart of institutional live birth coverage by workforce density is not shown here: it needs institutional delivery coverage by region, which the example dataset does not provide.

After the analysis

The regions with low inputs, and those whose inputs do not translate into services, are the findings to bring to the Planning for dissemination, decisions and action (p. 245) session, with the explanation agreed by the country team.

16.3 Private sector and RMNCAH-N services

The share of health facilities that are private, NGO or public, and the share of RMNCAH-N services provided by the private sector.

Health facilities belong to the government (public), to private owners, or to non-governmental organizations (NGO). This part shows the ownership of the facilities by region, from the facility data, and the share of some RMNCAH-N (p. 503) services that people receive from the private sector, from household surveys (p. 500).

> **At a glance**
>
> - **Measures:** the percent of facilities that are private, NGO and public, by region; the private share of care-seeking for child illness, caesarean sections and institutional deliveries, nationally and by area of residence.
> - **Good result:** not a target: a description that helps read the other results.
> - **In the app:** the ownership card and the **National** and **Urban/Rural** tabs of **Public/Private Share** (p. 409).

Why

Private facilities are counted in the facility density of the inputs (p. 228), but they may report fewer of their services to the health information system than public facilities. Where the private sector is large, services can look low in the facility data while people are receiving them, and the comparison of outputs by inputs (p. 235) has to take it into account.

What

Indicator	Numerator	Denominator	Interpretation
Share of private facilities	Number of private (for-profit) facilities	Number of health facilities	A large share means many services may happen outside public reporting
Share of NGO facilities	Number of NGO (not-for-profit) facilities	Number of health facilities	
Share of public facilities	100% minus the private and NGO shares		
Private share of a service (survey)	Prevalence of the service received in the private sector	Prevalence received in the public and private sectors	Shown for care-seeking for child illness, caesarean section and institutional delivery

Formula

$$S_{\text{private}} = \frac{N_{\text{private}}}{N_F} \times 100 \qquad S_{\text{NGO}} = \frac{N_{\text{NGO}}}{N_F} \times 100$$

$$S_{\text{public}} = 100 - S_{\text{private}} - S_{\text{NGO}}$$

$$Q = \frac{p_{\text{private}}}{p_{\text{public}} + p_{\text{private}}} \times 100$$

where

- N_{private}, N_{NGO} and N_F are the numbers of private, NGO and all health facilities in the latest year;
- Q is the private share of a service;
- p_{private} and p_{public} are the survey prevalence of the service received in the private and in the public sector.

Worked example: in a region with 300 facilities, of which 120 private and 3 NGO, the shares are 120 ÷ 300 × 100 = **40%** private, **1%** NGO and **59%** public.

Interpretation

- **Which regions have a large private sector?** Read their service rates with care: services in private facilities may be missing from the facility data.
- **Does the private share of services in the survey match the share of private facilities?** A large gap shows whether private facilities provide many services or few.
- **Is the private share different in urban and rural areas?** It usually is: private facilities tend to be concentrated in towns.
- Check that the year in the chart title matches your data before quoting it.

Example: ownership of health facilities by region

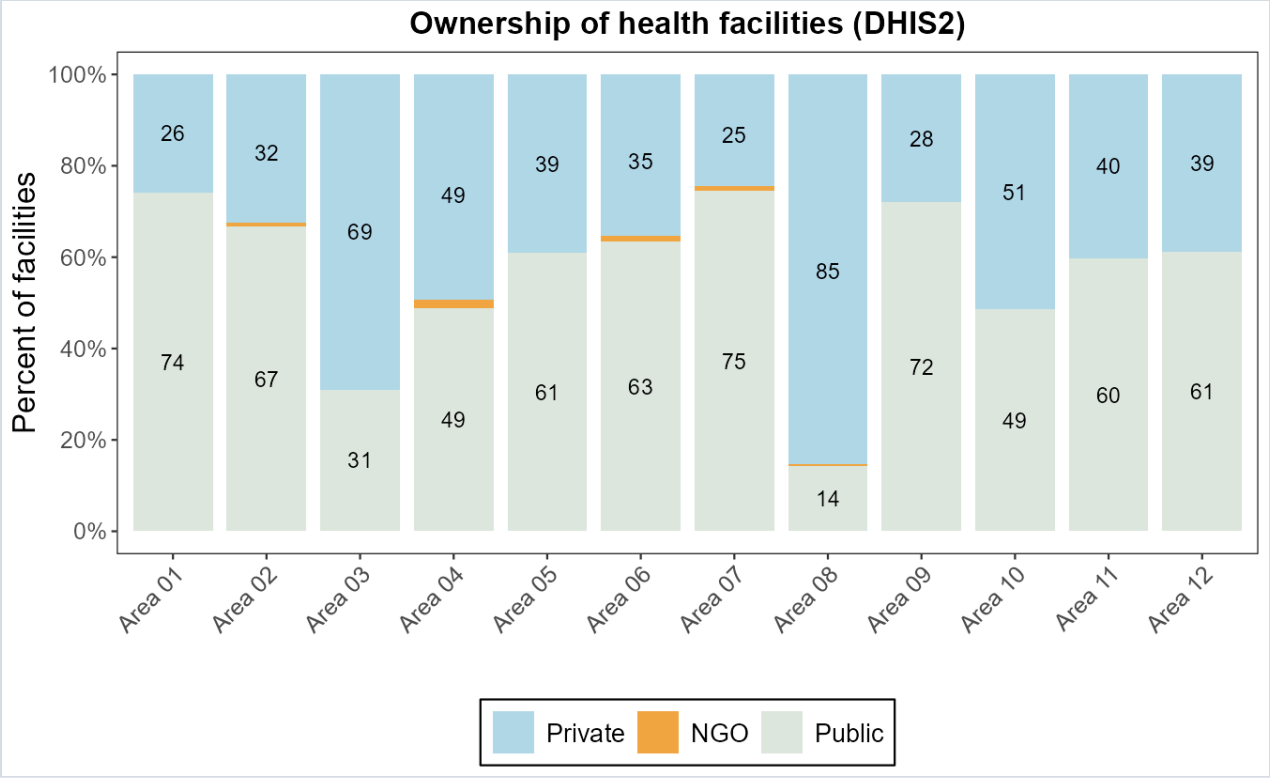


Figure 16.11 — *Percent of health facilities that are private, NGO and public, by Admin 1 region*

Reading the chart: one bar per Admin 1 (p. 496) region, divided into private (blue), NGO (orange, very thin) and public (grey) facilities, in percent.

What it shows: the private share ranges from **25%** to **85%**: in Area 08, 85% of the facilities are private, and in Area 03, 69%; in the other regions it is between 25% and 51%. NGO facilities are at most 2%. Area 08 is also the region with the highest facility density, and its service index is below the median (see Health systems outputs by inputs (p. 240)): part of the services of its many private facilities may not be in the facility data.

The survey-based charts (private share of services, national and urban or rural) are not shown: they need survey data on the sector of care, which the example dataset does not include.

After the analysis

The private sector share is part of the context for the findings brought to Planning for dissemination, decisions and action (p. 245).

CHAPTER 17

Planning for dissemination, decisions and action

The final workshop step: identifying who will use the findings, tailoring information products, and planning concrete actions to take the work forward.

The objective of the final session of the workshop is to identify concrete actions to take forward the work you have completed: at least two or three actions that build on it, including both the sharing of the analytical methods and engaging with the country platform (p. 498) to disseminate the key findings. As you have been preparing and interpreting your analyses, you have been encouraged to think about the full cycle of data-driven decision-making.

It is the last step of the end-to-end workflow (p. 43), after Health system performance (p. 225), and uses the findings of all the previous steps.

At a glance

- **Answers:** who will use the findings, for which questions, and which concrete actions will take the work forward?
- **Uses:** the analysis, interpretation and key findings from the previous steps.
- **Produces:** information products and dissemination strategies tailored to users, and actions for data use, capacity building and further analysis.
- **In the apps:** Reports (p. 415) and Exports, reports and AI (p. 411).

Why

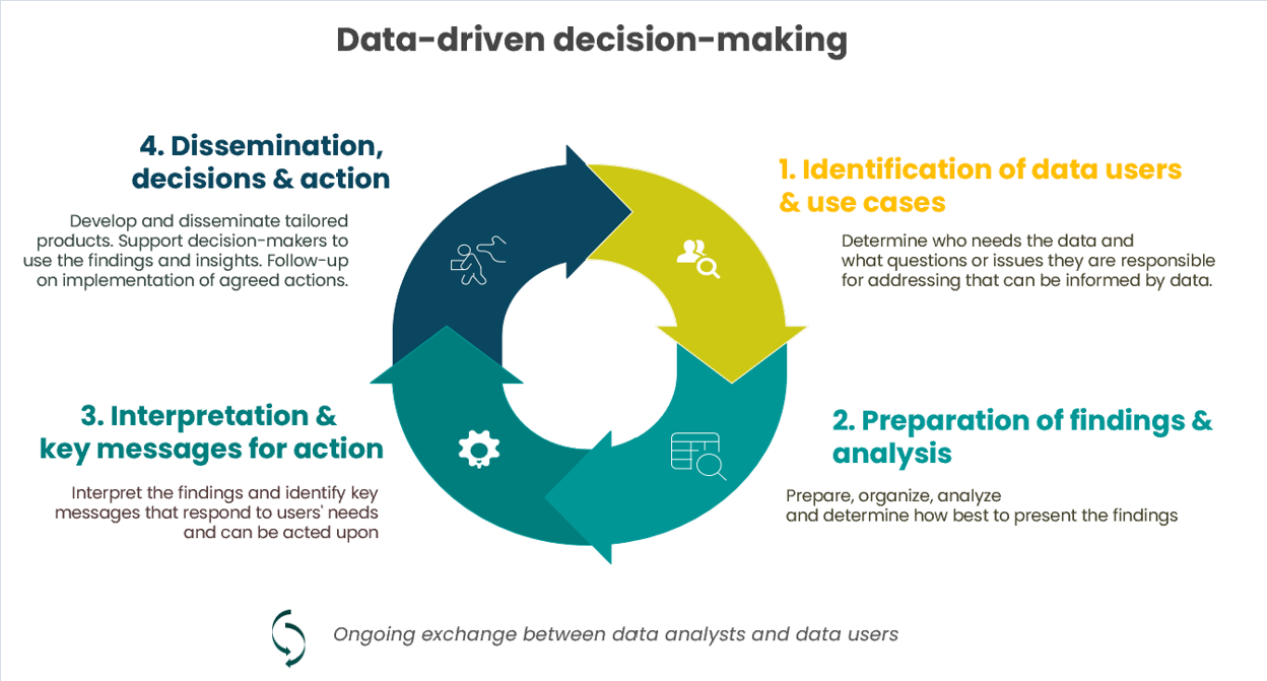


Figure 17.1 — *Cycle of data-driven decision-making: identification of data users and use cases, preparation of findings and analysis, interpretation and key messages for action, dissemination, decisions and action*

The cycle starts with identifying the data users and understanding the questions or issues they aim to address, and then preparing the analysis, interpretation and key findings that would be most useful for addressing those questions or issues.

Information products and dissemination strategies should be tailored to users' preferences, and supplemented with support to decision-makers in using the analysis and following up on agreed actions. A critical element is a clear connection between the workshop's approach and outcomes and the reviews of progress and performance led by the Ministry of Health: it should be reflected in analytical reports for annual reviews or for global reporting. At each stage, consider what will be most useful for your audience to understand, and the key messages you intend to communicate.

How the parts fit together

Start from the users: who they are and what they need decides which follow-on activities make sense and which products to prepare. The products carry the findings; the activities make sure they are discussed, understood and acted on, and that the methods stay in use after the workshop.

Doing it in the app

① Build a report

The Reports (p. 415) page in the **Output** section builds Word or PDF documents and PowerPoint or PDF slide decks from the loaded dataset. Standard reports contain prompts headed **Notes** or **Interpretations**: replace them with the interpretation agreed at your review meeting.

② Export charts, tables and datasets

Download single charts, tables and datasets, or ask the DataSuite chat assistant about the saved dataset. See Exports, reports and AI (p. 411).

Points to consider

- Who are the main audiences, and what insights might they be interested in?
- What questions could the analyses help answer, and what decisions or policies can they inform?
- Which platforms and meetings could you engage to reach these audiences, beyond a one-off presentation?
- How will this link to the Countdown country collaboration and to existing plans to build capacity for data analysis and use?

Common pitfalls

- Sharing a chart without its data or the denominator. Always state the denominator and adjustment used.
- Sharing the analysis once and stopping there: the findings are used when they feed routine forums where decisions are taken.
- Leaving interpretation prompts in a report: it is the most common reason a chartbook looks unfinished.

17.1 Identifying use cases and audiences

Who will use the analyses, for which questions and decisions, and through which platforms and routine forums.

Before you start developing your analyses, consider as a team how they could be used and who the main audiences will be. The answers decide what to emphasise in the findings, which products to prepare, and where to present them.

> **At a glance**
>
> - **Decides:** the audiences, the questions they need answered, and the platforms and forums that reach them.
> - **Good result:** each key finding linked to an audience, a decision it can inform, and a routine forum where it will be discussed.
> - **In the app:** the **Report language** and **Region of this report** of each report on **Reports** (p. 419) follow from the audience.

Why

To make the most of the workshop, it is important to do more than share the analysis as a one-off event. The analyses are used when they feed the routine forums where data are used and decisions are taken, and when they connect to the reviews of progress and performance led by the Ministry of Health.

What

Key considerations:

- **Identify use cases and audiences.** What insights might the audiences be interested in? What questions could the analyses help answer? What specific decisions or policies can the analyses help address in an evidence-informed way?
- **Use appropriate dissemination platforms.** Think about which platforms and meetings you could engage to reach the relevant audiences, such as the country platform (p. 498) and technical working groups.
- **Integrate into routine decision-making forums.** Feed the forums where data and analyses are used and decisions are taken, rather than a single presentation. This could include working sessions with your target audiences, not only to share the findings but also to explore some of the concepts learned during the workshop, strengthen their capacity to interpret these analyses themselves, and identify opportunities for action.
- **Link to Ministry of Health reviews and reporting processes.** Make a clear connection between the workshop's approach and outcomes and the reviews of progress and performance led by the Ministry of Health, reflected in analytical reports for annual reviews or for global reporting (for vaccines, the reviews led by the immunization programme, and global reporting such as WUENIC (p. 505)).

Also think about how this will link to your Countdown country collaboration and to existing plans to build capacity for data analysis and use.

Interpretation

- **Is there an audience for each key finding?** A finding nobody is responsible for acting on needs an audience before it is disseminated.
- **What will each audience understand?** At each stage, consider what will be most useful for them to understand and the key messages you intend to communicate.
- **Which forum, and when?** Prefer the meetings that already take decisions (annual reviews, planning cycles, technical working groups) over new events.
- **Who follows up?** Agree who supports the decision-makers in using the findings and follows up on agreed actions.

Next

The audiences and their questions decide the follow-on activities (p. 250) and the information products (p. 252) to prepare.

17.2 Follow-on activities

Examples of activities after the workshop: supporting data use and action, building capacity, and identifying needs for further analysis.

The final session of the workshop identifies at least two or three concrete actions that build on the work completed. These activities show how the analysis outputs can be translated into action, capacity strengthening and further research within the Countdown to 2030 framework. Choose those that fit your audiences (p. 248).

> **At a glance**
>
> - **Decides:** the actions that follow the workshop, who carries them out, and by when.
> - **Good result:** two or three concrete actions, including both sharing the analytical methods and engaging with the country platform (p. 498) to disseminate the key findings.
> - **In the app:** the standard reports of **Reports** **(p. 416)** and the downloads of **Exports, reports and AI** **(p. 411)** carry the findings into each activity.

Why

Beyond building the capacity of the analysts in the Ministry of Health and its programmes, the workshop and its analyses aim to improve the quality of the data reported, to increase country ownership of the estimates, and to encourage the adoption of similar processes at sub-national level. These aims are reached through what happens after the workshop.

What

Purpose of the activity	Examples
Supporting data use and action	Present findings to the country platform; feed strategies and reviews; write an analytical brief
Building capacity	Cascade training (p. 497); sensitize statistics staff; share the data quality assessment with HMIS (p. 499) and M&E staff
Identifying needs for further analysis	Implementation research (p. 500); further research on gender and equity

Supporting data use and action

- Present findings, interpretation and key insights to the country RMNCAH platform and related technical working groups for them to review and interpret. Support them to identify opportunities to take action based on data and evidence: for example, address gender or geographic inequities, test pilots with implementation research or evaluation, or advocate for policy change or resources.
- Include the analysis as part of the evidence base to inform the development of a new RMNCAH-N strategy, a mid-term review and an evaluation, if any of these are underway.
- Produce an article or analytical brief for targeted audiences with key learning, takeaways and proposed actions following the workshop, as an entry point for follow-up.

Building capacity

- Carry out cascade training at national and district levels on the concepts, interpretation and use of specific analyses, through the country Countdown collaboration.
- Hold sessions to sensitize national bureau of statistics staff on the Countdown methodology, to build ownership.
- Share the concepts and findings of the health facility data quality assessment (p. 57) with HMIS and M&E staff, to build their understanding and as an opportunity for strengthening routine systems for data quality.
- Build HMIS and M&E staff capacity to understand the relevance of gender and equity-related indicators and analyses.
- Build the capacity of staff to incorporate some of the new analytical techniques and approaches into dashboards, policy briefs, statistical reports and similar products.

Identifying needs for further analysis

- Identify opportunities for implementation research to address bottlenecks, or to inform the potential scale-up or replication of successes, based on findings from the analysis.
- Identify opportunities for further research or analysis to better understand the trends in gender and equity observed in the analysis, and to inform options for action.

Interpretation

- **Is each action concrete?** Name who does it, for which audience, and by when.
- **Do the actions cover both purposes of the session?** Sharing the analytical methods, and engaging with the country platform on the key findings.
- **Is the follow-up linked to existing processes?** Annual reviews, strategy development and the Countdown country collaboration keep the actions going after the workshop.

Next

The information products (p. 252) carry the findings into these activities.

17.3 Information products

The reports and charts that carry the findings to stakeholders, and how to prepare them before they are shared.

Information products are what the audiences receive: reports, slide decks, briefs, single charts and tables. They should be tailored to the users' preferences. The apps build standard reports from the loaded dataset, which you then complete with the interpretation agreed by the country team.

> **At a glance**
>
> - **Produces:** the synthesis chartbook (p. 504), a sub-national one-pager (p. 504) per region, other standard reports, and single charts and tables with their data.
> - **Good result:** every prompt answered, the data quality, adjustment and denominators stated, the language of the audience.
> - **In the app:** the standard reports of **Reports** (p. 416), and chart and table downloads in **Exports, reports and AI** (p. 411).

Why

The findings are only as useful as the products that carry them. A report with unanswered prompts, or a chart without its denominator, cannot be checked or trusted by the people who receive it.

What

Two of the apps' standard reports are the ones most stakeholders ask for:

- The **synthesis chartbook** is the complete analysis in one document for the review meeting.
- The **sub-national one-pager** gives each region its own statistical profile.

What each standard report contains, and how to build it, is in Reports (p. 416). Other products (an article, an analytical brief, a policy brief) are written by the team from these reports and charts; the apps provide no templates for them.

Preparing reports for stakeholders

1. Settle data quality, adjustment and the denominators (in the Vaxx app, the denominator) before you generate the synthesis, and state them in a text block on the first page.
2. Answer every interpretation prompt or delete it. Leaving prompts in is the most common reason a chartbook looks unfinished.
3. Fill in the cover: title, editors, date and logos of the partners.
4. Produce the report in the language of the audience, and use **Print layout** before you send a PDF.
5. Duplicate the one-pager for each region and share the PowerPoint files with the regional teams (for the Vaxx app, also the immunization managers).
6. Link findings to actions: see Supporting data use and action (p. 250).

Sharing charts and tables

Download the data with the plot when you share a chart. The Excel file carries the title and the numbers, so readers can check the figure. Always state the denominator and the adjustment used. See Exports, reports and AI (p. 411).

Interpretation

- **Does every chart say which denominator and adjustment it uses?** If not, add it before sharing.
- **Is every interpretation prompt answered with the team's agreed interpretation?** Delete the prompts you do not answer.
- **Does each region have its one-pager, and does each audience have a product in its language?**

Next

The products are presented in the forums chosen in Identifying use cases and audiences (p. 248), and followed up through the follow-on activities (p. 250).

PART III

DataSuite Apps

Overview of the DataSuite applications (Data Extractor, RMNCAH App, Vaxx App and Pooled App), what each one is for, and how data moves from one to the next.

CHAPTER 18

DataSuite applications

Overview of the DataSuite applications (Data Extractor, RMNCAH App, Vaxx App and Pooled App), what each one is for, and how data moves from one to the next.

Four apps, one analysis

The **Data Extractor** gets the data out of DHIS2 (p. 498). The **RMNCAH** and **Vaxx** apps check, adjust and analyse one country's data. The **Pooled** app puts several countries side by side.

Learn about how data moves between the apps (p. 257)

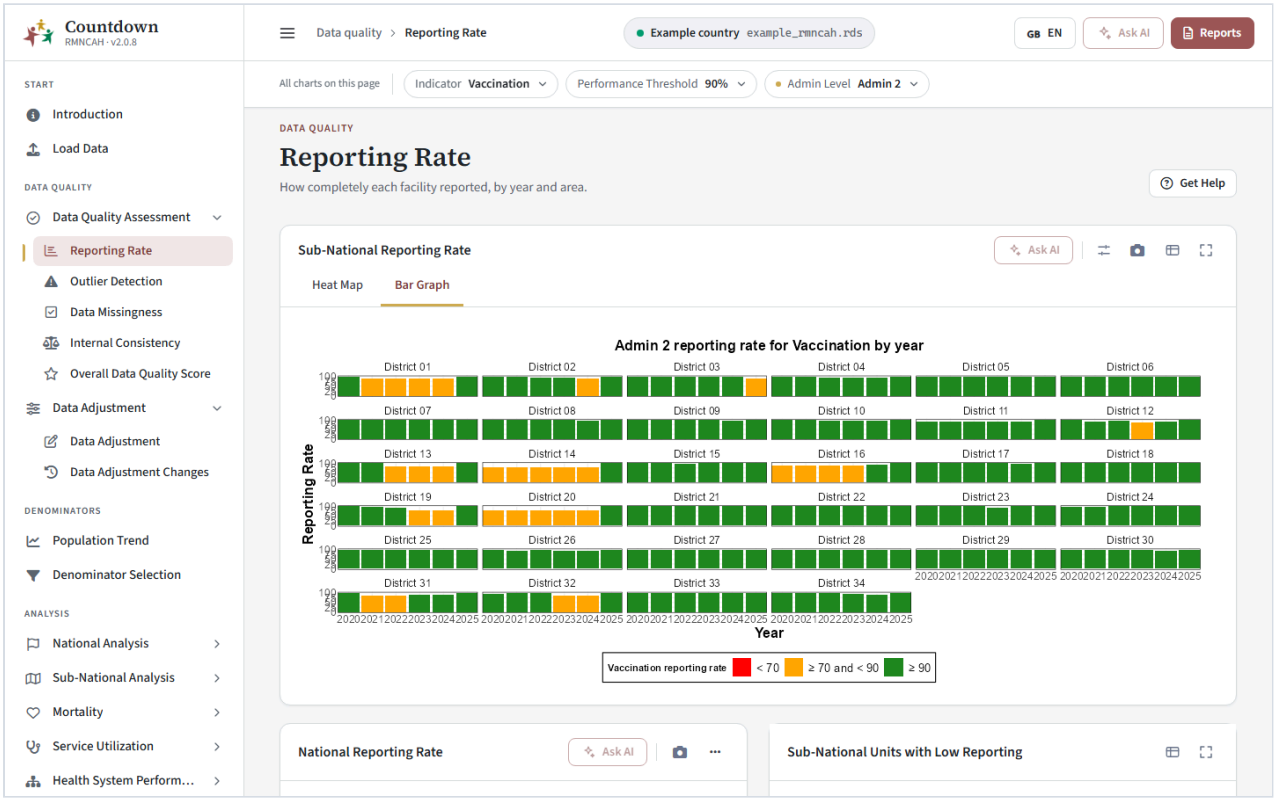


Figure 18.1 — *The RMNCAH app on the Reporting Rate page: sidebar, top bar, filter bar and a chart card*

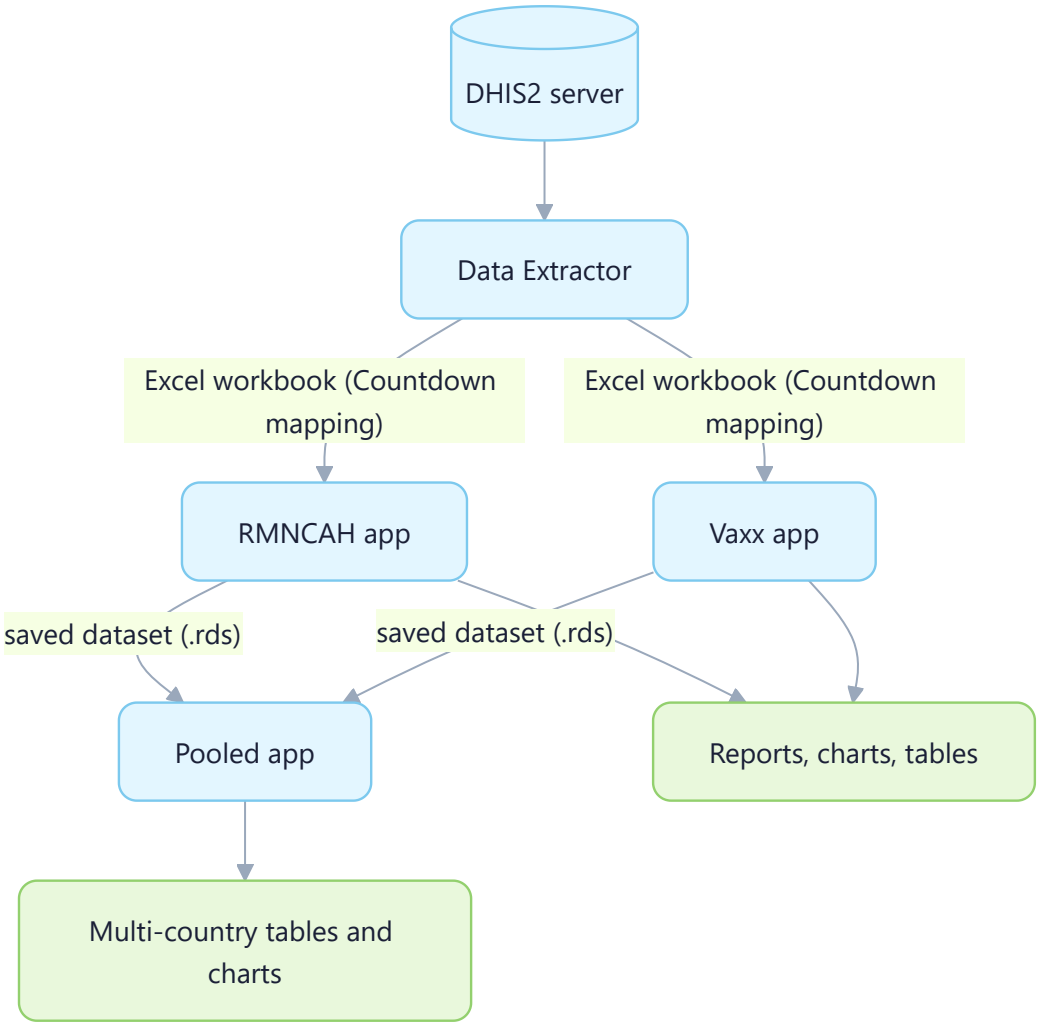
The apps

What each app reads and writes:

App	Use it to	Input	Output
Data Extractor (p. 269)	Map DHIS2 data to indicators and download it	A DHIS2 server	Excel workbook (or JSON)
RMNCAH App (p. 289)	Analyse reproductive, maternal, newborn, child and adolescent health	`.xls`, `.xlsx`, `.dta`, `.rds`	Saved dataset (`.rds`), reports, charts

Vaxx App (p. 289)	Analyse immunization coverage and inequalities	`.xls`, `.xlsx`, `.dta`, `.rds`	Saved dataset (`.rds`), reports, charts
Pooled App (p. 424)	Compare several countries or datasets	`.rds` files saved by RMNCAH or Vaxx	Pooled file (`.rds`), CSV, Excel
AI Assistant (p. 260)	Explain results, build reports and charts	The dataset open in an app	Answers with their sources

How data moves between the apps



Each app saves what it makes next to the file it opened: the extractor writes the workbook, the RMNCAH and Vaxx apps save the dataset (with every choice made in the analysis) as an .rds file (p. 503), and the Pooled app combines those `.rds` files into one pooled file. The DataSuite assistant works alongside the analysis apps.

CHAPTER 19

RMNCAH and vaccine analyses

What differs in the method between an RMNCAH analysis (RMNCAH app) and a vaccine analysis (Vaxx app), step by step, with the methodology part and the app page for each.

The RMNCAH app and the Vaxx app follow the same Countdown methodology (p. 42) on different indicators. An **RMNCAH analysis** covers reproductive, maternal, newborn, child and adolescent health; a **vaccine analysis** is narrowed to immunization. Most of the method is the same for both: the outlier rule, the adjustment rules and the national rates (pregnancy loss, twins, stillbirths, neonatal and post-neonatal mortality) do not change. The table lists only what differs, with the part of the methodology that explains it and the app page where it is done.

Step	RMNCAH analysis	Vaccine analysis	Method · In the app
Indicators	The full RMNCAH set: antenatal, delivery and postnatal care, stillbirths and deaths, a set of vaccines, outpatient and inpatient visits	ANC1, institutional deliveries and live births, and the vaccines, with dropout, zero-dose and under-vaccinated children	Indicator definitions (p. 509) · Differences at a glance (p. 293)
Survey values	Include ANC4, C-section and low birth weight	Include OPV1 and OPV3, without ANC4, C-section and low birth weight	National survey parameters (p. 54) · National Rates (p. 301)
Reporting completeness	The ANC, delivery, immunization and outpatient forms	The ANC, delivery and immunization forms	Completeness of reporting (p. 59) · Reporting Rate (p. 309)
Internal consistency	ANC1 with penta1, penta1 with penta3	The same, plus OPV1 with OPV3	Internal consistency (p. 67) · Internal Consistency (p. 317)
Annual data quality score	The mean of 7 indicators	The mean of 8 indicators (with OPV1 to OPV3); missing values and outliers are counted over the vaccines	Overall data quality score (p. 72) · Overall Data Quality Score (p. 320)
Adjustment factor k	One per service: ANC, delivery, vaccines and outpatient	One per service: ANC, delivery and vaccines	Adjusting for completeness of reporting (p. 81) · Choose the K-Factors (p. 327)
Denominators	Two: a maternal denominator and a vaccination denominator	One vaccination denominator for every indicator	Selecting the best denominator (p. 108) · Choose the denominator (p. 339)
National coverage	The RMNCAH indicators, family planning and the continuum of care	The vaccines, with dropout, zero-dose and under-vaccinated children	National coverage (p. 115) · Indicators available (p. 345)

Coverage targets	ANC4, institutional live births and the vaccines	The vaccines and dropout	Target values (p. 176) · Coverage Target (p. 349)
Equity	Maternal, newborn and vaccine indicators	Vaccines and dropout	Equity analysis (p. 154) · Equity Assessment (p. 355)
Bayesian coverage	Done	Not done	Bayesian modeling approach (p. 143) · Bayesian Analysis (p. 373)
Mortality	Done	Not done	Mortality (p. 179) · Mortality (p. 381)
Service utilization	Done	Not done	Service utilization (p. 204) · Service Utilization (p. 393)
Health system performance	Done	Not done	Health system performance (p. 225) · Health System Performance (p. 403)

A dataset saved by one app cannot be opened in the other: see Accepted files (p. 295).

CHAPTER 20

Leveraging AI for Analysis and Interpretation

How AI can support analysis and interpretation of routine health data, what it must not replace, how to write prompts, and what the DataSuite assistant does.

This page is about using AI well when you analyse and interpret routine health data: what it can take on, what stays with the analyst and the country team, and how to write a prompt that gets a useful answer. It ends with an example prompt for a gender analysis and with what the assistant in DataSuite knows and does. Read it before you use **Ask AI** in the apps (Exports and Ask AI (p. 413)).

What AI can and cannot do

Artificial intelligence (AI) can be a powerful tool to enhance the efficiency, depth, and accessibility of data analysis and interpretation. It can support analysts in rapidly exploring datasets, identifying patterns, summarizing trends, and generating initial visualizations and insights. AI can also help structure written explanations, translate technical findings for different audiences, and synthesize relevant evidence from large volumes of information. When used well, AI can enable analysts to spend less time on routine or time-intensive tasks and more time on critical thinking, such as contextualizing findings, building shared understanding, and tailoring messages to decision-makers' needs.

However, AI must be used thoughtfully, with sound professional judgment and clear accountability. It should not replace the analyst's role in interpreting results, assessing data limitations, or making context-specific judgments. AI-generated outputs may contain inaccuracies, overlook important nuances, or reflect biases in the underlying data. Therefore, outputs must always be reviewed, validated, and adapted.

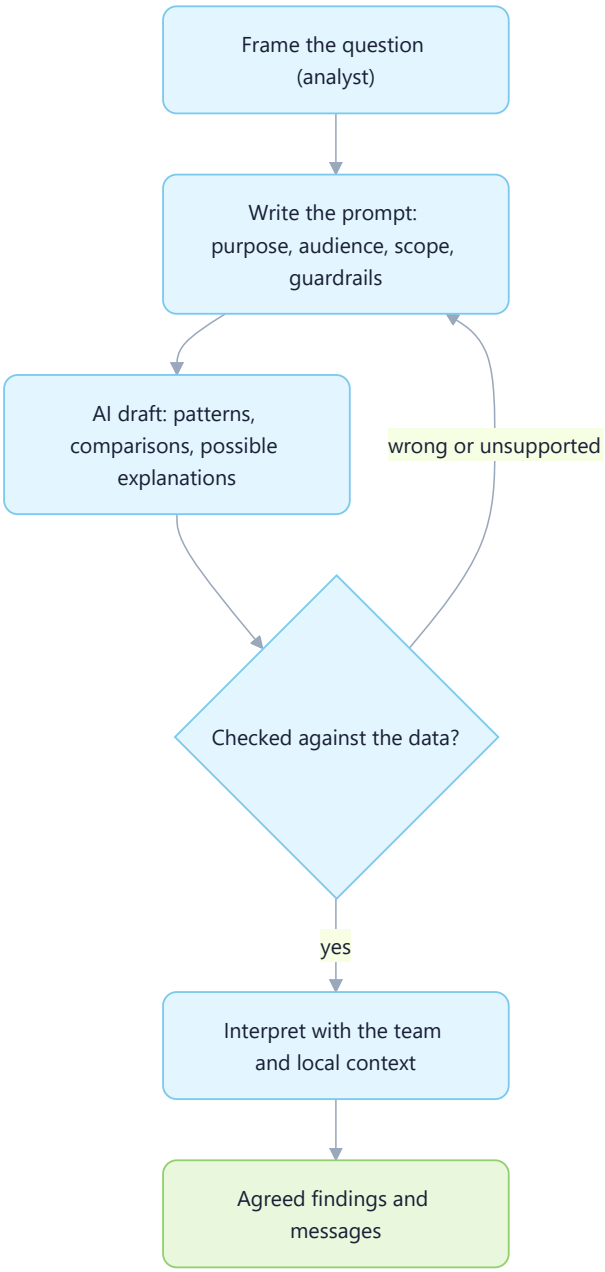
Collaboration between data producers, analysts, and data users to interpret findings is also essential to build shared understanding and ownership. Final interpretations, conclusions, and recommendations should always remain grounded in human expertise, subject-matter expertise, and an understanding of the local context. Therefore, AI should be considered a support tool that amplifies and strengthens analysis and interpretation, but is not a substitute for critical thinking, judgment, and shared decision-making.

AI can amplify	AI must not replace
Analysis – Understanding what the data shows	
Sharpening the definition of the problem or question that the analysis should respond to	**Framing the problem or question** that the analysis should respond to
Surfacing data quality issues (e.g., incompleteness, inconsistencies, outliers, etc.)	**Assessing data quality and fitness for use**, including reliability, comparability, limitations
Faster, deeper pattern detection; connecting signals from across sources	Deciding **what patterns to look for** in the data and why

Structured comparisons across time, geographies, subgroups, etc.	Determining **which comparisons are valid and relevant**
Generating **clear visual summaries**	**Selecting appropriate visualizations** and defining what should be shown and compared
Interpretation – Understanding what the data means	
Generating possible explanations and hypotheses for trends and patterns, testing assumptions	Using context, **collective knowledge and experience** to identify likely explanations; whether relationships are causal and actionable
Presenting sets of options and scenarios for prioritization	**Weighing up options**, deciding priorities; **owning decisions** and justifying choices
Surfacing counterarguments and bias checks	**Integrating stakeholder perspectives**, navigating power dynamics
Refining key messages and adapting language, tone, and formats for different audiences	**Judging what messages are most appropriate** and useful for each audience, **ensuring accuracy**

AI in the workflow

The analyst frames the question and owns the conclusions; AI drafts in between, and every draft is checked before it is used.



Effective AI prompts

What makes an effective prompt

- **Be clear on the purpose** – e.g., identify trends, show anomalies, interpret trends.
- **Define the audience** – e.g., analyst deep dive versus policymaker summary.
- **Specify geography, time, and scope** – e.g., country, administrative level, indicators, time period.
- **Provide interpretation guidance** – e.g., highlight gender and equity gaps, suggest possible explanations.
- **Specify the output format** – e.g., a slide deck, short narrative, comparison table.

- **Set guardrails** – e.g., stick to the data shown; flag uncertainty or data quality issues; avoid speculation beyond evidence; note assumptions and inferences.

> **Example of an effective prompt:**
>
> *Using only the data provided, summarize key trends, identify patterns and outliers at national level, and highlight gender and equity differences. Clearly distinguish between what the data shows (analysis) and what it may mean (interpretation). Explicitly identify data limitations, assumptions, and sources of uncertainty.*

AI prompts and gender analysis

Using AI prompts to support gender analysis and interpretation

Below is an example of a prompt that you can adapt and paste into the DataSuite Assistant to generate an initial gender analysis and interpretation. Please incorporate your country's data into the prompt.

The AI output is designed to provide a starting point only. It will highlight key assumptions and suggest reflection questions that your team can use to guide your discussions. All AI outputs should be carefully reviewed and refined based on your team's knowledge, expertise, and country context, focusing on the findings that are most relevant for your intended audience.

These prompts can be used to generate findings for your CAM analytical report or used to create a short summary for the chartbook.

Please reach out to Lior Idit Miller (lmiller9@worldbank.org) and/or Elizabeth Hazel (ehazel1@jhu.edu) for questions or support at the CAM or afterwards. More details are available on the MAGE project website <https://mageproject.org/>.

Prompt for country analytical teams

Using only the data points below, generate a gender-lens interpretation of maternal health in this country. The AI should provide an initial interpretation that requires country team review, validation, and refinement.

The interpretation should be brief and concise. For each section, the AI should flag assumptions and pose reflective questions for teams to consider based on their contextual knowledge.

Country data:

- **Country name:** [Add country name here]
- **Demand for family planning satisfied by modern contraceptives:** [Add country data here]
- **Unmet need for family planning:** [Add country data here]
- **ANC 4+ visits coverage:** [Add country data here]
- **ANC first trimester coverage:** [Add country data here]
- **Institutional delivery coverage:** [Add country data here]

- **Postnatal care coverage:** [Add country data here]
- **C-section rate:** [Add country data here]
- **Urban/rural institutional delivery gap (percentage points):** [Add country data here]
- **Maternal education institutional delivery gap (percentage points):** [Add country data here]

Output format:

Create a bold heading for this section called "Draft interpretation for country team review". Include this opening text: Access to essential maternal health services reflects broader gender inequities that shape women's autonomy over their reproductive health and their ability to seek care.

Organize the findings by subheading: Family planning; Antenatal care; Delivery and postnatal care; Equity. Then for each subheading include "Interpretation", "Assumptions for team validation", and "Reflection questions". Reflection questions should be limited to the specific assumptions mentioned and no more than 3.

Generate policy message options:

Based on the data provided, generate 4-5 bullet point illustrative policy messages that country teams can use as starting points for discussion. These should be:

- Directly linked to specific data points provided
- Policy-relevant and potentially actionable
- Framed through a gender and equity lens
- Concise (1-2 sentences each)

Each bullet should identify the data-driven finding and its gender/equity implication. Present these as suggestions rather than recommendations.

Include this closing text: These are illustrative examples intended to spark discussion. Country teams should validate for accuracy and relevance to current policy priorities, adapt language to local context, or develop alternative messages that better reflect their strategic objectives. Teams may select, combine, revise, or replace these options entirely.

Example of an AI output using a country's 2025 CAM data

> For illustrative purposes only.

Draft interpretation for country team review

Access to essential maternal health services reflects broader gender inequities that shape women's autonomy over their reproductive health and their ability to seek care.

Family planning

Interpretation: The country demonstrates relatively strong performance in modern contraceptive uptake (80.6% demand satisfied), suggesting progress in women's access to fertility control. However, the persistent 20% unmet need indicates that one in five women who wish to delay or prevent pregnancy cannot do so. This gap may reflect gendered barriers including limited decision-making power, partner opposition, or restricted access for certain groups such as adolescents and unmarried women.

Assumptions for team validation:

- The unmet need figure captures women across all age groups and marital statuses equally
- High contraceptive satisfaction rates reflect genuine choice rather than method limitations
- Partner/spousal consent is not a significant barrier to contraceptive access

Reflection questions:

1. Does the 20% unmet need concentrate among specific populations (adolescents, rural women, unmarried women) who face additional gender-related access barriers?
2. To what extent does male partner involvement—either supportive or restrictive—influence contraceptive uptake in the country?
3. Are current family planning services designed to accommodate women with limited autonomy in household decision-making?

Antenatal care

Interpretation: The significant gap between first trimester ANC coverage (27.6%) and ANC 4+ visits (62.4%) suggests women are entering care late in pregnancy. This pattern may reflect gendered constraints including women's limited control over household resources for transport, time poverty due to domestic responsibilities, or cultural norms around pregnancy disclosure. Late ANC initiation also occurs due to limited pregnancy diagnosis options at both individual and health system level. Late ANC initiation reduces opportunities for early risk detection and birth preparedness.

Assumptions for team validation:

- Late ANC initiation is primarily driven by demand-side barriers rather than supply-side availability
- Women have agency to seek care independently once pregnancy is confirmed
- The ANC 4+ figure reflects the recommended contact schedule quality

Reflection questions:

1. What gender-specific barriers (financial dependence, time constraints, need for permission) delay women's first ANC visit in the country?
2. Are there cultural practices around pregnancy concealment that disproportionately affect women's early care-seeking?
3. How do women's domestic and caregiving responsibilities influence their ability to attend multiple ANC visits?

Delivery and postnatal care

Interpretation: The country's high institutional delivery rate (95.2%) represents a significant achievement in ensuring women access skilled birth attendance. The C-section rate of 9.6% falls slightly below the WHO-recommended ranges (10-15%), suggesting somewhat appropriate access to emergency obstetric care without over-medicalization. Strong postnatal care coverage (84.1%) indicates continued engagement with the health system post-delivery. These figures suggest health system investments have successfully addressed major access barriers for facility-based care.

Assumptions for team validation:

- High institutional delivery rates reflect genuine access and quality rather than coercive policies
- The C-section rate is equitably distributed across wealth and geographic groups
- Postnatal care includes meaningful content addressing maternal and newborn health needs

Reflection questions:

1. Does the high institutional delivery rate mask variations in quality or respectful maternity care practices?
2. Is the C-section rate equitably available to women in rural areas facing obstetric emergencies?
3. Does PNC adequately address women's own health needs, or does it focus predominantly on newborn care?

Equity

Interpretation: The 25 percentage-point urban/rural gap in institutional delivery reveals significant geographic inequity, likely reflecting rural women's compounded disadvantages including distance to facilities, transport costs, and limited infrastructure. The 15 percentage-point education gap demonstrates how women's educational attainment—itself shaped by gender norms—influences maternal health outcomes. Less-educated women may face greater barriers including limited health literacy, reduced autonomy, and economic constraints that intersect with gender-based inequities.

Assumptions for team validation:

- The urban/rural gap primarily reflects supply-side factors (facility availability) rather than demand-side preferences
- Educational attainment serves as a proxy for women's empowerment and decision-making capacity
- These gaps have remained stable and are not already being addressed by recent interventions

Reflection questions:

1. For rural women who deliver outside institutions, what are the primary gendered barriers—distance, cost, permission, or preference for traditional care?
2. How does low educational attainment intersect with other dimensions of marginalization (poverty, age, parity) to compound barriers for specific groups of women?
3. Are existing programs specifically designed to reach women facing multiple intersecting disadvantages?

Suggested policy message options

Based on the data provided, the following illustrative policy messages are offered as starting points for country team discussion:

- **Closing the unmet need gap:** With 20% of women unable to access desired family planning despite 80.6% demand satisfaction, targeted interventions addressing gendered barriers—including partner engagement and adolescent-friendly services—could accelerate progress toward full reproductive autonomy.
- **Early ANC as a gender equity priority:** The 35 percentage-point gap between first trimester ANC (27.6%) and ANC 4+ coverage (62.4%) suggests women face early pregnancy barriers; addressing women's financial dependence and time poverty could improve early care initiation.
- **Sustaining delivery gains while addressing quality:** The country's 95.2% institutional delivery rate demonstrates strong access; the next frontier involves ensuring respectful, woman-centered care that upholds dignity and autonomy during childbirth.
- **Rural women's access imperative:** The 25 percentage-point urban/rural delivery gap indicates rural women face compounded geographic and gender barriers; community-based strategies and transport solutions could address these intersecting disadvantages.
- **Education-sensitive approaches:** The 15 percentage-point education gap in institutional delivery suggests less-educated women require tailored outreach that accounts for limited health literacy and potential constraints on autonomous decision-making.

These are illustrative examples intended to spark discussion. Country teams should validate for accuracy and relevance to current policy priorities, adapt language to local context, or develop alternative messages that better reflect their strategic objectives. Teams may select, combine, revise, or replace these options entirely.

The AI assistant in DataSuite

DataSuite's chat assistant works alongside the RMNCAH, Vaxx and Pooled apps. It is built to follow the guidance above: it supports your analysis, and your team stays responsible for the interpretation.

What it knows

- **Where you are.** It knows which app page you are on, which charts and tabs are on screen, and the filters you have set, so you can ask "what does this chart show?" or "which page am I on?".
- **Your dataset.** It answers from the results your dataset has saved -- the same numbers the app shows -- not from estimates of its own. It can look at other levels too, for example which districts drive a change in a national indicator.
- **This methodology.** It explains results with the Countdown methodology on this site. When the methodology does not cover a question, it says so rather than filling the gap.

What it tells you

Every answer that uses the methodology or your data ends with its **sources**: the documentation sections it relied on, as links, and the data it used -- a chart on screen, or a result saved in your dataset. When it calculates something the app does not show, it says so.

What it can do

- Explain a chart or a page, and what the results mean.
- Break a change down by region or district.
- Build a report: one of the standard reports, or a custom one you then edit in the Reports page.
- Make a custom chart that is saved with your dataset and redraws when the data changes, like the app's own charts.
- Open another dataset in its own tab, so each dataset keeps its own conversation context.

What it will not do

- **Change your dataset**, except to save a report or a chart you asked for.
- **Change what the app shows** (go to another page, set filters) unless your setting allows it. The setting *AI App Control* has three choices: ask each time (the default), allow, or off.

Where its files are

The assistant keeps its files for an analysis next to your dataset, in the `ai` folder of the dataset's `<name>.shiny-workspace` folder: the code it ran, conversation transcripts and the decisions it recorded. They stay with the data they are about.

CHAPTER 21

Data Extractor

Overview of the Data Extractor, which reads routine data from DHIS2 through your DataSuite sign-in, maps it to indicators and saves it as the workbook the RMNCAH and Vaxx apps load.

Get your DHIS2 data into a workbook

Say which DHIS2 (p. 498) data make up each indicator (a **mapping**), choose the periods and the level, and download. The workbook made with a Countdown mapping is the file the RMNCAH and Vaxx apps load.

Learn about how it works (p. 270)

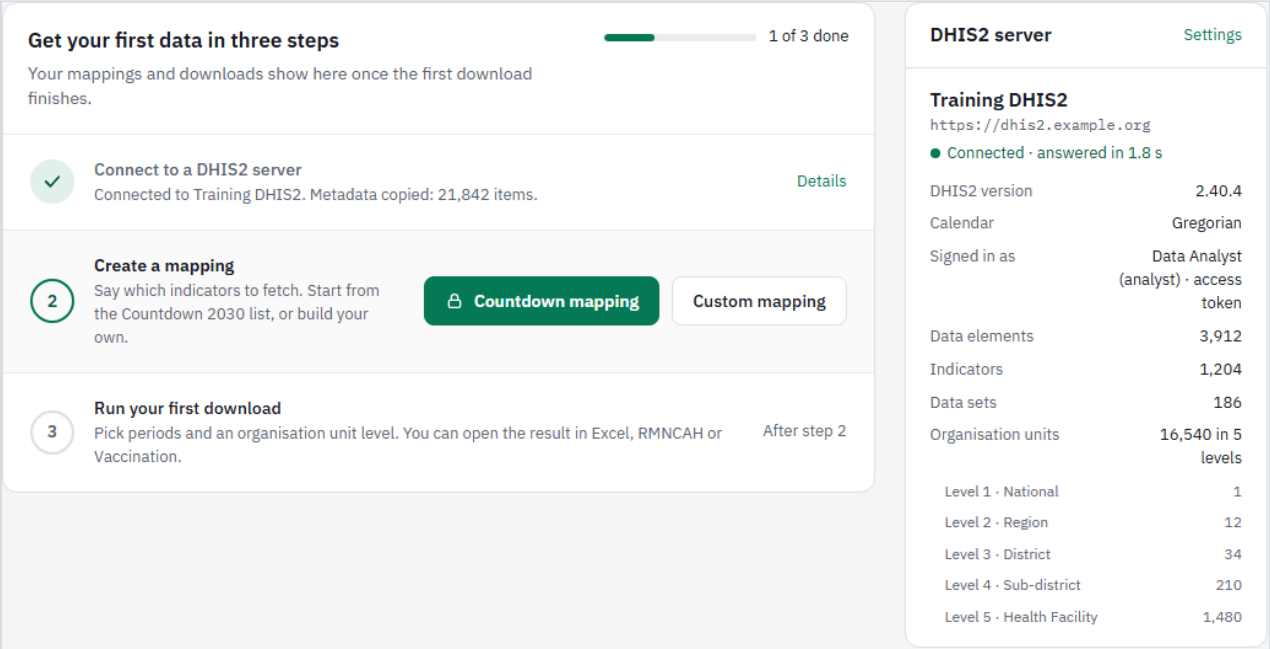


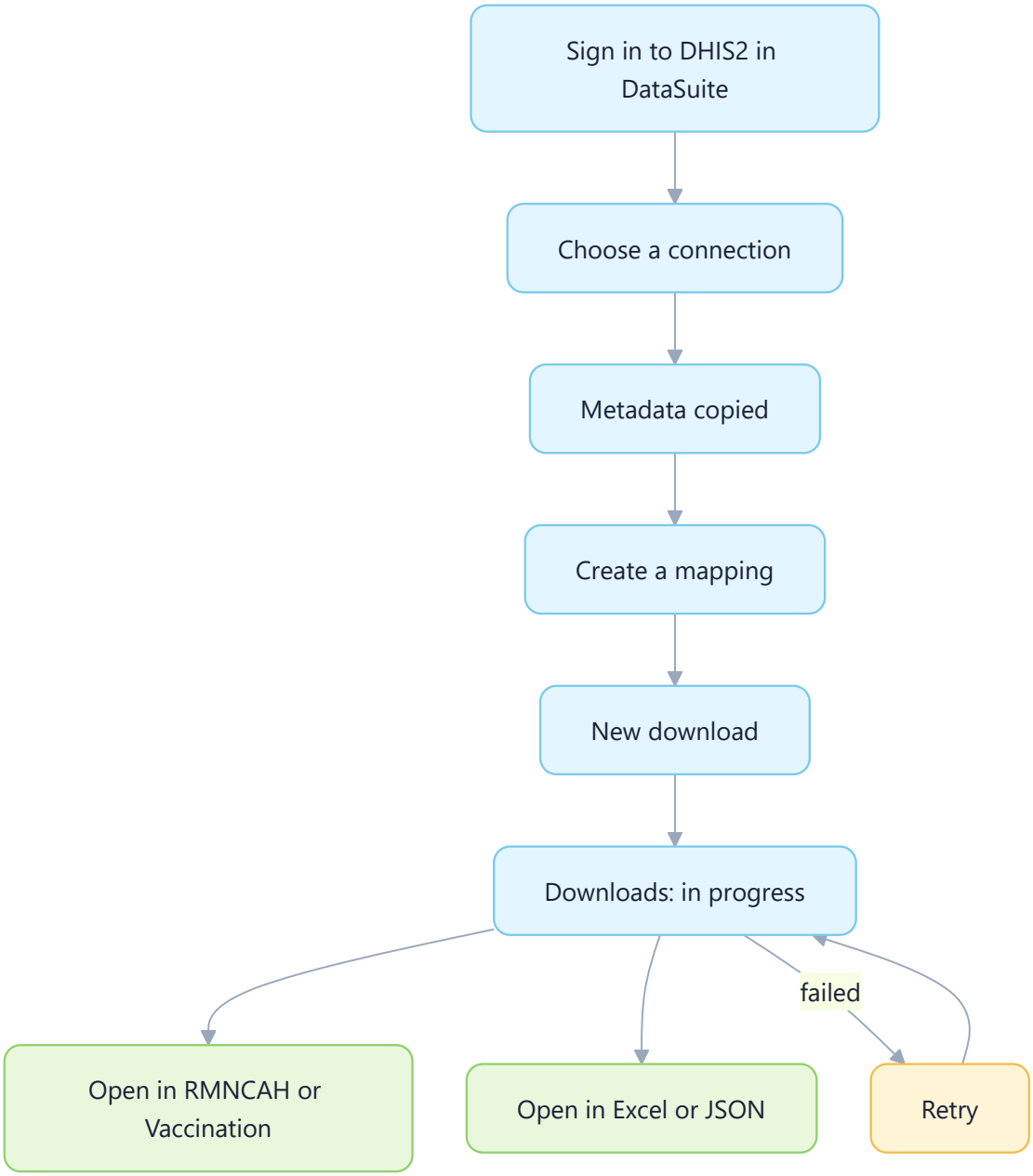
Figure 21.1 — *The Overview of a new connection: the three steps to a first download, and the DHIS2 server card*

Pages

Before you start: have a DHIS2 account on your country's server. Open the extractor from the Welcome page with **Open Data Extractor** (see First steps (p. 25)). The extractor is the first step of the Countdown analysis: its workbook is what the RMNCAH and Vaxx apps (p. 289) open.

> **Access**
> The extractor only reads from DHIS2, and only what your DHIS2 account may see. It cannot change or delete anything on the server, and it never sees your password or token: you sign in in DataSuite, which keeps them encrypted.

How it works



- **One connection at a time.** Each DHIS2 server you signed in to in DataSuite is a connection, with its own mappings, downloads and settings.
- **A copy of the metadata.** DataSuite keeps a copy of the server's data elements, indicators, data sets and organisation units, so searching while you map is instant.
- **Downloads in small requests.** A download is split into many small requests: more at once while the server keeps up, fewer when it struggles. You can pause and resume it, and retry it if it fails.

The workbook made with a Countdown mapping has the sheets the apps read: population, reporting completeness, services and health system data. What that data must contain: Analysis setup (p. 46) and the Indicator definitions (p. 509).

The extractor window

The sidebar shows the connection you are using (**Switch** chooses another), the four pages and, at the bottom, when the metadata was last copied, with **Sync metadata** to copy it again.

Page	What for
Overview	What needs attention, the latest downloads and the DHIS2 server. On a new connection: the three steps to a first download
Mappings	Create, edit, share and import mappings (Mapping indicators (p. 277))
Downloads	The downloads in progress and the history of finished and failed ones (Downloading data (p. 283))
Settings	The connection, the calendar, how downloads run, where files are saved, and when the metadata is refreshed

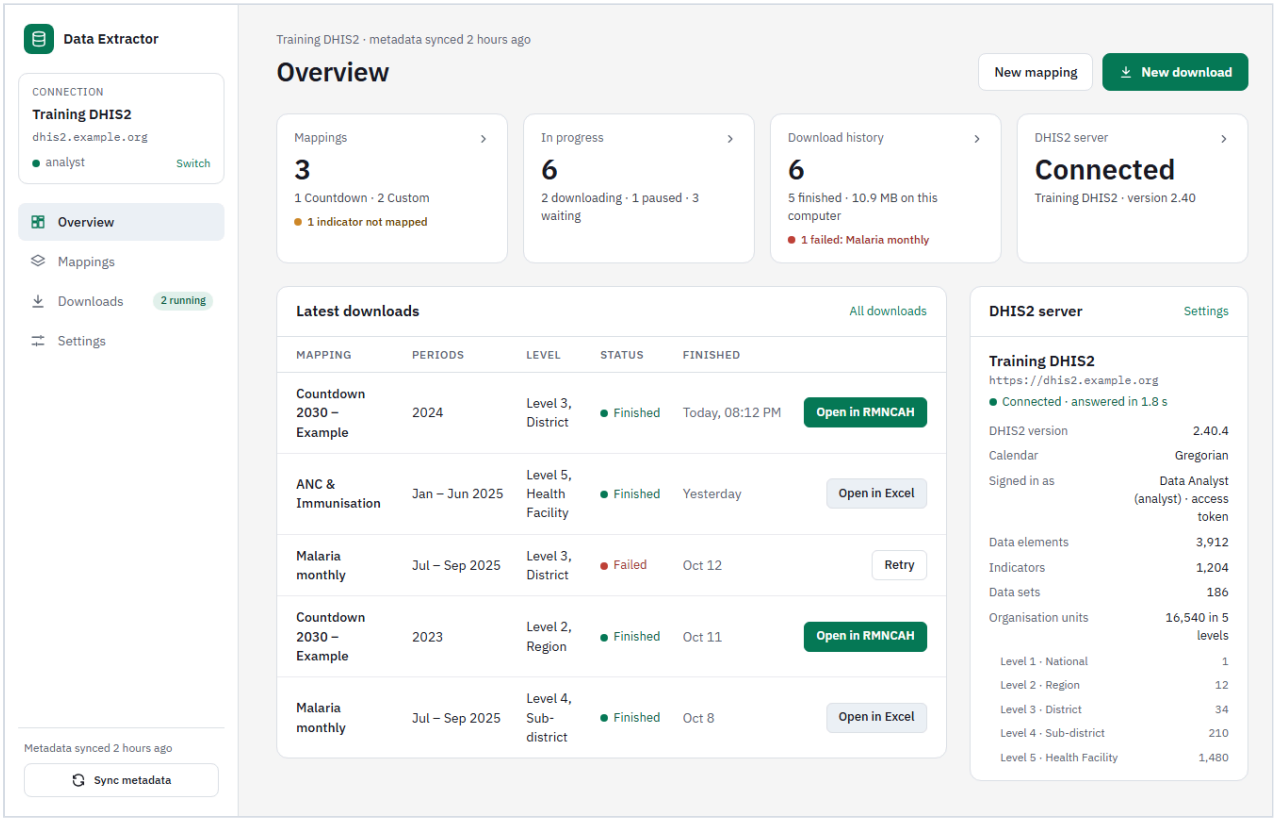


Figure 21.2 — *The Overview once there are mappings and downloads: the counts, the latest downloads and the server*

Common pitfalls

- **The wrong server.** Sign in to your country's production DHIS2 server, not a training or test copy, unless that is the one you mean to analyse.
- **Data you cannot see in DHIS2.** Access is set in DHIS2: if your account cannot see a data element or an area, the extractor cannot either. Ask your DHIS2 administrator.

- **A huge first download.** Try a new mapping on a short range first (for example **Last 12 months** at a high level), check the workbook, then download the full range.
- **A Custom mapping for the Countdown apps.** Only a download made with a **Countdown** mapping opens in the RMNCAH and Vaccination apps.

Next: DHIS2 login (p. 273).

21.1 DHIS2 login

How to sign in to your DHIS2 server in DataSuite, choose that connection in the Data Extractor and let it copy the server's metadata.

The extractor reads DHIS2 through a **connection**: a DHIS2 server you signed in to in DataSuite. You sign in once, in DataSuite's own dialog, and then choose that connection in the extractor. This is the first step, before you create a mapping (p. 277); after it, the extractor has a copy of the server's metadata to map with.

Before you start: have the address of your DHIS2 server and your DHIS2 username and password, or a personal access token.

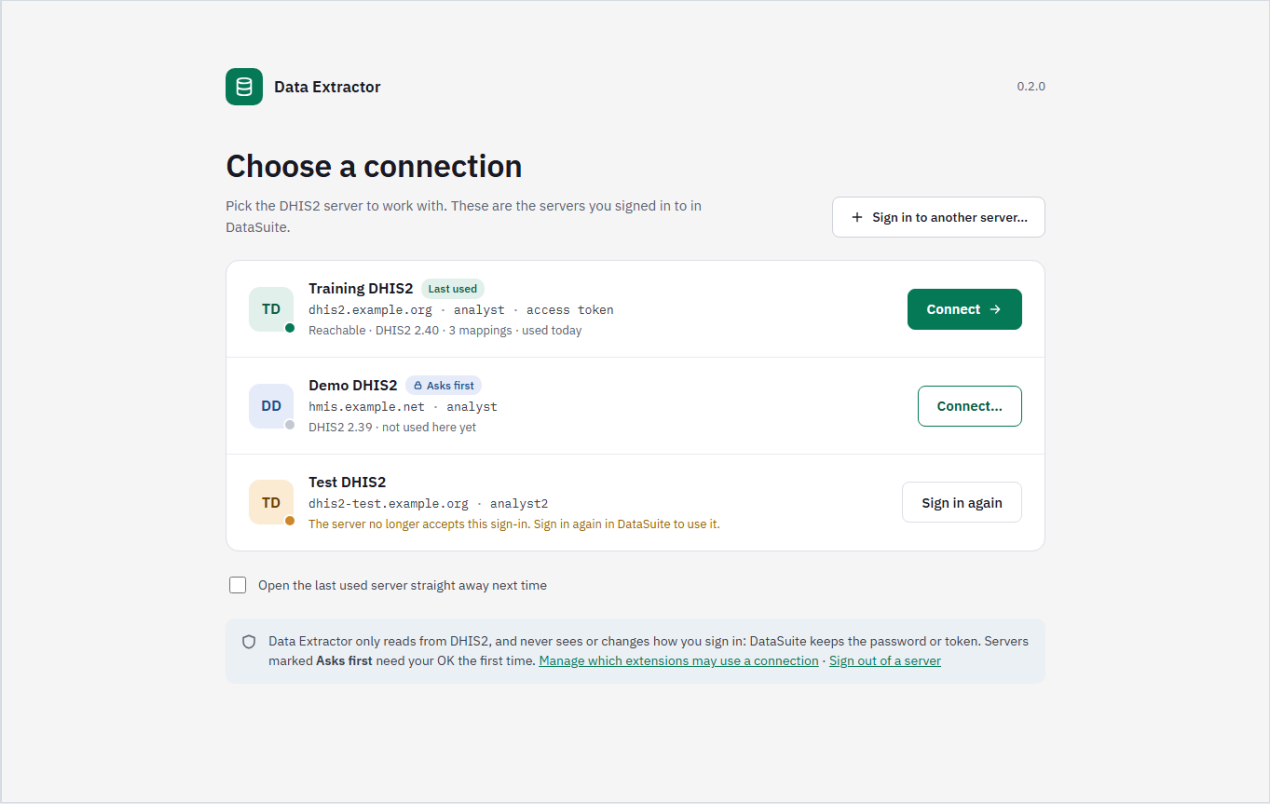
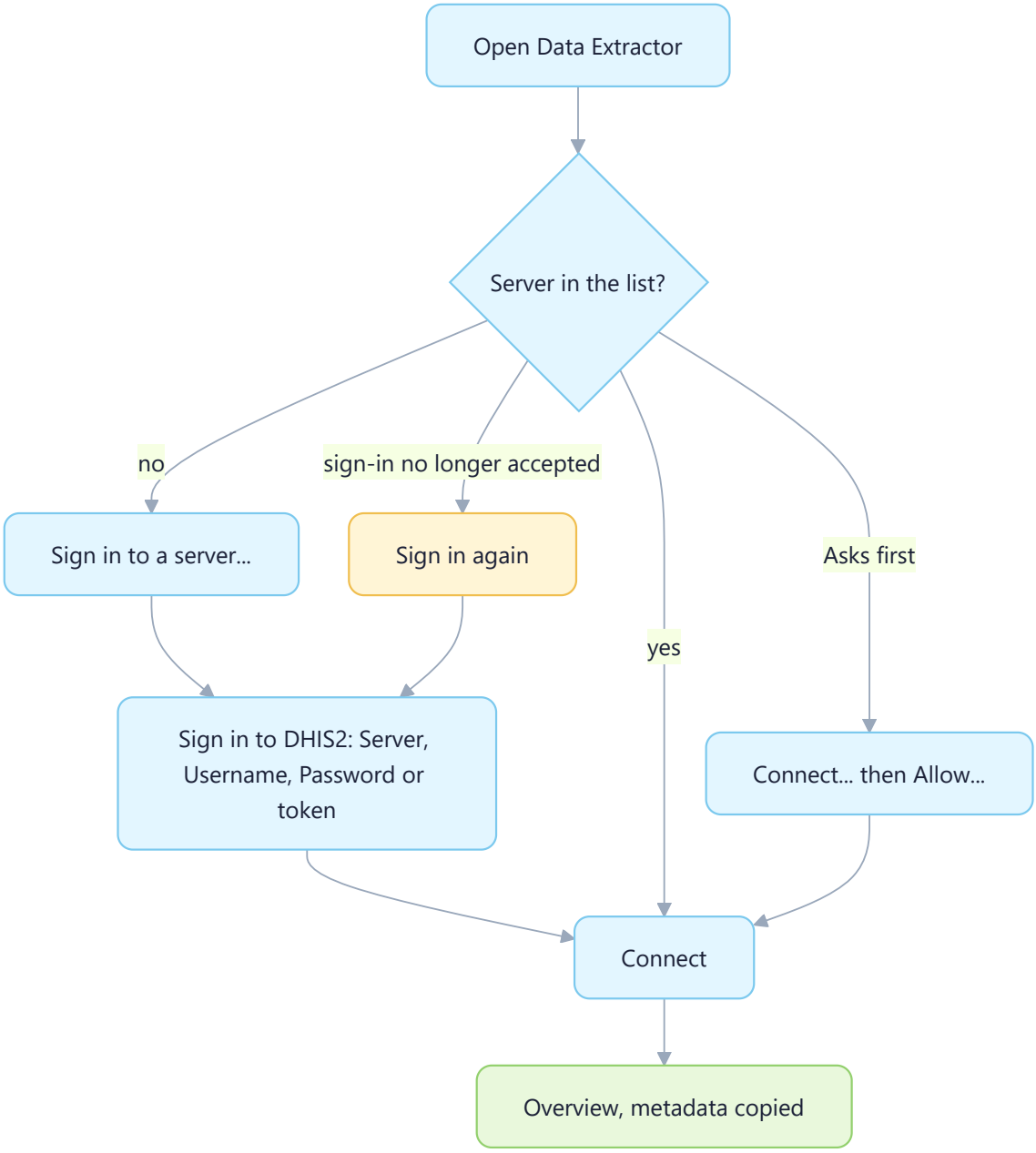


Figure 21.3 — *Choose a connection: the DHIS2 servers signed in to in DataSuite, each with its status*

Signing in



① **Open the extractor**

Click **Open Data Extractor** on the Welcome page (see First steps (p. 25)). It opens on **Choose a connection**, the list of the DHIS2 servers you signed in to in DataSuite. The first time, the extractor also brings over the mappings and downloads of the extractor that was built into earlier versions of DataSuite.

② Sign in to the server

If your server is not in the list, click **Sign in to a server…** (or **Sign in to another server…**). DataSuite's **Sign in to DHIS2** dialog opens:

- **Server**: the full address of the server, starting with `https://`, for example `https://dhis2.example.org`. Known servers are suggested as you type.
- **Username**: your DHIS2 username. Leave it empty if you use a personal access token.
- **Password or personal access token**: a personal access token is recommended (DHIS2 2.39 or later: **Profile** > **Edit profile** > **Personal access tokens**). It can be limited, can expire and can be revoked.

Click **Sign In**. DataSuite keeps the password or token encrypted; the extractor never sees it.

③ Choose the connection

Each server shows whether it answers, its DHIS2 version, its mappings and when you last used it. Click **Connect**.

- **Asks first**: the extractor has not used this connection yet. Click **Connect…**, read what allowing it means (it reads data and metadata with your account's access, and can't change anything), then **Allow…** and confirm in DataSuite.
- **The server no longer accepts this sign-in**: the password changed or the token expired. Click **Sign in again**.

Tick **Open the last used server straight away next time** to skip this list.

④ Let it copy the metadata

After you connect, DataSuite copies the server's metadata: data elements, indicators, data sets, category option combinations and organisation units. The **Overview** then shows how many items were copied, and the bottom of the sidebar shows when (**Metadata synced …**). The copy makes searching while you map instant, and it holds only what your account may see.

> **Access is set in DHIS2**
> If you cannot see a data set, a data element or an area in DHIS2, the extractor cannot see it either: it does not bypass DHIS2 permissions. Ask your DHIS2 administrator for access.

Keeping the metadata up to date

By default the copy is refreshed once a day. If something was added to DHIS2 since, click **Sync metadata** at the bottom of the sidebar, or **Settings** > **Metadata** > **Refresh now**. **Refresh the copy** in the same place sets when it refreshes by itself: **Each time the extractor opens**, **Once a day**, **Once a week** or **Only when I ask**.

Changing or leaving a connection

Only DataSuite changes a connection. **Settings** > **Connection** shows the server, the address, who is signed in and how, with **Test connection** to check that the server answers. To use another address, user or token, sign in again in DataSuite (**Sign in to a server…**): that makes a new connection, with its own mappings and downloads. To carry a mapping over, export it (**Mappings** > ⋯ > **Export to share**) and import it in the new connection. **Sign out…** and **Extension access…** are there too. **Switch**, in the sidebar, goes back to the list of connections.

Common pitfalls

- **The address of a page, not of the server.** Type the server's address (`https://dhis2.example.org`), not a link to a dashboard or app inside it. The message *The server's address starts with https://* means the address is not complete.
- **Invalid username or password.** Check them by signing in to DHIS2 in your browser, or use a personal access token.
- **Unable to reach the DHIS2 server.** Check your internet connection and the address; the server may also be down.
- **An item missing after signing in.** Click **Sync metadata**; if it is still missing, your account may not have access to it. More in the FAQs (p. 492).

Next: Mapping indicators (p. 277).

21.2 Mapping indicators

How to create a Countdown or Custom mapping in the Data Extractor, match each indicator to the DHIS2 data elements, indicators or data sets behind it, and share mappings.

A **mapping** says which DHIS2 data make up each indicator you download: for example, that "ANC 1st visit" is the sum of two data elements of the antenatal care form. You make it once per server, after signing in (p. 273), and reuse it for every download (p. 283). For the Countdown analysis, create a **Countdown** mapping: its workbook is the one the RMNCAH and Vaxx apps load.

Before you start: connect to your DHIS2 server and let the metadata copy finish (DHIS2 login (p. 273)).

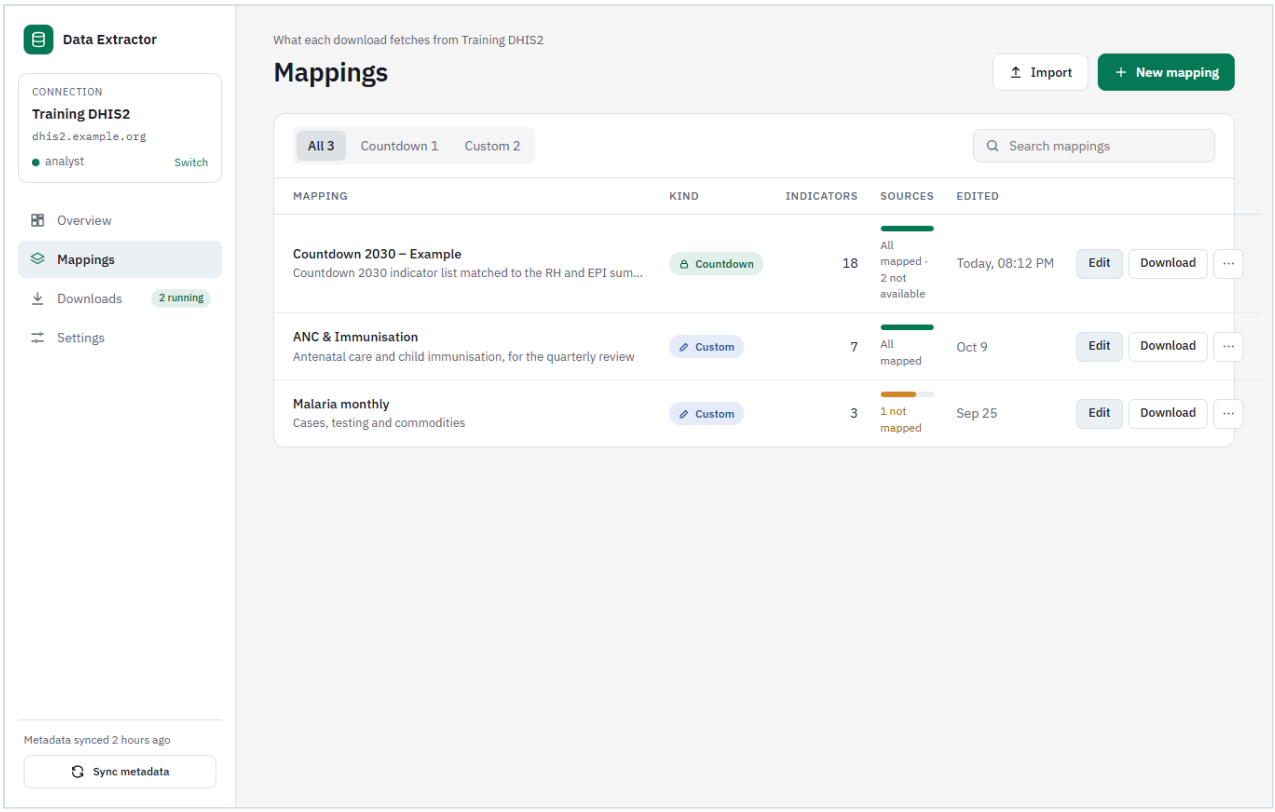
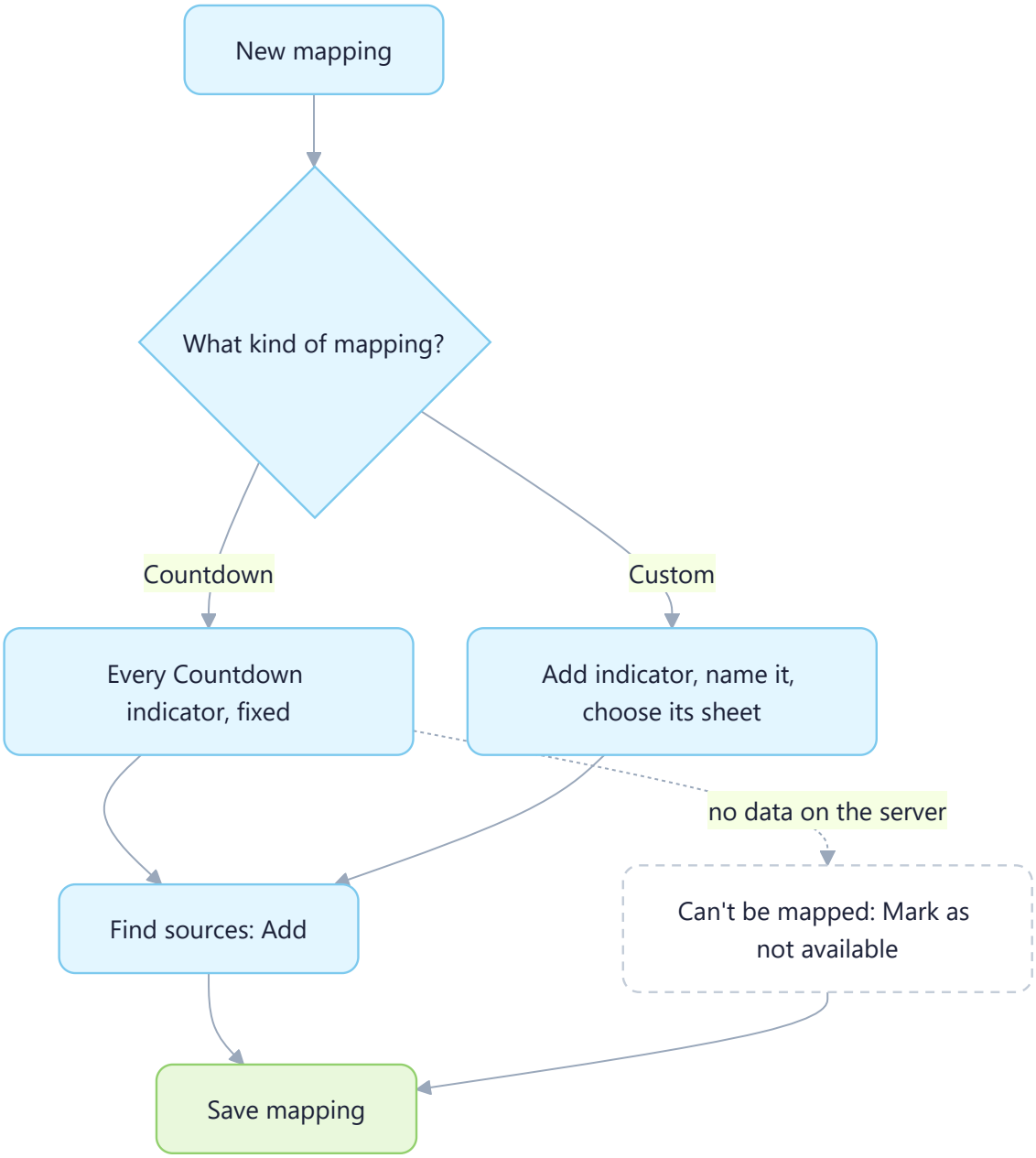


Figure 21.4 — *Mappings: each mapping with its kind, its indicators and how many are mapped*

Creating a mapping



① **Click New mapping**

On **Mappings** (or on the **Overview** of a new connection), click **New mapping**.

② Choose the kind

- **Countdown** starts from the Countdown 2030 indicator list. The indicators are fixed: you match each one to the data on your server. Its downloads come in the parts the Countdown apps read: population, reporting completeness and services (and the health system figures).
- **Custom** starts empty. You name your own indicators and choose the data behind each: for programme reviews, other health programmes and one-off analyses.

The kind cannot be changed later. Click **Next**.

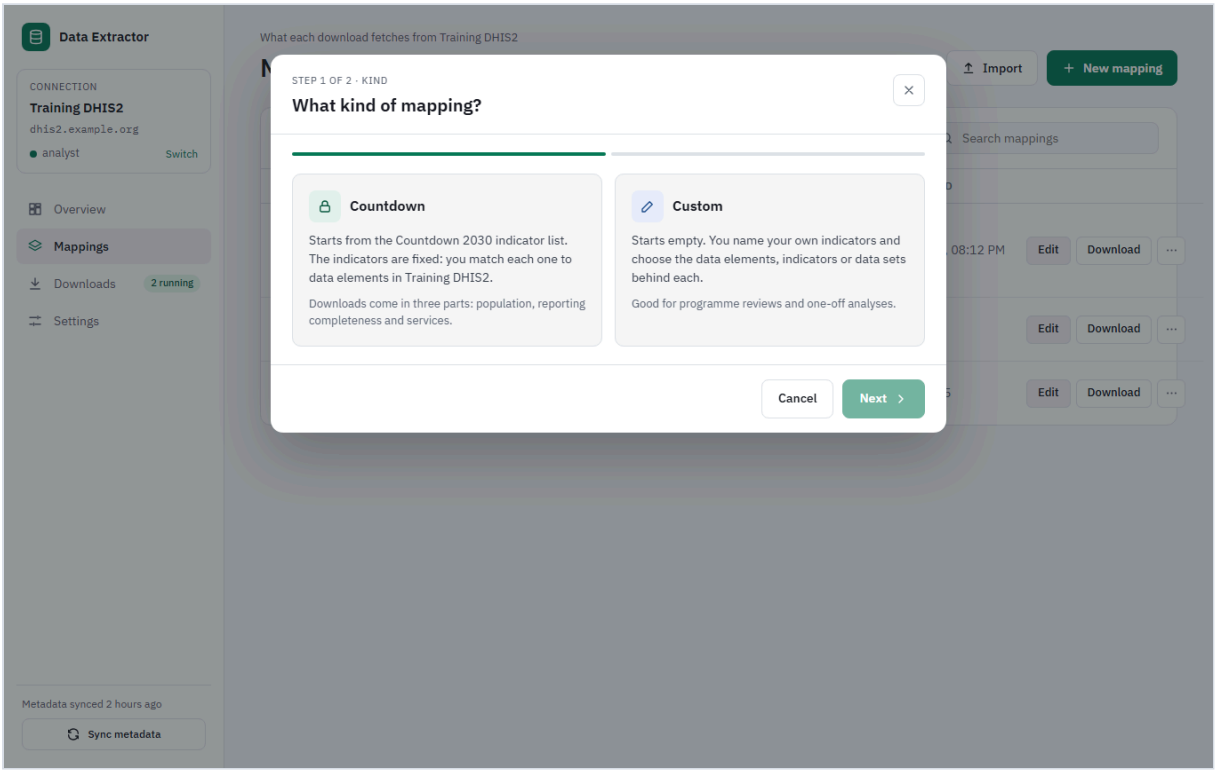


Figure 21.5 — *What kind of mapping? Countdown or Custom*

③ Name it

Give the mapping a **Name** and, if you like, a **Description**. The name is shown in **Mappings** and **New download** and used in file names; a Countdown mapping is named after the server to begin with. Click **Create mapping**: the mapping editor opens.

④ Match the indicators, then save

Select an indicator on the left, search for its data under **Find sources** on the right and click **Add** (details below). When you are done, click **Save mapping**. **Discard** leaves without saving.

The mapping editor

The editor has three columns: the indicators, the selected indicator, and **Find sources**. The header counts how many indicators are mapped; **Sheets in the Excel download** shows how the workbook will be organised.

A Countdown mapping

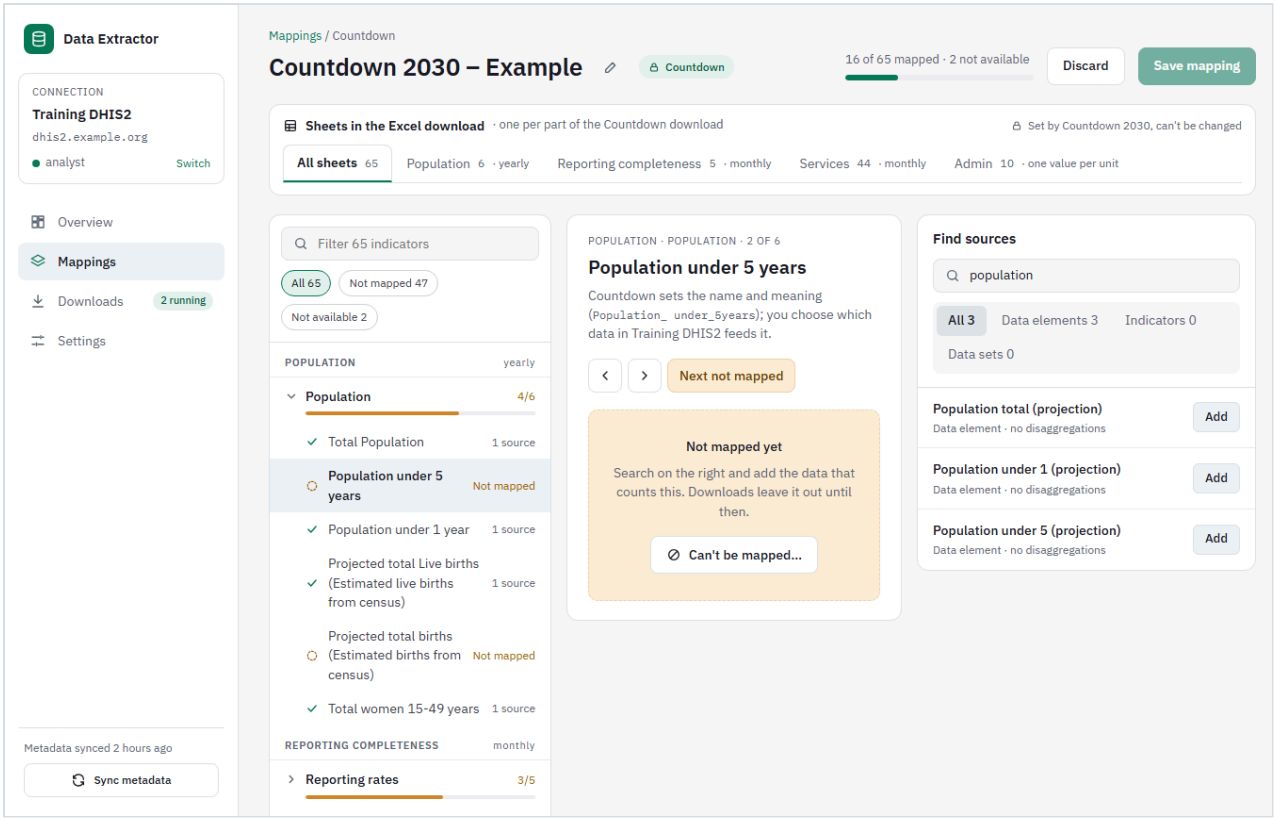


Figure 21.6 — *The editor of a Countdown mapping: the indicators by part, the selected indicator, and a search for its sources*

- The indicators are grouped by part, one sheet of the workbook each: **Population** (yearly), **Reporting completeness** (monthly), **Services** (monthly) and **Admin** (one value per unit, the health system figures). The sheets are set by Countdown 2030 and can't be changed.
- Countdown sets each indicator's name and meaning; you choose which data on your server feed it. **Not mapped** in the filter shows only those still to do, and **Next not mapped** jumps to the next one.
- If the server has no data for an indicator, click **Can't be mapped...**, choose why (**Not collected**, **No matching data** or **Other reason**), add a note if useful and click **Mark as not available**. It then counts as done, and the workbook keeps its column, empty. **Map it instead** undoes this.
- An indicator that is neither mapped nor marked as not available is left out of downloads until you map it.

A Custom mapping

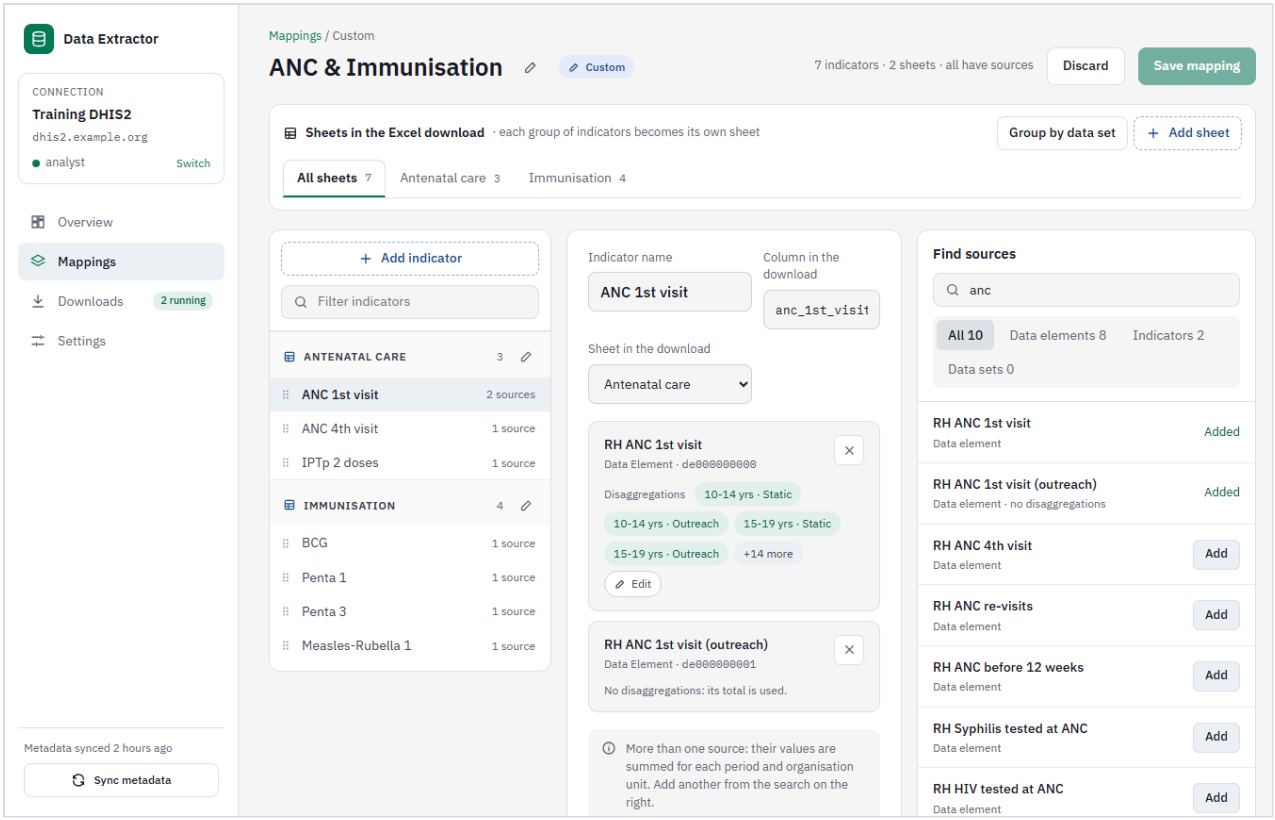


Figure 21.7 — *The editor of a Custom mapping: indicators on two sheets, a source with its disaggregations, and the search*

- **Add indicator** adds one. Give it an **Indicator name**; **Column in the download** follows the name until you change it; **Sheet in the download** chooses its sheet.
- **Add sheet** adds a sheet (each sheet becomes a sheet of the workbook); the pencil renames it. **Group by data set** puts each indicator on a sheet named after the data set of its first source.
- Drag indicators on the left to change the column order. **Delete indicator** removes the selected one.

Finding and adding sources

- Type part of a name, a code or a form number under **Find sources**. The results are split into **Data elements**, **Indicators** and **Data sets**; click **Add**.
- **A data element**: by default all its disaggregations (category option combinations) count. Click **Edit** on the source to choose them under **Which disaggregations count?** (**Include all**, **Exclude all**, or one by one).
- **A DHIS2 indicator** is added as the data elements it is calculated from.
- **A data set** gives its reporting rate, reports received and reports expected: use it for the reporting completeness indicators.
- **More than one source**: their values are summed for each period and organisation unit.

Sharing and reusing mappings

In the list on **Mappings**, each mapping shows its kind, its number of indicators and how many are mapped (**All mapped**, **N not mapped**, **N not available**). Its buttons:

- **Edit** opens the editor; **Download** opens New download (p. 283) with this mapping.
- ⋯ > **Export to share** saves the mapping as a `.json` file (the mapping only, never data or passwords); **Duplicate** makes a copy; **Delete…** deletes it (downloads made with it are kept).
- **Import** reads a shared `.json` file and checks it against your server first: which indicators it matches, which sources your server lacks (those indicators come in as not mapped). If you already have a mapping of that name, choose **Import as a new mapping** or **Replace mine**.

If the extractor closed while you were editing, **Mappings** offers the unsaved mapping again: **Keep and edit** or **Discard**.

Common pitfalls

- **The same data counted twice.** Adding a data element and a DHIS2 indicator built from it, or a total and its parts, sums them both. Check the sources of each indicator.
- **Disaggregations left in by mistake.** Some data elements mix age groups or places of service; use **Edit** to keep only those that match the indicator's definition (Indicator definitions (p. 509)).
- **Leaving indicators unmapped.** In a Countdown mapping, mark what the server does not collect as **not available** rather than leaving it, so you know it was checked.
- **Forgetting to save.** Changes are kept only when you click **Save mapping**.

Next: Downloading data (p. 283).

21.3 Downloading data

How to start a download in the Data Extractor (mapping, time range, organisation unit level), follow it in Downloads, and open the result in the RMNCAH or Vaccination app, in Excel or as JSON.

A download fetches the values of a mapping's indicators from DHIS2 for the periods and the organisation unit level you choose. It is the last step in the extractor, after mapping the indicators (p. 277); the result is the workbook you open in the RMNCAH or Vaxx app (p. 295), or in Excel. Large downloads are split into many small requests, so they can run for a while, be paused and be picked up later.

Before you start: have a saved mapping with at least one indicator mapped (Mapping indicators (p. 277)).

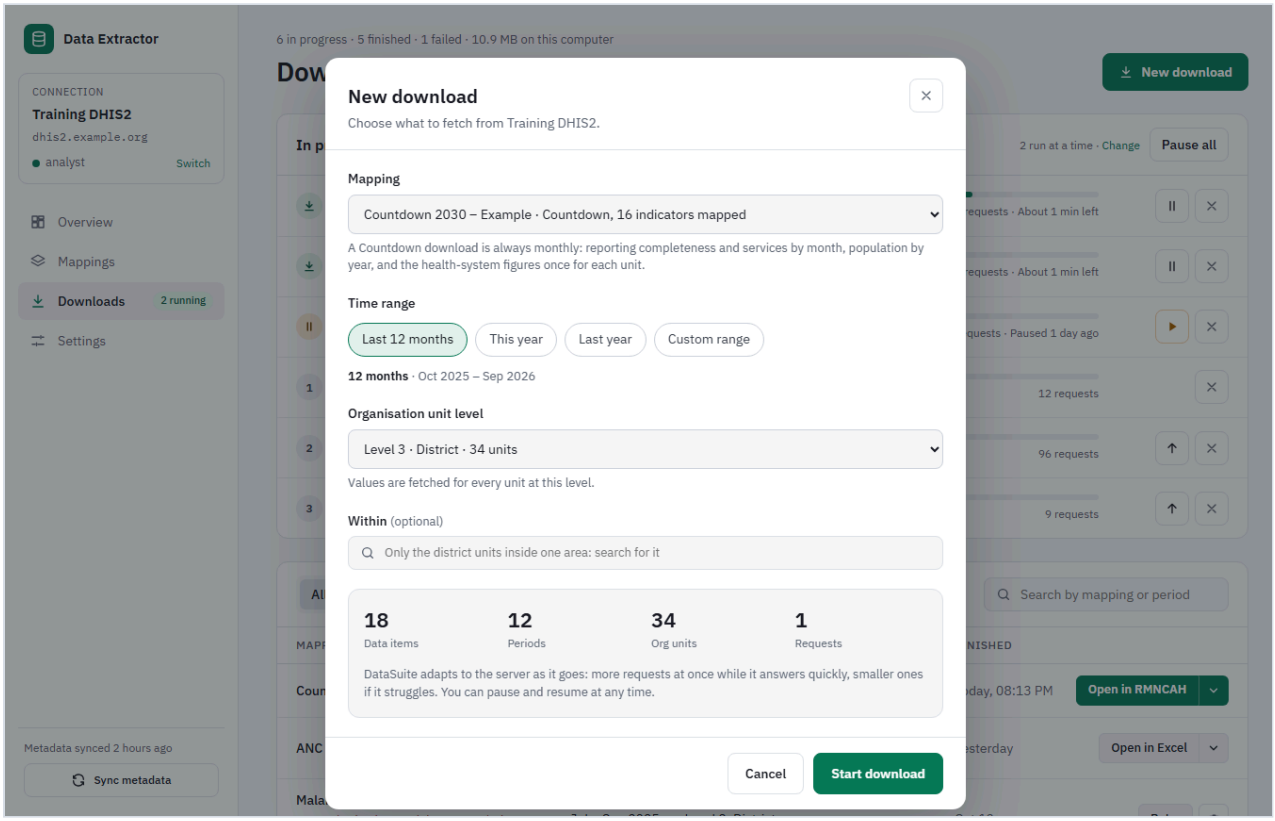


Figure 21.8 — *New download: the mapping, the time range, the organisation unit level, and what the download will take*

Starting a download

1. **Click New download**
   On **Downloads** or **Overview**, or **Download** next to a mapping on **Mappings**.

2. **Choose the mapping**
   Choose it under **Mapping**. A Countdown download is always monthly: reporting completeness and services by month, population by year, and the health system figures once for each unit.

③ Choose the time range

Click **Last 12 months**, **This year**, **Last year** or **Custom range**. With **Custom range**, pick the first and the last month under **From month** and **To month**; with a Custom mapping you can **Pick by Monthly** or **Yearly**. On a server that uses the Ethiopian calendar, **Calendar** switches between **Gregorian** and **Ethiopian** months. The line under the buttons shows how many months the range holds.

④ Choose the organisation unit level

Under **Organisation unit level**, choose the level the values are fetched for, for example the district level: every unit at that level is downloaded. **Within** (optional) limits the download to the units inside one area: search for the area and select it.

⑤ Check the estimate, then start

The box at the bottom counts the **Data items**, **Periods**, **Org units** and **Requests**, and lists any indicators left out because they are not mapped yet. A very large number of requests means a long download. Click **Start download**: the extractor opens **Downloads**.

> **Countdown download**
> For the Countdown analysis, download the last five complete years (**Custom range**, from January of the first year to December of the last) at the district level (often Level 3).

Following a download

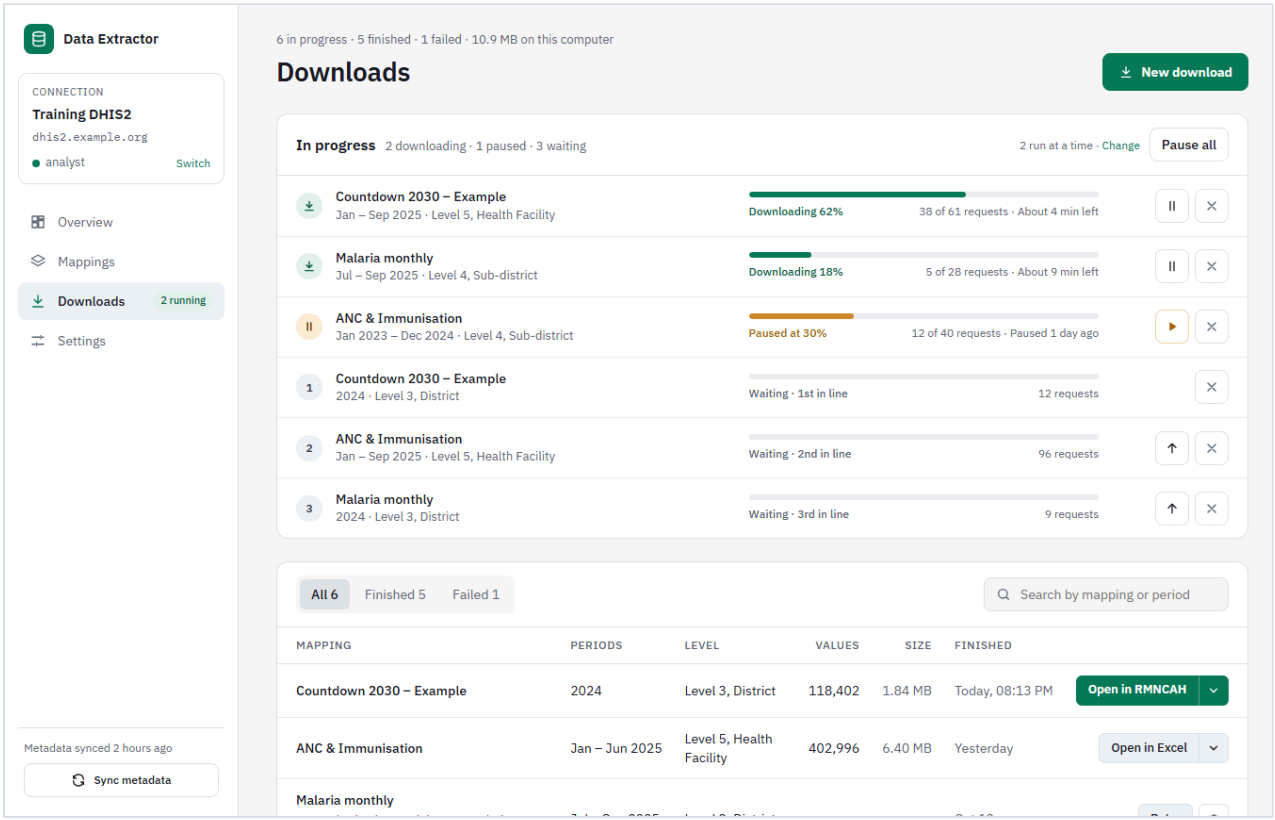


Figure 21.9 — *Downloads: the downloads in progress, running, paused and waiting, and the history*

In progress lists the downloads that are running, paused or waiting their turn, each with its progress, the requests done and the time left. A few run at the same time (2 by default; **Change** opens Settings), the others wait in line.

Button	What it does
Pause / Resume	Pauses a running download, or resumes a paused one where it stopped
Move up	Moves a waiting download up the line
Cancel	Stops the download and moves it to the history as failed; what it fetched so far is kept for a retry
Pause all / **Resume all**	Pauses or resumes every download

Downloads keep running while you work on other pages or in other tabs.

Opening the result

When a download finishes it moves to the history, under **In progress**: **All**, **Finished** or **Failed**, with a search by mapping or period. Each finished download shows its periods, level, number of values and size.

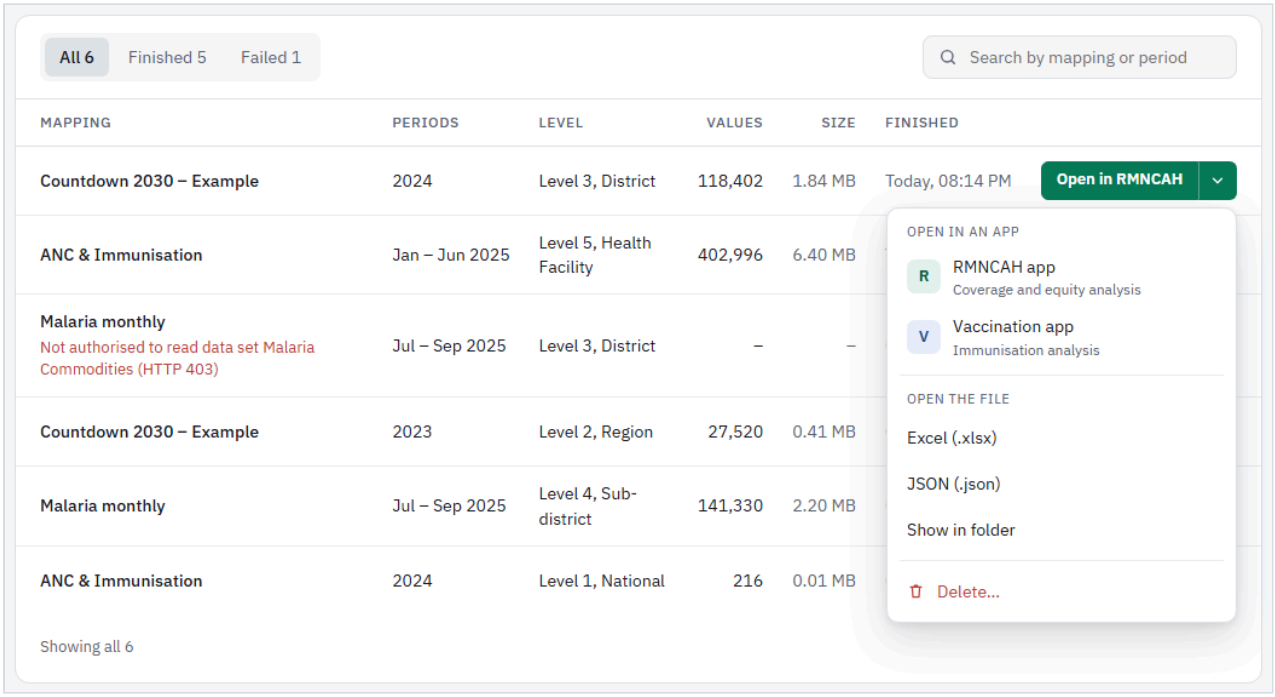


Figure 21.10 — *The ways to open a finished download: in an app, as a file, or in its folder*

- **Open in RMNCAH**: a download made with a Countdown mapping opens straight in the RMNCAH app (the arrow beside it offers the **Vaccination app**, the Vaxx app). The workbook is written to `Documents\Countdown data\<server>`, and the app saves its dataset next to it. This needs the Countdown Analytics extension; **Get the analysis apps** shows it if it is missing.
- **Open in Excel** or **JSON (.json)**: writes the file to the folder of the settings and opens it.
- **Show in folder** opens that folder; **Delete…** deletes the downloaded data from this computer.

A failed download shows why under its name; click **Retry** to run it again.

The Countdown workbook has the sheets `Service_data`, `Population_data`, `Reporting_completeness` and `Admin_data`, with every column the apps expect, in the template's order, empty where the mapping has no indicator for it. A Custom mapping's workbook has one sheet per sheet of the mapping. What the apps need in it: Analysis setup (p. 46).

Where files are saved

A finished download is written as a file only when you open it, into the folder of **Settings** > **Files** > **Save to** (by default `Documents\DataSuite\Downloads`).

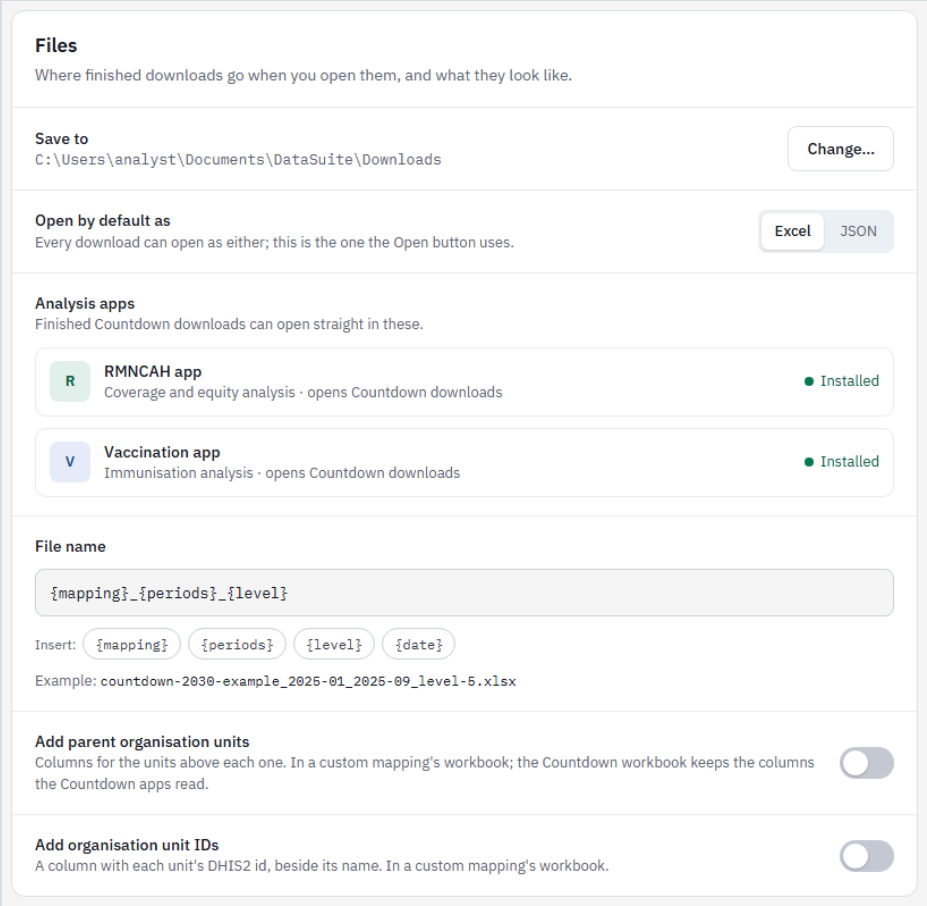


Figure 21.11 — *Settings > Files: the folder, the default format, the analysis apps and the file name*

- **Open by default as**: **Excel** or **JSON**, the format the Open button uses.
- **File name**: a pattern built from `{mapping}`, `{periods}`, `{level}` and `{date}`; the example under it shows the result.
- **Add parent organisation units** and **Add organisation unit IDs**: extra columns in a Custom mapping's workbook (the Countdown workbook keeps the columns the apps read).
- On a server that uses the Ethiopian calendar, **Settings** > **Calendar & periods** chooses whether periods are picked and written in the Ethiopian or the Gregorian calendar.

If the server is slow

Settings > **Downloads** sets how hard the extractor works the server. Lower these if the server is shared or slow.

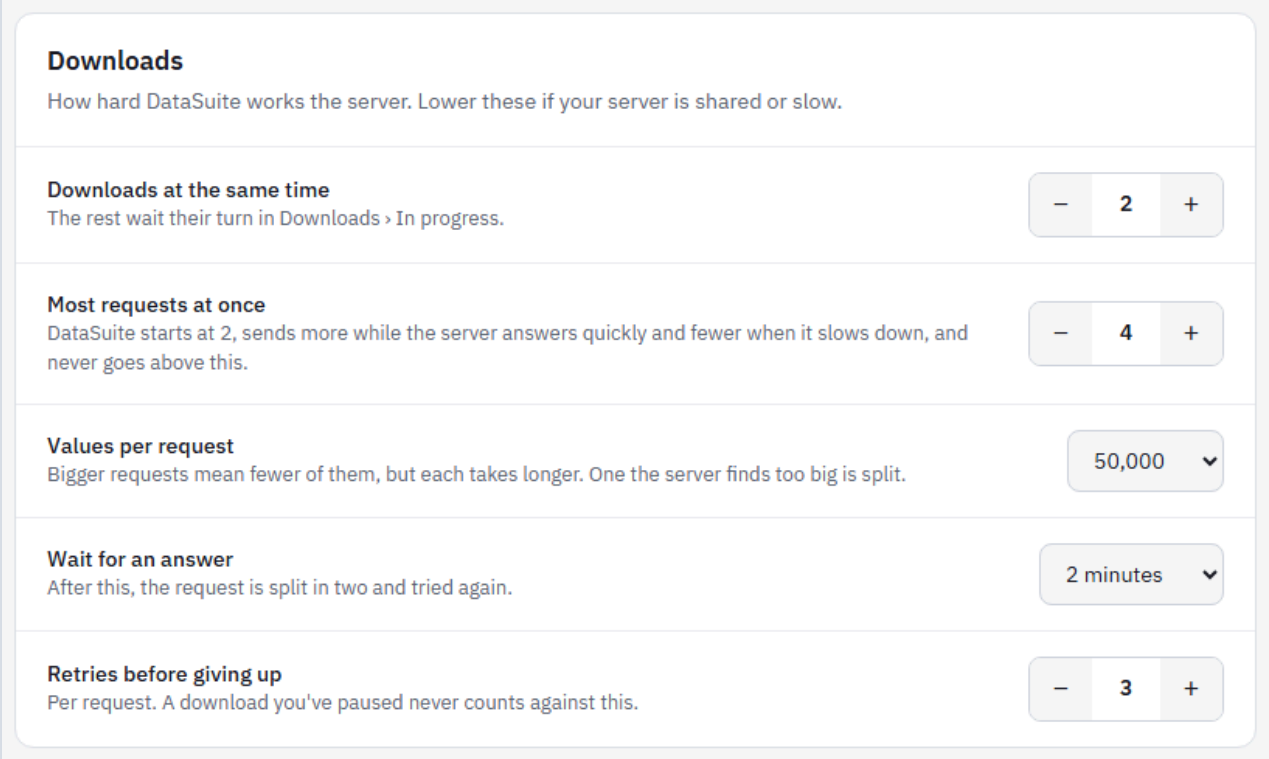


Figure 21.12 — *Settings > Downloads: downloads at the same time, requests at once, values per request, wait and retries*

Setting	Default	What it means
Downloads at the same time	2	The others wait their turn in **In progress**
Most requests at once	4	The extractor starts at 2, sends more while the server answers quickly, and never goes above this
Values per request	50,000	Bigger requests mean fewer of them, but each takes longer; one the server finds too big is split
Wait for an answer	2 minutes	After this, the request is split in two and tried again
Retries before giving up	3	Per request; a paused download never counts against this

Click **Save changes** at the top of Settings.

Common pitfalls

- **A download that fails with "Not authorised".** Your account cannot read one of the mapping's data sets or data elements. Ask your DHIS2 administrator, or remove that source from the mapping, then **Retry**.
- **Indicators left out.** The estimate lists indicators that are not mapped yet: they are not in the download. Map them, or mark them as not available, first.

- **A very low level over many years.** Facility level over five years can mean thousands of requests. Download the district level, or fewer years at a time.
- **Opening a Custom download in the apps.** Only a Countdown mapping's download opens in the RMNCAH and Vaccination apps. More in the FAQs (p. 492).

Next: load the workbook in the app: Load Data (p. 295).

CHAPTER 22

RMNCAH and Vaxx apps

The RMNCAH and Vaxx apps: what each is for, who uses which, a tour of the app window, the page groups in the order you use them, and how the two apps differ.

Analyse routine health facility data

Both apps take the workbook from the Data Extractor (p. 269), check and adjust the data, choose the denominators, and estimate coverage, inequalities and trends. **RMNCAH** covers maternal, newborn and child health; **Vaxx** covers immunization. RMNCAH Vaxx

Learn about the analysis order *(p. 291)*

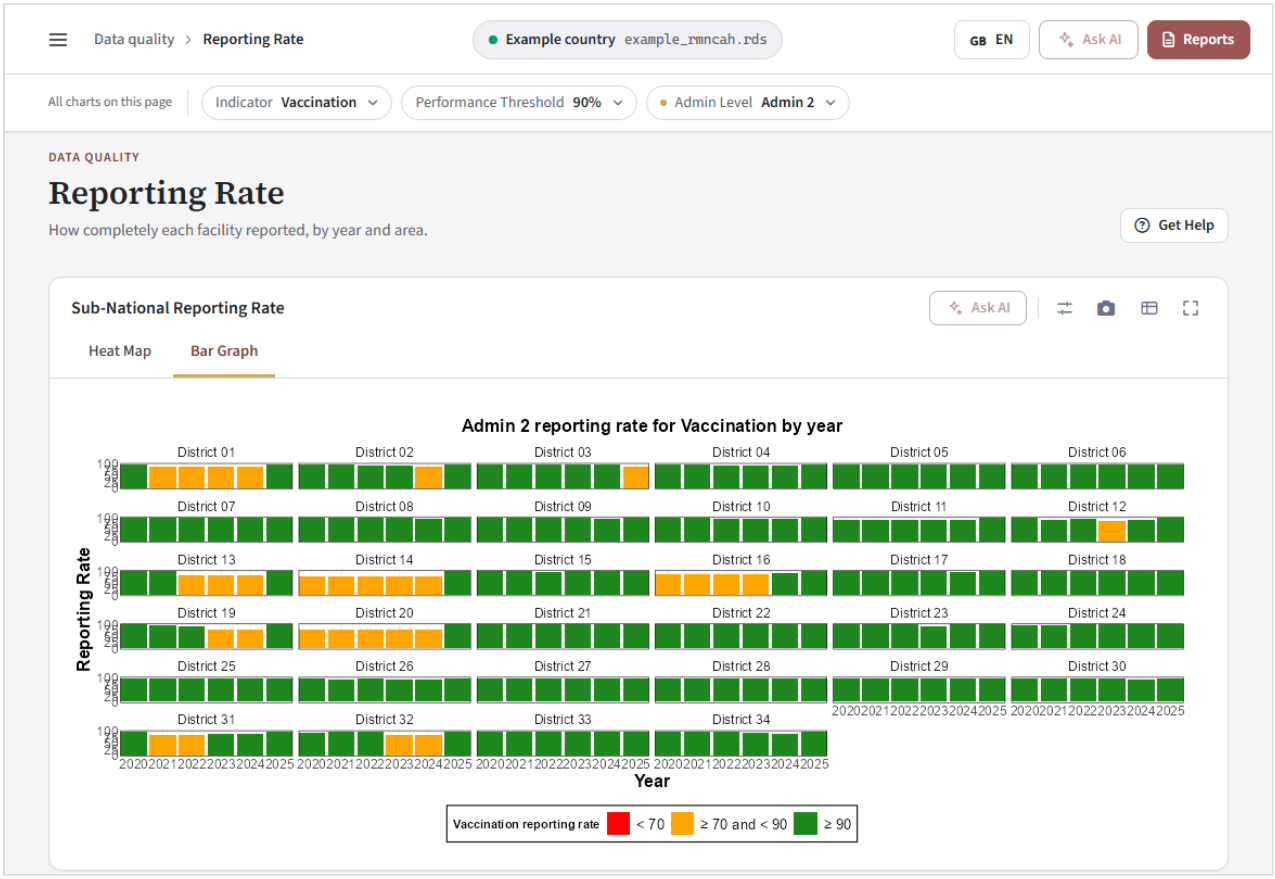


Figure 22.1 — *The RMNCAH app on the Reporting Rate page: top bar, filter bar, page header and a chart card*

Page groups

Where each guide's pages sit in the sidebar:

Sidebar	Pages	Guide
Start	Introduction, Load Data	Load Data (p. 295)

Data quality > Data Quality Assessment	Reporting Rate, Outlier Detection, Data Missingness, Internal Consistency, Overall Data Quality Score	Data Quality Assessment (p. 306)
Data quality > Data Adjustment	Data Adjustment, Data Adjustment Changes	Data Adjustment (p. 322)
Denominators	Population Trend, Denominator Selection	Denominators (p. 333)
Analysis > National Analysis	National Coverage, Coverage Target, Continuum of Care (RMNCAH), National Inequality (Routine Data, Survey Data)	Coverage, equity and targets (p. 342), Advanced analysis (p. 371)
Analysis > Sub-National Analysis	Sub-National Coverage, Sub-National Inequality, Coverage Target	Sub-national analysis (p. 359)
Analysis > Mortality (RMNCAH)	Institutional Mortality, Mortality Mapping, Mortality Completeness	Mortality (p. 381)
Analysis > Service Utilization (RMNCAH)	Service Utilization DQA, National Utilization, Sub-National Utilization, MCH Preventive vs Curative Index	Service utilization and health system (p. 391)
Analysis > Health System Performance (RMNCAH)	National Health System, Sub-National Health System, PHC Performance, Public/Private Share	Service utilization and health system (p. 391)
Analysis > Bayesian Analysis (RMNCAH)	National Analysis, Sub-National Analysis	Advanced analysis (p. 371)
Top bar	Reports	Reports (p. 415)

The two apps

- **RMNCAH** covers reproductive, maternal, newborn, child and adolescent health (p. 503): antenatal care, deliveries, postnatal care, vaccination, outpatient and inpatient visits, deaths. On top of coverage and equity it has pages for mortality, service use, the health system and modelled (Bayesian) estimates.
- **Vaxx** covers immunization only: the routine vaccines, dropout (p. 498), and zero-dose and under-vaccinated (p. 505) children. Its indicator list, survey inputs and pages are narrowed to vaccination.

Who uses which

- **RMNCAH**: the country team preparing a Countdown review or an annual health sector review: monitoring and evaluation officers, HMIS (p. 499) and DHIS2 (p. 498) data managers, maternal, newborn and child health programme managers, and analysts.
- **Vaxx**: immunization programme managers and national immunization analysts who need routine coverage that is checked and comparable with surveys, nationally and by district.

Neither app needs any programming: everything is done with pages, filters, charts and tables.

The app window

Both apps have the same window.

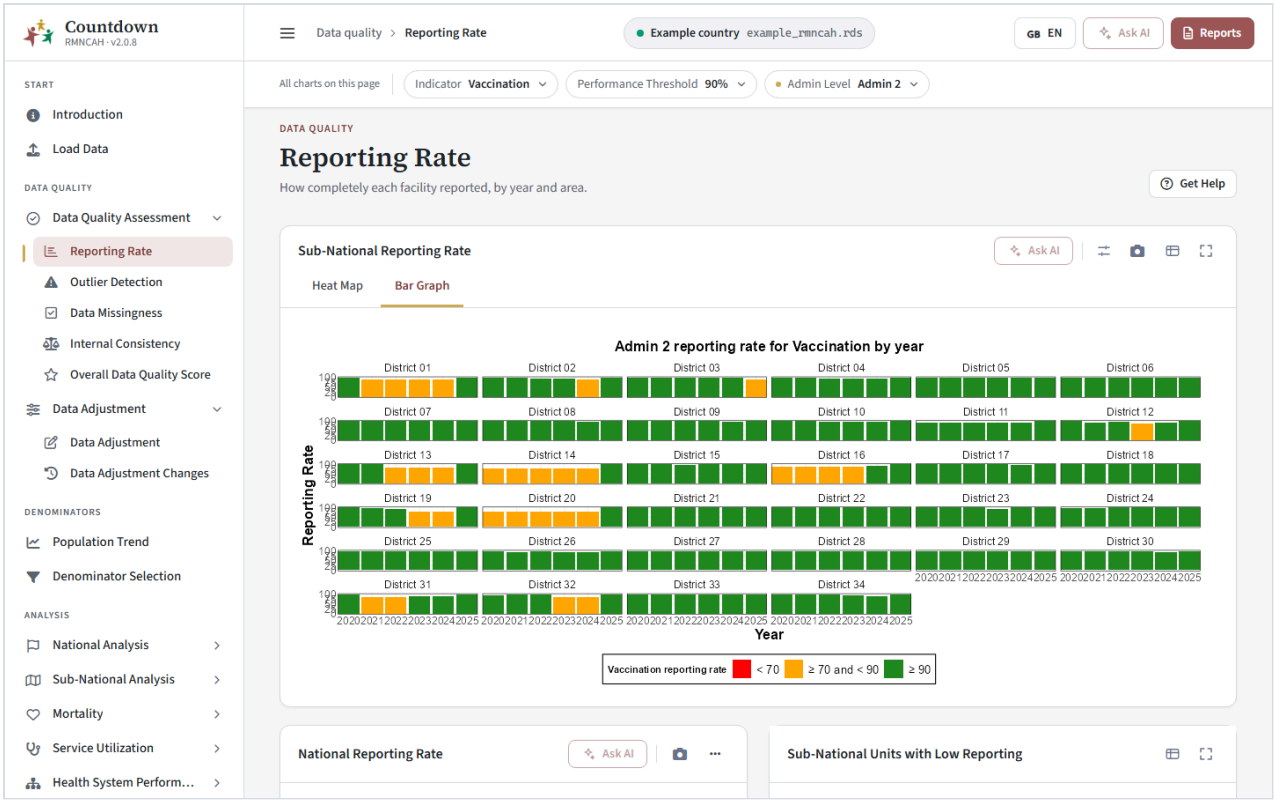
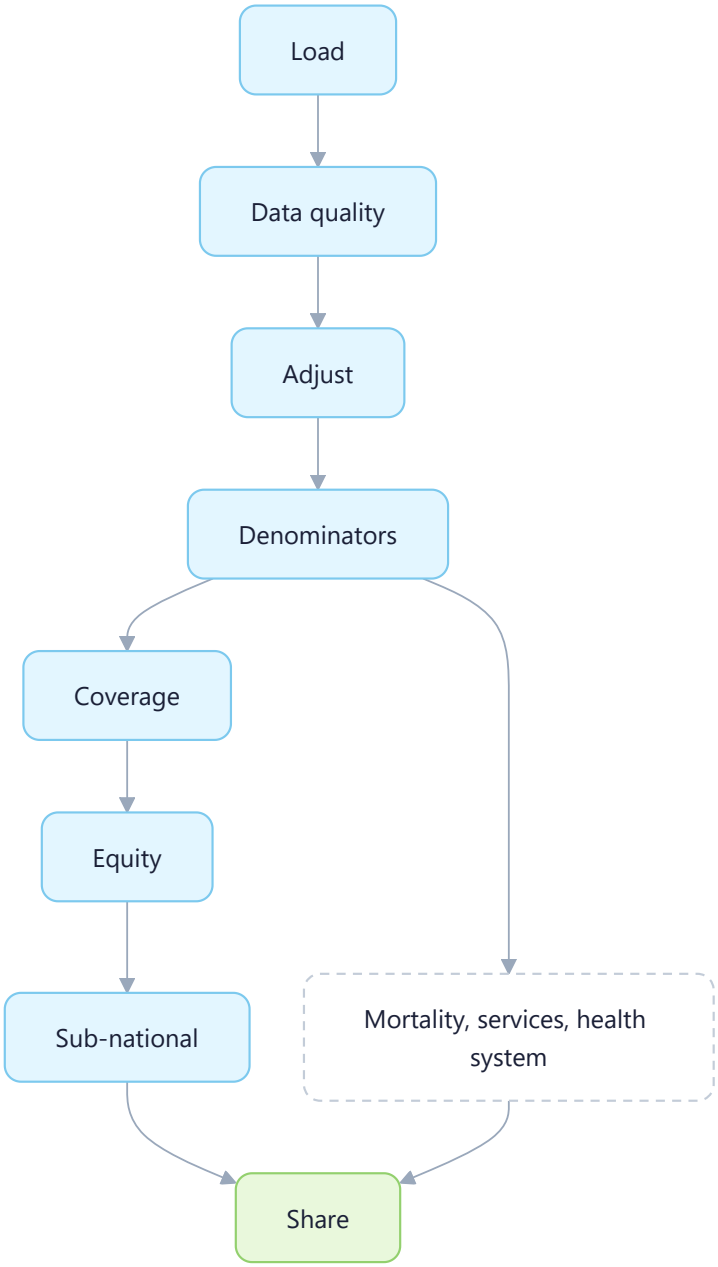


Figure 22.2 — *The RMNCAH app on the Reporting Rate page: sidebar, top bar, page header, filter bar and a chart card*

- **Sidebar** (left): the pages, in sections: **Start**, **Data quality**, **Denominators** and **Analysis**. A group with an arrow opens to show its pages. Pages you cannot use yet are greyed out: the Denominators and Analysis pages open once the data is loaded and adjusted. Clicking one shows *Load and finish a dataset first to unlock this page*.
- **Top bar**: where you are (section and page), the dataset chip (country and file name), the language switch (English, French, Portuguese), **Ask AI** and **Reports**. See Exports, reports and AI (p. 411) and Reports (p. 415).
- **Page header**: the section, the page title and a one-line subtitle. **Get Help** opens the page's guide on this site. Some pages also have **Generate Report**. Analysis pages show the denominators in use.
- **Filter bar**: **All charts on this page**, then chips such as **Indicator**, **Admin Level** or **Performance Threshold**. Click a chip to choose; the choice changes every card on the page.
- **Cards**: each chart or table sits in a card with a title, often tabs (one per indicator or view), **Ask AI**, and toolbar buttons to customise the chart, download it as an image, download its data (Excel) and expand it. A table may have its own control marked *This table only*.

Analysis order

Each step uses the result of the one before. The guide explains the pages; the method page explains why and how to read the results.



① Load the data

Load Data. Guide: Load Data (p. 295). Method: Analysis setup (p. 44).

② Assess data quality

The **Data Quality Assessment** pages. Guide: Data Quality Assessment (p. 306). Method: Data quality assessment (p. 57).

③ Adjust the data

Data Adjustment and **Data Adjustment Changes**. This step opens the pages that follow. Guide: Data Adjustment (p. 322). Method: Data adjustment (p. 75).

④ Choose the denominators

Population Trend and **Denominator Selection**. Guide: Denominators (p. 333). Method: Denominators (p. 95).

⑤ Estimate coverage

National Coverage, **Coverage Target**, and in RMNCAH **Continuum of Care** and **Bayesian Analysis**. Guides: Coverage, equity and targets (p. 342), Advanced analysis (p. 371). Method: Coverage (p. 115).

⑥ Look at equity

National Inequality (Routine Data, Survey Data). Guide: Coverage, equity and targets (p. 342). Method: Equity (p. 153).

⑦ Go sub-national

The **Sub-National Analysis** pages. Guide: Sub-national analysis (p. 359). Method: Sub-national analysis (p. 165).

⑧ Mortality, service use and the health system (RMNCAH)

The **Mortality**, **Service Utilization** and **Health System Performance** pages. Guides: Mortality (p. 381), Service utilization and health system (p. 391). Methods: Mortality (p. 179), Service utilisation (p. 204), Health system (p. 225).

⑨ Share the results

Reports, and the download buttons on every card. Guides: Reports (p. 415), Exports, reports and AI (p. 411). Method: Dissemination (p. 245). **Ask AI** helps read a chart: AI assistant (p. 260).

Differences at a glance

	RMNCAH	Vaxx
Indicators	Antenatal care, deliveries, postnatal care, vaccination, outpatient and inpatient visits, deaths	Routine vaccines, plus first antenatal visit, institutional deliveries and live births; dropout, zero-dose and under-vaccinated
Reporting Rate services	ANC, Institutional Delivery, Vaccination, OPD Visits	ANC, Institutional Delivery, Vaccination
Internal Consistency	ANC1 and Penta1, Penta1 and Penta3, a custom pair	Also OPV1 and OPV3, IPV1 and IPV2
Data Adjustment groups	Antenatal care, Delivery and postnatal care, Vaccination, Outpatient visits, Inpatient admissions	Antenatal care, Delivery and postnatal care, Vaccination
Denominators	A maternal and a vaccination denominator	A vaccination denominator
Extra pages	Continuum of Care, Mortality, Service Utilization, Health System Performance, Bayesian Analysis	None
Saved dataset	`name_rmncah.rds`	`name_vaccine.rds`

> A dataset saved by one app cannot be opened in the other: the app stops with *This dataset was built for the 'vaccine' indicator group, but this app is for 'rmncah'*. (or the other way round).

Indicator definitions: Indicator definitions (p. 509).

22.1 Load Data

How to open the RMNCAH or Vaxx app on a data file and take it through the Load Data wizard, step by step, to a dataset ready for analysis.

Load Data is the first page of both apps. Its wizard reads your data file, checks it, asks for the national survey values and saves the dataset that every other page uses. RMNCAH Vaxx

Before you start: have the workbook from the Data Extractor (p. 269), or a dataset the app saved earlier. What the files must contain: Analysis setup (p. 44).

Open the app

On the DataSuite start page choose the RMNCAH or Vaxx app (see First steps (p. 21)). DataSuite asks for a data file and opens the app on it. The app reads the file at once: **Load Data** shows a green **Upload successful** banner with the file name.

Accepted files

File	What happens
`.xlsx`, `.xls`	A workbook with the required sheets (Analysis setup > HMIS data (p. 46)). The wizard starts at Upload Data.
`.dta`	A Stata file with the data already merged. The wizard starts at Upload Data.
`.rds`	A dataset saved by the app (p. 503). It was checked when it was first loaded, so the app shows the **Ready for analysis** summary instead of the wizard.

Any other file type stops with *Upload failed: Unsupported file format*.

When you finish the wizard, the app saves a copy of the dataset next to your file: the file name with `_rmncah.rds` added (Vaxx: `_vaccine.rds`). The next time you open the same Excel or Stata file, the app opens that saved copy instead, with all your settings. To load a changed workbook from scratch, delete the saved copy first.

Ready for analysis

When the dataset is already loaded (a saved `.rds`, or after you finish the wizard), Load Data shows a summary instead of the wizard.

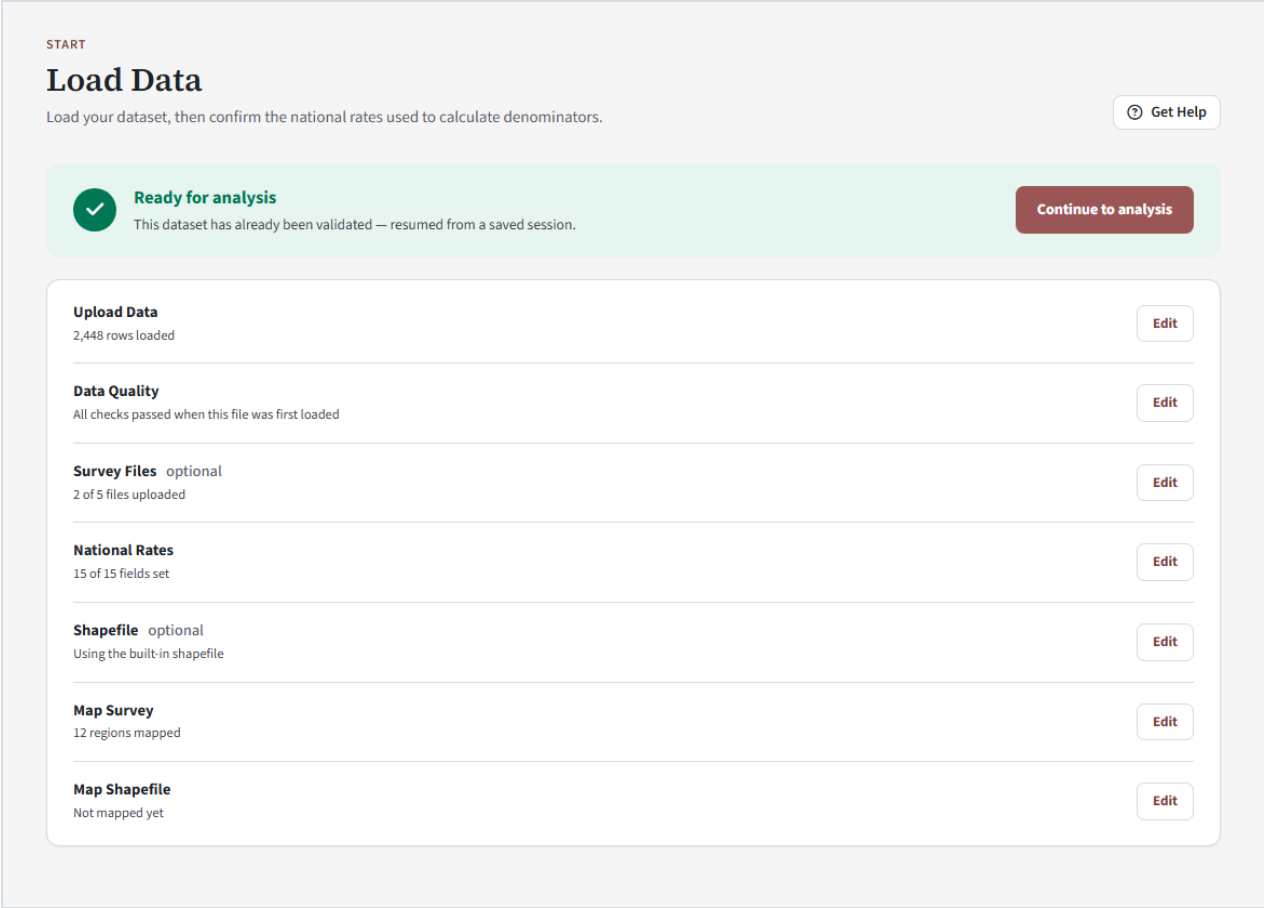
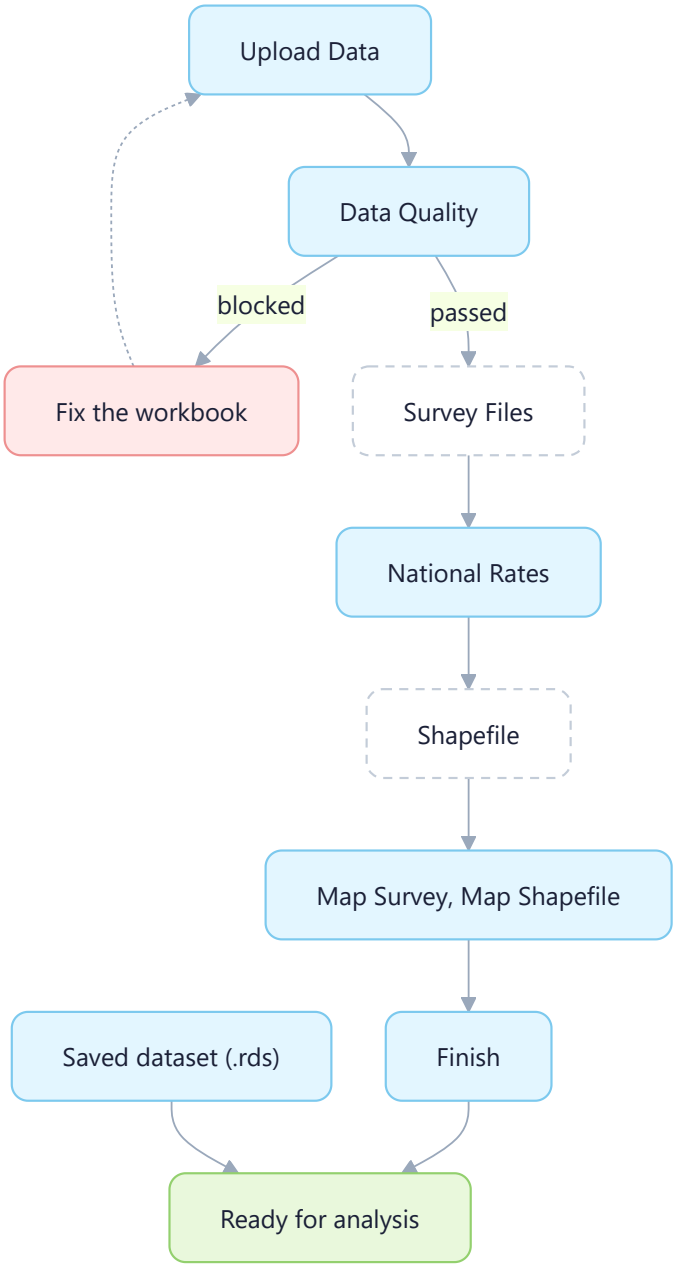


Figure 22.3 — *Load Data showing Ready for analysis, with one row per wizard step and an Edit button on each*

- The banner says whether the dataset was *resumed from a saved session* or has just been saved.
- Each row is one wizard step with what it holds, for example *2,448 rows loaded* or *15 of 15 fields set*.
- **Edit** reopens that step. **← Back to overview** returns to the summary.
- **Continue to analysis** opens Reporting Rate (p. 309).

The Load Data wizard

A new Excel or Stata file goes through seven steps. The step rail at the top shows them: a tick marks a finished step, a padlock a step you cannot open yet. The button at the bottom right moves on: **Continue to …** after a step you completed, **Skip to …** past an optional step you left as it is. **← Back** returns to the step before.


Upload Data
Data Quality
blocked
passed
Fix the workbook
Survey Files
National Rates
Shapefile
Map Survey, Map Shapefile
Saved dataset (.rds)
Finish
Ready for analysis

① Upload Data

The **Dataset** card shows the file that was read. **Download Master Dataset** saves the merged data as a Stata file (`master_dataset.dta`).

Additional reference data (optional) replaces the app's built-in reference data for the country with your own `.dta` files: **Upload UN Estimates data**, **Upload WUENIC estimates** and **Upload UN Mortality estimates**. Leave them empty to use the built-in UN estimates (p. 504), WUENIC (p. 505) and UN mortality estimates (p. 504).

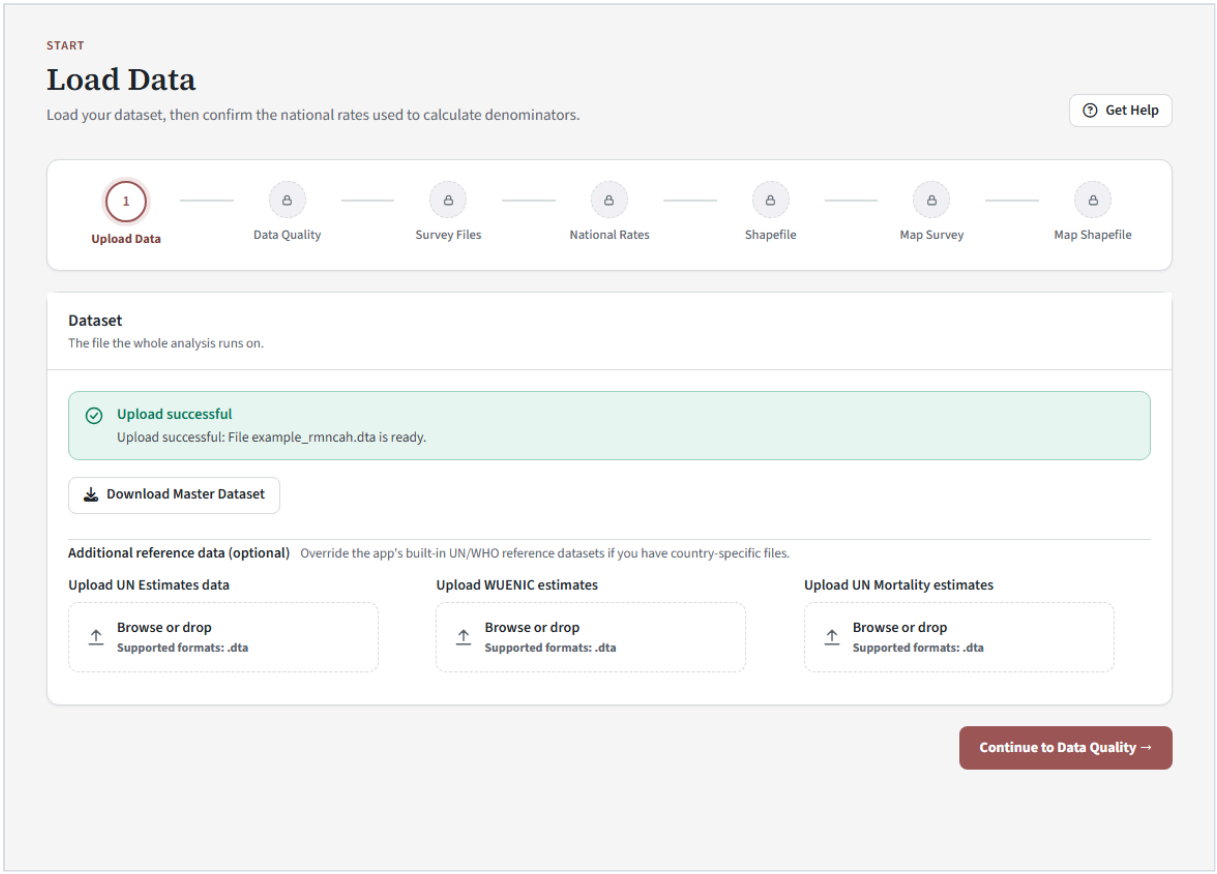


Figure 22.4 — *The Upload Data step: the Dataset card with Upload successful and the three reference data boxes*

Press **Continue to Data Quality →**.

② Data Quality

The checks (p. 498) run by themselves when the file is read. A banner gives the result: *All checks passed — no blocking issues found*, or *Some checks found issues that need fixing before you can continue*.

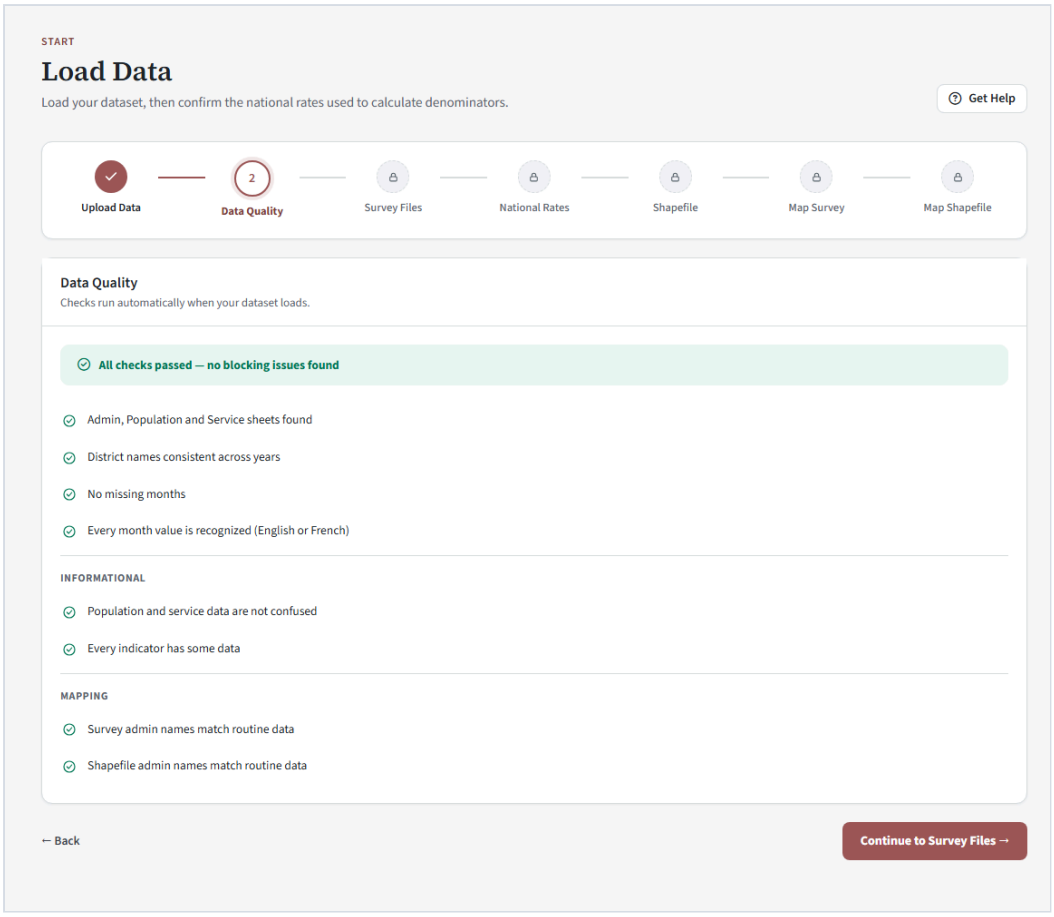


Figure 22.5 — *The Data Quality step: every check passed*

The list has three parts:

- The first rows are blocking checks (p. 497), for example *Admin, Population and Service sheets found* or *No missing months*. A failed one shows a warning sign and says what is wrong. You cannot continue until every blocking check passes.
- **Informational** checks only warn, for example *Every indicator has some data*.
- **Mapping** checks say whether the region names of your survey files and map match your data. You fix them in the mapping steps.

To fix a blocking check, correct the workbook and load it again. What each check tests: Analysis setup > Checks on the uploaded data (p. 48).

③ Survey Files (optional)

The **Survey data folder** card has one box per survey file (`.dta`): **National data**, **Regional data**, **Area data**, **Maternal education data** and **Wealth index quintile data**. Upload the ones you have; leave the others empty.

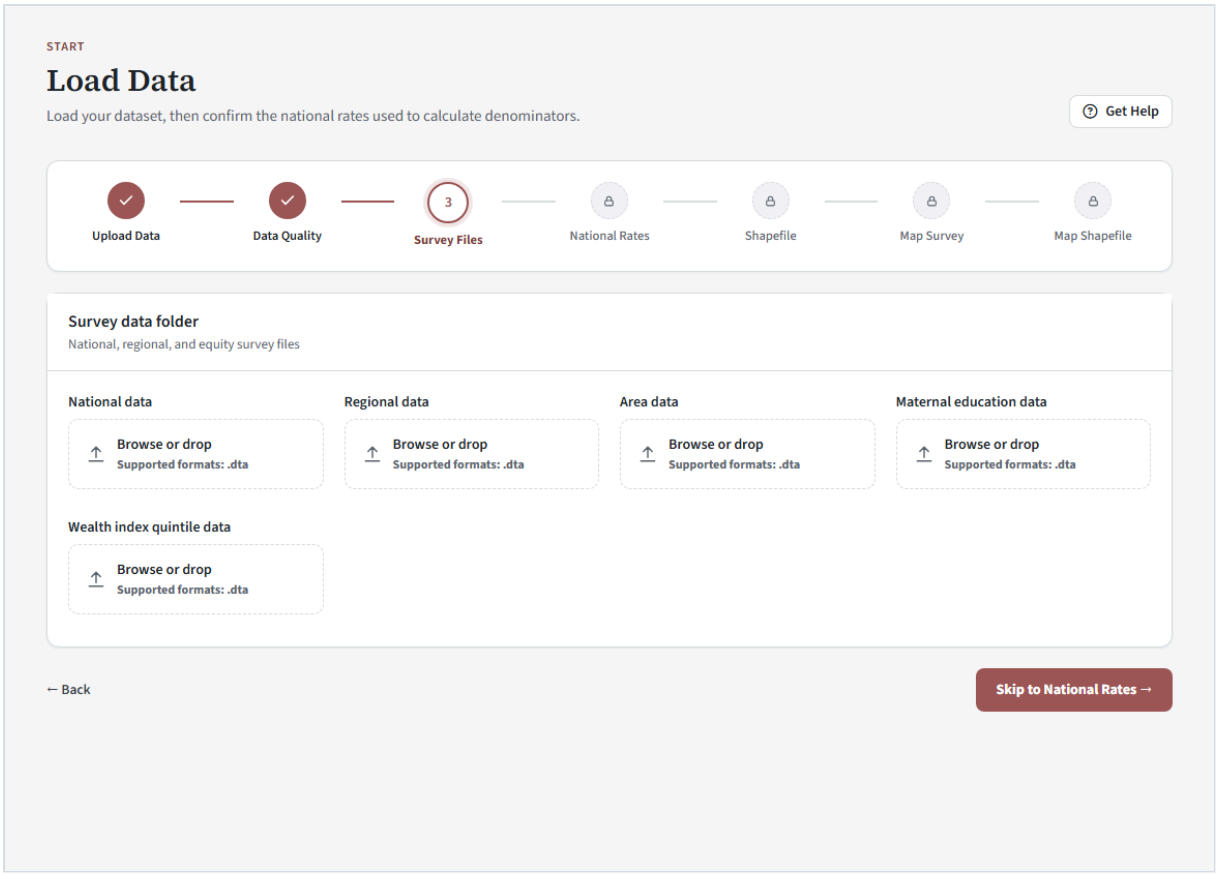


Figure 22.6 — *The Survey Files step with the five upload boxes*

What each file is used for, and what happens without it: Analysis setup > Survey files and shapefile (p. 50).

④ National Rates

The **National rates for denominator calculation** card holds the survey values the app needs to build the denominators. Every field is required. An empty field says *Not set*. *Needed for denominators*., and the bottom of the page shows *Required fields remaining* next to a disabled **Continue** button.

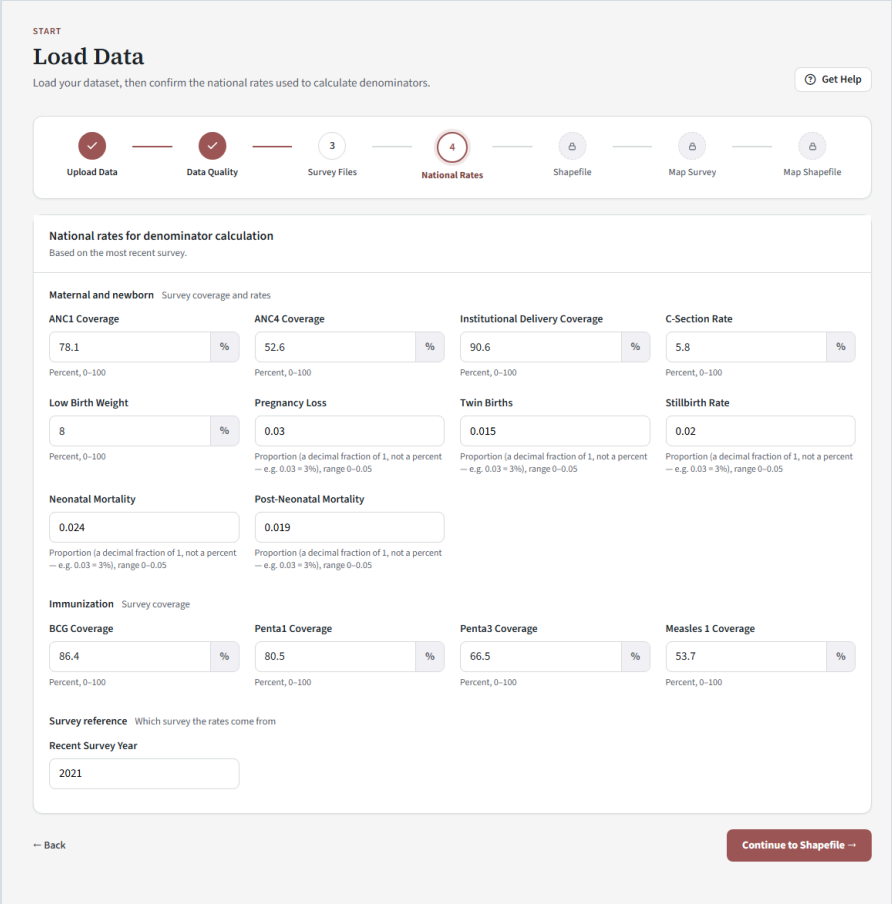


Figure 22.7 — *The National Rates step with every field filled in*

- **Maternal and newborn:** ANC1 Coverage, ANC4 Coverage, Institutional Delivery Coverage, C-Section Rate and Low Birth Weight (percent, 0–100), then Pregnancy Loss, Twin Births, Stillbirth Rate, Neonatal Mortality and Post-Neonatal Mortality (proportions: 0.03 means 3%).
- **Immunization:** BCG, Penta1, Penta3 and Measles 1 Coverage (percent).
- **Survey reference: Recent Survey Year**, and **Survey Data Start Year** once a national survey file is loaded.

If you uploaded a national survey file, the values are filled in for you and a box says *Pre-filled from your survey*. Check them all the same. What each value means and where to find it: Analysis setup > National survey parameters (p. 55).

⑤ Shapefile (optional)

The app draws its maps with a built-in shapefile (p. 503) for the country, and the step says *Using the built-in shapefile*. To use your own map, drop the `.shp`, `.dbf`, `.shx` and `.prj` files in **Upload a custom shapefile**, choose the column that holds the region names (*Which column holds the region (admin-1) names?*) and press **Use this shapefile**.

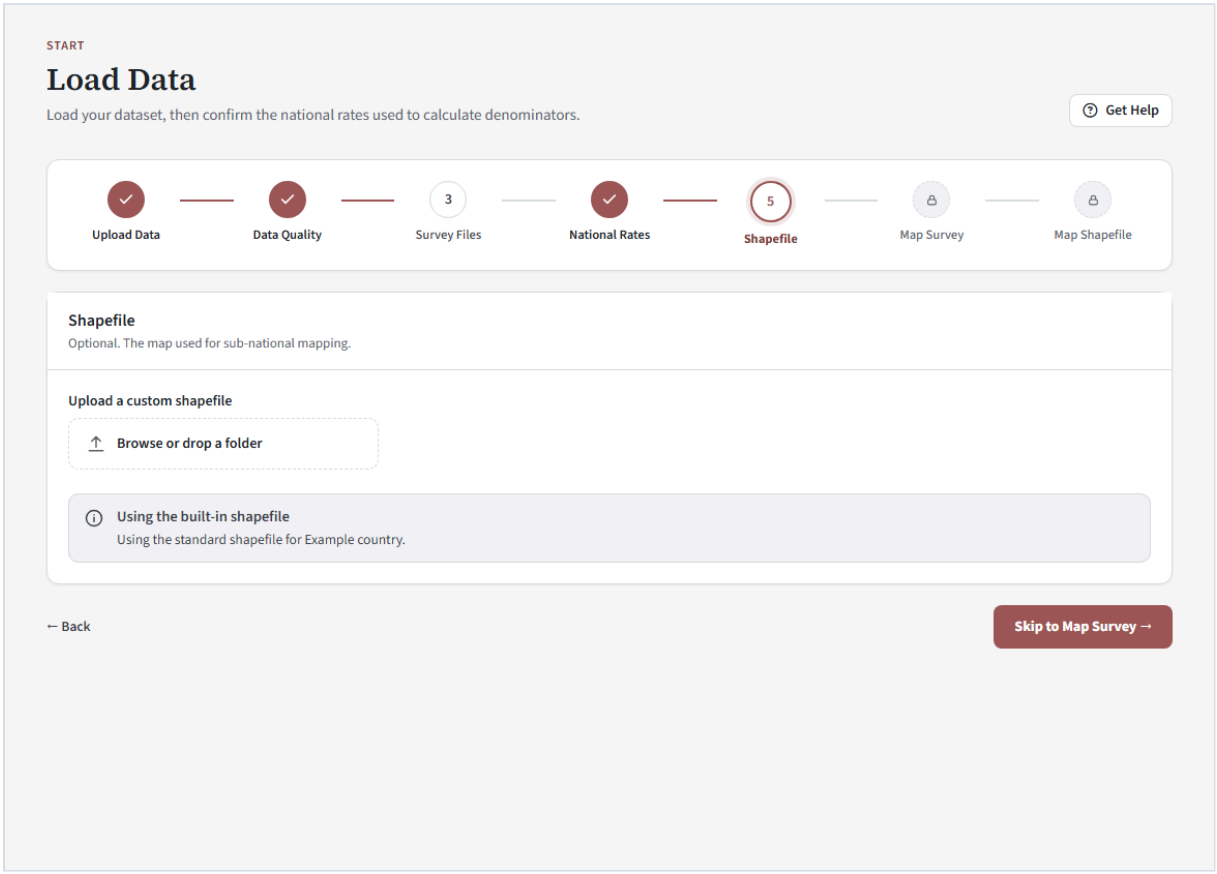


Figure 22.8 — *The Shapefile step using the built-in shapefile*

⑥ Map Survey and Map Shapefile

These two steps match the region names of the survey files (**Map Survey**) and of the map (**Map Shapefile**) to the Admin 1 (p. 496) names in your data. Press **Match Region Names**. The app matches the names it can and tags each match with the reason (*Case differs*, *Accent differs*, *Approximate spelling match*). Choose a name for any region still on *Select…*, then press **Save Mapping**.

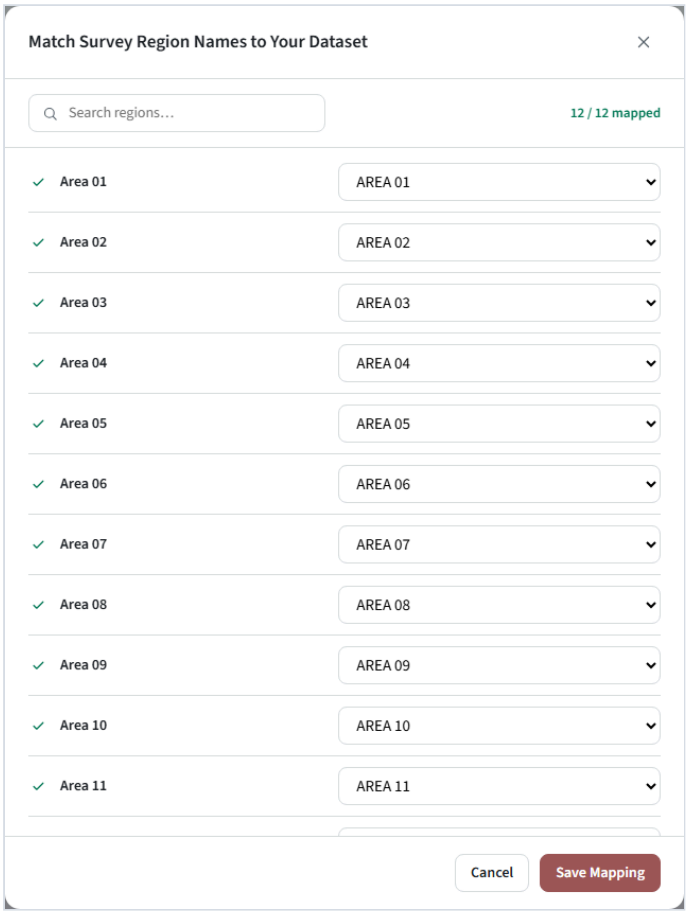


Figure 22.9 — *The dialog that matches survey region names: twelve regions matched automatically*

Save once even if nothing needs correcting: until you do, the step shows *Review needed* and you cannot continue. After saving, a box says how the names were matched. **Clear mapping** starts again.

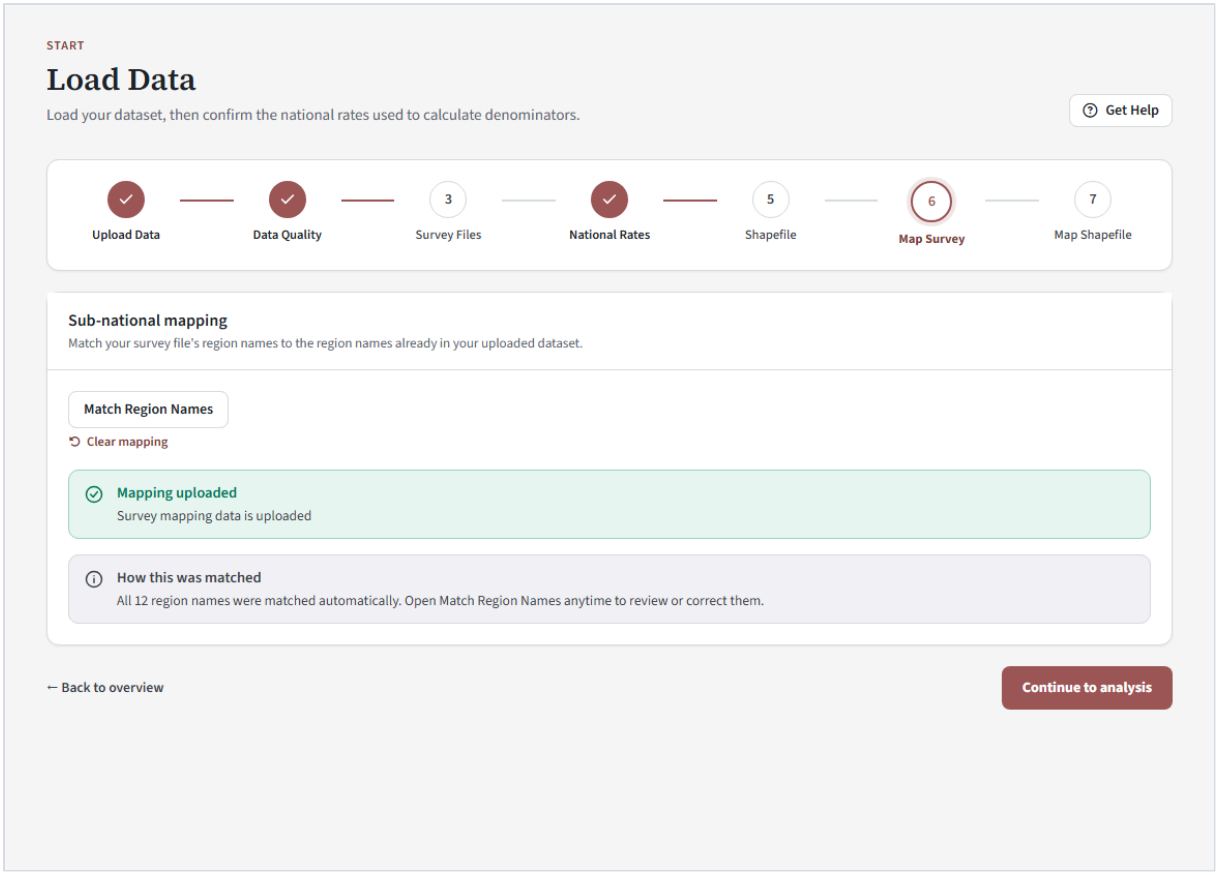


Figure 22.10 — *The Map Survey step after saving: Mapping uploaded, and how the names were matched*

⑦ Finish Loading Data

On the last step the button reads **Finish Loading Data**. The app merges the data, saves the dataset (see Accepted files) and shows Ready for analysis with *Your dataset has been saved and is ready for analysis*. Press **Continue to analysis**.

> **Vaxx app**
> The steps are the same. The mapping button reads **Map Survey Data**.

Method: Analysis setup (p. 44).

Troubleshooting

Message	When you see it	What to do
Upload failed: Unsupported file format.	The file is not `.xls`, `.xlsx`, `.dta` or `.rds`.	Export the data again as an Excel workbook.
Upload failed: The file provided does not exist	The file was moved or deleted after you picked it.	Open the app again and pick the file where it is now.
Upload failed: missing sheet(s) and the sheet names	The workbook lacks a required sheet.	Add the sheets listed, with their exact names (HMIS data (p. 46)).
Upload failed: Key Columns … missing in …	A sheet lacks a column the app joins on, such as the district, year or month.	Add the column to that sheet and load the workbook again.

Upload failed: Sheet … is empty.	A required sheet has no rows.	Fill the sheet, or export the workbook again.
Upload failed: This dataset was built for the 'vaccine' indicator group, but this app is for 'rmncah'. (or the other way round)	You opened a dataset saved by the other app.	Open it in the app that saved it, or load the original workbook.
Some checks found issues that need fixing before you can continue	A blocking check failed on the Data Quality step. Its row says what is wrong, for example *The country name "…" in Admin_data is not recognized.* or *District(s) present in the data but missing from Admin_data: …*	Correct the workbook and load it again.
Complete the earlier steps first	You clicked a locked step in the rail.	Finish, or skip, the current step first.
Required fields remaining	A National Rates field is still empty, or a mapping is not saved.	Fill in every field, or save the mapping.
Upload failed: … under a survey file or reference data box	That `.dta` file could not be read.	Check that it is the right file for that box.
Some survey region names don't match your dataset (or *Some shapefile region names don't match your dataset*)	Region names in the survey files or the map differ from your data.	Match them in Map Survey or Map Shapefile.
Review needed	The region matches have not been saved yet.	Press **Match Region Names**, then **Save Mapping**.
Load and finish a dataset first to unlock this page.	You opened another page before finishing Load Data.	Finish the wizard.

The detail after *Upload failed:* and in a failed check comes from the file check itself and is in English.

Common pitfalls

- Typing percentages as proportions, or the other way round, in National Rates: coverage fields take 0–100, rates take proportions such as 0.03.
- Editing the workbook and opening it again: the app opens the saved copy, not your changes. Delete the `_rmncah.rds` (or `_vaccine.rds`) copy first.
- Not saving the mapping: Map Survey and Map Shapefile need one save even when every name already matches.

Pitfalls in preparing the data: Analysis setup > Common pitfalls (p. 45).

Next: Data quality (p. 306).

CHAPTER 23

Data Quality Assessment

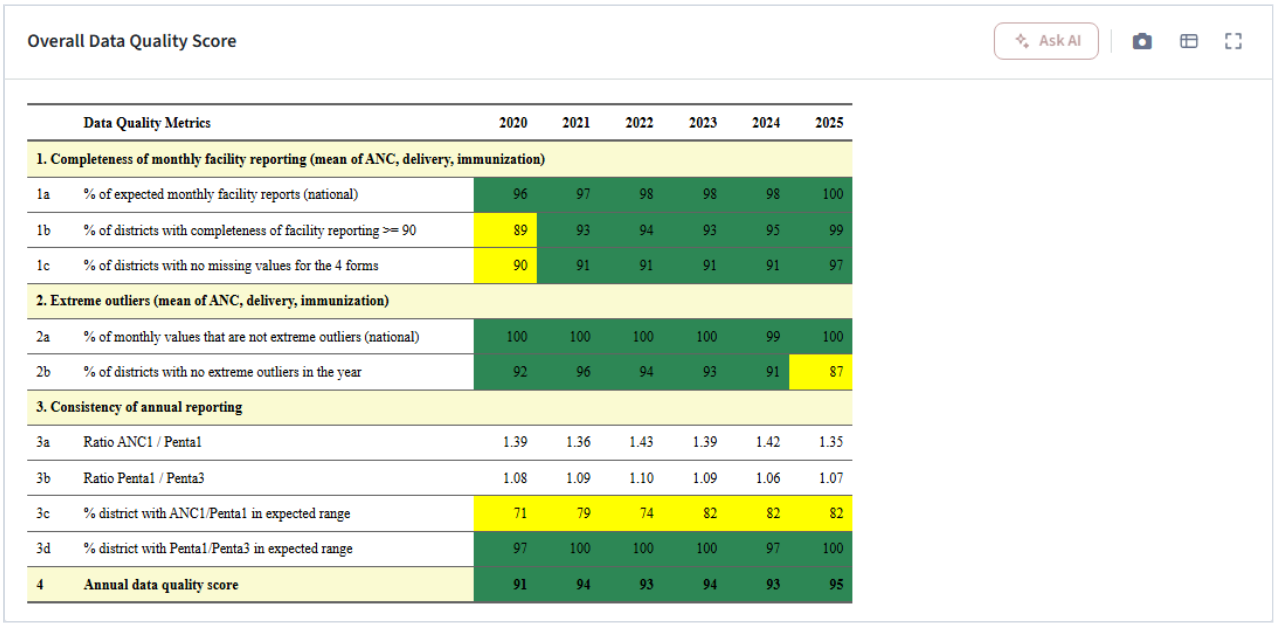
The five Data Quality Assessment pages (Reporting Rate, Outlier Detection, Data Missingness, Internal Consistency, Overall Data Quality Score) that show whether the facility data can be trusted.

Can these numbers be trusted?

Before any coverage is calculated, five pages show how complete, plausible and consistent the facility data are, and where they are not. What you find guides the choices on Data Adjustment (p. 322).

RMNCAH Vaxx

Learn about the data quality assessment *(p. 57)*

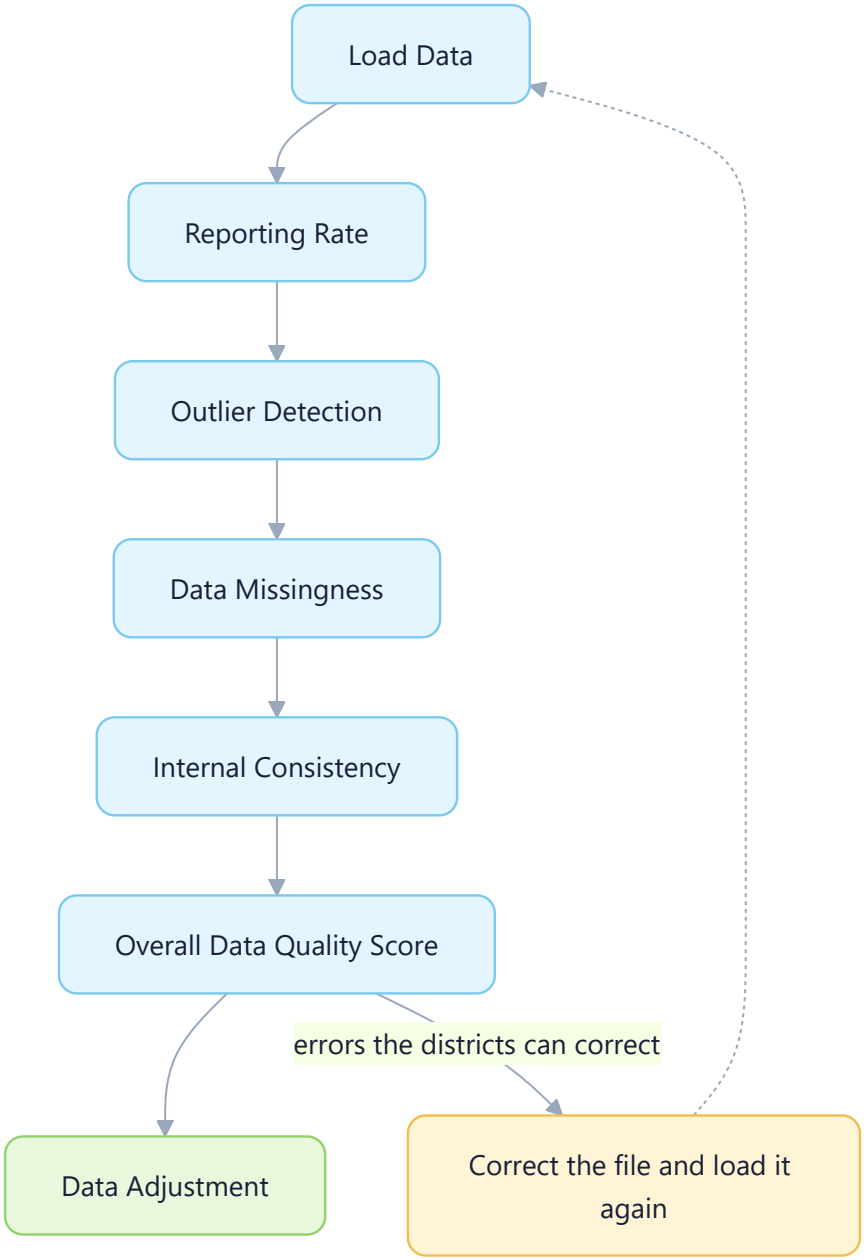


	Data Quality Metrics	2020	2021	2022	2023	2024	2025
1. Completeness of monthly facility reporting (mean of ANC, delivery, immunization)							
1a	% of expected monthly facility reports (national)	96	97	98	98	98	100
1b	% of districts with completeness of facility reporting >= 90	89	93	94	93	95	99
1c	% of districts with no missing values for the 4 forms	90	91	91	91	91	97
2. Extreme outliers (mean of ANC, delivery, immunization)							
2a	% of monthly values that are not extreme outliers (national)	100	100	100	100	99	100
2b	% of districts with no extreme outliers in the year	92	96	94	93	91	87
3. Consistency of annual reporting							
3a	Ratio ANC1 / Penta1	1.39	1.36	1.43	1.39	1.42	1.35
3b	Ratio Penta1 / Penta3	1.08	1.09	1.10	1.09	1.06	1.07
3c	% district with ANC1/Penta1 in expected range	71	79	74	82	82	82
3d	% district with Penta1/Penta3 in expected range	97	100	100	100	97	100
4	**Annual data quality score**	**91**	**94**	**93**	**94**	**93**	**95**

Figure 23.1 — *The Overall Data Quality Score card: one row per check, one column per year*

Pages

Before you start: finish Load Data (p. 295). Review the pages in the sidebar order and come back to them whenever you load a corrected file.



Most pages have a filter bar at the top, **All charts on this page**: it changes every card on the page. A control marked *This table only* changes one table. Every card can be downloaded or explained with **Ask AI** (Exports and Ask AI (p. 411)).

Common pitfalls

- Changing the Performance Threshold to make the data look better: it is stored in the dataset and changes the score and the adjustment too. Set it once.
- Reading the charts with the wrong filter: the filter bar changes every card, and a region chosen in the **Admin 1** chip stays chosen.
- Expected ratios that look wrong: check the survey values in National Rates (p. 301).

- Moving on without asking the districts: errors they can correct at the source are better fixed in the file than by the adjustment.

Pitfalls in reading the results: Data quality assessment > Common pitfalls (p. 58).

Next: Data Adjustment (p. 322).

23.1 Reporting Rate

Answers how completely the facilities sent their monthly reports, by service, area and year.

The **Reporting Rate** page answers the first question about the facility data: did the facilities send their monthly reports? It is the first Data Quality Assessment page to open after loading the data. Where reporting is low, the reported numbers count fewer services than were given, so what you see here guides the k-factors you set later on Data Adjustment (p. 324). RMNCAH Vaxx

Before you start: finish Load Data (p. 295).

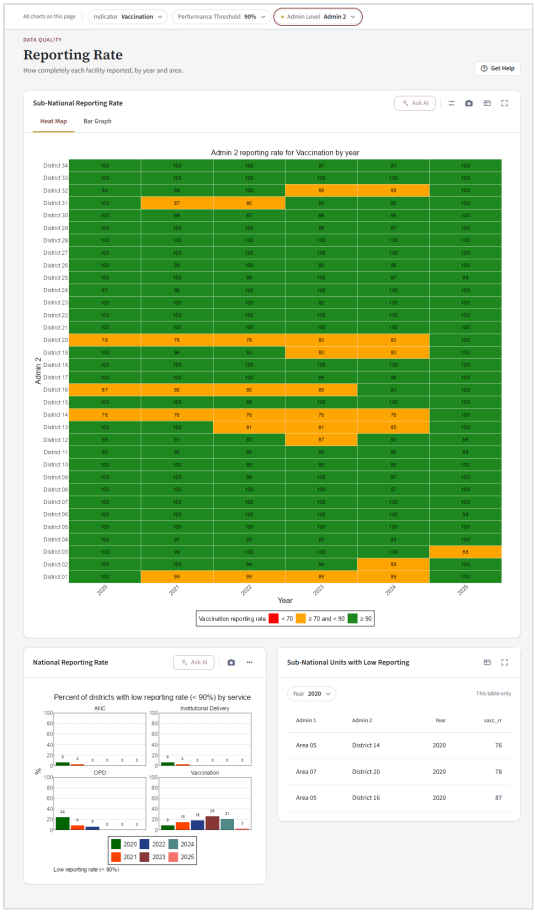


Figure 23.2 — *The Reporting Rate page for vaccination by district: the heat map, the national chart and the districts with low reporting*

Using the page

1. **Choose the service and the level**
   Pick the **Indicator**, the service whose reports are counted, and the **Admin Level**. Leave the **Performance Threshold** at its value unless your team has agreed on another one.

2. **Look for weak areas and years**
   On the **Heat Map**, red and orange cells are the areas and years where reporting fell below the threshold. The **National Reporting Rate** chart shows whether this is getting better or worse over time.

③ List the areas to follow up

In **Sub-National Units with Low Reporting**, choose a **Year** to list the areas below the threshold. These are the areas to discuss with the teams that report.

④ Save or explain a chart, then move on

The card toolbar downloads a chart or its data, and **Ask AI** explains a card (Exports and Ask AI (p. 411)). Repeat for each **Indicator**, then open Outlier Detection (p. 312).

Controls

Control	Choices	Default and notes
Indicator	ANC, Institutional Delivery, Vaccination, OPD Visits	ANC. The service whose reports are counted.
Performance Threshold	A percentage	90%. The level a district must reach (p. 498). Stored in the dataset and used again by the Overall Data Quality Score and Data Adjustment.
Admin Level	Admin 1, Admin 2	Admin 1. With Admin 1, an **Admin 1** chip (**Select All** or one region) shows the districts of one region.
Year (*This table only*)	The years in the data	The first year.

Cards

Sub-National Reporting Rate — the reporting rate (p. 503) of each area by year. **Heat Map**: one cell per area and year, red below 70, orange from 70 to the threshold, green above. **Bar Graph**: the same rates as bars. A mostly green map means the facilities reported well. A row of red or orange cells points to an area with a reporting problem; a column of them, to a year with one.

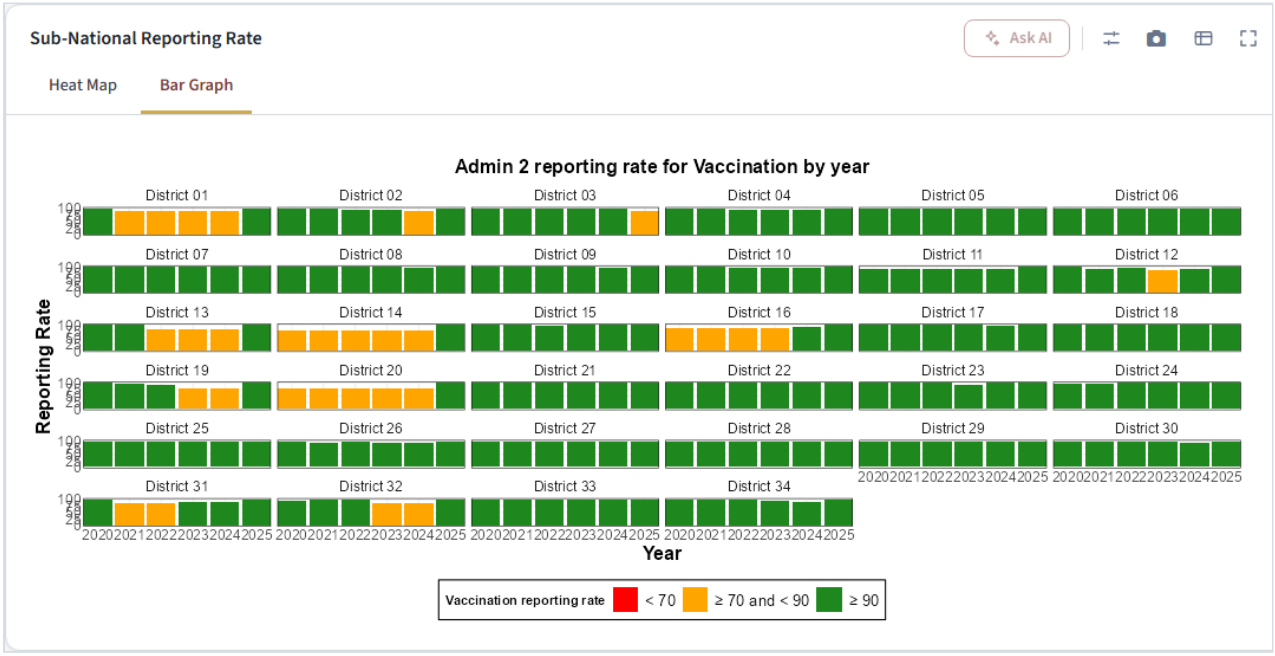


Figure 23.3 — *The Bar Graph tab: the reporting rate of each area by year*

National Reporting Rate — the percentage of districts below the threshold, one panel per service, by year. With one region chosen it becomes **District Reporting Rate in** that region. Here low bars are good: the closer to zero, the fewer districts fall short. Bars that grow from year to year mean reporting is getting worse.

Sub-National Units with Low Reporting — the areas below the threshold in the chosen year, with their rate. The shorter the list, the better; an area that appears year after year needs follow-up.

> Vaxx app
> The **Indicator** choices are ANC, Institutional Delivery and Vaccination.

Common pitfalls

- Checking one service only: each service has its own reporting form, so ANC can be well reported while delivery is not. Look at every **Indicator**.
- Lowering the **Performance Threshold** so that fewer districts look poor: it is saved in the dataset and changes the Overall Data Quality Score and the adjustment too.
- Reading the table as if it covered every year: its **Year** control changes the table only, and the charts keep showing all years.

Method: Completeness of monthly facility reporting (p. 60).

23.2 Outlier Detection

Answers which monthly values are so far from the usual ones that they may be data entry errors.

The **Outlier Detection** page finds the monthly values that are so far from a district's usual values that they may be data entry errors, such as an extra zero. Open it after Reporting Rate (p. 309). The flagged months are the ones the outlier correction on Data Adjustment (p. 324) changes, so use this page to judge whether they are errors or real changes. RMNCAH Vaxx

Before you start: finish Load Data (p. 295).

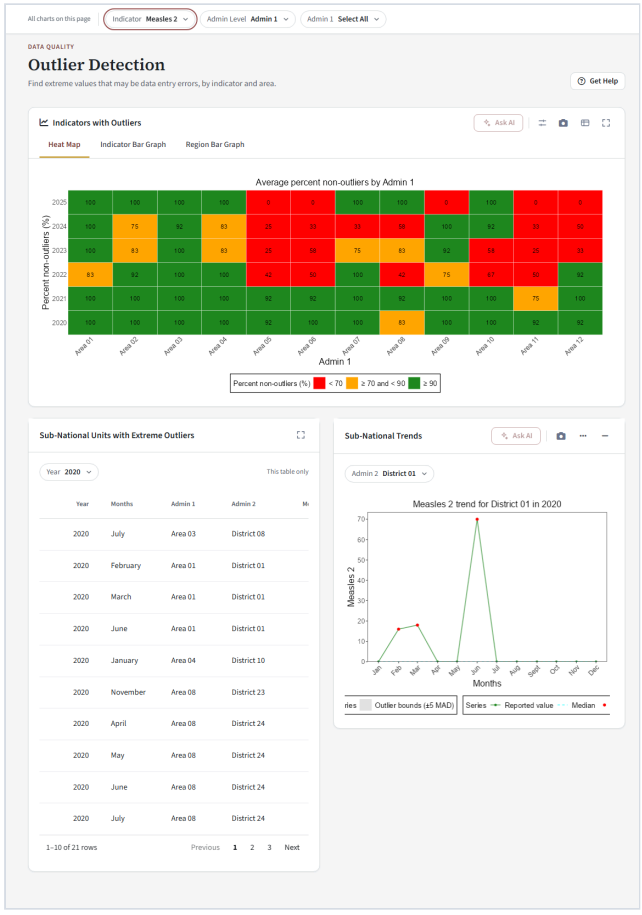


Figure 23.4 — *The Outlier Detection page for measles 2: the heat map, the flagged months and one district's trend*

Using the page

1. **Choose the indicator**

   Pick the **Indicator** and the **Admin Level**. The page starts on the first indicator in the list, so change it to the one you want to check.

2. **Find the areas with outliers**

   On the **Heat Map**, low percentages mark the areas and years with many extreme values. The two bar graph tabs compare indicators and regions.

③ Check the flagged months

In **Sub-National Units with Extreme Outliers**, choose a **Year** and compare each reported value with the district's median.

④ Look at one district's trend

In **Sub-National Trends**, pick a district in the **Admin 2** chip to see whether a flagged month is a single spike or part of a lasting change. Then open Data Missingness (p. 315).

Controls

Control	Choices	Default and notes
Indicator	Every indicator in the data	The first in the list.
Admin Level	Admin 1, Admin 2	Admin 1, with the **Admin 1** region chip.
Year (*This table only*)	The years in the data	The first year.

Cards

Indicators with Outliers — Heat Map: the percentage of monthly values that are not extreme outliers (p. 499), by area and year. **Indicator Bar Graph**: the same by indicator and year. **Region Bar Graph**: the same by region and year. Values close to 100% are good. The lower the value, the more months were flagged.



Figure 23.5 — *The Indicator Bar Graph tab: monthly values without outliers, by indicator and year*

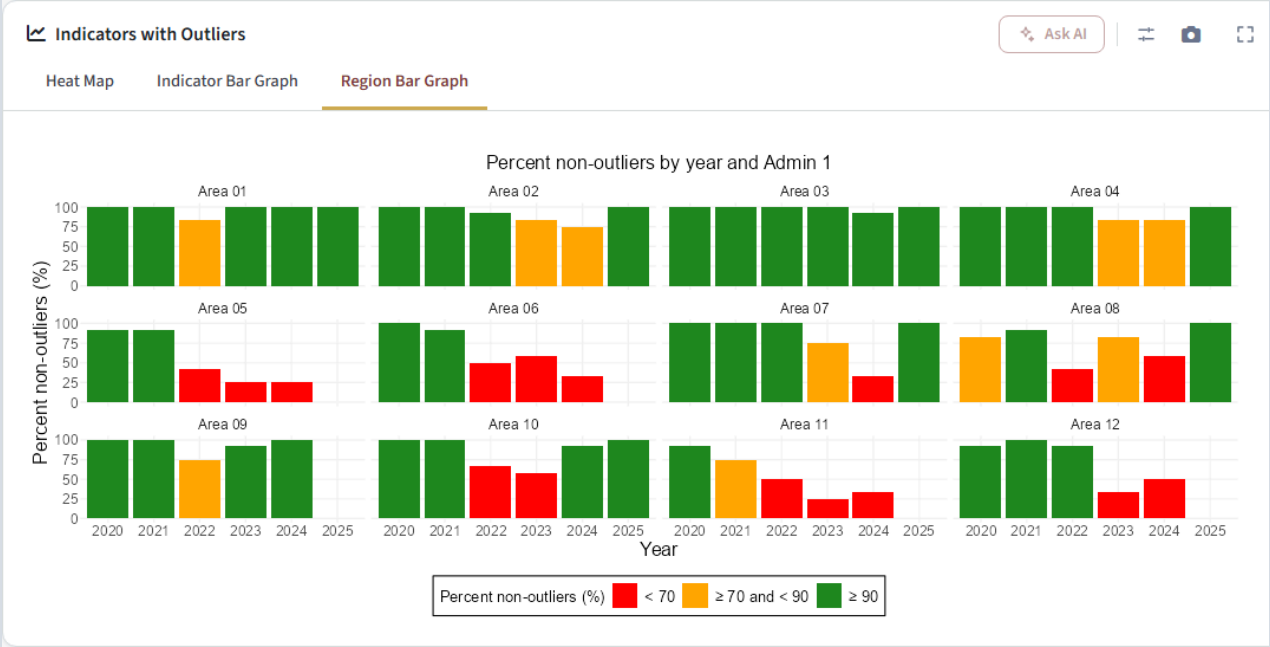


Figure 23.6 — *The Region Bar Graph tab: monthly values without outliers, by region and year*

Sub-National Units with Extreme Outliers — every flagged month of the chosen year: area, district, month, the reported value, the district's median (p. 501) and its MAD (p. 501). A value many times the median often means a typing error; a value just outside the normal range may be real.

Sub-National Trends — pick a district (**Admin 2** chip) to see its months for the year: the reported values, the median and the band of normal values (*Outlier bounds (±5 MAD)*). Points outside the band are the outliers. One point far outside points to an error; several months at a new level point to a real change, such as a campaign or a new facility.

Common pitfalls

- Treating every flagged month as an error: a vaccination campaign or a stock-out can cause a real peak or dip. Check with the district before you rely on the correction.
- Reading the page with the first indicator still selected: the table and the trend show one indicator only.
- Reading few outliers as good data: a district with many empty months has few values to flag. Read this page together with Data Missingness (p. 315).

Method: Extreme outliers (p. 65).

23.3 Data Missingness

Answers which indicators and areas have months with no value.

The **Data Missingness** page shows which indicators and areas have months with no value. Open it after Outlier Detection (p. 312). A missing month makes the yearly total too low, so the list of empty months tells you what to ask the districts about and what the missing-value correction on Data Adjustment (p. 324) will fill in. RMNCAH Vaxx

Before you start: finish Load Data (p. 295).

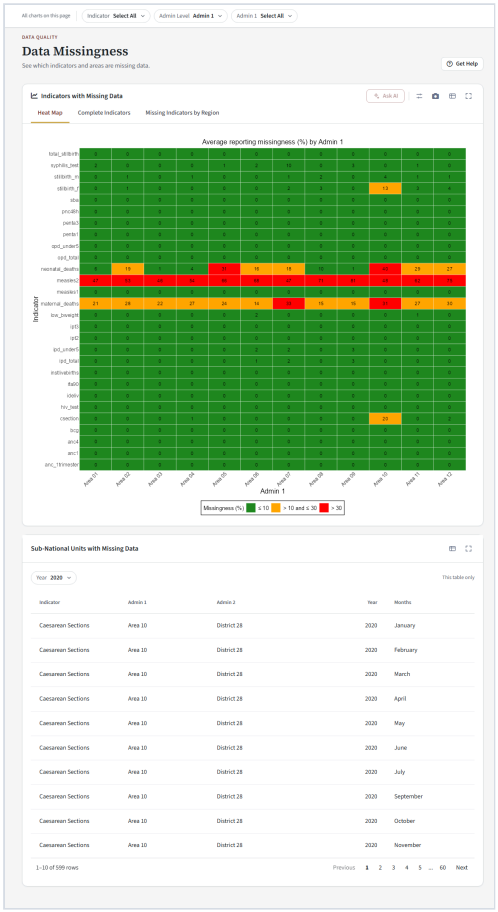


Figure 23.7 — *The Data Missingness page: the heat map and the list of district-months with no value*

Using the page

1. **Start with every indicator**
   Leave **Indicator** on **Select All** to see all indicators at once, and choose the **Admin Level**.

2. **Find the gaps**
   On the **Heat Map**, look for the areas and years with a high share of missing values. The **Complete Indicators** tab shows which indicators have gaps.

3. **Look at one indicator**
   Pick an **Indicator** to see its own pattern; an average over all indicators can hide one with many gaps.

④ **List the empty months**

In **Sub-National Units with Missing Data**, choose a **Year** to list each district and month with no value. Then open Internal Consistency (p. 317).

Controls

Control	Choices	Default and notes
Indicator	**Select All** or one indicator	Select All.
Admin Level	Admin 1, Admin 2	Admin 1, with the **Admin 1** region chip.
Year (*This table only*)	The years in the data	The first year.

Cards

Indicators with Missing Data — Heat Map: the share of missing values (p. 502) by area and year; shares close to zero are good. **Complete Indicators**: the percentage of districts with no missing month, by indicator and year; close to 100% is good. **Missing Indicators by Region**: the same by region, over time; a line that drops shows a region whose reporting is getting worse.

Sub-National Units with Missing Data — every indicator, district and month with no value in the chosen year. Each row is a month to check with the district: was a report sent but not entered, or was the service not offered?

Common pitfalls

- Stopping at **Select All**: the average over all indicators can look good while one indicator has many gaps.
- Expecting the table to show every year: it shows the year chosen in its **Year** control.
- Relying on the adjustment to fill every gap: a filled month is an estimate. Where the districts can send the real numbers, correct the file and load it again.

Method: Missing monthly values (p. 62).

23.4 Internal Consistency

Answers whether related indicators agree: more first antenatal visits than first pentavalent doses, and more first doses than third.

The **Internal Consistency** page checks whether related indicators agree with each other. Most women who start antenatal care have a baby who later gets the first pentavalent dose, and most children who get the first dose get the third, so these counts should keep a steady ratio. Open the page after Data Missingness (p. 315). A ratio out of range, or a district far from the others, points to over- or under-reporting of one of the two indicators. RMNCAH Vaxx

Before you start: finish Load Data (p. 295), including the survey values in National Rates (p. 301), which give the expected ratios.

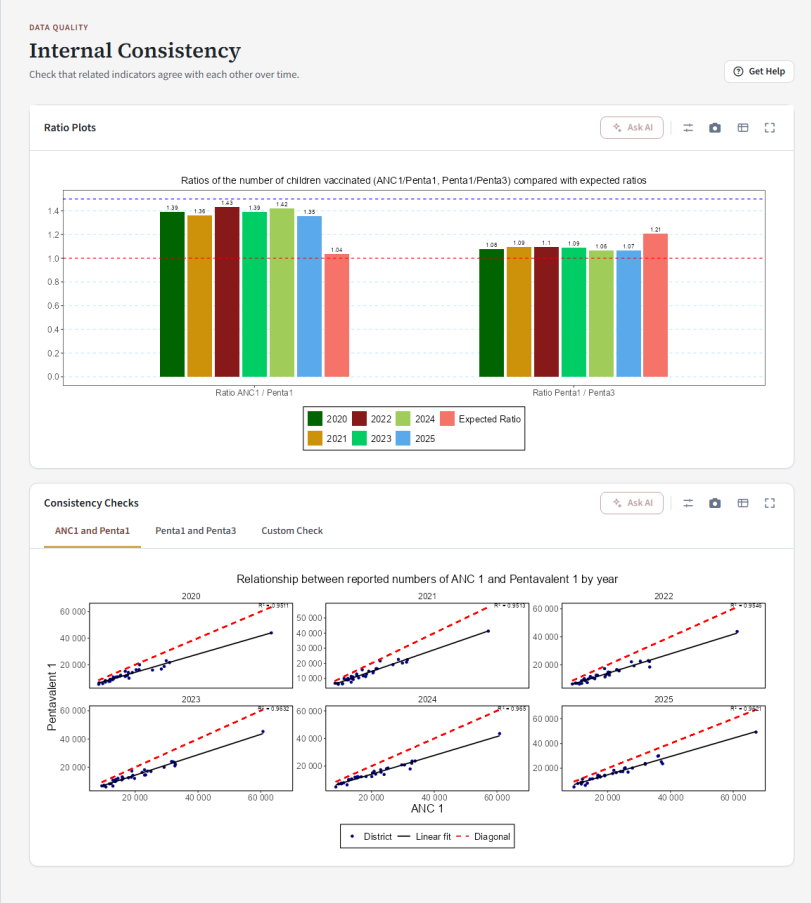


Figure 23.8 — *The Internal Consistency page: the yearly ratios against the expected ratio, and ANC1 against penta1 by district*

Using the page

This page has no filter bar.

① Check the national ratios

In **Ratio Plots**, see whether each year's ratio falls between the dashed lines around the expected ratio.

② **Find the districts that disagree**

In **Consistency Checks**, open each tab and look for dots far from the line of the other districts.

③ **Compare other indicators**

The **Custom Check** tab lets you choose the indicator on each axis, for any pair you expect to move together.

④ **Move on**

Download a chart or ask about it from the card toolbar (Exports and Ask AI (p. 411)), then open Overall Data Quality Score (p. 320).

Cards

Ratio Plots — the national ratios ANC1 / Penta1 and Penta1 / Penta3 for each year, next to the expected ratio (p. 499), with dashed lines for the acceptable range. The expected ratios come from the survey values in National Rates (p. 301). A ratio inside the dashed lines is consistent. Outside them, one of the two counts is too high or too low that year.

Consistency Checks — one chart per year with one dot per district, a linear fit (p. 500) (black line) and the diagonal (red dashed line). Tabs: **ANC1 and Penta1**, **Penta1 and Penta3**, and **Custom Check**, where you choose the two indicators. Dots close to the black line show districts that report both indicators in the same way. A dot far from it is a district whose numbers disagree and are worth checking.

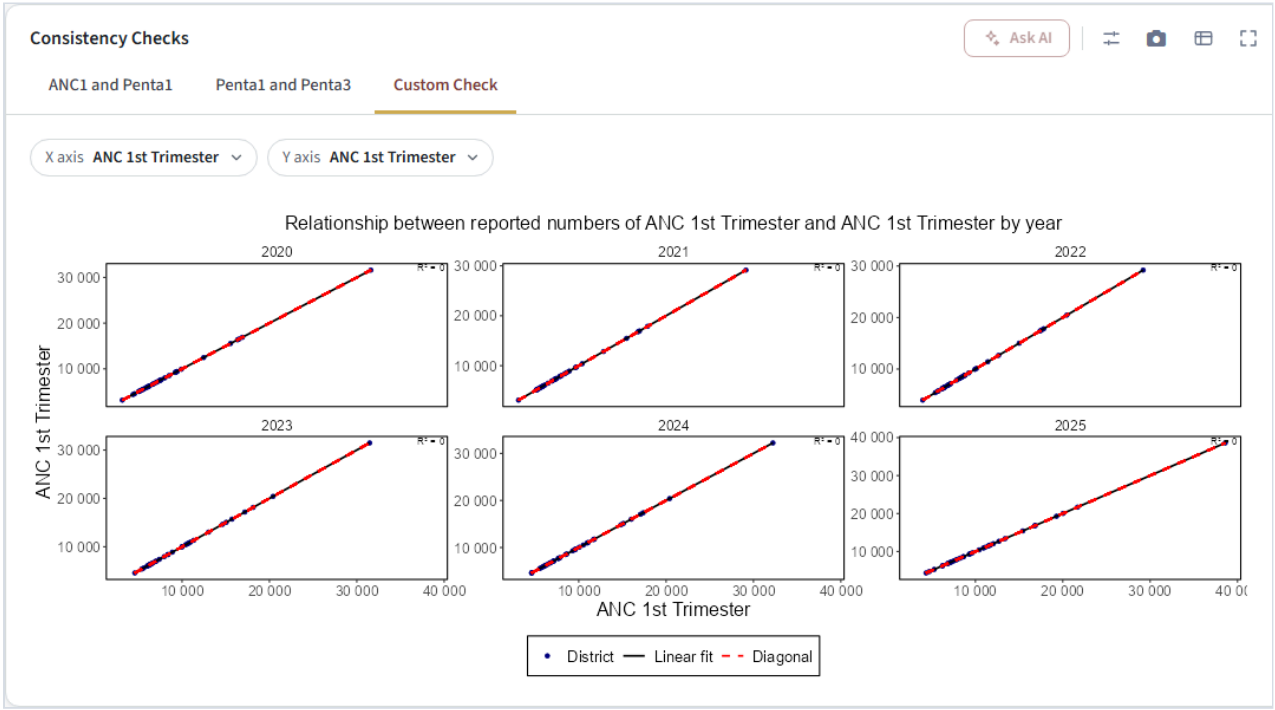


Figure 23.9 — *The Custom Check tab, where you choose the indicators on each axis*

> **Vaxx app**
> Consistency Checks also has **OPV1 and OPV3** and **IPV1 and IPV2** tabs.

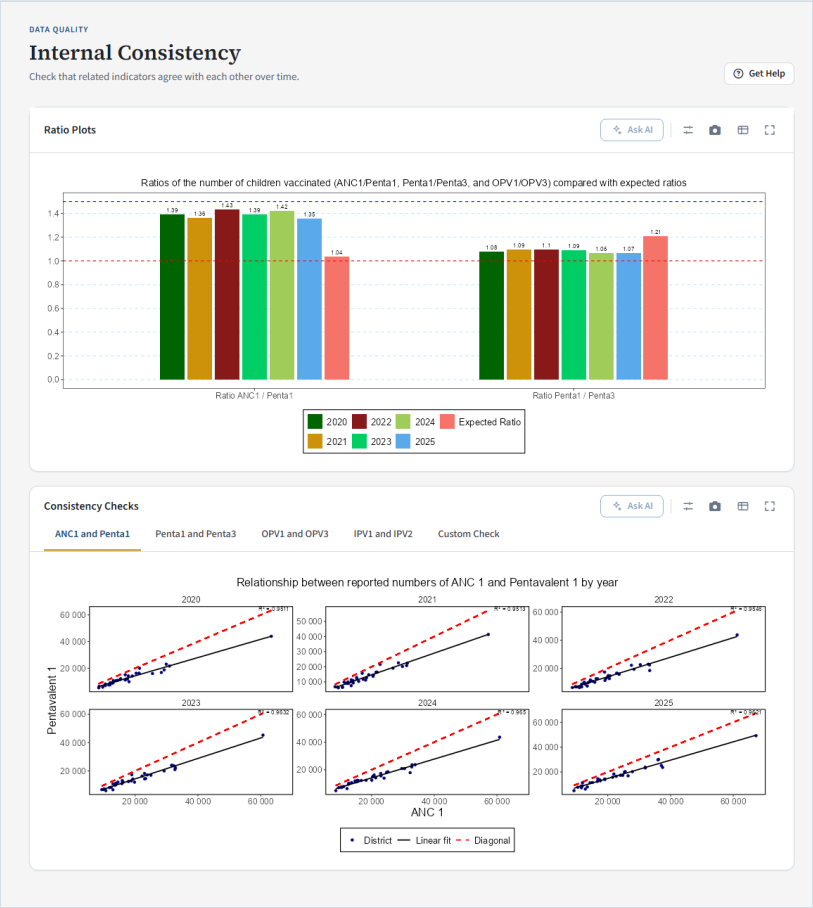


Figure 23.10 — *Vaxx: the Internal Consistency page*

Common pitfalls

- An expected ratio that looks wrong: it comes from the survey values entered in National Rates (p. 301). Correct them there.
- Reading one year only: a ratio out of range in a single year is often a reporting problem that year; one out of range every year may mean the expected ratio does not fit.
- Choosing two unrelated indicators in **Custom Check**: the dots then scatter, and the chart says nothing about data quality.

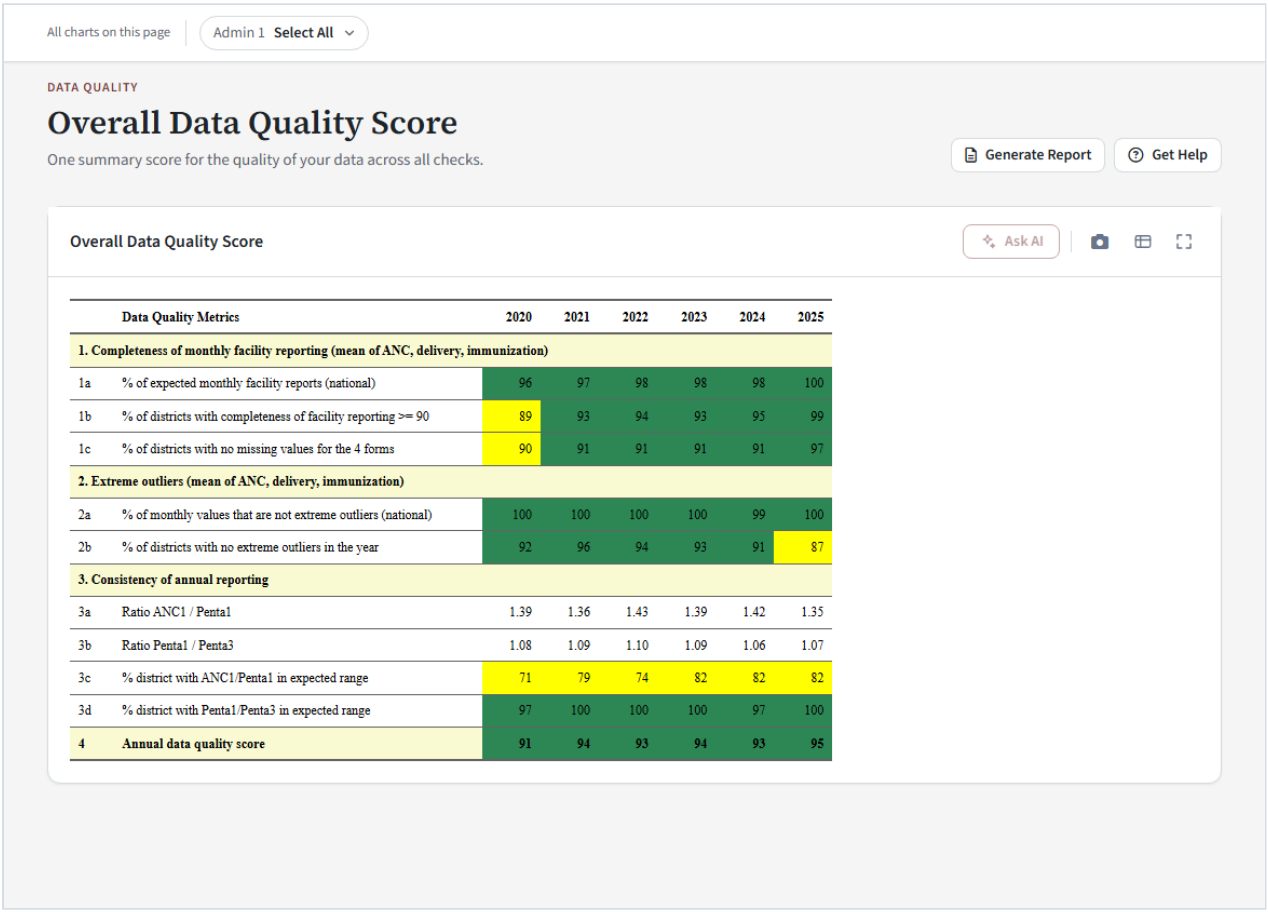
Method: Consistency of annual reporting (p. 68).

23.5 Overall Data Quality Score

Answers how good the data are each year, in one table.

The **Overall Data Quality Score** page sums up the data quality checks in one table, with a column for each year. Open it last, after the other Data Quality Assessment pages: it shows at a glance which years and which checks fall short, and it is the table to share when you report on data quality. What it shows then guides your choices on Data Adjustment (p. 324). RMNCAH Vaxx

Before you start: review Reporting Rate (p. 309), whose **Performance Threshold** colours this table.



	Data Quality Metrics	2020	2021	2022	2023	2024	2025
1. Completeness of monthly facility reporting (mean of ANC, delivery, immunization)							
1a	% of expected monthly facility reports (national)	96	97	98	98	98	100
1b	% of districts with completeness of facility reporting >= 90	89	93	94	93	95	99
1c	% of districts with no missing values for the 4 forms	90	91	91	91	91	97
2. Extreme outliers (mean of ANC, delivery, immunization)							
2a	% of monthly values that are not extreme outliers (national)	100	100	100	100	99	100
2b	% of districts with no extreme outliers in the year	92	96	94	93	91	87
3. Consistency of annual reporting							
3a	Ratio ANC1 / Penta1	1.39	1.36	1.43	1.39	1.42	1.35
3b	Ratio Penta1 / Penta3	1.08	1.09	1.10	1.09	1.06	1.07
3c	% district with ANC1/Penta1 in expected range	71	79	74	82	82	82
3d	% district with Penta1/Penta3 in expected range	97	100	100	100	97	100
4	**Annual data quality score**	**91**	**94**	**93**	**94**	**93**	**95**

Figure 23.11 — *The Overall Data Quality Score table: one column per year, green where the value meets the threshold*

Using the page

1. **Choose the area**

   Leave the **Admin 1** chip on **Select All** for the national table, or pick a region.

2. **Read the score row**

   Row 4, the annual score, tells you how good each year is overall.

3. **Find the checks that fall short**

   Look for yellow and red cells in the rows above it, then go back to the page of that check to see which areas cause it.

4 **Share the result**

Generate Report in the page header opens the data quality report. Then go to Data Adjustment (p. 324).

Controls

Control	Choices	Default and notes
Admin 1	**Select All** (national) or one region	Select All.

Cards

Overall Data Quality Score — one column per year. The app numbers the rows: 1a to 1c for completeness of reporting, 2a and 2b for extreme outliers (p. 499), 3a to 3d for consistency, and 4 for the annual data quality score (p. 502). Green cells meet the threshold, yellow ones are from 70% up to the threshold, and red ones are below 70%. The two ratio rows, 3a and 3b, show the ratio itself and have no colour. A column that is mostly green is a year you can analyse with confidence; a year with several red cells needs care when you read its coverage.

> Vaxx app
> The consistency rows also cover OPV1 and OPV3: rows 3a to 3c are the ratios, and rows 3f to 3h the percentages of districts within the expected range.

Common pitfalls

- Reading the national table only: a good national score can hide a region with poor data. Check the regions in the **Admin 1** chip.
- Changing the **Performance Threshold** on Reporting Rate to turn cells green: the threshold is saved in the dataset and also changes the adjustment.
- Taking the score as a pass or fail: it tells you where to look, and the other pages tell you why.

Method: Annual data quality score (p. 73) (what each row is: the score table (p. 73)).

CHAPTER 24

Data Adjustment

How to choose the data kept for analysis and the adjustment settings on Data Adjustment, adjust the data, and check the result on Data Adjustment Changes.

Turn the reported numbers into the analysis dataset

Choose which data to keep, set how incomplete reporting, outliers and missing months are corrected, then check what changed. The problems found in the data quality (p. 306) review decide the settings.

RMNCAH Vaxx

Learn about data adjustment *(p. 75)*

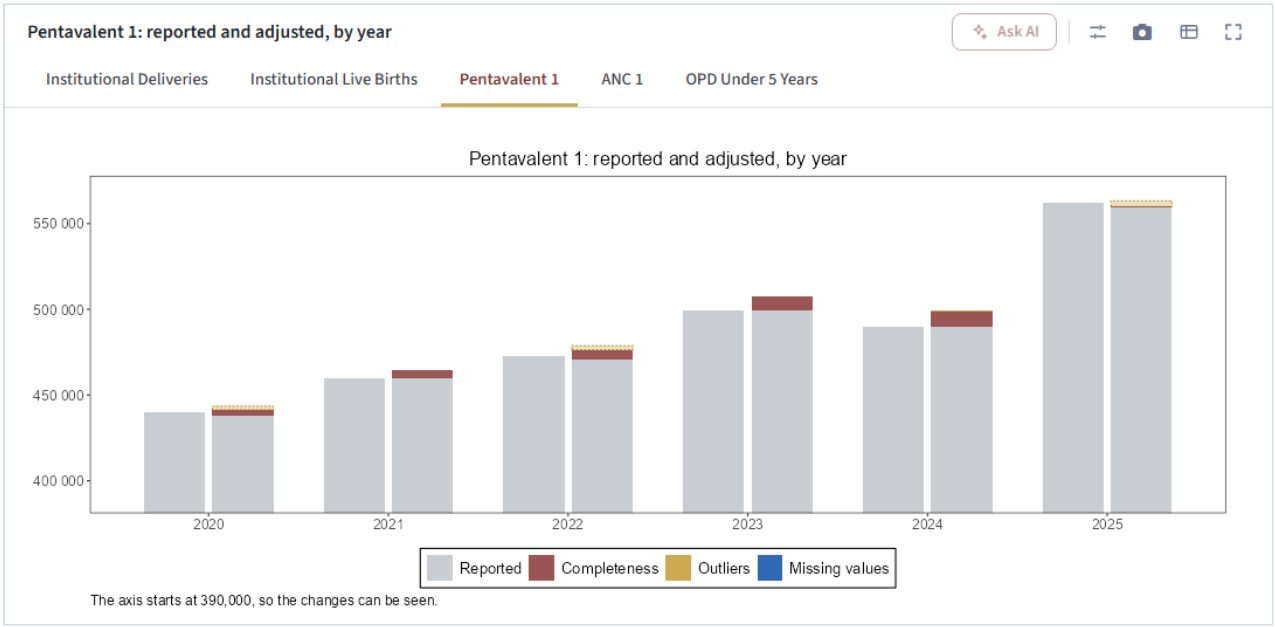
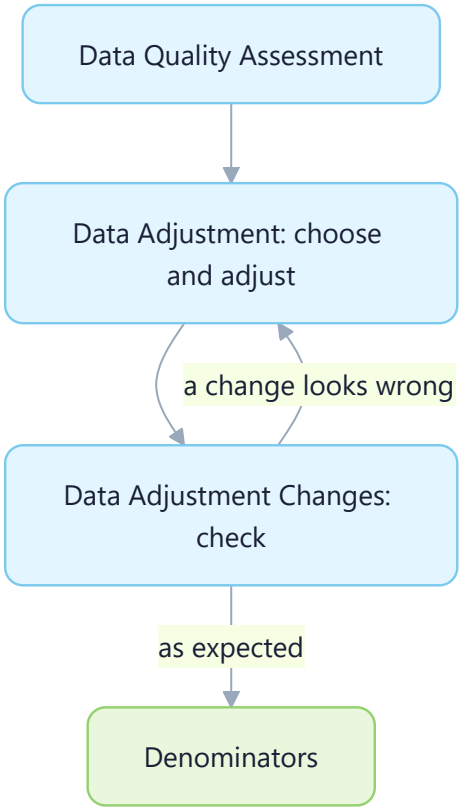


Figure 24.1 — *Data Adjustment Changes: reported and adjusted numbers by year, coloured by adjustment step*

Pages

Before you start: finish Load Data (p. 295) and review the data quality pages.



> The pages under Denominators and Analysis, and Reports, use the adjusted data. They stay locked in the sidebar until the data has been adjusted at least once.

Common pitfalls

- Editing the settings and not pressing **Adjust data again**: the analysis pages keep the previous adjustment.
- Removing a region or district while the national denominator is a UN estimate for the whole country: national coverage comes out too low. Choose the denominator with that in mind (Denominators (p. 333)).
- Changing the settings after the denominators were chosen: check Denominator Selection (p. 338) again.

Pitfalls in choosing k and removing data: Data adjustment > Common pitfalls (p. 78).

Next: Denominators (p. 333).

24.1 Data Adjustment

Sets what data the analysis keeps and how the numbers are adjusted, then applies it.

The **Data Adjustment** page turns the reported numbers into the data the analysis uses. You choose which years and areas to keep, and how to correct incomplete reporting, extreme outliers and missing months; **Adjust data** then applies it. Open it after the Data Quality Assessment (p. 306) pages, whose results guide each choice. Until the data has been adjusted once, the Denominators and analysis pages stay locked in the sidebar. RMNCAH Vaxx

Before you start: review the Data Quality Assessment (p. 306) pages, in particular Reporting Rate (p. 309).

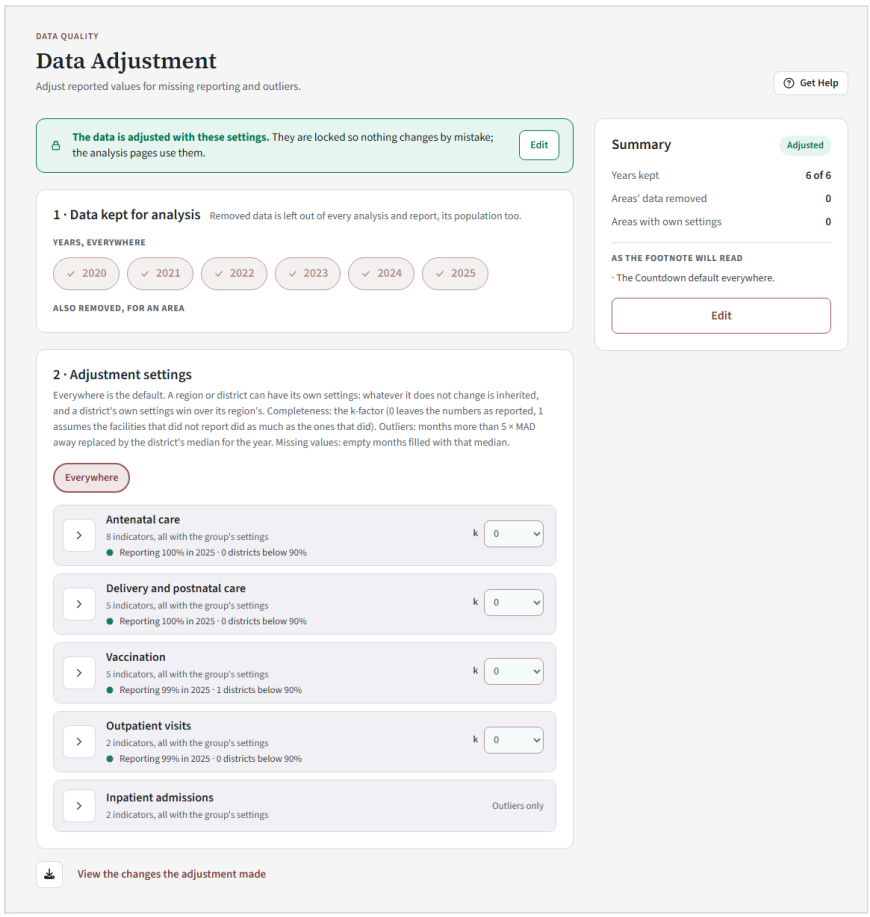
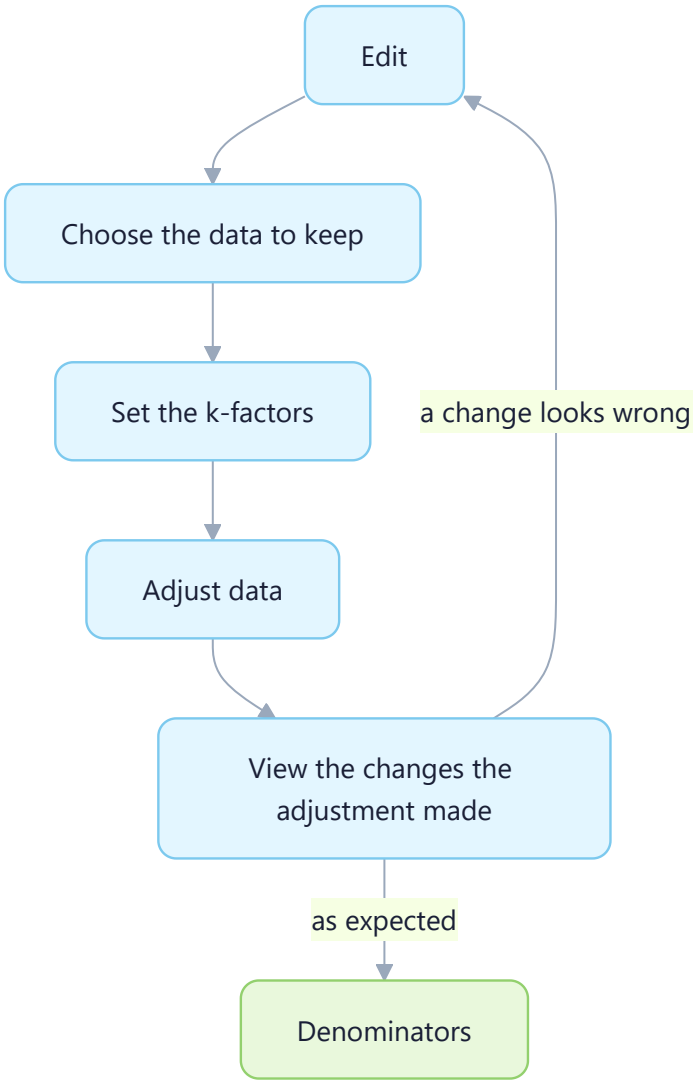


Figure 24.2 — *The Data Adjustment page, locked: the years kept, the adjustment settings by service group, and the Summary*

Using the page



1. **Press Edit**
   The banner turns to *Editing*. *The analysis pages keep the adjustment already made until you press Adjust data again*. **Cancel changes** drops your edits.

② Choose the data to keep

Untick a year under **Years, everywhere** to leave it out everywhere. To leave out one region or district, press **Remove data for an area**, pick it (or type a district's name), choose **Every year** or the years, and press **Remove its data**.

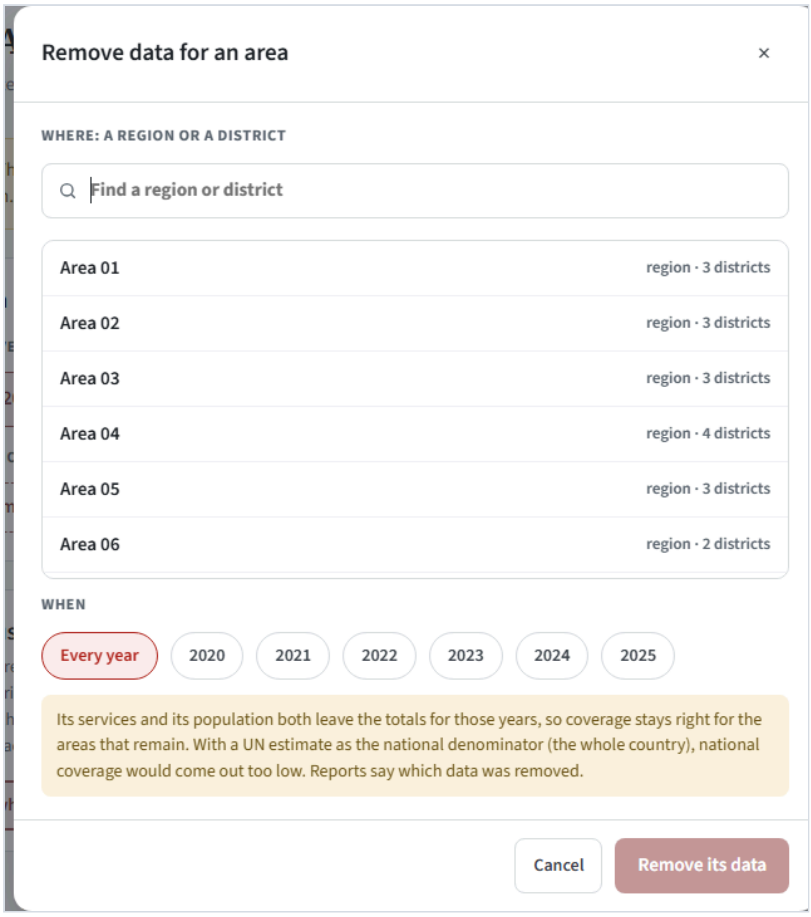


Figure 24.3 — *The Remove data for an area dialog: where, when, and what removing does*

When to remove data, and what it does to coverage: Data adjustment > Remove years (p. 93).

③ Set the k-factors

Choose a **k** for each service group. To give one indicator its own value, or to switch off its outlier or missing-value correction, open the group with its arrow. To give a region or district its own settings, press **Settings for a region or district**, pick it and press **Give it its own settings**; what it does not change is inherited.

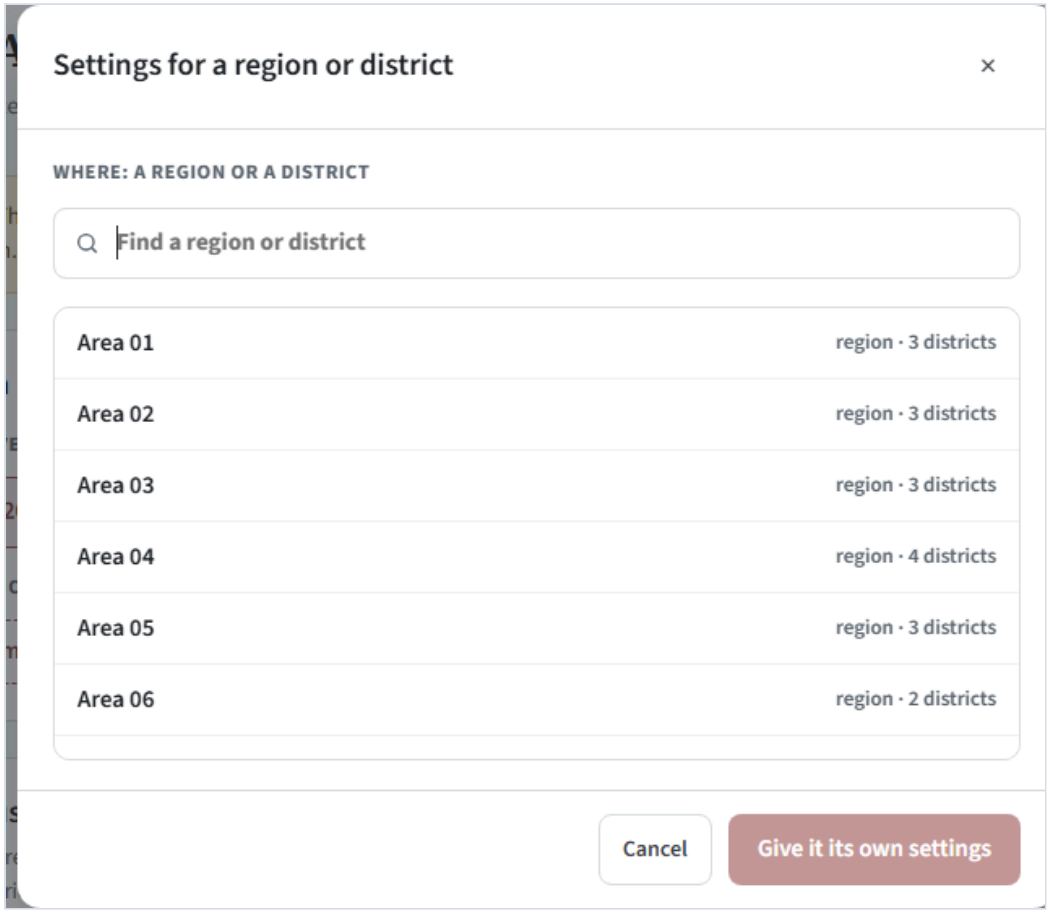


Figure 24.4 — *The Settings for a region or district dialog*

What each value assumes and how to choose: Choosing k (p. 81).

④ Press Adjust data again

The Summary button reads **Adjust data** the first time and **Adjust data again** after that. The page shows *Adjusting…*, then *The data is adjusted with these settings.* and locks again. If it fails, the message *The data could not be adjusted:* gives the reason.

⑤ Check the changes

Press **View the changes the adjustment made** to open Data Adjustment Changes (p. 331).

Controls

Control	Choices	Default and notes
Edit	—	Unlocks the settings. Once the data is adjusted they are locked, so nothing changes by mistake.

Years, everywhere	One chip per year in the data	Every year kept (ticked). Untick a year to leave it out.
Remove data for an area	A region or district; **Every year** or some years	Nothing removed.
Everywhere / **Settings for a region or district**	Everywhere, or a region or district you add	Everywhere. A district's own settings win over its region's.
k (one per service group)	0, 0.25, 0.5, 0.75, 1	The value stored in the dataset.
Per indicator (open a group with its arrow)	Completeness (k): the group's value or its own; **Outliers** on or off; **Missing values** on or off	The group's k; both corrections on.
Back to the Countdown default	—	Puts every setting back to the default.
Adjust data / **Adjust data again**	—	Applies the settings to the data.

Cards

1 · Data kept for analysis — the years and areas the analysis keeps. *Removed data is left out of every analysis and report, its population too.* Under **Also removed, for an area** each removal is listed as *area (region or district) · when · removed from the analysis*. Check that it lists exactly what you meant to leave out.

2 · Adjustment settings — one row per service group: **Antenatal care**, **Delivery and postnatal care**, **Vaccination**, **Outpatient visits** and **Inpatient admissions** (outliers only). Each row gives the number of indicators, the group's reporting rate (p. 503) in the last year (*Reporting 99% in 2025 · 1 districts below 90%*) and its k-factor (p. 500). The arrow at the left opens the group's indicators, with the number of months each correction would change. A group with high reporting changes little whatever its k; a group with many districts below the threshold is where the choice of k matters most.

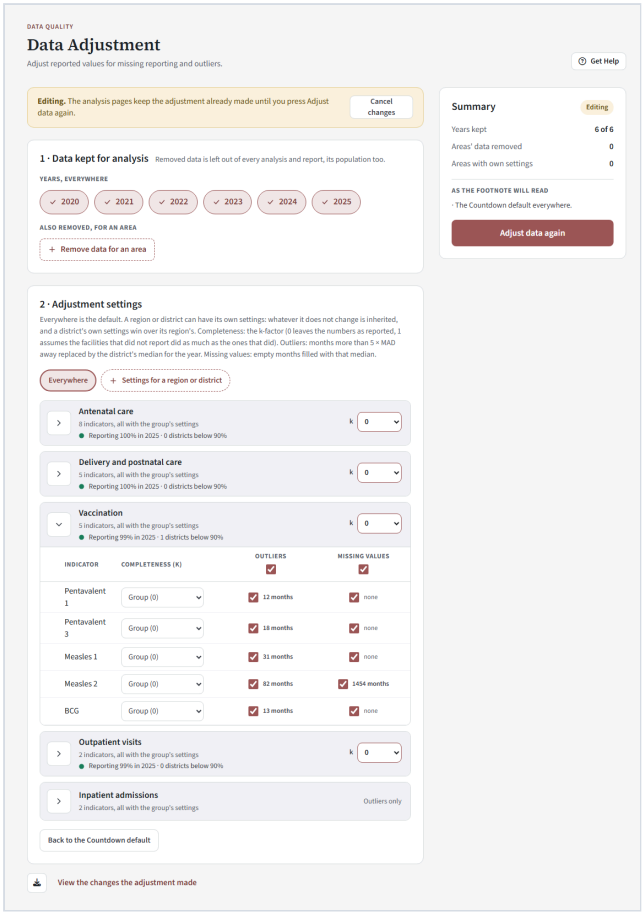


Figure 24.5 — *Editing: the Vaccination group open, with each indicator's k, outlier and missing-value settings*

Summary — the status (*Adjusted*, *Editing* or *Not adjusted*), **Years kept** (for example *6 of 6*), **Areas' data removed**, **Areas with own settings**, and the text the reports' footnote will print (**As the footnote will read**). Its button reads **Edit**, **Adjust data** or **Adjust data again**. Before you move on, the status should read *Adjusted*.

Below the cards, **Download Adjusted Dataset** saves the adjusted data as a Stata file (`master_adj_dataset.dta`) and **View the changes the adjustment made** opens Data Adjustment Changes (p. 331).

> **Vaxx app**
> There are three service groups: **Antenatal care**, **Delivery and postnatal care** and **Vaccination** (16 indicators).

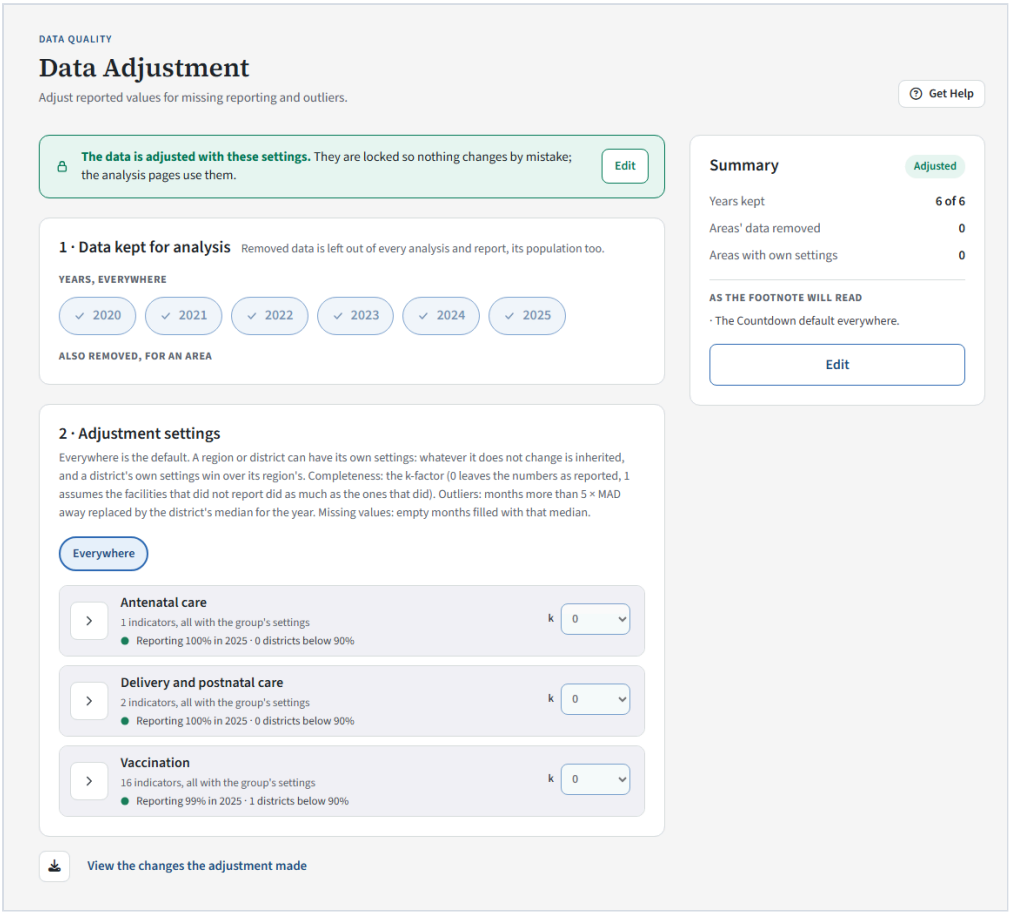


Figure 24.6 — *Vaxx: the Data Adjustment page with its three service groups*

Common pitfalls

- Leaving the page while it reads *Editing*: your changes are not used until you press **Adjust data again**, and **Cancel changes** drops them.
- Removing a year everywhere when only one area had a problem that year: use **Remove data for an area** instead, so the other areas keep their data.
- Pressing **Back to the Countdown default** to undo one change: it puts every setting back. To undo one indicator's own setting, use the ↺ button next to it (*Back to the inherited setting*).

Method: Data adjustment (p. 75) (how the adjustments fit together (p. 76)).

24.2 Data Adjustment Changes

Shows, indicator by indicator, how much each adjustment step changed the reported numbers.

The **Data Adjustment Changes** page shows what the adjustment did: for each indicator and year, the reported number, the adjusted number, and how much of the change came from each step. Open it right after pressing **Adjust data** on Data Adjustment (p. 324). Use it to check that the changes are the size you expected before you choose the denominators, and to explain them in your report. RMNCAH Vaxx

Before you start: adjust the data on Data Adjustment (p. 324). Before the first adjustment the page shows the rules that **Adjust data** will apply.

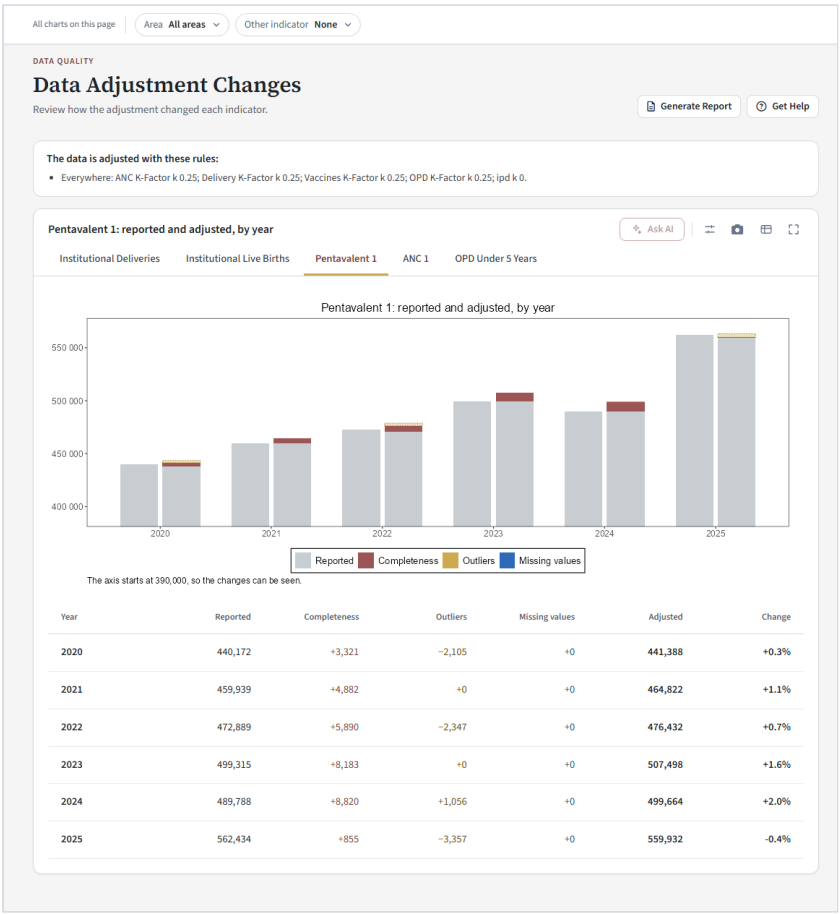


Figure 24.7 — *The Data Adjustment Changes page for pentavalent 1: reported and adjusted numbers by year, and the change made by each step*

Using the page

1. **Check the rules**

   The first card lists the settings in use. Check that they are the ones you chose.

2. **Pick an indicator and an area**

   Open an indicator tab, or choose one in **Other indicator**. Leave **Area** on **All areas** for the national picture, or pick a region or district.

3 **Read the chart and the table**

Compare the reported and adjusted bars, see which step made the change, and read the **Change** in percent in the table.

4 **Report or go back**

Generate Report in the page header opens the adjustment report. If a change looks wrong, go back to Data Adjustment (p. 324) and change its settings; otherwise go on to Denominators (p. 333).

Controls

Control	Choices	Default and notes
Area	**All areas**, a region or a district	All areas.
Other indicator	**None** or any other adjusted indicator	None. Picking one replaces the tab's indicator.

Cards

The data is adjusted with these rules: — the settings in use, as the footnote prints them (*The data is not adjusted yet. These are the rules Adjust data will apply:* before the first adjustment).

Institutional Deliveries: reported and adjusted, by year (the title follows the indicator) — one tab per indicator: **Institutional Deliveries**, **Institutional Live Births**, **Pentavalent 1**, **ANC 1** and **OPD Under 5 Years**. The chart puts the reported number next to the adjusted number for each year; the adjusted bar is split by step: **Completeness**, **Outliers** and **Missing values**. A step that is switched off is marked *(Off)*. The axis does not start at zero, so small changes show; a note under the chart gives where it starts. The table gives, per year, the **Reported** number, the change from each step, the **Adjusted** number and the **Change** in percent. Where reporting was good the two bars are almost the same. A large **Completeness** part means low reporting in that year; a large **Outliers** part means some extreme values were replaced, which is worth checking on Outlier Detection (p. 312).

Generate Report in the page header opens the adjustment report.

> Vaxx app
> The tabs are **Institutional Live Births**, **BCG**, **Pentavalent 1** and **Measles 1**.

Common pitfalls

- Reading the bars by their height alone: the axis does not start at zero, so a small change can look large. Read the **Change** in percent in the table.
- Forgetting that **Other indicator** replaces the tab: set it back to **None** to see the tab's own indicator again.
- Checking the national total only: a change that is small for the country can be large for one district. Pick the areas with low reporting and those you gave their own settings.

Method: how to read the comparison, The effect of the adjustments (p. 76); what to report, Reporting the adjustment (p. 78).

CHAPTER 25

Denominators

How to compare the population sources on Population Trend and choose the denominators on Denominator Selection, since the denominator shapes every coverage estimate.

Choose what coverage is divided by

Coverage is only as good as its target population (p. 504). Compare the population sources, then compare coverage under each denominator with the survey and pick the best. Every coverage page uses your choice. RMNCAH Vaxx

Learn about choosing the denominator (p. 95)

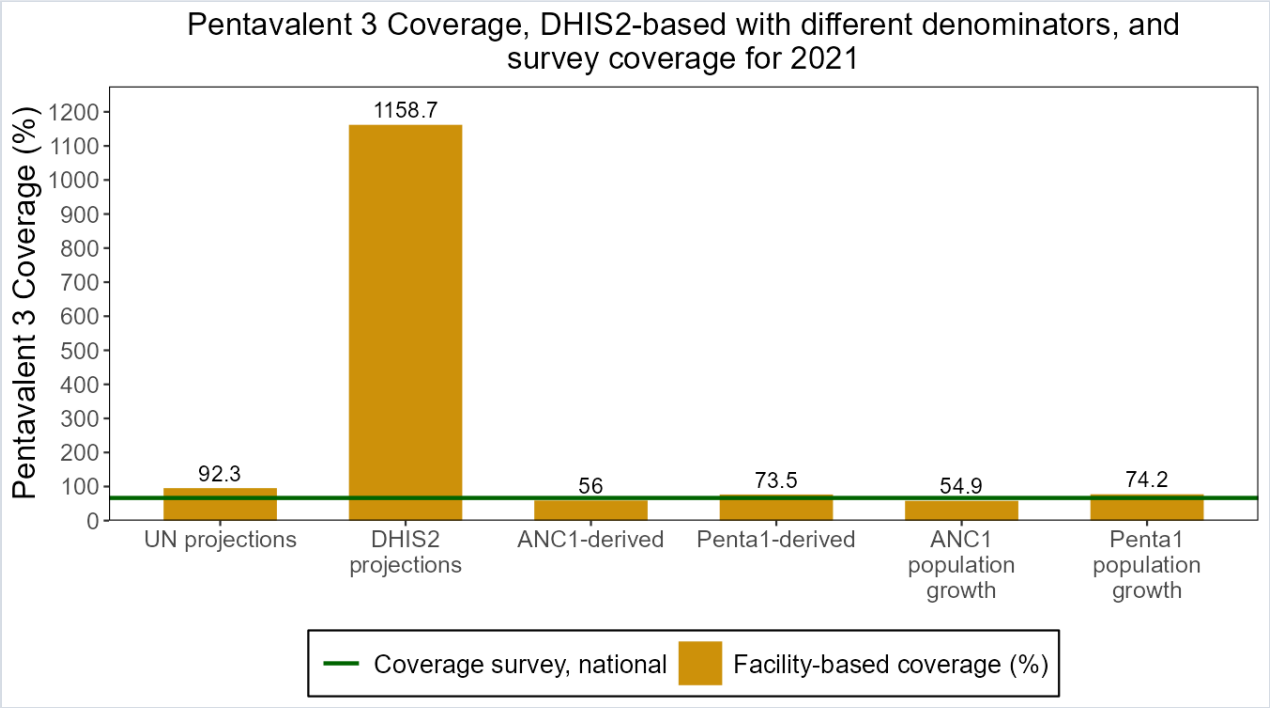
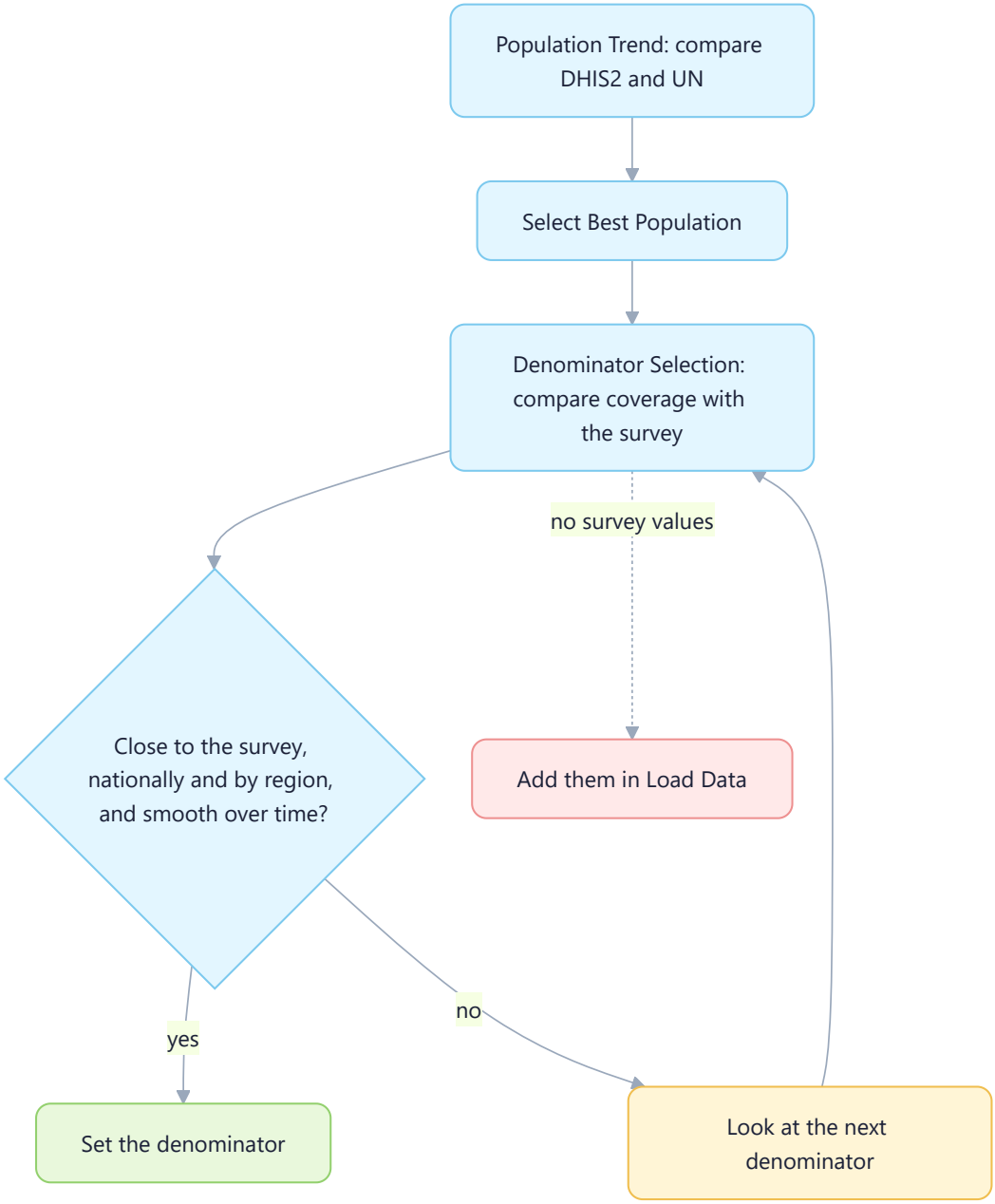


Figure 25.1 — *Pentavalent 3 coverage with each denominator, against the national survey coverage*

Pages

Before you start: adjust the data on Data Adjustment (p. 324), and enter the survey values in National Rates (p. 301).



Common pitfalls

- Choosing a denominator for each chart: it is chosen once, here, for the whole analysis. The analysis pages show it in their header.
- Changing the denominator later without reviewing the pages that use it and the figures already exported.
- Empty survey comparison: the survey coverage values are missing in National Rates (p. 301).
- Changing the data adjustment after choosing: compare the cards again.

Pitfalls in choosing the denominator: Denominator assessment and selection > Common pitfalls (p. 96).

Next: Coverage, equity and targets (p. 342).

25.1 Denominator Assessment

Answers how far the DHIS2 population projections agree with the UN projections. In the sidebar this page is called Population Trend.

The **Denominator Assessment** page compares the population figures in your DHIS2 (p. 498) data with the UN projections (p. 504), year by year: total population, births and children under 1. It is the first Denominators page, called **Population Trend** in the sidebar, and you open it after adjusting the data. It tells you how far the DHIS2 figures can be trusted, and you leave it having chosen the DHIS2 population the population growth denominators are built from. RMNCAH Vaxx

Before you start: adjust the data on Data Adjustment (p. 324); until then the page is locked.

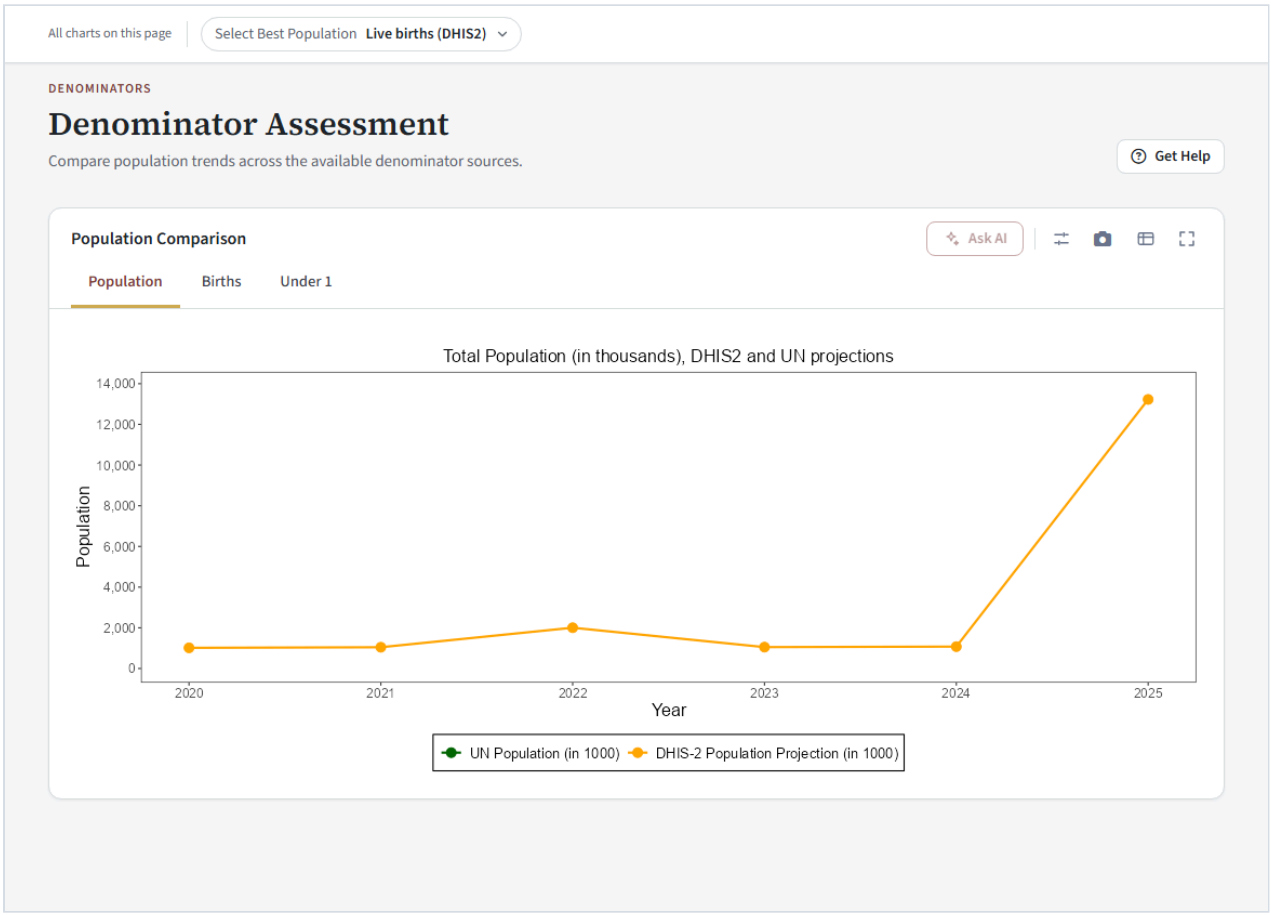


Figure 25.2 — *The Denominator Assessment page: total population, DHIS2 and UN projections, by year*

Using the page

1. **Compare the total population**
   On the **Population** tab, see how close the DHIS2 line is to the UN line, and whether it moves smoothly from year to year.

2. **Compare births and children under 1**
   Do the same on the **Births** and **Under 1** tabs. On **Births**, DHIS2 has two lines: live births and total births.

③ Choose the population

Set **Select Best Population** in the filter bar to the DHIS2 population that looks most reliable. The choice is saved in the dataset.

④ Move on

Download a chart or ask about it from the card toolbar (Exports and Ask AI (p. 411)), then open Denominator Selection (p. 338).

Controls

Control	Choices	Default and notes
Select Best Population	Live births (DHIS2), Total births (DHIS2), Total population (DHIS2), Under 1 year (DHIS2)	The value stored in the dataset. The DHIS2 population used to build the population growth denominators.

Cards

Population Comparison — the UN projection and the DHIS2 projection, in thousands, by year. Tabs: **Population** (total population), **Births** (DHIS2 live births and total births against UN births) and **Under 1**. Two lines that stay close and move in parallel mean the DHIS2 figures are consistent with the UN ones. A gap that widens over time, or a sudden jump in the DHIS2 line, shows a problem in the projections: coverage based on them will drift or jump too.

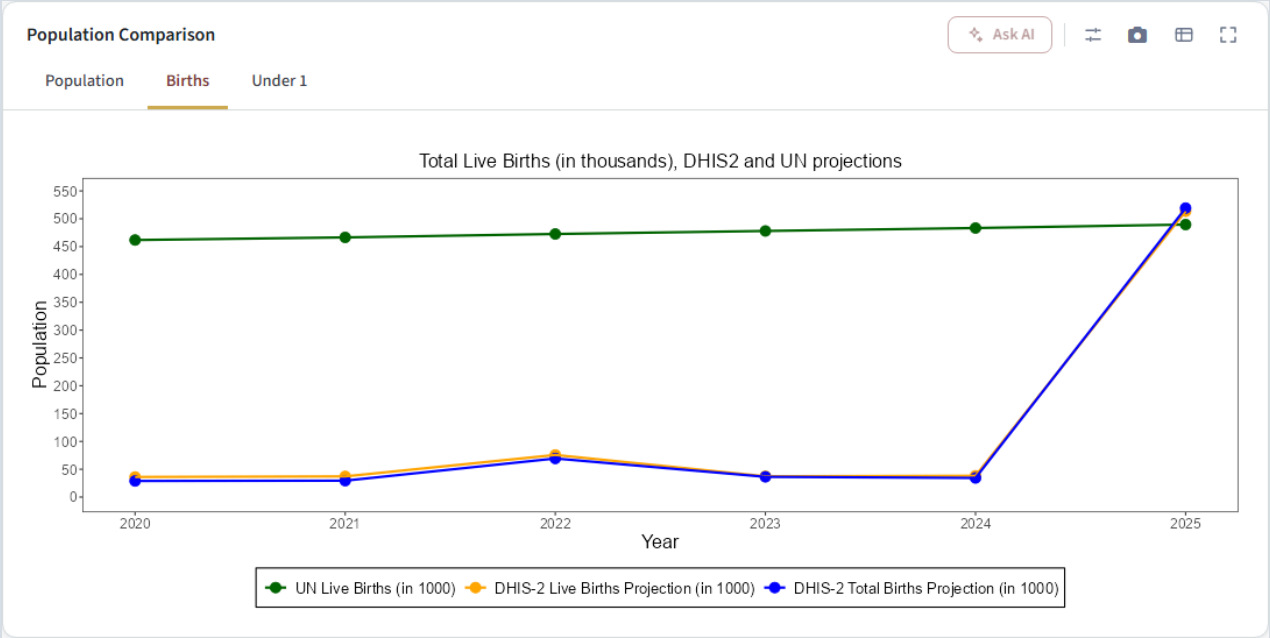


Figure 25.3 — *The Births tab: DHIS2 live births and total births against the UN projection*

Common pitfalls

- Expecting **Select Best Population** to set the denominator: it only chooses the population the population growth denominators grow from. The denominator itself is chosen on Denominator Selection (p. 338).

- Taking the UN line as the truth: it is a projection too. The page shows where the two sources disagree, not which one is right.
- Comparing one tab only: the total population can agree while births or children under 1 do not.

Method: how to choose the population, and what gaps or breaks between the lines mean: Quality of the DHIS2 population projections (p. 100).

25.2 Denominator Selection

Answers which denominator gives coverage closest to the survey, region by region and over time, and records your choice.

The **Denominator Selection** page compares the coverage each denominator gives with the coverage measured by the household survey (p. 500), nationally, by region and over time, and records the denominator you choose. Open it after Denominator Assessment (p. 335). The choice is made once, here, and every coverage, equity, target and mortality page uses it. RMNCAH Vaxx

Before you start: adjust the data on Data Adjustment (p. 324), and enter the survey coverage values in National Rates (p. 301).

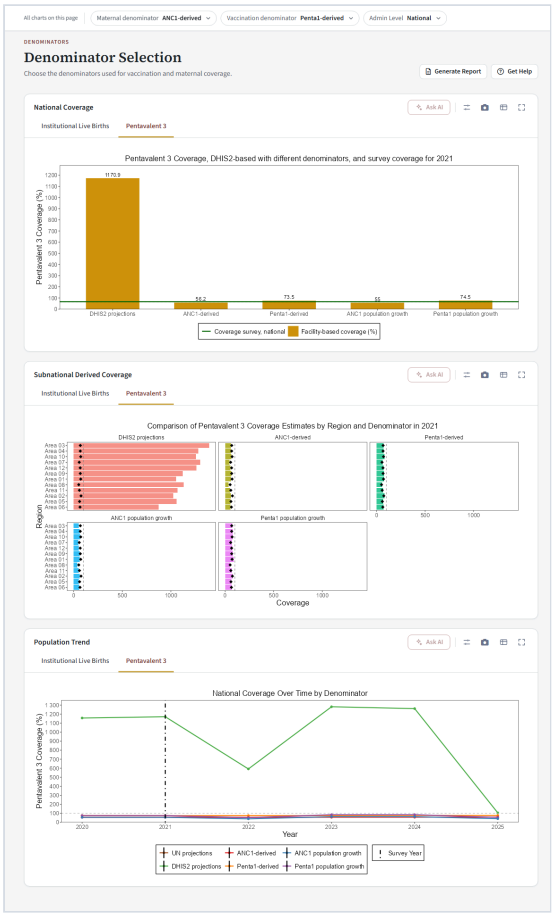


Figure 25.4 — *The Denominator Selection page for pentavalent 3: national coverage by denominator, coverage by region, and coverage over time*

Using the page

① Compare the cards

For each indicator tab, look for the denominator whose coverage is closest to the survey nationally, matches the regional pattern, and moves smoothly over time. Method: Comparison of national estimates (p. 110), Subnational comparison in the survey year (p. 110), Trends in national estimates (p. 110).

② Choose the denominator

Set **Maternal denominator** and **Vaccination denominator** in the filter bar. The choice is saved in the dataset straight away and used by every analysis page. Note why you chose it. How to weigh the three cards: Final selection and synthesis (p. 110).

③ Report the choice

Generate Report in the page header opens the denominator report. Then go on to Coverage, equity and targets (p. 342).

Controls

Control	Choices	Default and notes
Maternal denominator	DHIS2 projections, ANC1-derived, Penta1-derived, Penta1 population growth, ANC1 population growth	The value stored in the dataset. Used for maternal indicators.
Vaccination denominator	The same five	The value stored in the dataset. Used for vaccination indicators.
Admin Level	National, Admin 1, Admin 2 (with a region chip)	National. Changes the National Coverage and Population Trend cards.

Denominator options

The **UN projections** appear on the charts for comparison but cannot be chosen. **DHIS2 projections** are the population figures in the data. **ANC1-derived** and **Penta1-derived** are built from the facility counts (p. 499) and the survey coverage. **ANC1 population growth** and **Penta1 population growth** carry those forward with the population growth (p. 502) of the population you set on Denominator Assessment. What each one is and how it is computed: Denominator options (p. 103) and Population growth options (p. 106).

Cards

National Coverage — for the survey year, the facility-based coverage (p. 499) under each denominator (bars) against the national survey coverage (line). Tabs: **Institutional Live Births** and **Pentavalent 3**. Without survey coverage values the card shows *No coverage to show yet* and *Survey coverage estimates are missing. Add them on Load Data to compare denominators against the survey.*, with a **Go to Load Data** button. The bar whose top is closest to the line is the denominator that agrees best with the survey. A bar far above 100% means that denominator is too small.

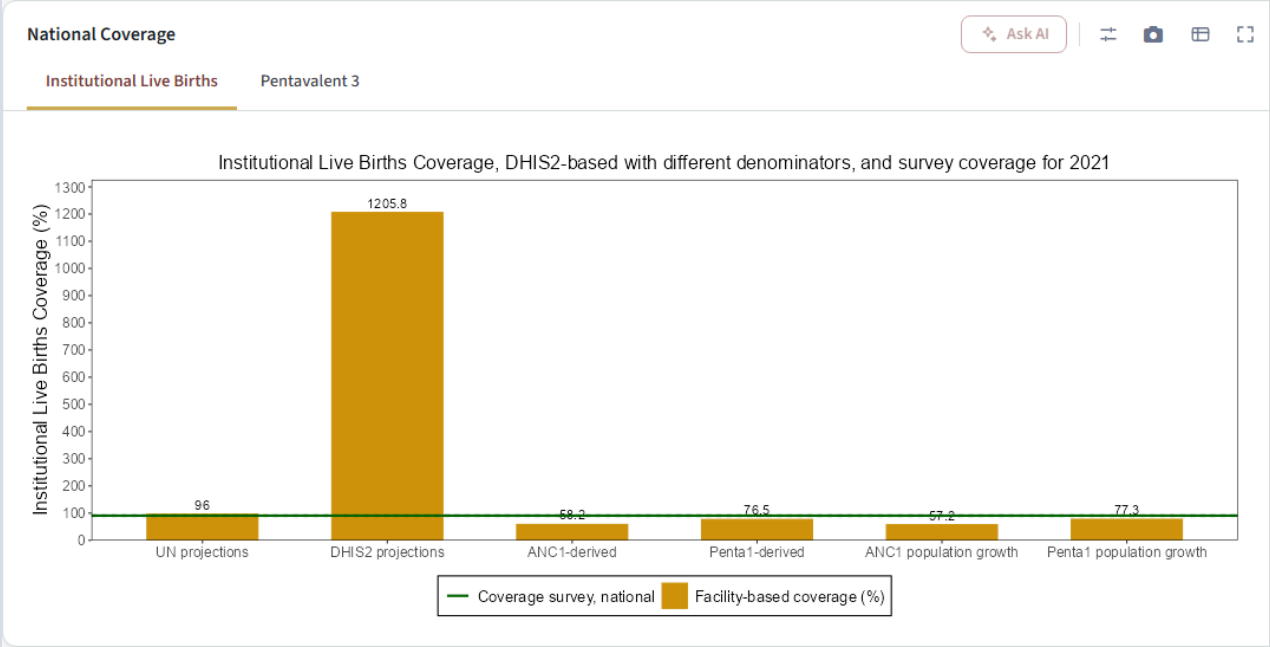


Figure 25.5 — *National Coverage for institutional live births: coverage under each denominator against the survey line*

Subnational Derived Coverage — for the survey year, one panel per denominator with the coverage of every region (bars) and the region's survey value (dot). Same tabs. The best panel is the one where the bars sit close to the dots in most regions, not only on average.

Population Trend — national coverage (or that of the area chosen in Admin Level) by year, one line per denominator, with the survey year marked. Same tabs. A good denominator gives a smooth line; a sudden jump or drop with no change in services comes from the denominator, not from the services.

Generate Report in the page header opens the denominator report.

> Vaxx app
> There is only a **Vaccination denominator**. The National Coverage tabs are **Institutional Live Births**, **BCG**, **Pentavalent 3** and **Measles 1**; the other two cards have **Pentavalent 1**, **Pentavalent 3** and **Measles 1**.

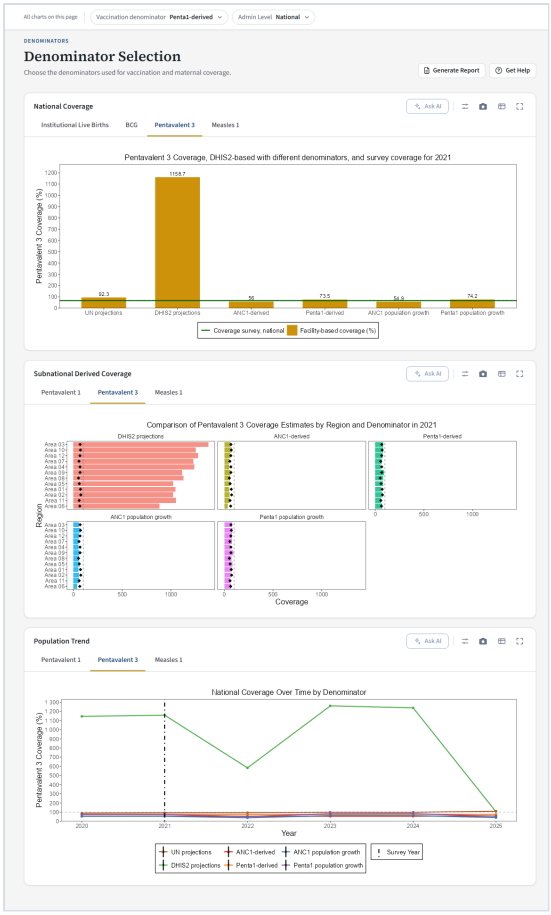


Figure 25.6 — *Vaxx: the Denominator Selection page with one vaccination denominator*

Common pitfalls

- Choosing from the National Coverage card alone: a denominator can match the survey nationally and still be far off in the regions or jump over time. Read all three cards.
- Choosing from one tab only: check every indicator tab before setting a denominator.
- An empty National Coverage card: the survey coverage values are missing. Use **Go to Load Data** to add them.

Method: Denominator assessment and selection (p. 95).

CHAPTER 26

National Analysis

The National Analysis pages, National Coverage, Coverage Target, National Inequality and Equity Assessment, which show coverage, targets and inequalities for the whole country from the adjusted data.

Coverage for the whole country

See how facility coverage compares with surveys, how many areas reach the global targets, and how far areas and population groups are from the national figure. This is where you first read the results.

RMNCAH Vaxx

Learn about coverage estimation *(p. 115)*

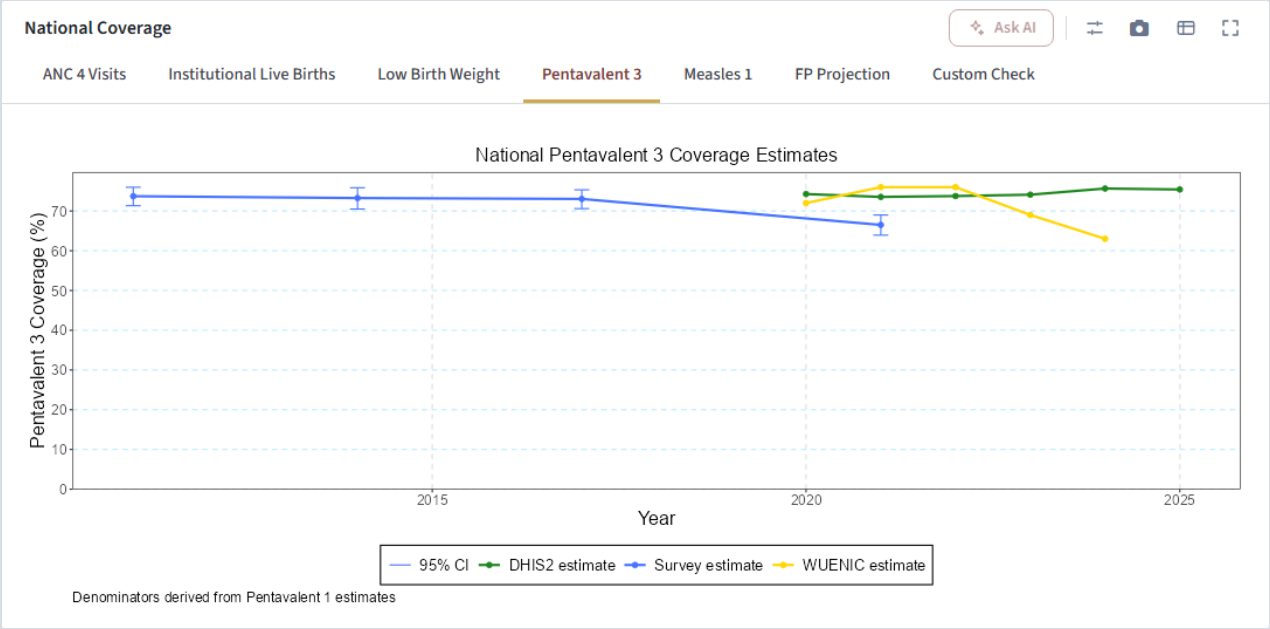


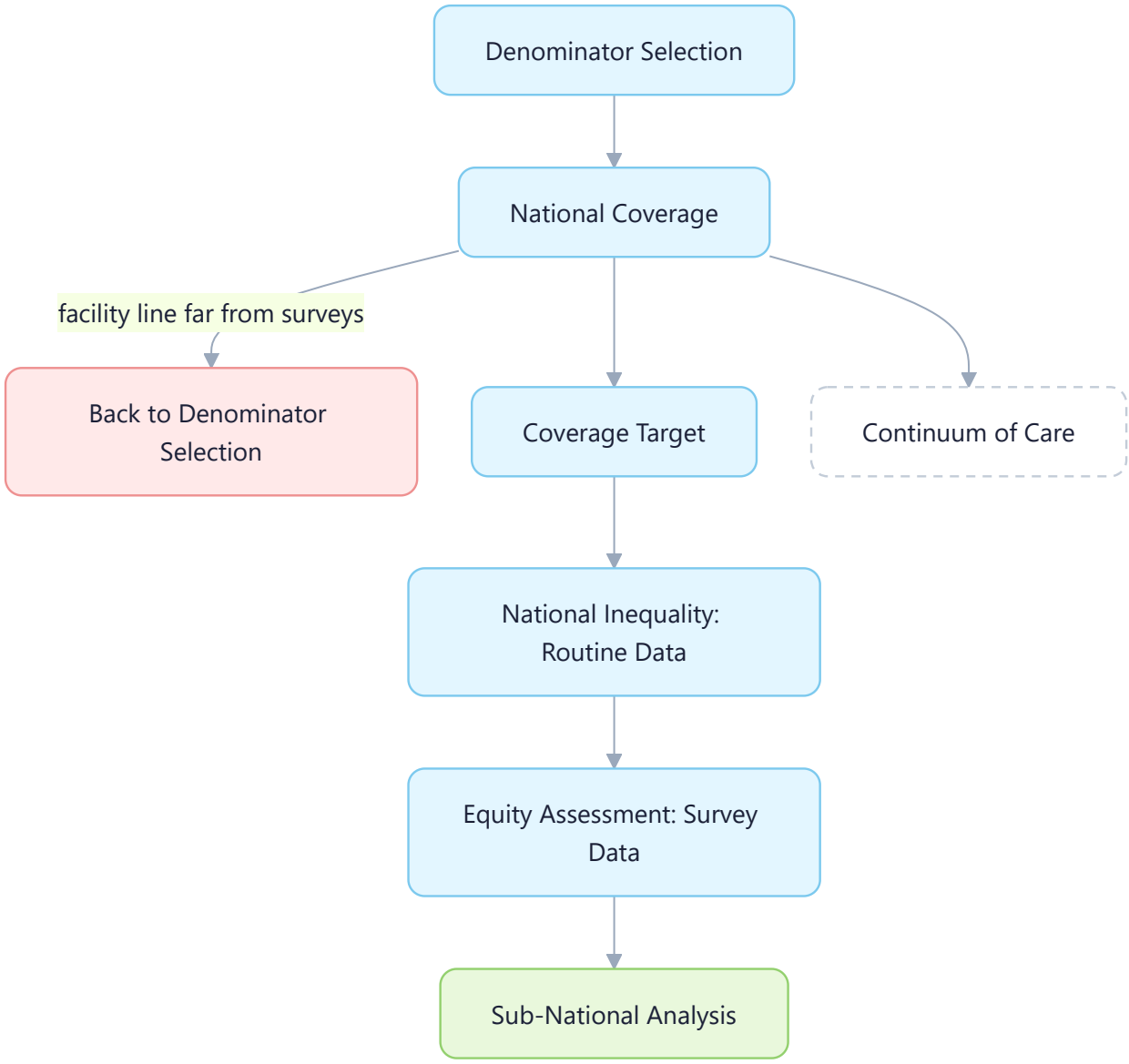
Figure 26.1 — *The National Coverage card: facility, survey and WUENIC estimates of Pentavalent 3 coverage by year*

Pages

Before you start: load the data, run Data Adjustment (p. 324) and choose the denominators on Denominator Selection (p. 338). For the survey lines and the Equity Assessment page, load the survey files in the Survey Files step of Load Data (p. 295).

In the sidebar the group has **National Coverage**, **Coverage Target**, **Continuum of Care** (RMNCAH) and **National Inequality**, which opens two pages: **Routine Data** and **Survey Data**. Every page shows a **Denominator** strip under its title: the denominators the charts use, set on the Denominator Selection page.

How the pages fit together



National Coverage tells you whether the facility data can be trusted against the surveys. Coverage Target and National Inequality then look behind the national figure, at the areas; Equity Assessment looks at groups of people. Continuum of Care (RMNCAH only) follows the steps of care.

Continuum of Care

This RMNCAH page follows coverage along the care pathway, nationally and by region. It is described with the other advanced analyses: Continuum of Care (p. 376) (its **Get Help** button opens that section).

Common pitfalls

- **No survey line, or an error on Equity Assessment.** The survey estimates come from the survey files of the Survey Files step in Load Data (p. 295). Without the national file the coverage charts show the DHIS2 line only; without the area, education or wealth files the Equity Assessment chart shows an error instead of the equiplot. Load the files, then come back.
- **An empty chart.** The indicator has no data for that view, or no unit reaches the target (Coverage Target). Pick another tab or indicator.
- **Results changed.** The coverage pages follow the denominators set on Denominator Selection. Check the Denominator strip and the chart caption before you compare two charts.
- **Skipping National Coverage.** If the facility line is far from the surveys, the targets and inequality pages inherit the same problem.

Pitfalls in reading the results: National coverage (p. 117), Subnational analysis (p. 167) and Equity assessment (p. 155).

Next: Sub-National Analysis (p. 359).

26.1 National Coverage

How does coverage from facility data (DHIS2) compare with the survey and WUENIC estimates, year by year?

National Coverage answers one question: does the coverage calculated from the facility data (DHIS2 (p. 498)) tell the same story as the household surveys (p. 500) and, for vaccines, the WUENIC (p. 505) estimates? It is the first page of the National Analysis group, opened once the data are adjusted and the denominators chosen. If the facility line agrees with the surveys, you can rely on it for the years the surveys do not cover; if not, go back to the denominators or the data quality pages before you go further. RMNCAH Vaxx

Before you start: run Data Adjustment (p. 324), choose the denominators on Denominator Selection (p. 338), and load the survey files in Load Data (p. 295) to see the survey line.

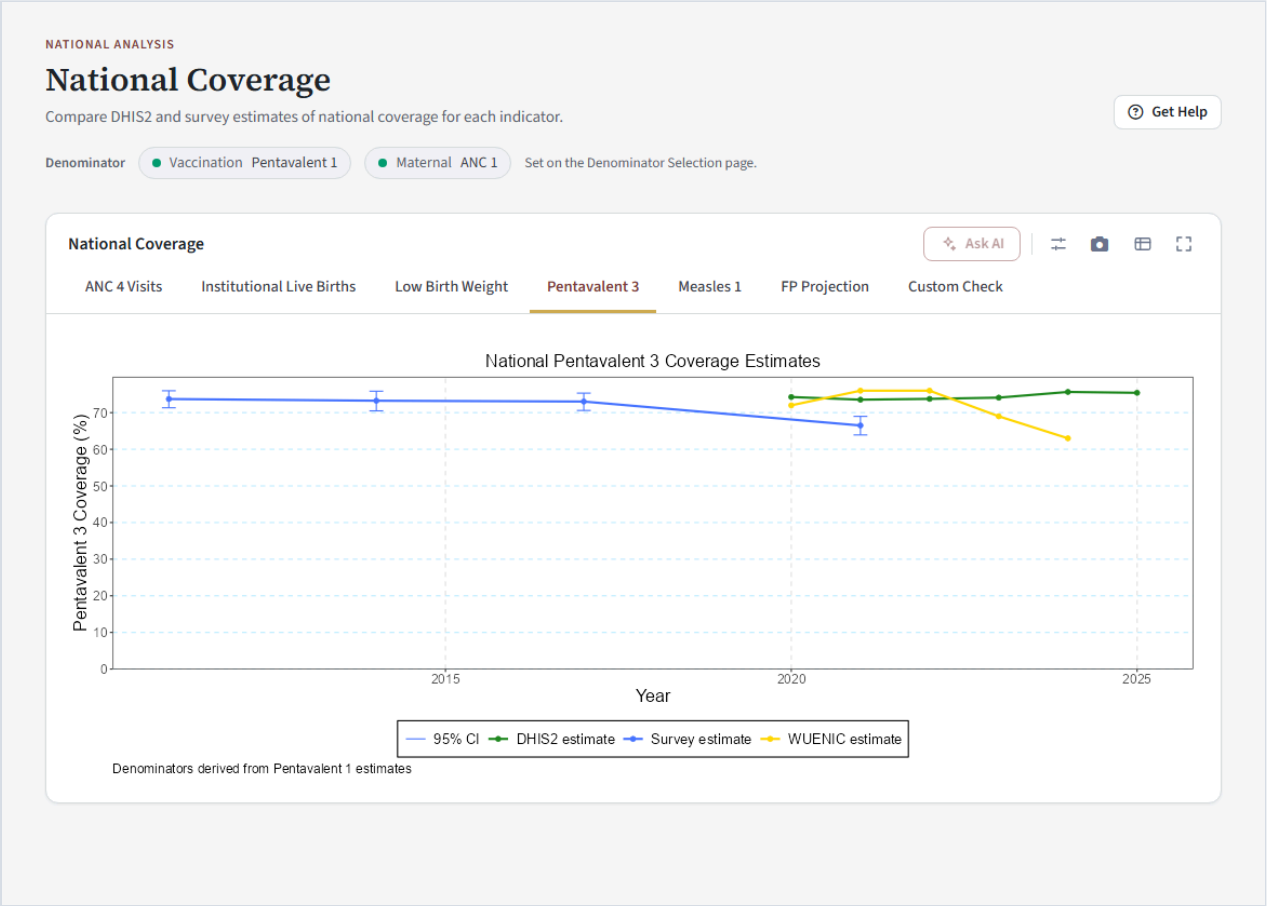


Figure 26.2 — *The National Coverage page with the Pentavalent 3 tab open*

Using the page

① Check the denominators

The **Denominator** strip under the title shows the denominators the charts use. Change them on Denominator Selection, not here.

2. **Pick an indicator tab**
   Start with **ANC 4 Visits**, then go through the tabs. For any other indicator, open **Custom Check** and choose it in the **Indicator** chip.

3. **Compare the lines**
   Look at where the green DHIS2 line sits against the blue survey points and their error bars, year by year.

4. **Keep or share the chart**
   Use the card toolbar: **Ask AI** for a reading of the chart, or download it (see Exports, reports and AI (p. 411)). Then go on to Coverage Target (p. 349).

Controls

Control	Choices	Default/notes
Indicator tabs	ANC 4 Visits, Institutional Live Births, Low Birth Weight, Pentavalent 3, Measles 1, FP Projection, Custom Check	ANC 4 Visits
Indicator (Custom Check tab only)	Any analysis indicator, from ANC 1st Trimester to Syphilis Test	ANC 1st Trimester

The page has no filter bar: it always shows the national level.

Cards

National Coverage — one chart per tab, titled "National … Coverage Estimates". The green line is the **DHIS2 estimate**, computed from the adjusted data and the denominator named in the caption ("Denominators derived from … estimates"). The blue line is the **Survey estimate**, with its **95% CI** as error bars, one point per survey year. For vaccines a yellow **WUENIC estimate** line is added.

How to read it: a good result is a green line that passes through or close to the survey error bars and follows the same trend. A green line far above the survey points, or above 100%, often points to a denominator that is too small; a line far below can mean missing reports or a denominator that is too large.

The **Custom Check** tab adds an **Indicator** chip, so you can draw the same chart for any other indicator.

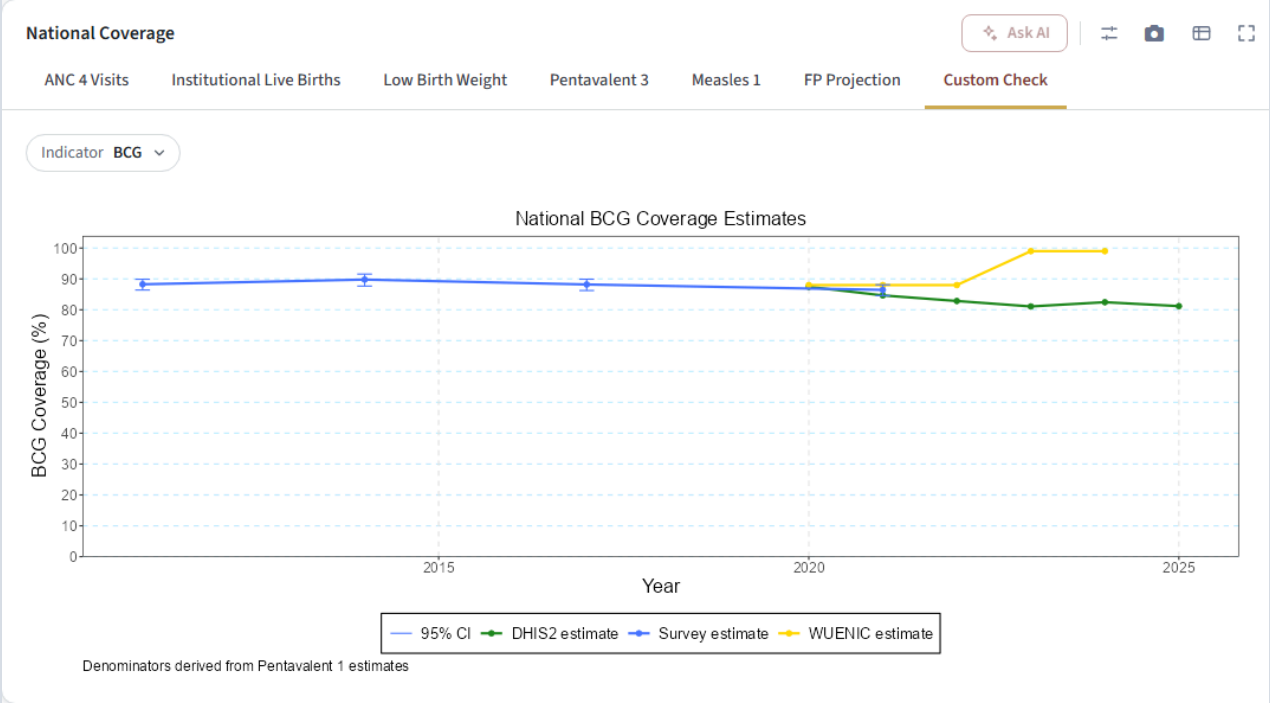


Figure 26.3 — *The Custom Check tab, here for BCG*

The **FP Projection** tab draws the family planning projection instead of a coverage chart: see FP Projection (p. 379).

Vaxx app

The tabs are ANC 4 Visits, Institutional Live Births, Low Birth Weight, Pentavalent 3, Measles 1 and Custom Check (no FP Projection). The Custom Check list holds the vaccination indicators: every dose of BCG, IPV, Measles, OPV, PCV, Pentavalent and Rotavirus, the four dropout rates, **Under-vaccinated** and **Zero-dose**. The Denominator strip shows the vaccination denominator only.

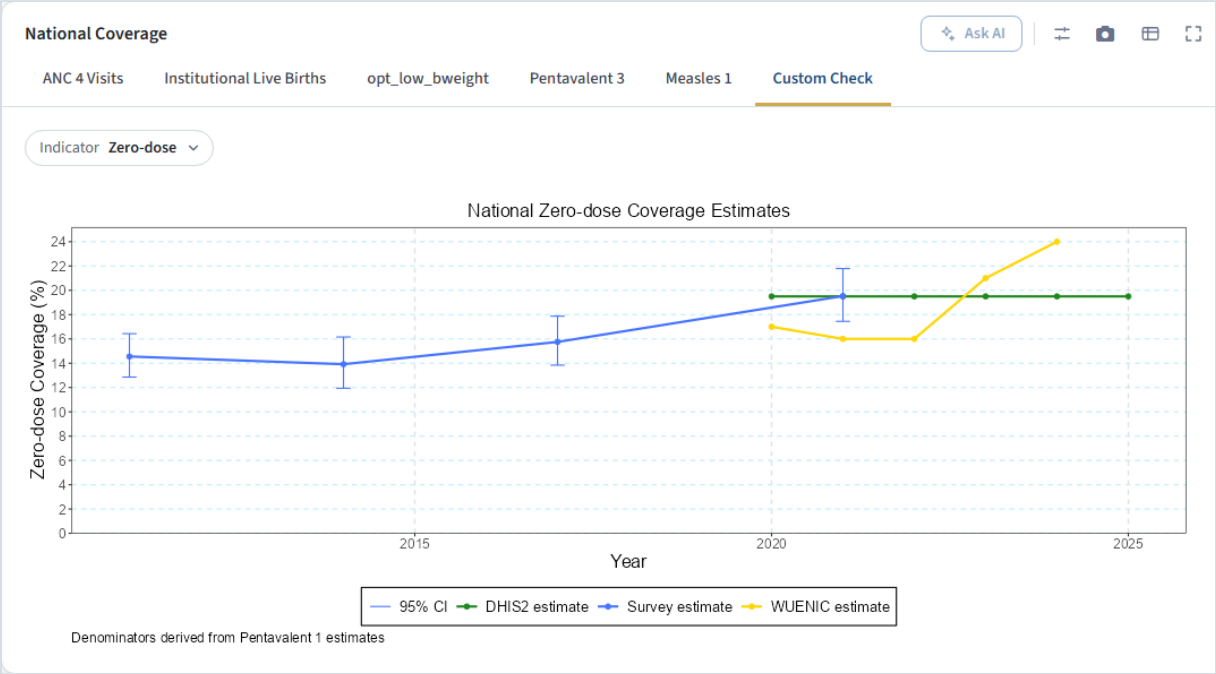


Figure 26.4 — *Vaxx app: Custom Check with Zero-dose*

Common pitfalls

- **Only the green line.** No survey file was loaded, or the survey has no value for this indicator. Load the national survey file in the Survey Files step of Load Data.
- **Reading one survey point as a trend.** One survey year gives one level, not a direction. Compare the shape of the green line with the WUENIC line (vaccines) or with later surveys.
- **Comparing charts made with different denominators.** Check the caption of each chart before you compare them.

Method: Comparison with surveys and targets (p. 135); each indicator in Antenatal care (p. 120), Delivery care (p. 124) and Immunization (p. 128).

Next: Coverage Target (p. 349).

26.2 Coverage Target

What share of regions or districts reach the global coverage targets each year, and which ones are they?

Coverage Target shows how many areas of the country reach the global coverage targets each year, and lists the ones that do. Open it after National Coverage (p. 345): the national figure can hide areas far below the target. Use it to see whether the share of areas on target is growing, and which regions or districts are already there. RMNCAH Vaxx

Before you start: run Data Adjustment (p. 324) and choose the denominators on Denominator Selection (p. 338).

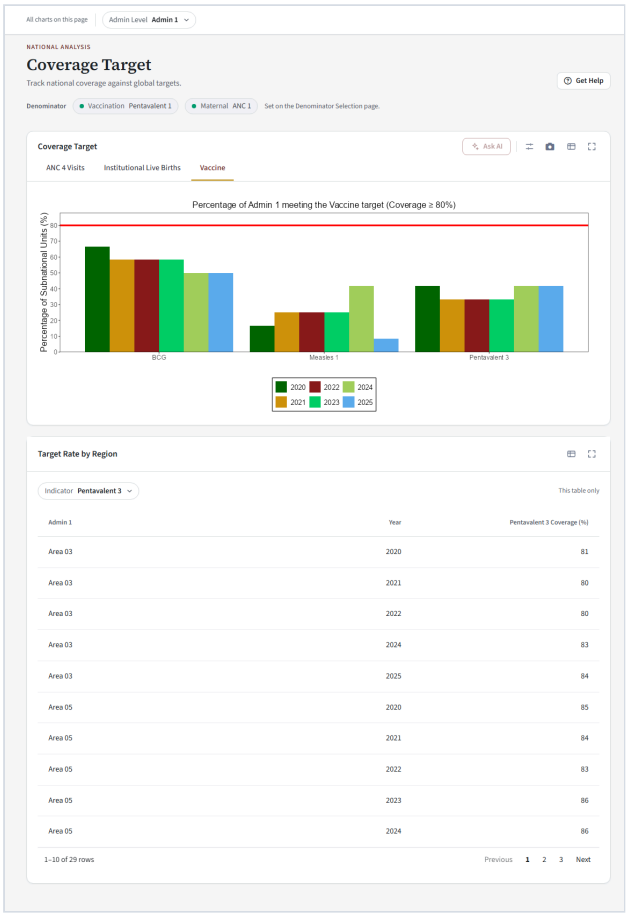


Figure 26.5 — *The Coverage Target page with the Vaccine tab and the table for Pentavalent 3*

Using the page

1. **Choose the level**

   In the filter bar, pick **Admin 1** to count regions or **Admin 2** to count districts. The choice changes every chart and the table.

2. **Read the bars**

   Open each tab and compare the bars from year to year against the target in the chart title.

3 **List the areas on target**

In the table, choose the indicator in its **Indicator** chip: the table lists the areas and years that reach the target. Download it from the table toolbar if you need the list.

4 **Look inside one region**

To count the districts of one region, go to Coverage Target (p. 368) in Sub-National Analysis. Otherwise go on to National Inequality (p. 352).

Controls

Control	Choices	Default/notes
Admin Level (all charts on this page)	Admin 1, Admin 2	Admin 1: the share of Admin 1 (p. 496) units; Admin 2: the share of Admin 2 (p. 496) units (districts)
Indicator tabs	ANC 4 Visits, Institutional Live Births, Vaccine	ANC 4 Visits
Indicator (table only)	Any indicator in the dataset	ANC 1st Trimester; pick the indicator you want to list

Cards

Coverage Target — bars give the percentage of units meeting the target (p. 502), one bar per year. The title states the rule, for example "Percentage of Admin 1 meeting the Vaccine target (Coverage ≥ 80%)", and a red line runs across the chart. The **Vaccine** tab has one group of bars each for BCG, Measles 1 and Pentavalent 3.

How to read it: bars that grow from year to year mean more areas reach the target; a good result is most areas on target in the latest year. Low bars, or bars that fall, show that the national figure hides many areas below the target.

Target Rate by Region (at Admin 1) or **Districts Meeting the Coverage Target** (at Admin 2) — the units and years where the indicator chosen in the table's **Indicator** chip reaches the target, with its coverage. This chip changes this table only. The table can be downloaded and expanded from its toolbar.

How to read it: the areas listed are the ones on target; an area that appears in early years but not in the latest one has fallen below the target.

Vaxx app

The tabs are **Vaccine** (BCG, Measles 1, Pentavalent 3) and **Dropout** (Penta 1 to Penta 3 and Penta 3 to Measles 1). For dropout the target is a rate below the threshold, so a high bar means many areas keep their dropout low. The table's Indicator chip starts at ANC 1.

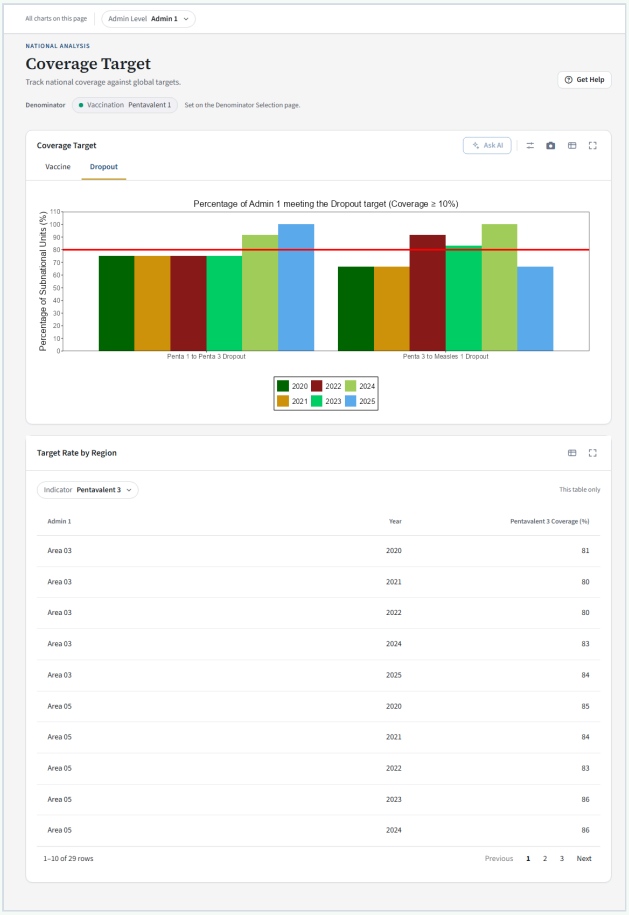


Figure 26.6 — *Vaxx app: the Dropout tab*

Common pitfalls

- **The table stays empty.** Its Indicator chip starts on an indicator that has no target (ANC 1st Trimester, or ANC 1 in Vaxx). Choose the indicator of the tab you are reading.
- **An empty chart.** No unit reaches the target for that tab. Try the other admin level.
- **Comparing Admin 1 and Admin 2 bars.** A share of regions and a share of districts are not the same measure: districts are smaller and vary more.

Method: Coverage targets (p. 175) (the thresholds: Settings (p. 176)).

Next: National Inequality (p. 352).

26.3 National Inequality

How far is each region from the national coverage, year by year, using the facility data?

National Inequality shows how far each region is from the national coverage, year by year, from the facility data. In the sidebar it is **National Inequality > Routine Data**. Open it after Coverage Target (p. 349): it tells you whether the gap between regions is narrowing, and the map shows where the low-coverage regions are. RMNCAH Vaxx

Before you start: run Data Adjustment (p. 324) and choose the denominators on Denominator Selection (p. 338). The map uses the shapefile of the Load Data page.

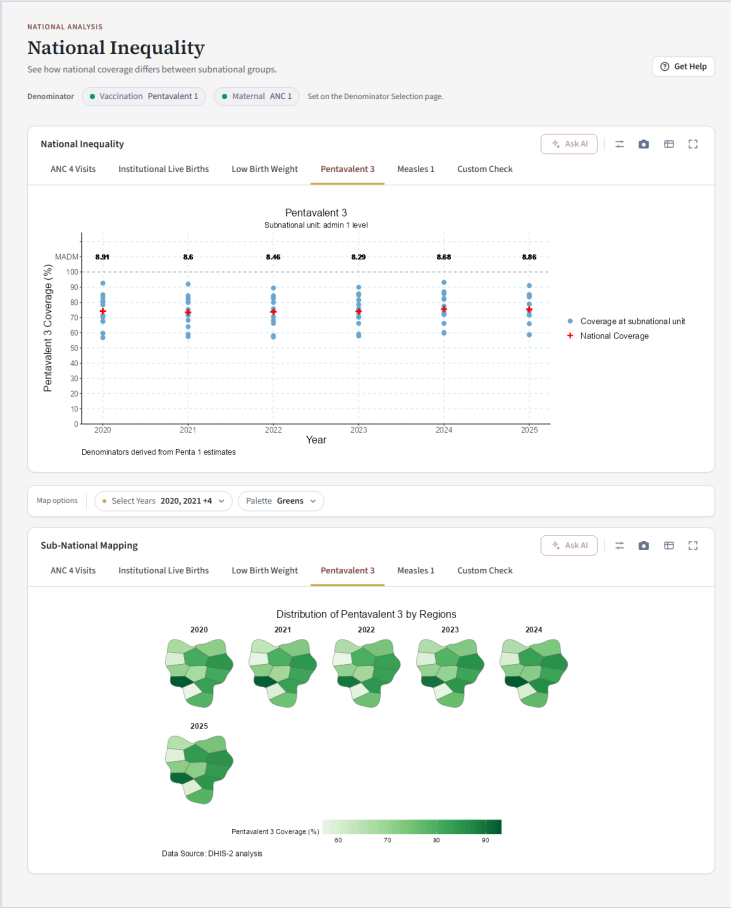


Figure 26.7 — *The National Inequality page, Pentavalent 3 tab, with the map below*

Using the page

1. **Pick an indicator**
   Choose a tab on the **National Inequality** card. Each card has its own tabs.

2. **Read the spread**
   See how widely the dots spread around the red cross, and how the MADM (p. 501) row changes from year to year.

3 **Find the regions on the map**

Pick the same tab on the **Sub-National Mapping** card. In **Map options**, choose the years to draw in **Select Years** and the colours in **Palette**.

4 **Look inside a region**

To see the districts of a region, go to Sub-National Inequality (p. 365). Otherwise go on to Equity Assessment (p. 355).

Controls

Control	Choices	Default/notes
Indicator tabs (each card)	ANC 4 Visits, Institutional Live Births, Low Birth Weight, Pentavalent 3, Measles 1, Custom Check	ANC 4 Visits
Select Years (map options)	Each year of the data, or All years	All years; one small map per year
Palette (map options)	Greens, Blues, Reds, Purples	Greens

Cards

National Inequality — each blue dot is the coverage of one Admin 1 unit in a year ("Coverage at subnational unit"); the red cross is the national coverage. The row at the top gives the MADM for each year: the bigger it is, the further the units are from the national figure. The caption names the denominator.

How to read it: a good result is dots close to the red cross and an MADM that falls over the years. Dots spread far apart, or an MADM that grows, mean some regions are being left behind; a dot alone far below the others is a region to look at first.

Sub-National Mapping — "Distribution of … by Regions": one map per selected year, each area shaded by its coverage, with the scale below. It has its own tabs; pick the same indicator in both cards to compare them.

How to read it: darker areas have higher coverage, lighter areas lower coverage (read the values on the scale). Compare the maps of the first and last years to see which regions moved.

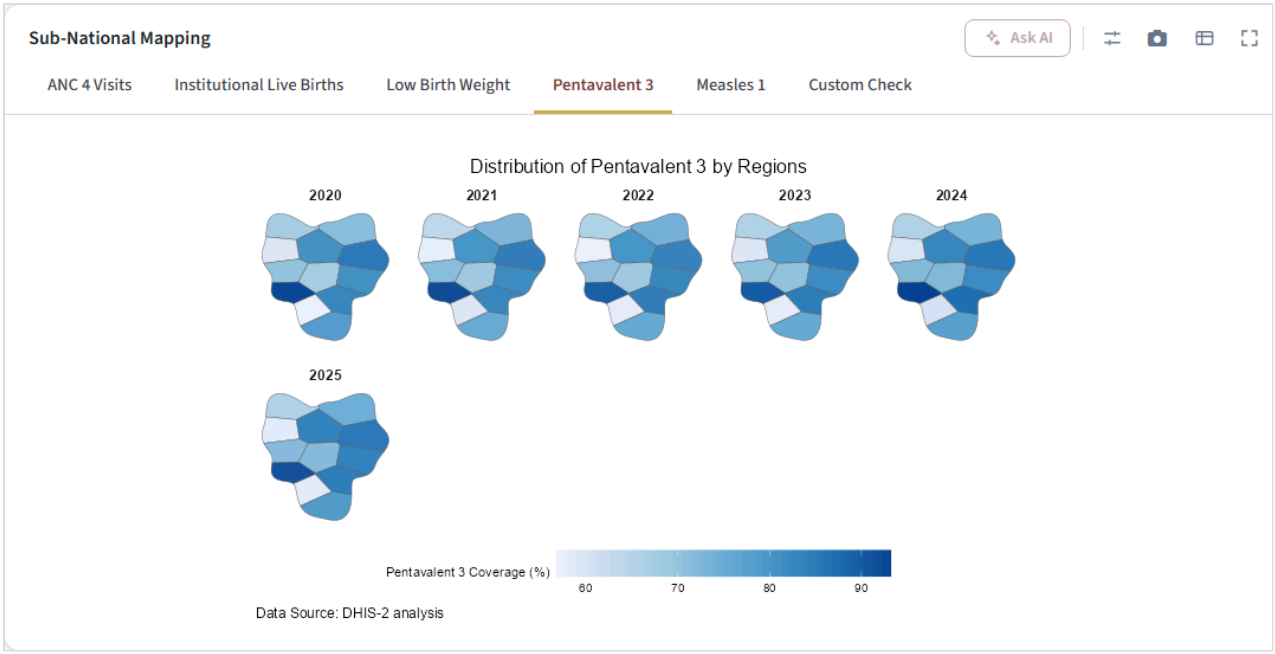


Figure 26.8 — *The Sub-National Mapping card with the Blues palette*

Common pitfalls

- **The two cards show different indicators.** Each card keeps its own tab. Pick the same tab in both before you compare them.
- **Areas with no colour on the map.** Their names in the shapefile do not match the region names of the dataset. Match them in the Map Shapefile step of Load Data (p. 295).
- **A region far from the others.** Before you call it a gap in coverage, check its reporting on Reporting Rate (p. 309).

Method: Subnational inequality (p. 172) (the MADM: Formula (p. 172); reading the chart: Interpretation (p. 173)).

Next: Equity Assessment (p. 355).

26.4 Equity Assessment

Who is being left behind? Survey coverage by area of residence, mother's education or wealth quintile.

Equity Assessment asks who is being left behind. It compares coverage between groups of the population, by area of residence, mother's education or wealth quintile (p. 505), from the household surveys (p. 500). In the sidebar it is **National Inequality > Survey Data**. Open it after National Inequality (p. 352), which compares places; this page compares groups of people, and tells you whether the gap between them is closing from one survey to the next. RMNCAH Vaxx

Before you start: load the area, education and wealth survey files in the Survey Files step of Load Data (p. 295). This page does not use the facility data or the denominators.

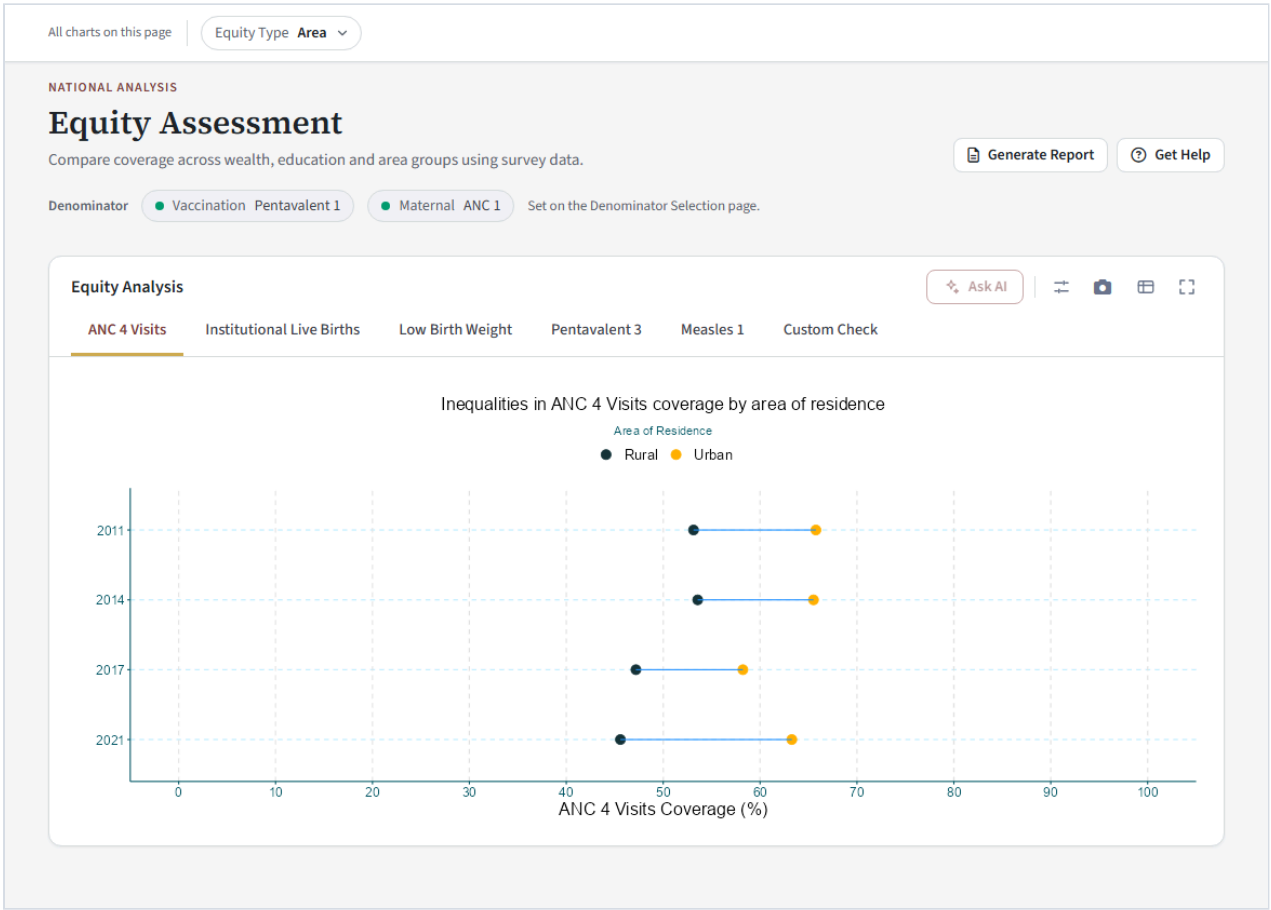


Figure 26.9 — *The Equity Assessment page, Area, ANC 4 Visits*

Using the page

1. **Choose the groups**
   In the filter bar, pick the **Equity Type**: **Area**, **Maternal Education** or **Wealth Quintile**.

2. **Pick an indicator tab**
   Go through the tabs, or open **Custom Check** to choose another indicator.

③ Compare the gaps

Read the length of each line, and compare the latest survey with the earlier ones.

④ Add it to a report

Generate Report in the page header builds the national inequality report from these pages (see Reports (p. 415)). Then go on to Sub-National Analysis (p. 359).

Controls

Control	Choices	Default/notes
Equity Type (all charts on this page)	Area, Maternal Education, Wealth Quintile	Area
Indicator tabs	ANC 4 Visits, Institutional Live Births, Low Birth Weight, Pentavalent 3, Measles 1, Custom Check	ANC 4 Visits

Cards

Equity Analysis — an equiplot (p. 499) titled "Inequalities in … coverage by …": one row per survey year, one dot per group (Rural and Urban; No education, Primary, Secondary or higher; Q1 to Q5), joined by a line. The longer the line, the bigger the gap between the groups.

How to read it: a good result is a short line that gets shorter from one survey to the next, with every group moving right. A long line, or one that grows, shows a group left behind; look at which group sits furthest left (often rural, no education or the poorest quintile, Q1).

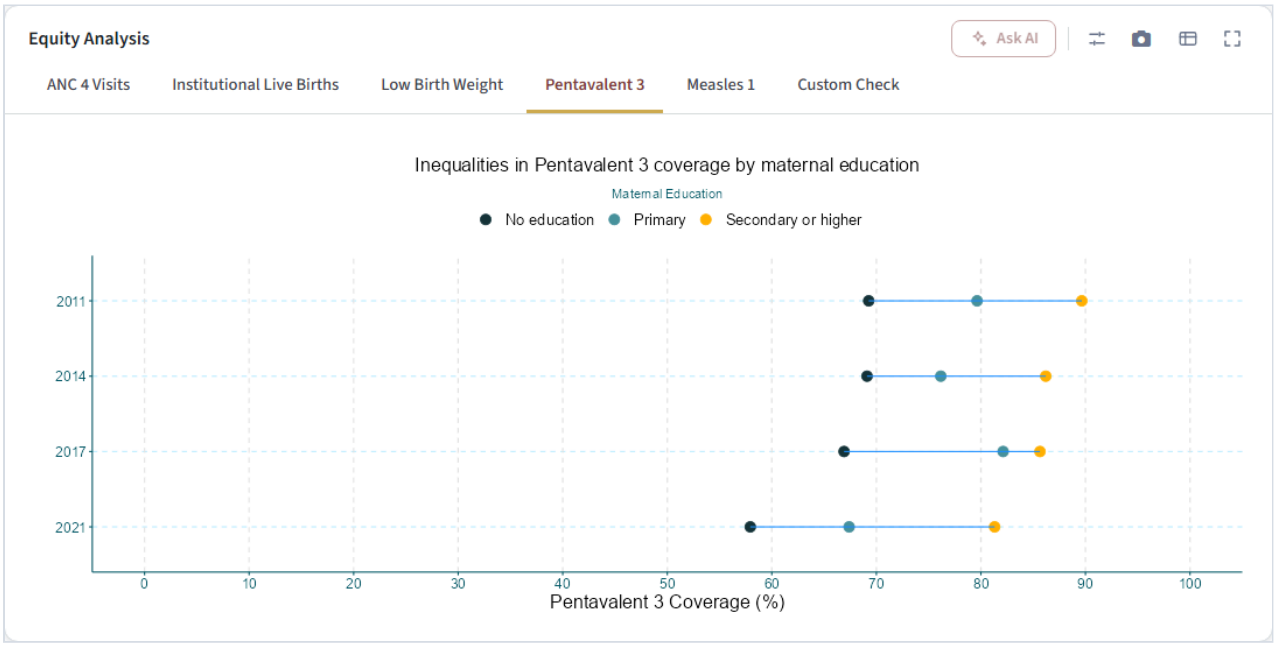


Figure 26.10 — *Maternal Education, Pentavalent 3*

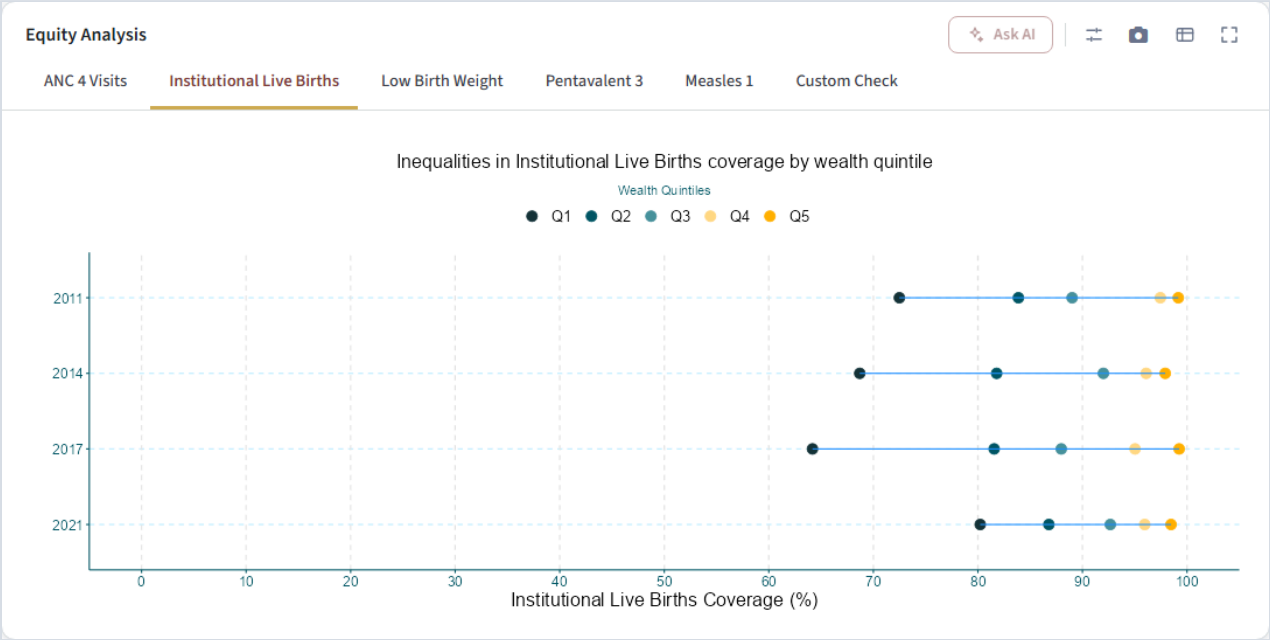


Figure 26.11 — *Wealth Quintile, Institutional Live Births*

Vaxx app

The tabs are **Pentavalent 3**, **Measles 1**, **Penta 1 to Penta 3 Dropout**, **Penta 3 to Measles 1 Dropout** and Custom Check. For a dropout, a dot further right means more children drop out, so the group furthest right is the one left behind.

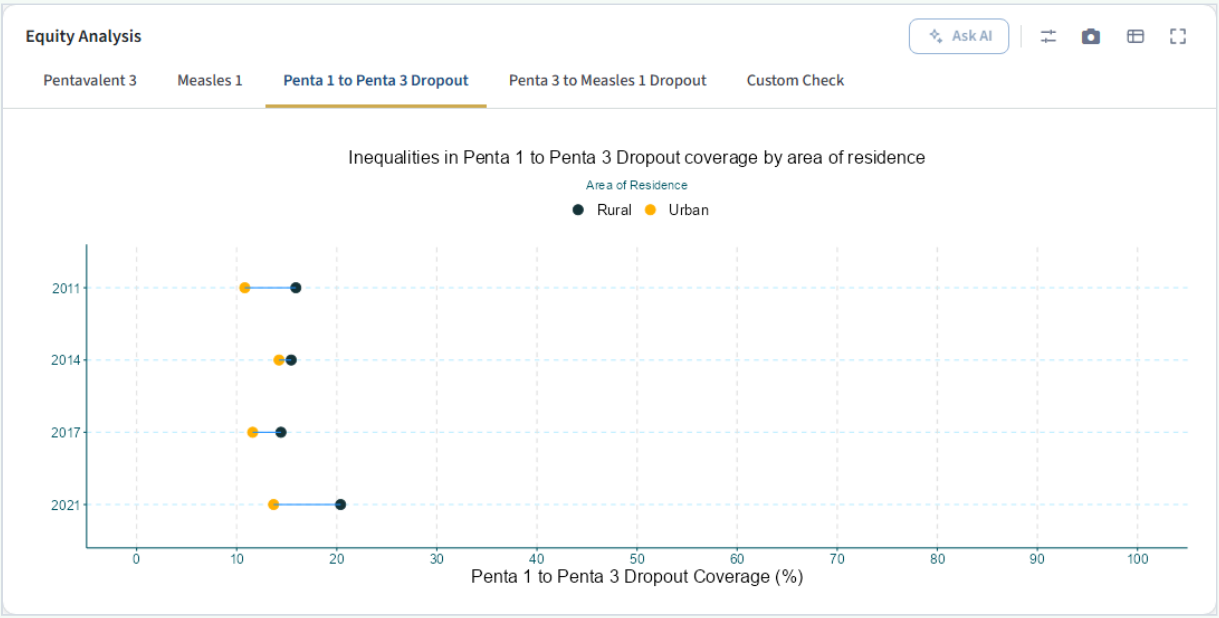


Figure 26.12 — *Vaxx app: Penta 1 to Penta 3 Dropout by area*

Common pitfalls

- **An error instead of the chart.** The survey file for this Equity Type is not loaded, or it has no value for this indicator. Load it in the Survey Files step of Load Data, then come back.

- **Reading one survey as a trend.** With a single survey row you see the gap, not whether it is closing.
- **Comparing with the facility charts.** These figures come from the surveys only; they do not change with the denominators or the data adjustment.

Method: Equity assessment (p. 153), one page per Equity Type: Area of residence (p. 163), Maternal education (p. 161), Wealth quintile (p. 157).

Next: Sub-National Analysis (p. 359).

CHAPTER 27

Sub-National Analysis

The Sub-National Analysis pages, Sub-National Coverage, Sub-National Inequality and Coverage Target, which repeat the national analysis for regions (Admin 1) and districts (Admin 2).

Find the areas behind the national figure

Repeat the national analysis for one region (Admin 1 (p. 496)) or one district (Admin 2 (p. 496)). Start with the regions that stood out on National Analysis (p. 342): programmes usually act at region and district level. RMNCAH Vaxx

Learn about sub-national analysis (p. 165)

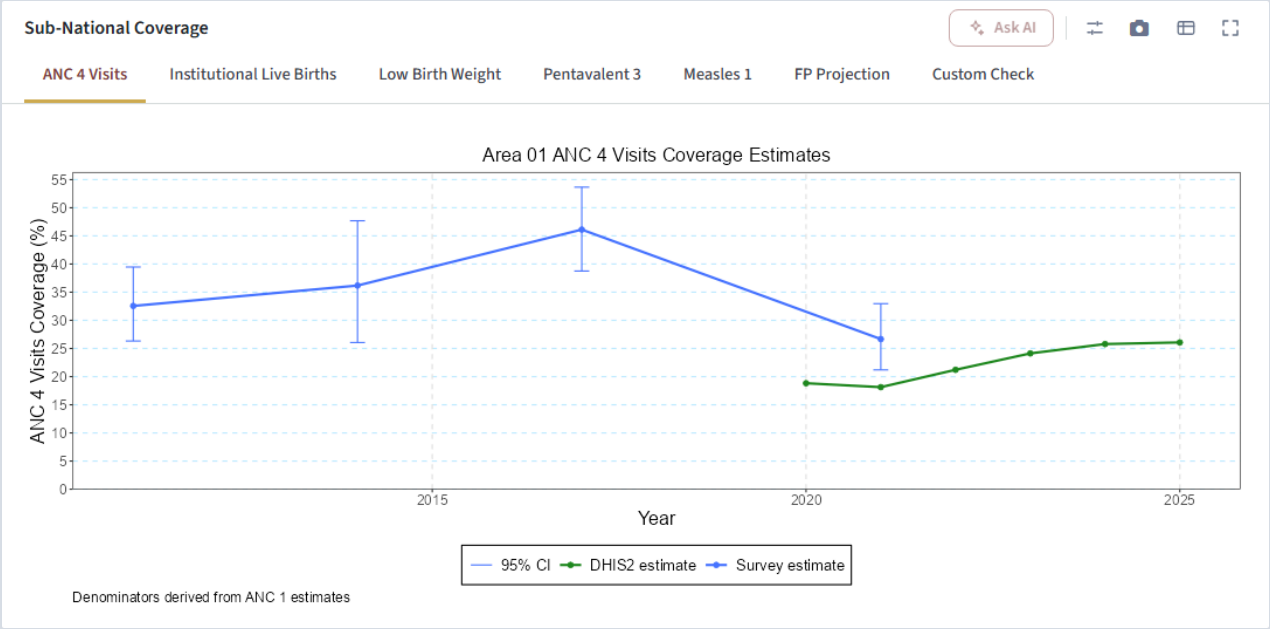


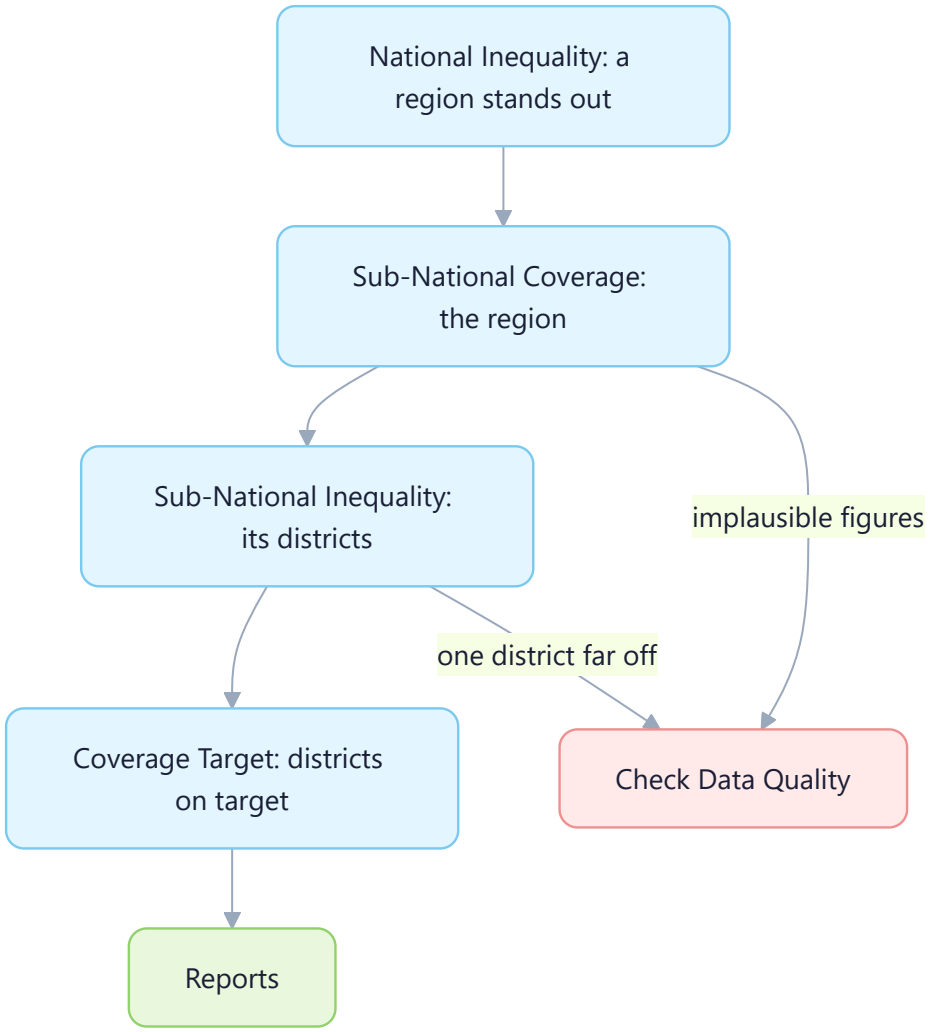
Figure 27.1 — *The Sub-National Coverage card: facility and survey coverage of four antenatal visits in one area*

Pages

Before you start: run Data Adjustment (p. 324) and choose the denominators on Denominator Selection (p. 338); load the survey files in Load Data (p. 295) to see the regional survey line.

Each page has a filter bar at the top, **All charts on this page**: its chips change every chart on the page. The **Denominator** strip under the title shows the denominators in use.

How the pages fit together



Start from a region, see whether its facility figures agree with its survey, then look at its districts: how far apart they are, and how many reach the targets. A figure that looks implausible sends you back to the data quality pages, not forward.

Common pitfalls

- **Judging a region without its reporting rate.** Low coverage can be low reporting. Check the area on Reporting Rate (p. 309) first.
- **Reading the survey line as a district figure.** At Admin 2 it is the survey estimate of the region.
- **An implausible district** (far above 100%, or a sudden jump): check it on the Data quality (p. 306) pages; to test another denominator, go back to Denominator Selection (p. 338).
- **An empty Coverage Target chart or table.** No district of that region reaches the target, or the table's Indicator chip is on an indicator without a target. Pick another region or indicator.

Pitfalls in reading the results: Subnational analysis > Common pitfalls (p. 167).

Next: Advanced analysis (p. 371) (RMNCAH) or Reports (p. 415).

27.1 Sub-National Coverage

How does facility coverage in one region or district compare with the survey estimates for that region?

Sub-National Coverage draws the National Coverage chart for one region or one district: the coverage from the facility data next to the survey estimate of the region. Open it after the National Analysis pages, for the areas that stood out there (a region far from the national figure, or below the target). It tells you whether an area's facility figures agree with its survey, and how its coverage has moved over the years. RMNCAH Vaxx

Before you start: run Data Adjustment (p. 324), choose the denominators on Denominator Selection (p. 338), and load the regional survey file in Load Data (p. 295) to see the survey line.

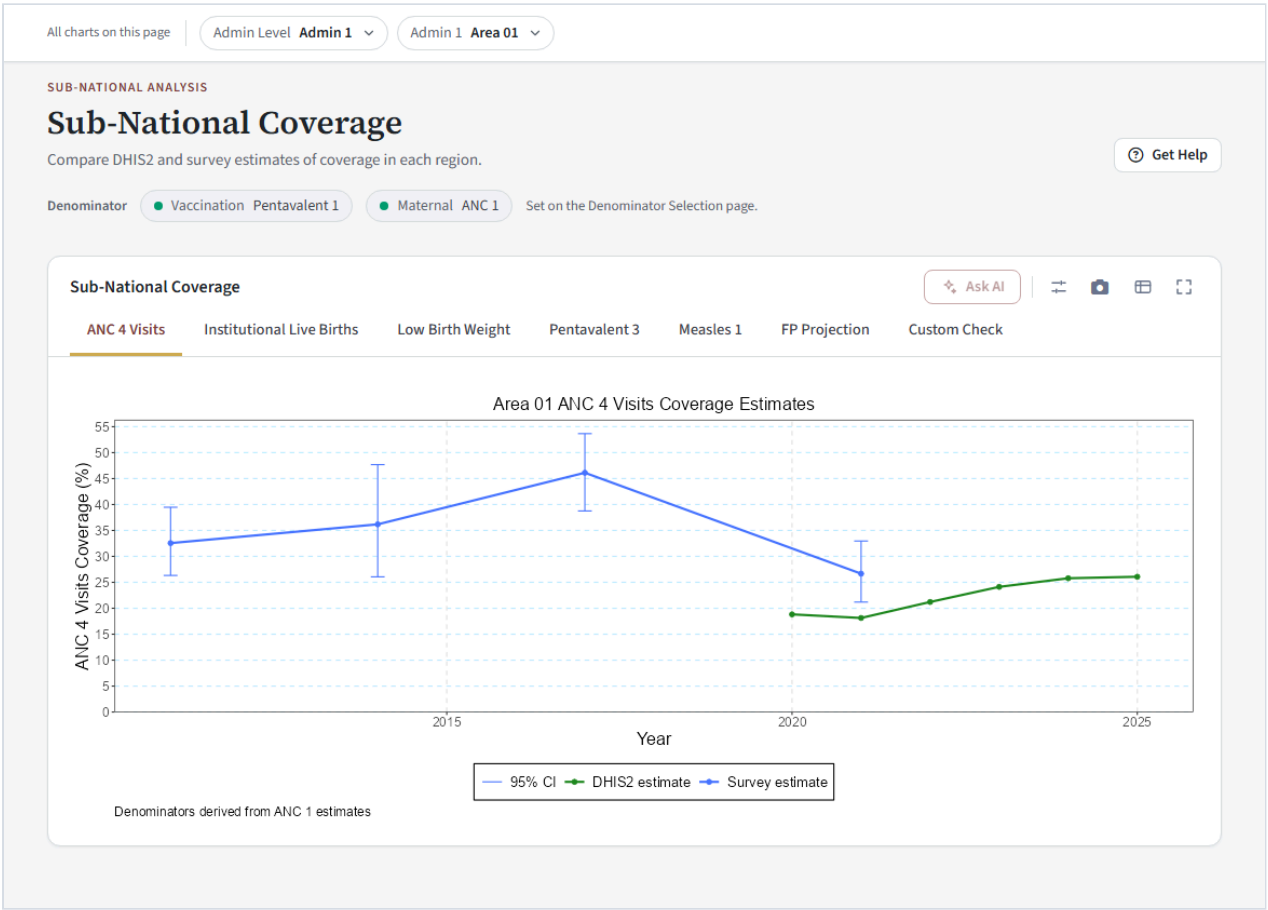


Figure 27.2 — *Sub-National Coverage for Area 01, ANC 4 Visits*

Using the page

1. **Choose the area**
   In the filter bar, pick **Admin 1** for a region or **Admin 2** for a district, then choose the area in the chip next to it.

2. **Pick an indicator tab**
   Go through the tabs, or open **Custom Check** to choose another indicator.

3 **Compare the lines**

Compare the green DHIS2 line with the blue survey points of the region and their error bars.

4 **Move on**

Use the card toolbar to ask AI or download the chart. Then go on to Sub-National Inequality (p. 365).

Controls

Control	Choices	Default/notes
Admin Level	Admin 1, Admin 2	Admin 1
Admin 1 (at Admin 1)	Each region	The first region
Admin 2 (at Admin 2)	Each district	The first district; replaces the Admin 1 chip
Indicator tabs	ANC 4 Visits, Institutional Live Births, Low Birth Weight, Pentavalent 3, Measles 1, FP Projection, Custom Check	ANC 4 Visits; Custom Check adds an **Indicator** chip

Cards

Sub-National Coverage — "... Coverage Estimates" for the chosen area: the green **DHIS2 estimate** from the adjusted data, and the blue **Survey estimate** of the region with its **95% CI**. The caption names the denominator. The FP Projection tab shows the national projection (see FP Projection (p. 379)).

How to read it: as on the national page, a good result is a green line within or near the survey error bars. A green line far above 100%, or jumps from one year to the next, usually come from the area's data or denominator rather than real changes in coverage.

At Admin 2 the chart shows one district. Surveys are not representative of districts, so the survey line is the one of the district's region.

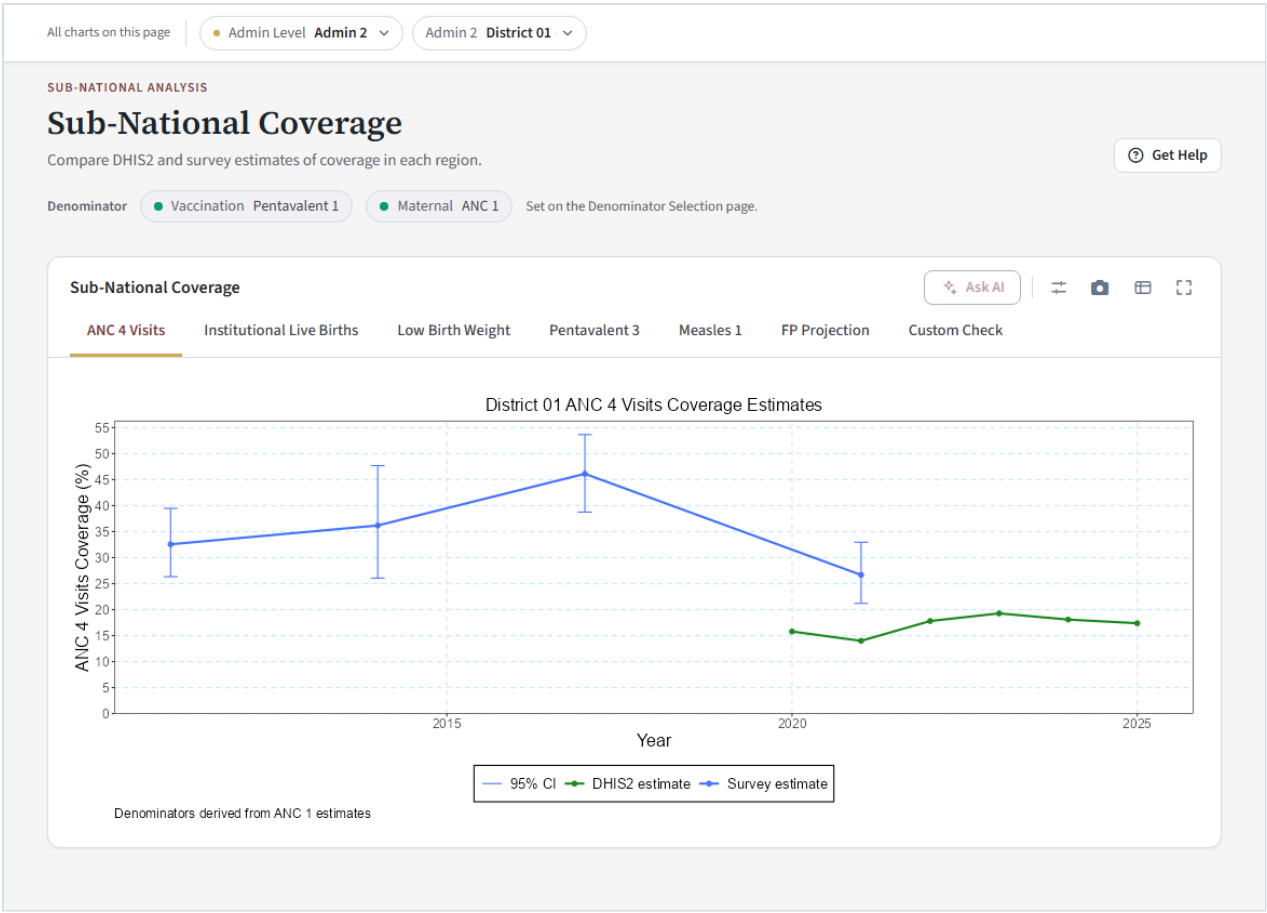


Figure 27.3 — *Admin 2: one district*

> **Vaxx app**
> The tabs and the Custom Check list are those of National Coverage (p. 345) in the Vaxx app (no FP Projection).

Common pitfalls

- **Reading the survey line as a district figure.** At Admin 2 it is the survey estimate of the whole region.
- **No survey line.** The regional survey file is not loaded, or its region names do not match the dataset. Check the Survey Files and Map Survey steps of Load Data.
- **Judging an area without its reporting rate.** Low coverage can be low reporting: check the area on Reporting Rate (p. 309).

Method: Subnational coverage (p. 168) (why district figures are noisier: What (p. 168); reading the chart: Interpretation (p. 170)).

Next: Sub-National Inequality (p. 365).

27.2 Sub-National Inequality

Inside one region, how far is each district from the region's coverage? At Admin 2, how far is each district from the national coverage?

Sub-National Inequality shows how far each district is from the coverage of its region or, at Admin 2, from the national coverage. Open it after Sub-National Coverage (p. 362), for a region that looked low or uneven on National Inequality (p. 352). It tells you whether the districts of the region move together or whether some are being left behind. RMNCAH Vaxx

Before you start: run Data Adjustment (p. 324) and choose the denominators on Denominator Selection (p. 338).

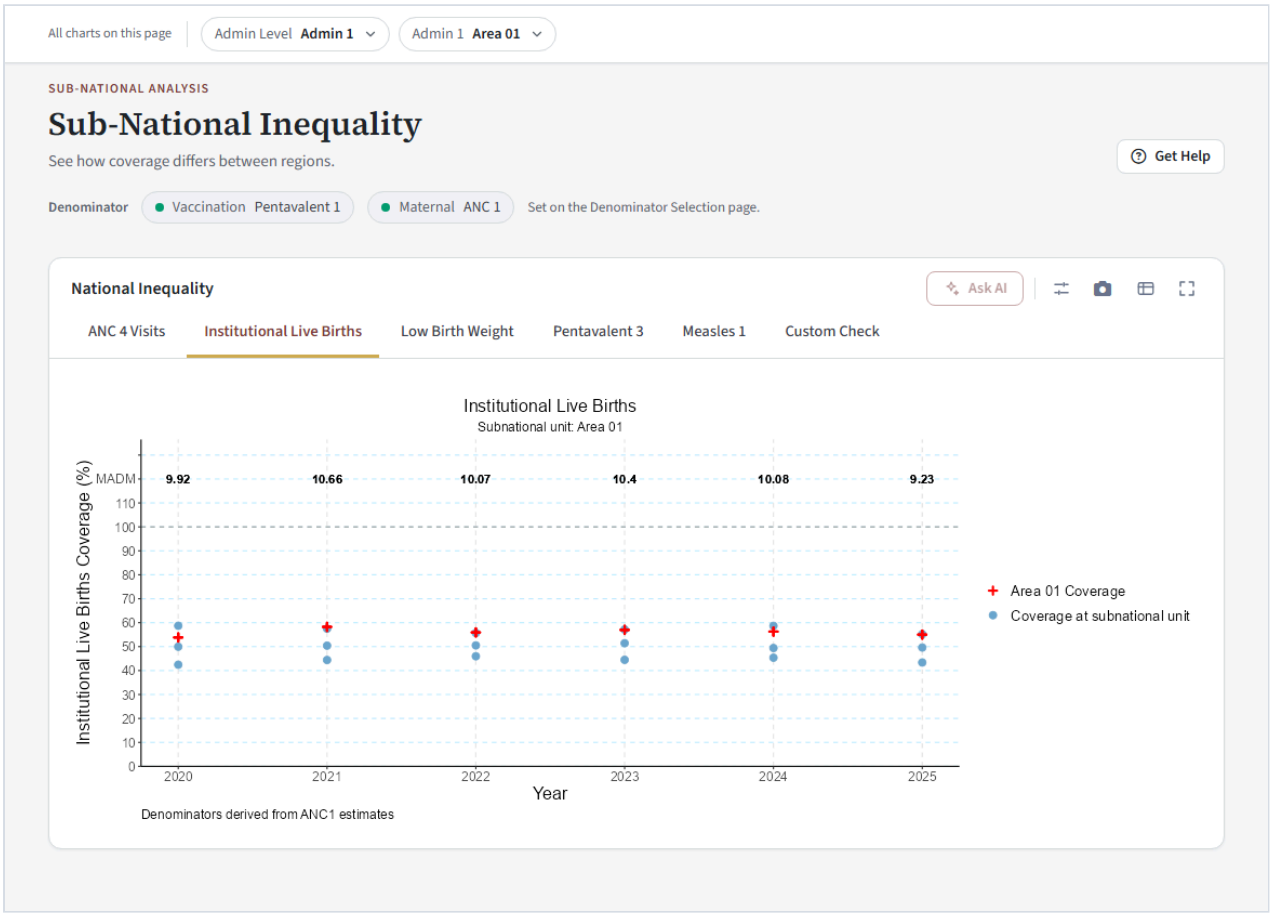


Figure 27.4 — *Sub-National Inequality for the districts of Area 01, Institutional Live Births*

Using the page

1. **Choose the level and region**

   In the filter bar, keep **Admin 1** and choose the region in the **Admin 1** chip to see its districts, or pick **Admin 2** to see every district of the country.

2. **Pick an indicator tab**

   Go through the tabs, or open **Custom Check** to choose another indicator.

3 **Read the spread**

See how far the dots spread around the red cross, and how the MADM (p. 501) row changes over the years.

4 **Move on**

Note the districts far below the others, then go on to Coverage Target (p. 368).

Controls

Control	Choices	Default/notes
Admin Level	Admin 1, Admin 2	Admin 1
Admin 1 (at Admin 1)	Each region	The first region; the chart shows its districts
Indicator tabs	ANC 4 Visits, Institutional Live Births, Low Birth Weight, Pentavalent 3, Measles 1, Custom Check	ANC 4 Visits

Cards

National Inequality (the card keeps this title) — each blue dot is one district in a year ("Coverage at subnational unit"); the red cross is the coverage of the chosen region ("… Coverage"). The subtitle names the units ("Subnational unit: Area 01"), and the top row gives the MADM for each year. At **Admin 2** the region chip goes away: the chart shows every district of the country against the national coverage ("Subnational unit: district level").

How to read it: a good result is dots gathered close to the red cross and an MADM that falls over the years. A wide spread, or a growing MADM, means some districts fall behind; a single dot far from the rest may be a data problem rather than a real gap.

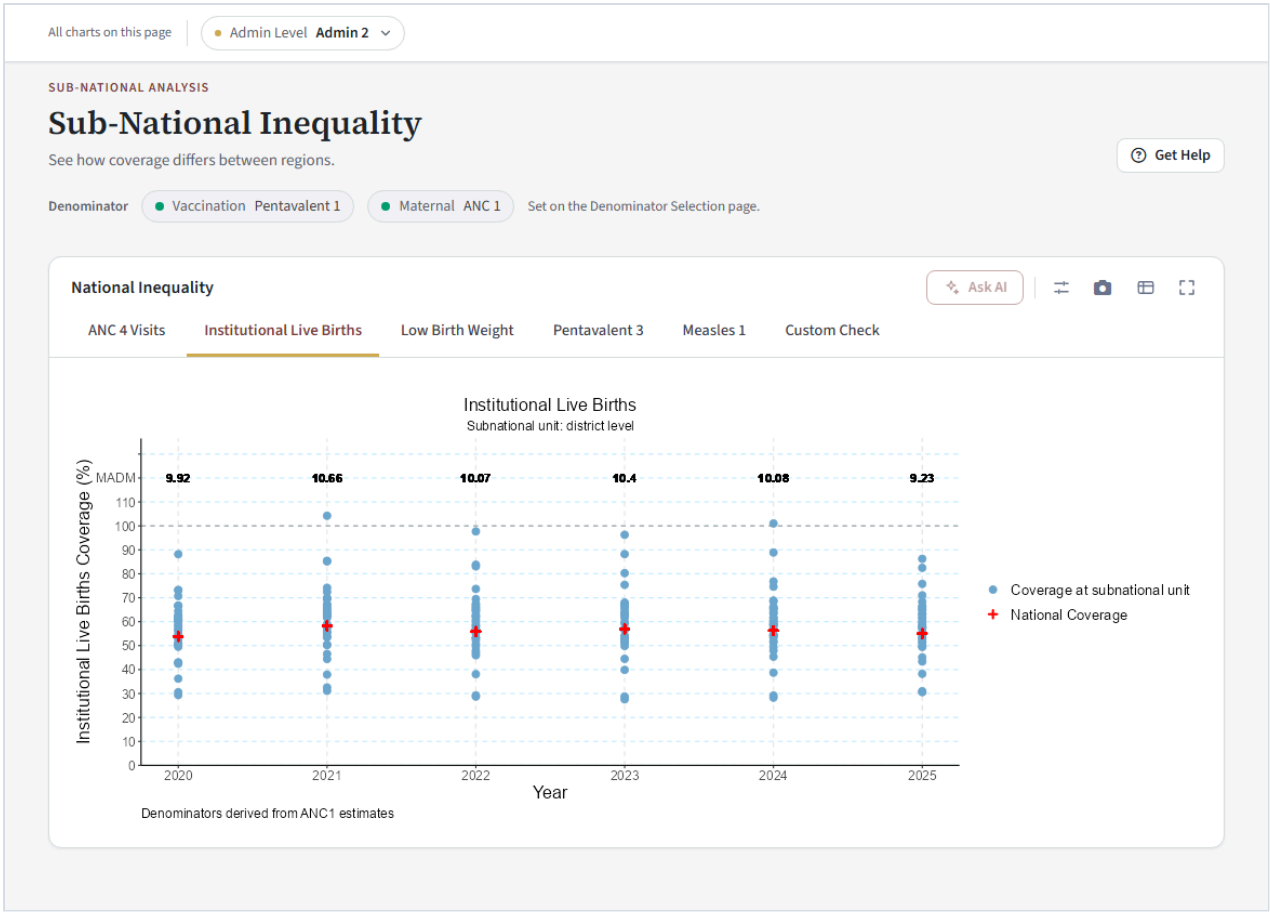


Figure 27.5 — *Admin 2: all districts against the national coverage*

Common pitfalls

- **Expecting district names on the chart.** The dots are not labelled. To see which districts are low, use Coverage Target (p. 368) or open the chart's data from the card toolbar.
- **A dot far above 100%.** That district's denominator or data are off. Check it on the Data quality (p. 306) pages before you read the spread.
- **Comparing Admin 1 and Admin 2 MADM.** At Admin 1 the districts are compared with their region, at Admin 2 with the country: the figures are not the same measure.

Method: Subnational inequality (p. 172) (the MADM: Formula (p. 172); reading the chart: Interpretation (p. 173)).

Next: Coverage Target (p. 368).

27.3 Coverage Target

What share of the districts of one region reach the global coverage targets each year, and which districts are they?

This Coverage Target page counts the districts of one region that reach the global coverage targets each year, and lists them. Open it after Sub-National Inequality (p. 365), for a region you want to look at district by district. It tells you how much of the region is on target, and which districts are not yet there. RMNCAH Vaxx

Before you start: run Data Adjustment (p. 324) and choose the denominators on Denominator Selection (p. 338).

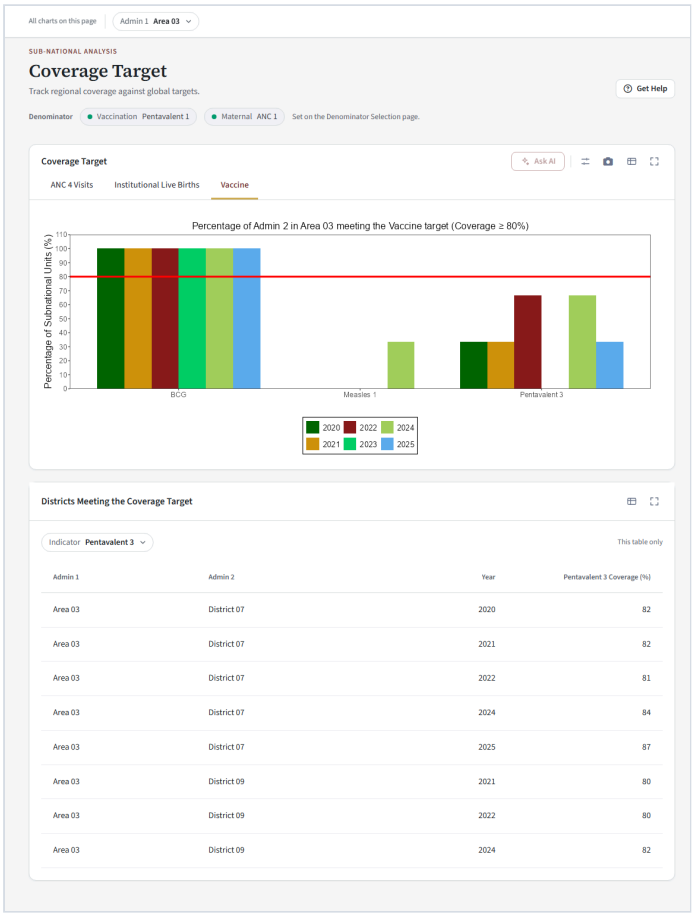


Figure 27.6 — *Coverage Target for the districts of Area 03, Vaccine tab*

Using the page

1. **Choose the region**
   In the filter bar, pick the region in the **Admin 1** chip. The page always counts its districts.

2. **Read the bars**
   Open each tab and compare the bars from year to year against the target in the chart title.

3. **List the districts on target**
   In the table, choose the indicator in its **Indicator** chip. Districts missing from the list are the ones below the target. Download the table from its toolbar if you need it.

4. **Move on**
   Go on to Advanced Analysis (p. 371) (RMNCAH) or build a report on Reports (p. 415).

Controls

Control	Choices	Default/notes
Admin 1	Each region	The first region; the units are always its districts
Indicator tabs	ANC 4 Visits, Institutional Live Births, Vaccine	ANC 4 Visits
Indicator (table only)	Any indicator in the dataset	ANC 1st Trimester; pick the indicator you want to list

Cards

Coverage Target — "Percentage of Admin 2 in … meeting the … target": one bar per year, with a red line across the chart. The targets are those of the national Coverage Target (p. 349) page.

How to read it: bars that grow over the years mean more districts of the region reach the target. With few districts in a region, one district more or less moves the bar a lot, so look at the trend rather than one year.

Districts Meeting the Coverage Target — the districts and years where the indicator chosen in the table's **Indicator** chip reaches the target, with its coverage; download it from the table toolbar.

How to read it: compare the list with the region's districts. A district listed in early years but not in the latest one has fallen below the target.

> Vaxx app
> The tabs are **Vaccine** (BCG, Measles 1, Pentavalent 3) and **Dropout** (Penta 1 to Penta 3, Penta 3 to Measles 1), as on the national page. For dropout, a district meets the target when its rate is below the threshold.

Common pitfalls

- **The table stays empty.** Its Indicator chip starts on an indicator without a target (ANC 1st Trimester, or ANC 1 in Vaxx). Choose the indicator of the tab you are reading.
- **An empty chart.** No district of that region reaches the target. Pick another tab, or compare with another region.
- **A district on target with coverage far above 100%.** Its denominator is probably too small: check it before you count it as a success.

Method: Coverage targets (p. 175) (the thresholds: Settings (p. 176); reading the percentages: Interpretation (p. 176)).

Next: Advanced Analysis (p. 371) (RMNCAH) or Reports (p. 415).

CHAPTER 28

Advanced Analysis

The RMNCAH app's Bayesian Analysis pages (national and sub-national), the Continuum of Care page and the FP Projection tab, which build on the adjusted data, the survey files and the selected denominators.

One best estimate per year, and where people drop out

A **Bayesian** model combines surveys and routine data into one estimate per year, with its uncertainty. The **Continuum of Care** follows coverage along the care pathway, and the **FP Projection** shows the family planning trend. RMNCAH

Learn about the Bayesian approach (p. 143)

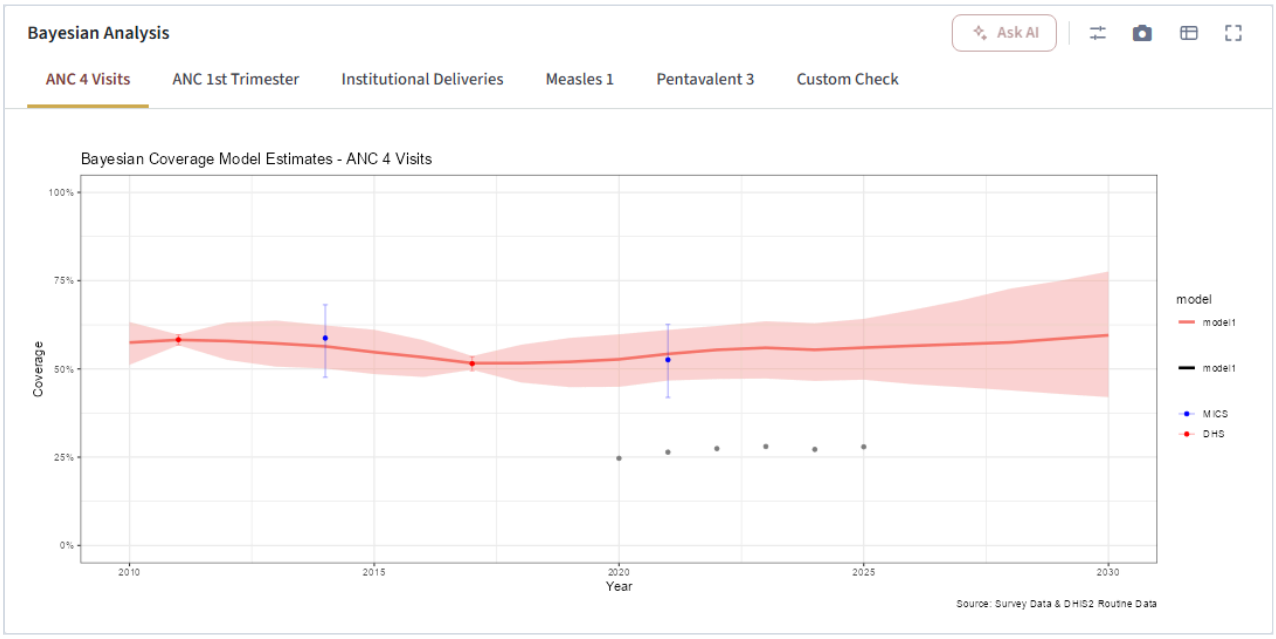
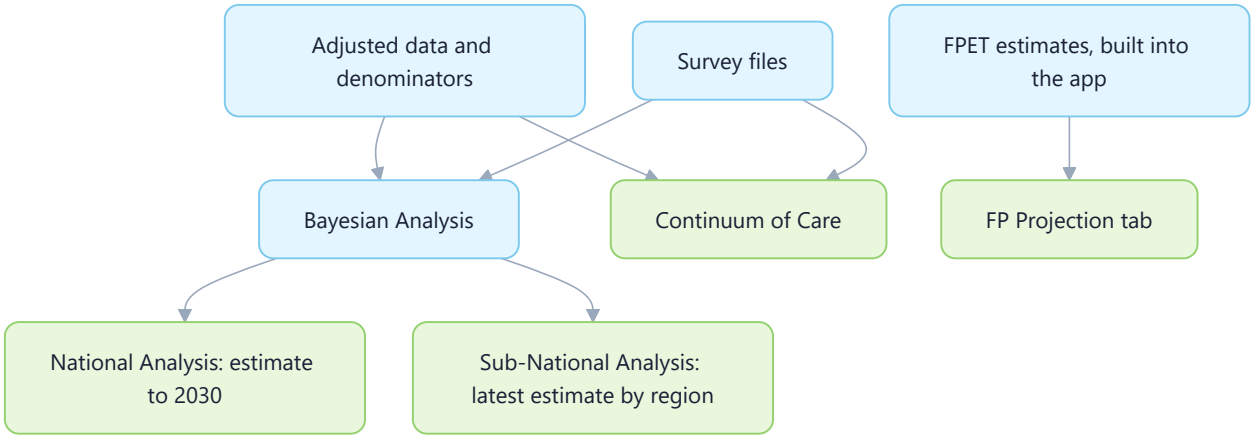


Figure 28.1 — *The Bayesian Analysis card: modelled coverage with its uncertainty band, survey points and routine data*

Pages

Before you start: run Data Adjustment (p. 324), choose the denominators on Denominator Selection (p. 338) and load the survey files in Load Data (p. 295): the Bayesian model and the survey bars need them.

What each page uses



The Bayesian model and the continuum profile need both the facility data and the surveys; the continuum heatmap uses the facility data only. The FP Projection uses estimates that come with the app.

Common pitfalls

- **The Bayesian page seems stuck on "Calculating...".** The first time a page is shown, the app fits a model for each tab in the background, which can take several minutes, and the app can be slow meanwhile. Wait; the models are then kept in the dataset and show at once next time.
- **A message about extra packages.** The Bayesian model needs extra packages: "The Bayesian model needs extra packages (a large download; a few minutes to install). They are installed only when you use the model." In DataSuite, click **Install the Bayesian model** and restart the app when DataSuite offers it; outside DataSuite, run the command shown.
- **An error in a Bayesian tab.** The model needs survey values for that indicator; without them (for example no institutional delivery estimate in the survey file) the tab cannot be drawn.
- **No survey bars in the continuum profile.** Load the survey files in Load Data.
- **Results changed.** These pages build on the denominators set on Denominator Selection; a new denominator means a new Bayesian model.

Pitfalls in reading the results: Bayesian modeling approach (p. 144) and National coverage (p. 117).

Next: Mortality (p. 381).

28.1 Bayesian Analysis

What is the most likely coverage each year once survey and routine data are combined, how sure is it, and where is it heading?

Bayesian Analysis combines the surveys and the routine facility data into one coverage estimate per year, with a range that says how sure it is and a short forecast to 2030. Open it after the National Analysis (p. 342) pages, when the facility line and the survey points disagree or the surveys are years apart. It gives you the single best figure to report, and its uncertainty. The sidebar group **Bayesian Analysis** has two pages, **National Analysis** and **Sub-National Analysis**; both are titled Bayesian Analysis. RMNCAH

Before you start: run Data Adjustment (p. 324), choose the denominators on Denominator Selection (p. 338) and load the survey files in Load Data (p. 295).

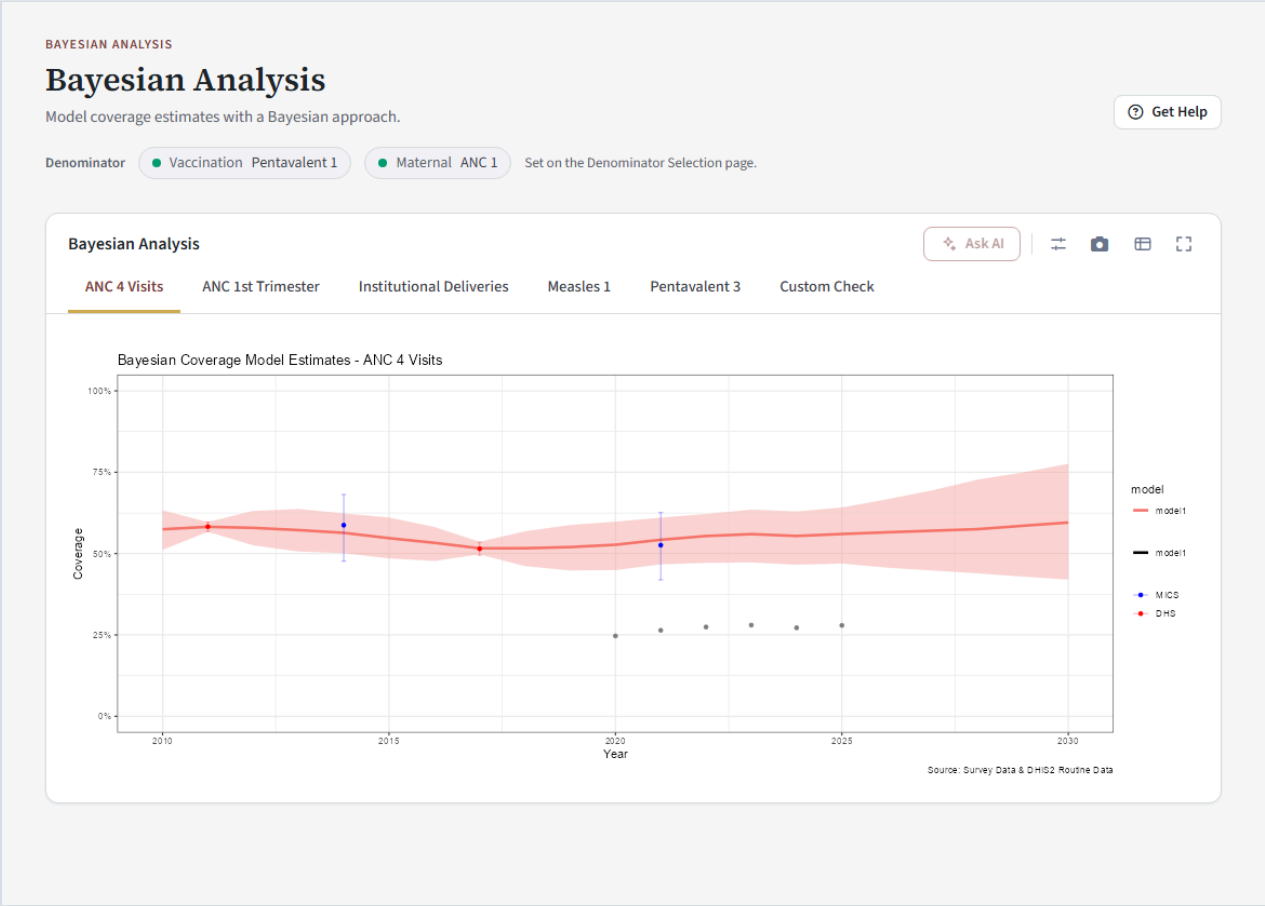


Figure 28.2 — *Bayesian Analysis, National Analysis, ANC 4 Visits*

Using the page

1. **Open the page and wait**

   The first time, the app fits a Bayesian model (p. 496) for each tab in the background. This can take several minutes; the models are then kept in the dataset.

2. **Pick an indicator tab**
   Choose one of the five indicators the model covers.

3. **Read the estimate and its band**
   Follow the red line year by year, and see how wide the shaded band is and where the survey and routine points sit.

4. **Compare the regions**
   Open **Sub-National Analysis** in the sidebar group to see the latest estimate for each region. Then go on to Continuum of Care (p. 376).

Controls

Control	Choices	Default/notes
Indicator tabs	ANC 4 Visits, ANC 1st Trimester, Institutional Deliveries, Measles 1, Pentavalent 3	ANC 4 Visits; the model covers these five indicators

Cards

Bayesian Analysis — on **National Analysis**, "Bayesian Coverage Model Estimates - ...": the red line is the model's best estimate of coverage each year, the shaded band its uncertainty range (p. 504), to 2030. Coloured dots are the surveys (DHS, MICS) with their confidence intervals (p. 497); grey points are the coverage from the routine data. The source note reads "Survey Data & DHIS2 Routine Data".

How to read it: a narrow band means the sources agree and the estimate is firm; a band that widens after the last survey shows the forecast is less sure. Grey points far from the red line mean the routine data tell a different story from the surveys.

On **Sub-National Analysis** the chart shows the latest estimate for each Admin 1 (p. 496) unit, sorted, with its uncertainty range.

How to read it: regions at the bottom have the lowest estimated coverage. Two regions whose ranges overlap may not really differ.

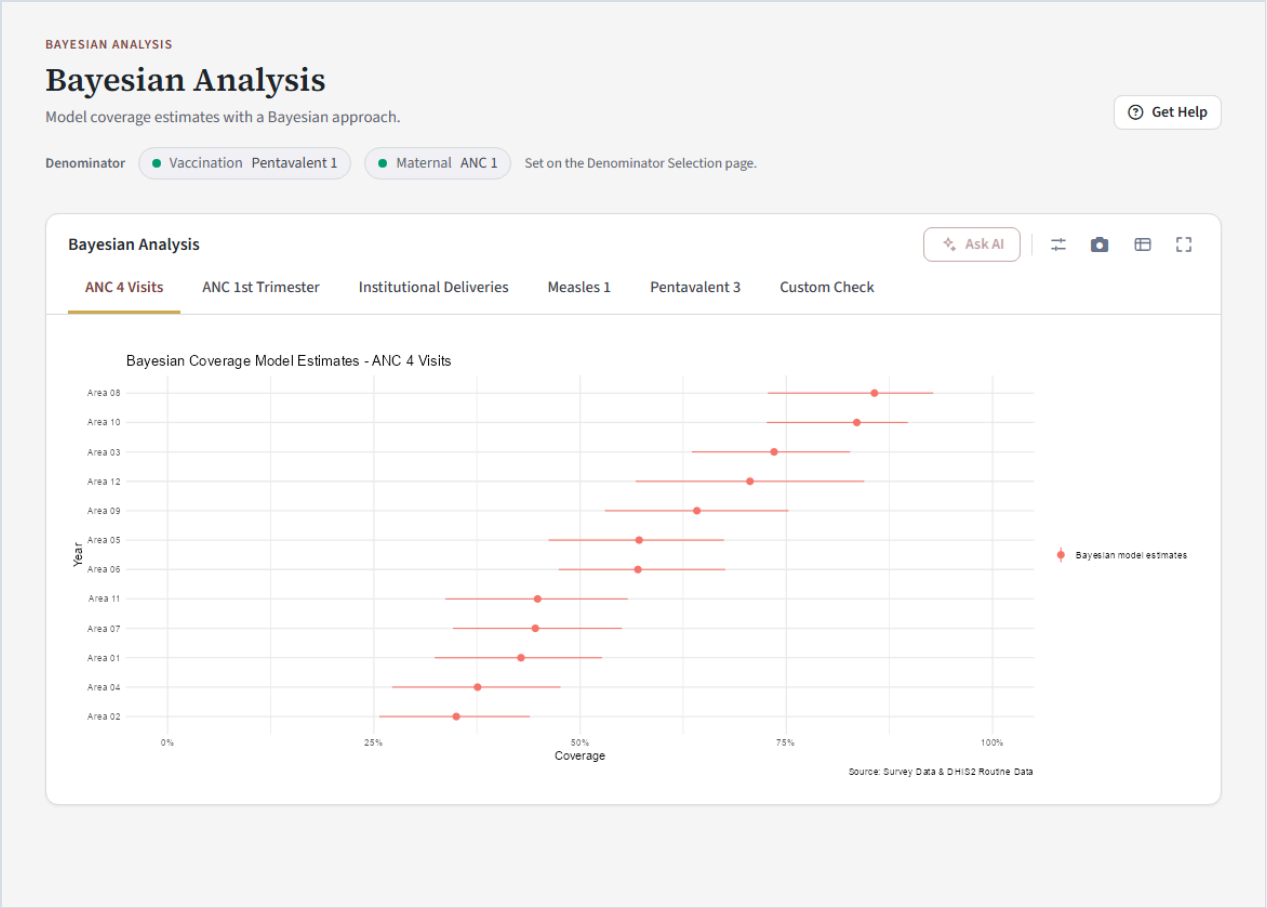


Figure 28.3 — *Bayesian Analysis, Sub-National Analysis, ANC 4 Visits*

Common pitfalls

- **"Calculating..." for a long time.** The models are fitted the first time each tab is shown. Wait; the next time they show at once.
- **An error in one tab.** The model needs survey values for that indicator. Without them the tab cannot be drawn: check the survey files.
- **Reading the forecast as a target.** The line after the last data year is a projection of the past trend, with a wide band.

Method: how the model uses each source, Survey data (p. 148) and Routine data (p. 150); reading the estimates, Estimates and short-term forecasts (p. 146).

Next: Continuum of Care (p. 376).

28.2 Continuum of Care

Where along the care pathway do women and children drop out, nationally and in each region?

Continuum of Care follows coverage along the care pathway (p. 497): from the first antenatal visit to the postnatal check for mothers, from BCG to Measles 1 for children. It answers where along the pathway women and children drop out, nationally and in each region. The page sits in the **National Analysis** group of the sidebar; open it after National Coverage (p. 345), once each indicator has been checked on its own. RMNCAH

Before you start: run Data Adjustment (p. 324) and choose the denominators on Denominator Selection (p. 338); load the survey files in Load Data (p. 295) for the survey bars.

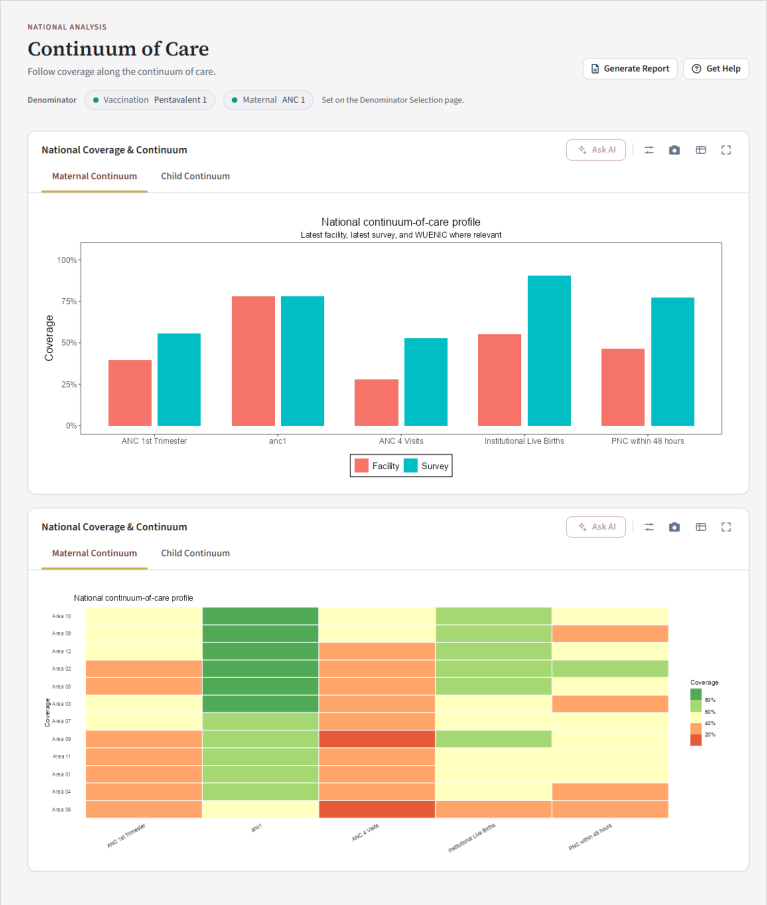


Figure 28.4 — *The Continuum of Care page, Maternal Continuum*

Using the page

① **Choose the pathway**

On each card, pick **Maternal Continuum** or **Child Continuum**. The two cards have their own tabs.

② **Find the biggest drop**

On the profile, look for the step where the bars fall the most from the step before.

3. **Find the regions behind it**
   On the heatmap, read down the column of that step to see which regions are lowest.

4. **Add it to a report**
   Generate Report in the page header builds the national coverage report (see Reports (p. 415)).

Controls

Control	Choices	Default/notes
Tabs (each card)	Maternal Continuum, Child Continuum	Maternal Continuum

Cards

Both cards are titled **National Coverage & Continuum** and have their own tabs.

Profile (top card) — "National continuum-of-care profile": for each step of the pathway, bars for the latest facility value, the latest survey value and, for vaccines, the WUENIC (p. 505) value. The maternal steps are ANC 1st Trimester, ANC 1, ANC 4 Visits, Institutional Live Births and PNC within 48 hours; the child steps BCG, Penta 1, Pentavalent 3 and Measles 1.

How to read it: in a good continuum the bars stay high from one step to the next. A sharp fall between two steps (for example from ANC 1 to ANC 4 Visits) shows where women or children are lost; a large gap between the facility and survey bars of one step points to that indicator's data or denominator.



Figure 28.5 — *The profile card, Child Continuum*

Heatmap (bottom card) — the same steps (columns) for each Admin 1 unit (rows), from facility data: green is high coverage, yellow intermediate, orange and red low. Read across a row for one area, down a column for one step.

How to read it: a row that turns orange or red after the first steps is a region where people drop out; a column that is red in many rows is a step that is weak across the country.

Common pitfalls

- **No survey bars.** The survey files are not loaded. Load them in the Survey Files step of Load Data.
- **Reading the heatmap as survey data.** It uses the facility data only, so a region with poor reporting looks low.
- **Comparing the two cards on different tabs.** Each card keeps its own tab; set both to the same pathway.

Method: Continuum of care (p. 137) (What (p. 137), Interpretation (p. 138)).

Next: FP Projection (p. 379).

28.3 FP Projection

How are modern contraceptive use and demand satisfied changing, according to the FPET estimates, from 2010 to 2030?

FP Projection shows how the use of modern contraception, and the share of demand it meets, have changed since 2010 and where they are heading to 2030, from the FPET (p. 499) estimates. It is the **FP Projection** tab of National Coverage (p. 345) (and of Sub-National Coverage, where it shows the same national projection). Open it with the other national coverage charts to add family planning to the picture; the estimates come with the app, so there is nothing to load or set. RMNCAH

Before you start: nothing; the tab works as soon as a dataset is loaded.

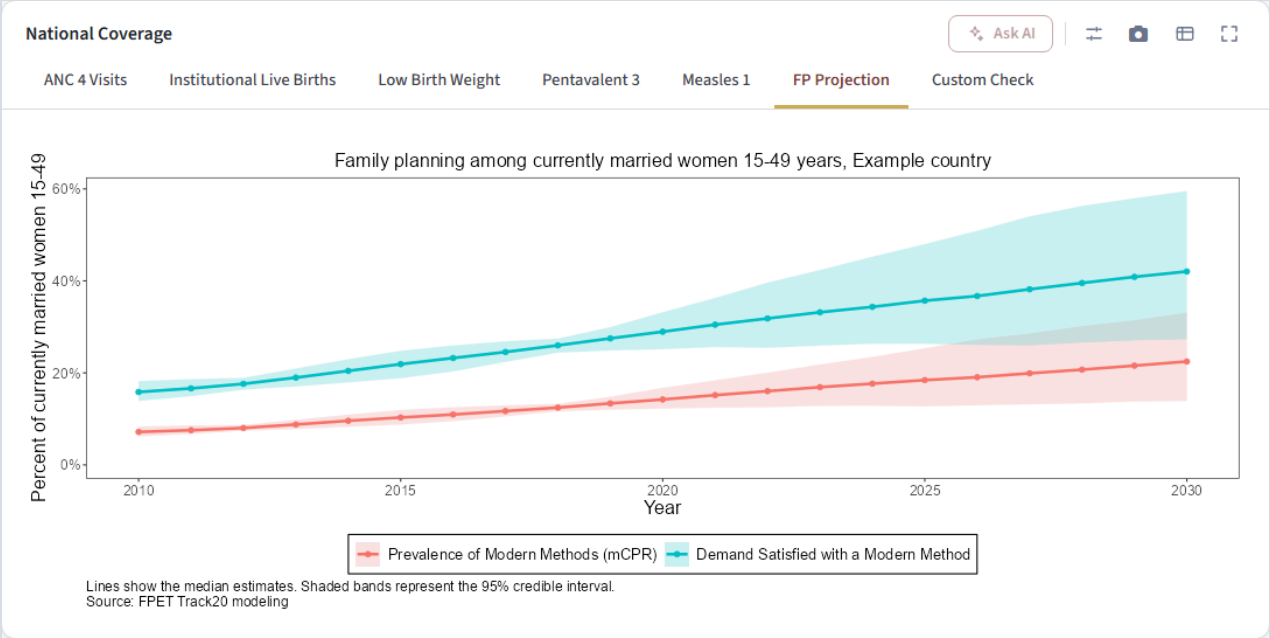


Figure 28.6 — *The FP Projection tab*

Using the page

1. **Open the tab**
   On **National Coverage**, pick the **FP Projection** tab.

2. **Read the two lines**
   Follow the two lines from 2010 to today, then the projection to 2030, and note how wide the bands become.

3. **Keep the chart**
   Use the card toolbar to ask AI or download the chart. Then go on to Mortality (p. 381).

Cards

National Coverage card, FP Projection tab — "Family planning among currently married women 15-49 years": the Prevalence of Modern Methods (mCPR) (p. 502) and the Demand Satisfied with a Modern Method (p. 498), from 2010 to 2030. Lines are the median estimates and the shaded bands the 95%

credible interval (source: FPET Track20 modelling).

How to read it: both lines rising is the progress you hope to see; demand satisfied rising faster than mCPR means more of the women who want to avoid a pregnancy use a modern method. Bands that widen in the later years show that the projection is less certain the further it goes.

Common pitfalls

- **Reading the projection as observed data.** The years after the latest survey are model projections, not measurements.
- **Expecting a regional chart.** On Sub-National Coverage the tab still shows the national projection.
- **Looking for it in the Vaxx app.** The tab is in the RMNCAH app only.

Method: Family planning (p. 132) (where the figures come from: What (p. 132); reading the chart: Interpretation (p. 133)).

Next: Mortality (p. 381).

CHAPTER 29

Mortality

The Institutional Mortality, Mortality Mapping and Mortality Completeness pages for maternal deaths, stillbirths and neonatal deaths recorded in facilities.

Deaths recorded in health facilities

These pages look at maternal deaths (p. 501), stillbirths (p. 504) and neonatal deaths before discharge (p. 498). Can the facility death data be used, and how many deaths do facilities miss? RMNCAH

Learn about the mortality analysis (p. 179)

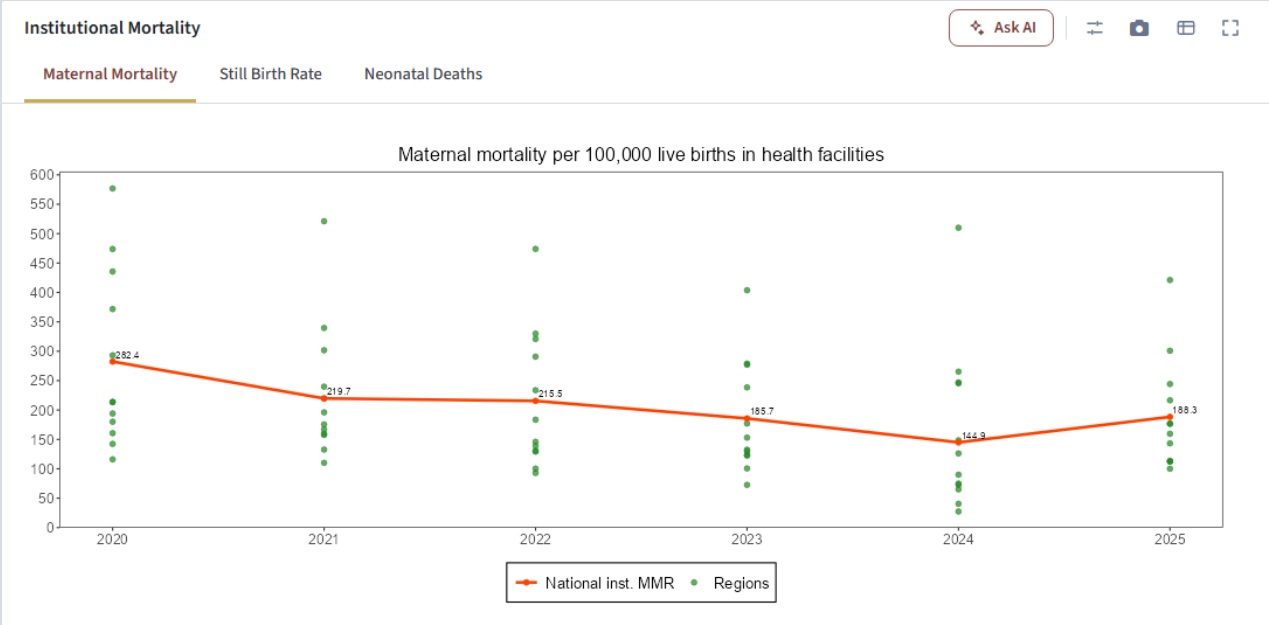


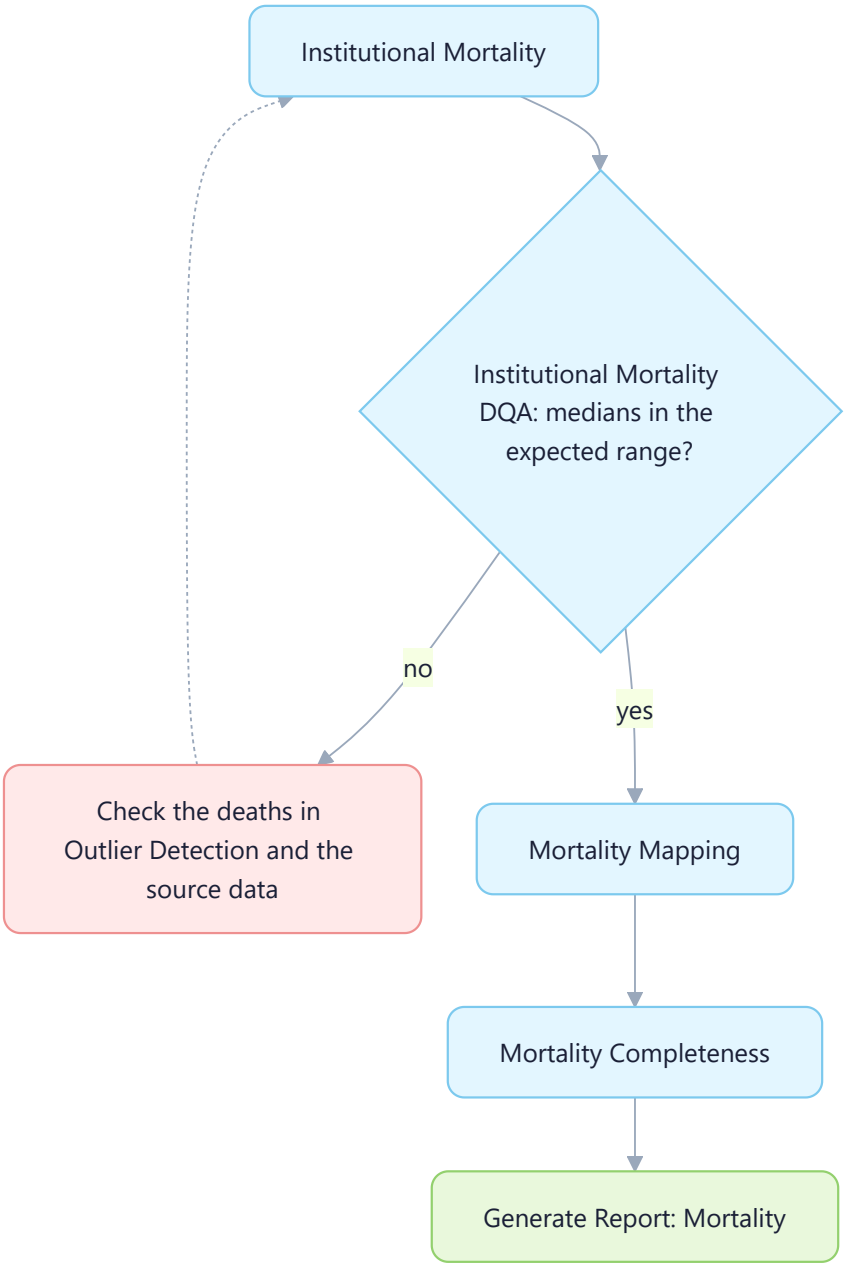
Figure 29.1 — *The Institutional Mortality card: maternal mortality in facilities, nationally and one dot per region, by year*

Pages

Before you start: load the data, run Data Adjustment (p. 324) and choose the denominators on Denominator Selection (p. 338). The pages use the reported deaths: Data Adjustment does not change them.

How the pages fit together

Read the three pages in order. The plausibility card on the first page decides whether the rest is worth reading: implausible figures are fixed in the source data first.



Common pitfalls

- Expecting Data Adjustment to correct the death counts. It does not: look up a suspicious value in Outlier Detection (p. 312) and correct it in the source data.
- Quoting a rate from the **Institutional Mortality** card without reading the **Institutional Mortality DQA** card.
- Reading a low mortality rate as good news. In facility data, a low rate more often means deaths are not recorded.
- Changing the denominators after reading Mortality Completeness: the completeness changes with them.

- Reading Mortality Completeness against out-of-date UN estimates. Load newer ones with **Upload UN Mortality estimates** in the Upload Data (p. 298) step if you have them.

Pitfalls of the method: Mortality > Common pitfalls (p. 181).

Next: Service utilization and health system (p. 391).

29.1 Institutional Mortality

How high is mortality in facilities, how does it change over the years, and are the figures plausible?

This page shows how many mothers, stillborn babies and newborns die in health facilities, year by year and region by region, and whether those numbers are believable. It is the first page of the Mortality group. Read it before you quote any facility mortality rate: if the plausibility checks fail, deaths are probably missing, and you check the data before going on to Mortality Mapping (p. 387). RMNCAH

Before you start: run Data Adjustment (p. 324) and choose the denominators on Denominator Selection (p. 338).

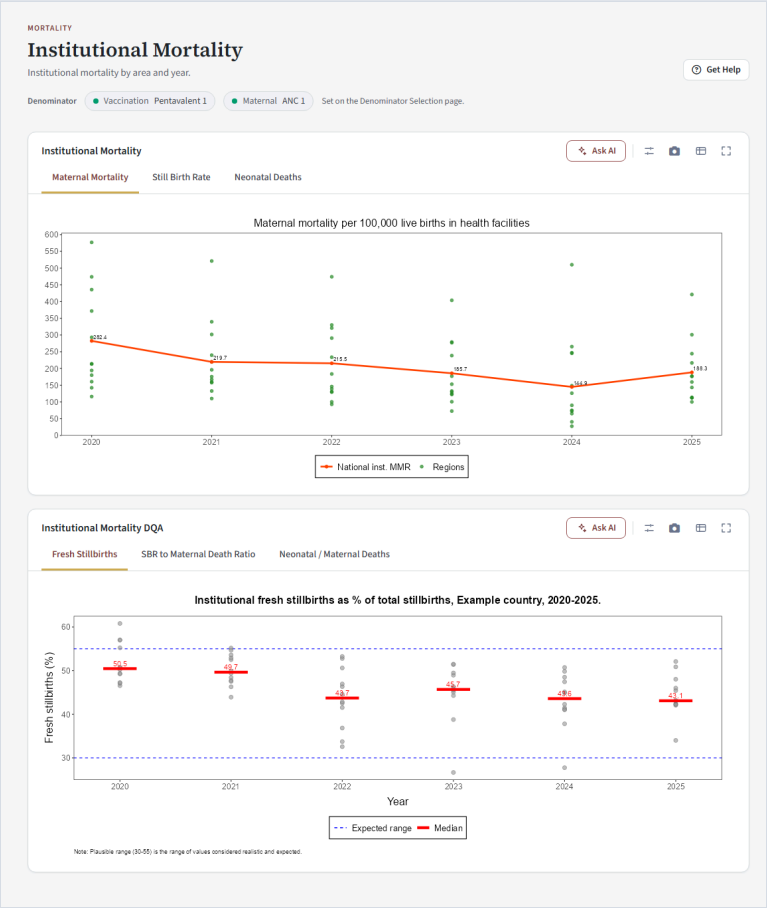


Figure 29.2 — *The Institutional Mortality page: rates by year and the plausibility card*

Using the page

① Look at the levels

In the **Institutional Mortality** card, go through the tabs **Maternal Mortality**, **Still Birth Rate** and **Neonatal Deaths**. Note the national line and the regions far from the others.

② Check that the figures are plausible

In the **Institutional Mortality DQA** card, go through **Fresh Stillbirths**, **SBR to Maternal Death Ratio** and **Neonatal / Maternal Deaths**. Note the years where the red median bar is outside the dashed lines.

③ **Decide what to do**

If the medians sit in the expected range, go on to Mortality Mapping (p. 387). If not, look up the regions and years in Outlier Detection (p. 312) and in the source data. **Ask AI** on a card explains the chart; **Download Data** saves the figures.

Controls

This page has no filters. The **Denominator** row under the title shows the denominators chosen on Denominator Selection; it is the same on every analysis page.

Cards

Institutional Mortality: the level of institutional mortality (p. 500), by year. Each dot is a region and the line is the country. Tabs:

- **Maternal Mortality**: maternal deaths per 100,000 live births in facilities.
- **Still Birth Rate**: stillbirths per 1,000 births in facilities.
- **Neonatal Deaths**: neonatal deaths before discharge (p. 498) per 1,000 live births in facilities.

How to read it: regions close to each other and a steady national line are what you hope to see. A region far below the others, or a sudden drop in one year, usually means deaths were not recorded, not that care improved. A single dot far above the others may be a data entry error.

Institutional Mortality DQA: checks that the death figures are plausible. Each chart shows the regions as grey dots, their **Median** as a red bar and the **Expected range** as two dashed lines; a note under the chart gives the range. Tabs:

- **Fresh Stillbirths**: fresh stillbirths (p. 499) as a percentage of all stillbirths.
- **SBR to Maternal Death Ratio**: stillbirths per maternal death.
- **Neonatal / Maternal Deaths**: neonatal deaths per maternal death.

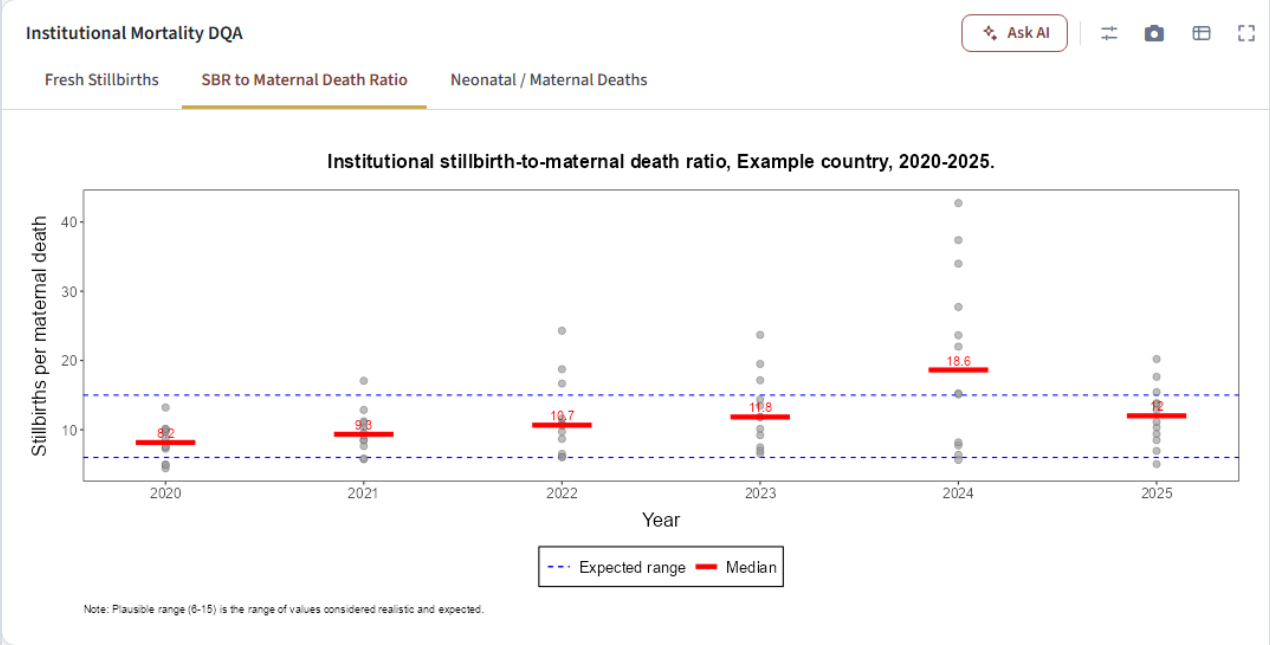


Figure 29.3 — *The SBR to Maternal Death Ratio tab: regions, median and expected range by year*

How to read it: a median between the dashed lines every year is a good sign. A median above the range on the two ratio tabs usually means maternal deaths are missing; a median outside the range on **Fresh Stillbirths** suggests fresh and macerated stillbirths are mixed up. Regions outside the range in every year are the first to check.

Common pitfalls

- Quoting a rate from the first card before reading the DQA card.
- Taking the regions with the lowest rates as the best performers. In facility data, a very low rate more often means missing deaths.
- Reading a change in a region with very few births as a trend: small numbers swing widely.

Method: Institutional maternal mortality ratio (p. 183), Institutional stillbirth rate (p. 186), Ratio of stillbirths to maternal deaths (p. 199), Neonatal deaths before discharge (p. 190).

29.2 Mortality Mapping

Where in the country is institutional mortality high or low?

This page puts institutional maternal mortality and the stillbirth rate on a map, one map per year, so you can see where in the country they are high or low. Open it after Institutional Mortality (p. 384), once the figures look plausible. Use it to spot regions with unusually low rates, which often point to under-reporting, before you look at completeness. RMNCAH

Before you start: read Institutional Mortality (p. 384).

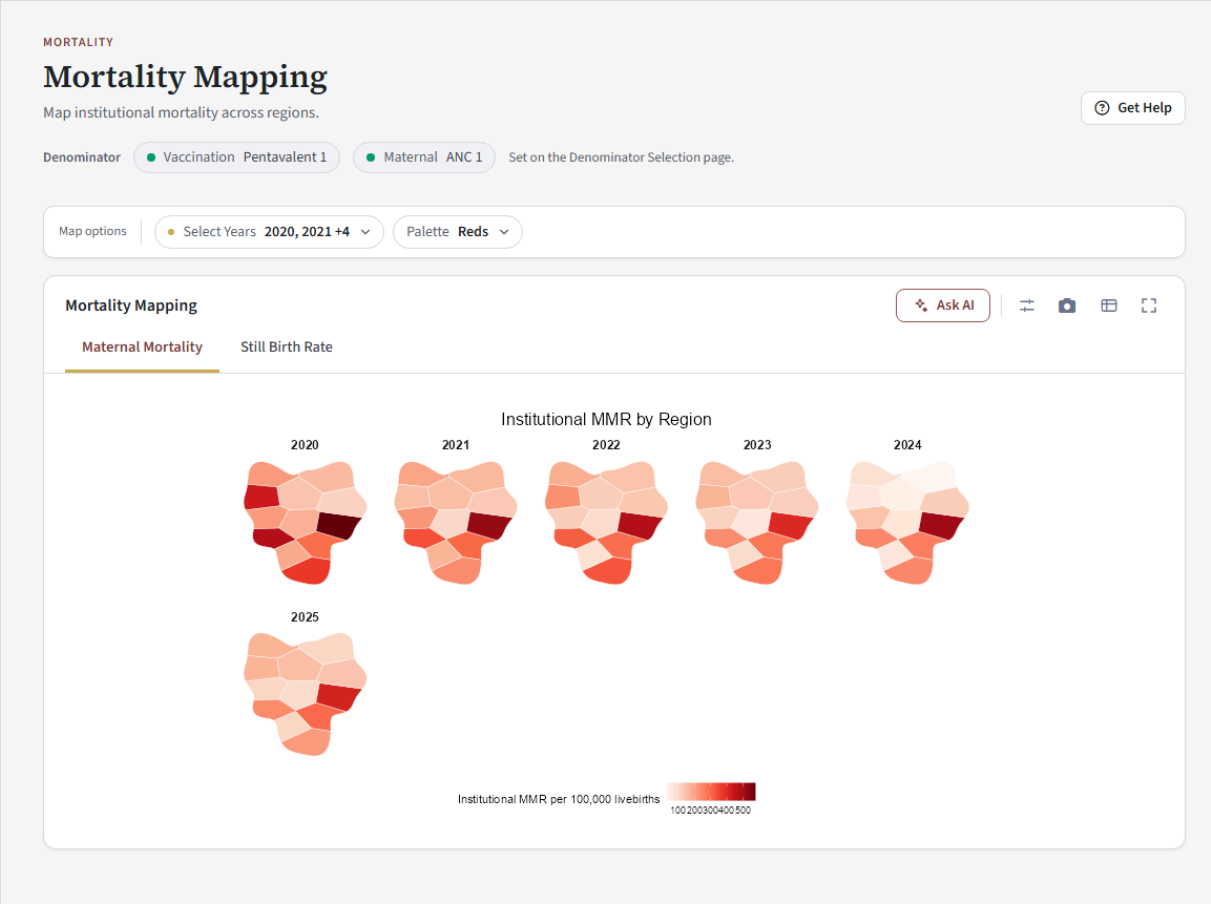


Figure 29.4 — *The Mortality Mapping page: one map of institutional maternal mortality per year*

Using the page

1. **Choose the years**

   In **Map options**, keep all years or pick the ones you want to compare in **Select Years**. Change **Palette** if you need other colours.

2. **Read the maps**

   Open **Maternal Mortality**, then **Still Birth Rate**. Compare each region from one year to the next, and the regions with each other.

3 **Note the regions to check**

Write down the regions that stay pale in every year, or that change colour sharply. Then go on to Mortality Completeness (p. 389). **Download Plot** saves the maps as an image.

Controls

Control	Choices	Default and notes
Select Years (Map options)	One or more years of the data	All years; one map per year
Palette (Map options)	Reds, Greens, Blues, Purples	Reds

Cards

Mortality Mapping: one map per selected year, each region coloured by its rate. Tabs **Maternal Mortality** and **Still Birth Rate**. The legend under the maps gives the scale.

How to read it: darker regions have higher rates. A pattern that stays the same from year to year is easier to trust. A pale region among darker ones, or a region that turns dark in one year only, deserves a look at its data: regions with few births swing widely, and a region that records its deaths well can look worse than one that does not.

Common pitfalls

- Reading a dark region as the region with the worst care. It may simply record its deaths better.
- Reading the colours without the legend: check the scale before you call a region high or low.

Method: Institutional maternal mortality ratio > Interpretation (p. 183).

29.3 Mortality Completeness

What share of deaths do facilities report, compared with the deaths expected from the UN mortality estimates?

This page estimates what share of maternal deaths and stillbirths the facilities report, by comparing the reported rates with the UN mortality estimates (p. 504). It is the last page of the Mortality group, after Institutional Mortality (p. 384) and Mortality Mapping (p. 387). You leave it knowing how far the facility rates can be trusted, and you can create the Mortality report from here. RMNCAH

Before you start: choose the denominators on Denominator Selection (p. 338); the page stays empty until they are set.

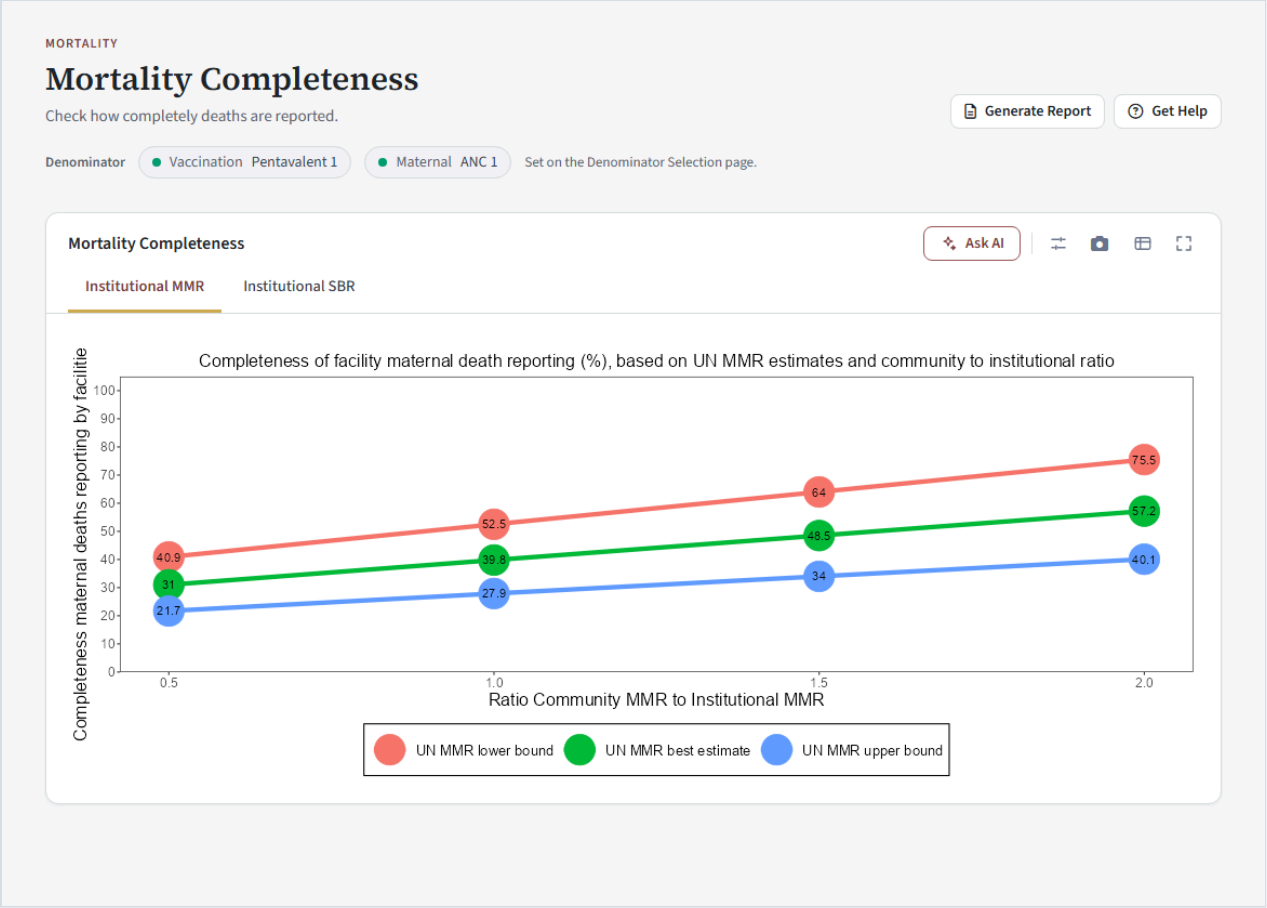


Figure 29.5 — *The Mortality Completeness page: completeness of maternal death reporting under three UN estimates*

Using the page

1. **Open a tab**
   Institutional MMR for maternal deaths, **Institutional SBR** for stillbirths.

2. **Pick the scenarios that fit**
   Choose the line that matches what you know of the true mortality (lower bound, best estimate or upper bound) and the ratios of community to facility mortality that are realistic for your country. Read the completeness at those points only.

③ **Create the report**

Press **Generate Report** to create the **Mortality** report, or **Download Data** to keep the figures.

Controls

This page has no filters. **Generate Report**, next to **Get Help**, creates the **Mortality** standard report and opens it on the Reports (p. 416) page.

Cards

Mortality Completeness: the completeness of facility death reporting, in percent. The horizontal axis is the assumed ratio of community to institutional mortality (p. 497), from 0.5 to 2. There is one line for each UN estimate: lower bound, best estimate and upper bound. Each circle gives the completeness for that scenario. Tabs **Institutional MMR** (maternal deaths) and **Institutional SBR** (stillbirths).

How to read it: completeness close to 100% in the scenarios that fit your country means facilities report most deaths. Low values in every plausible scenario mean the facility rates are too low and death reporting needs strengthening. A value above 100% means that scenario does not fit: facilities cannot report more deaths than occur.

The page uses the UN estimates built into the app, or the file you loaded with **Upload UN Mortality estimates** in the Upload Data (p. 298) step of Load Data. It also uses the denominators chosen on Denominator Selection.

Common pitfalls

- Quoting the whole range of circles. Narrow it to the scenarios that fit your country first.
- Changing the denominators afterwards: the completeness changes with them, and so does the report.
- Using the built-in UN estimates when the country has newer ones: load them in the Upload Data step.

Method: Completeness of facility death reporting (p. 194) (how it is computed) and its interpretation (p. 195) (choosing the scenario).

CHAPTER 30

Service Utilization and Health System Performance

The Service Utilization and Health System Performance pages, which look at outpatient and inpatient care for children and at facilities, beds and workforce.

Beyond coverage: service use and the health system

Service Utilization shows how much outpatient (OPD (p. 502)) and inpatient (IPD (p. 500)) care children receive. **Health System Performance** relates services to facilities, beds, health workers and the private sector. Use them to explain coverage gaps. RMNCAH

Learn about service utilization (p. 204)

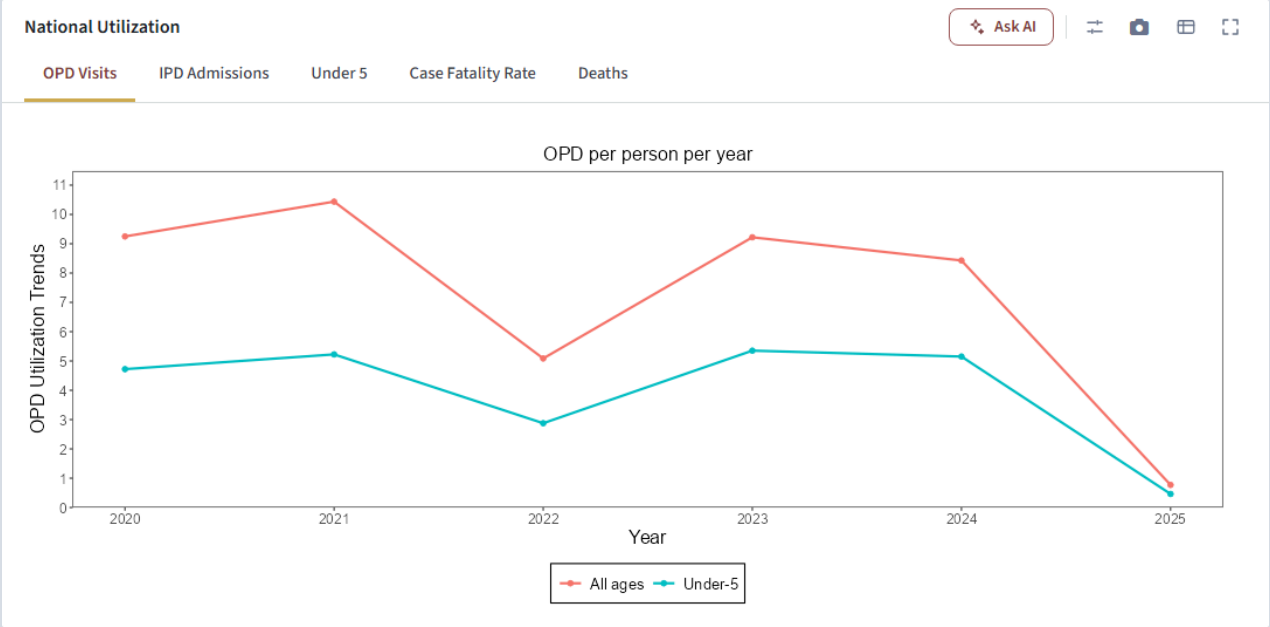


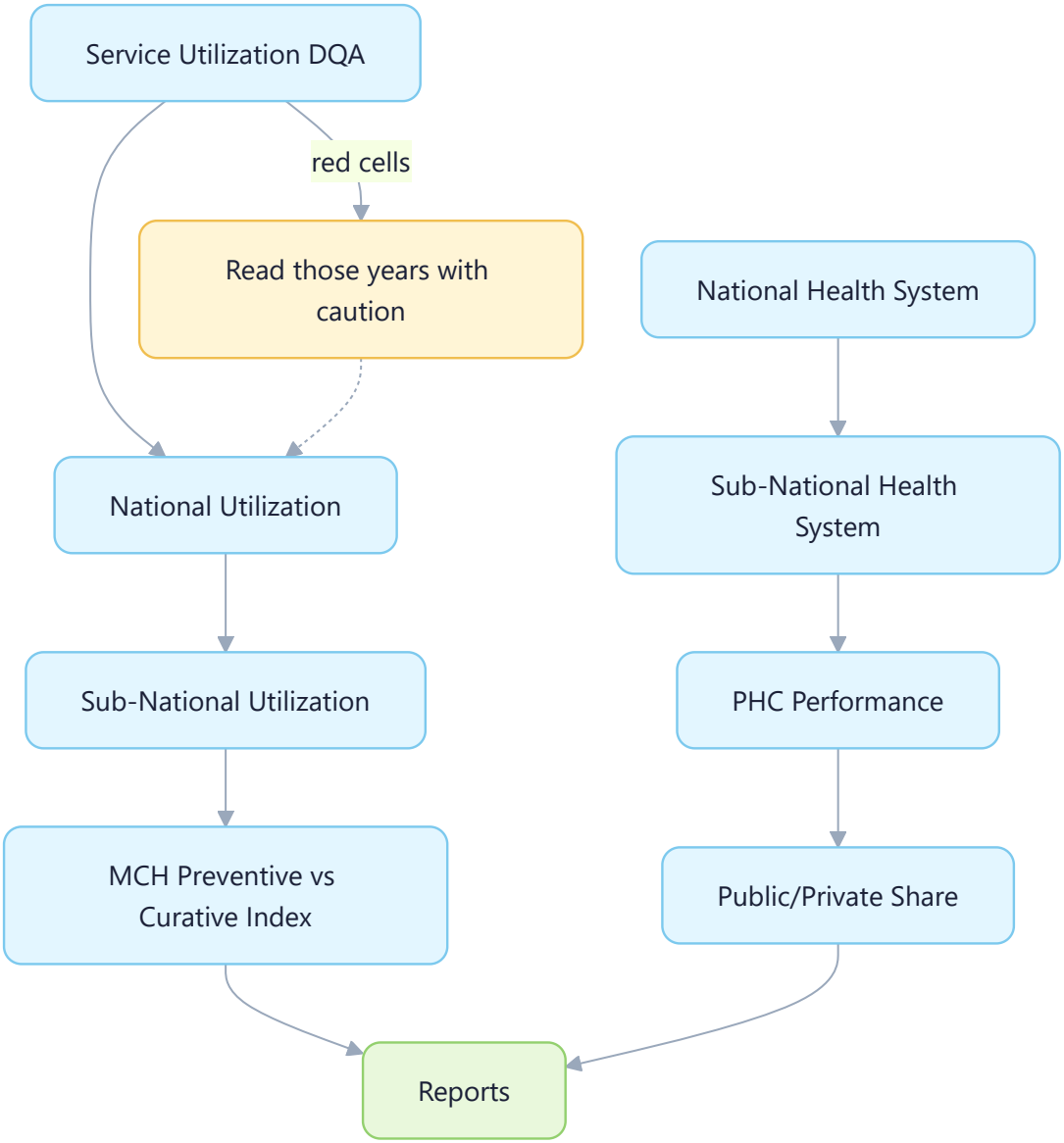
Figure 30.1 — *The National Utilization card: outpatient visits per person per year, for all ages and under 5*

Pages

Before you start: load the data and run Data Adjustment (p. 324); these pages read the adjusted data. Set the reporting threshold on Reporting Rate (p. 309) first: the Service Utilization DQA uses it.

How the pages fit together

Start each group with its national page. In Service Utilization, the data quality table comes first: it tells you which years of OPD and IPD data you can read.



Common pitfalls

- Opening these pages before Data Adjustment: they read the adjusted data.
- Reading the utilization trends before the **Service Utilization DQA** table.
- Leaving a region selected on Service Utilization DQA when you want national figures: choose **Select All**.
- Reading a jump in a rate as a change in service use without checking the population behind it: the rates are per person or per child.
- Quoting the year in the ownership chart title without checking it against your data.

Pitfalls of the method: Service utilization > Common pitfalls (p. 206) and Health system performance > Common pitfalls (p. 227).

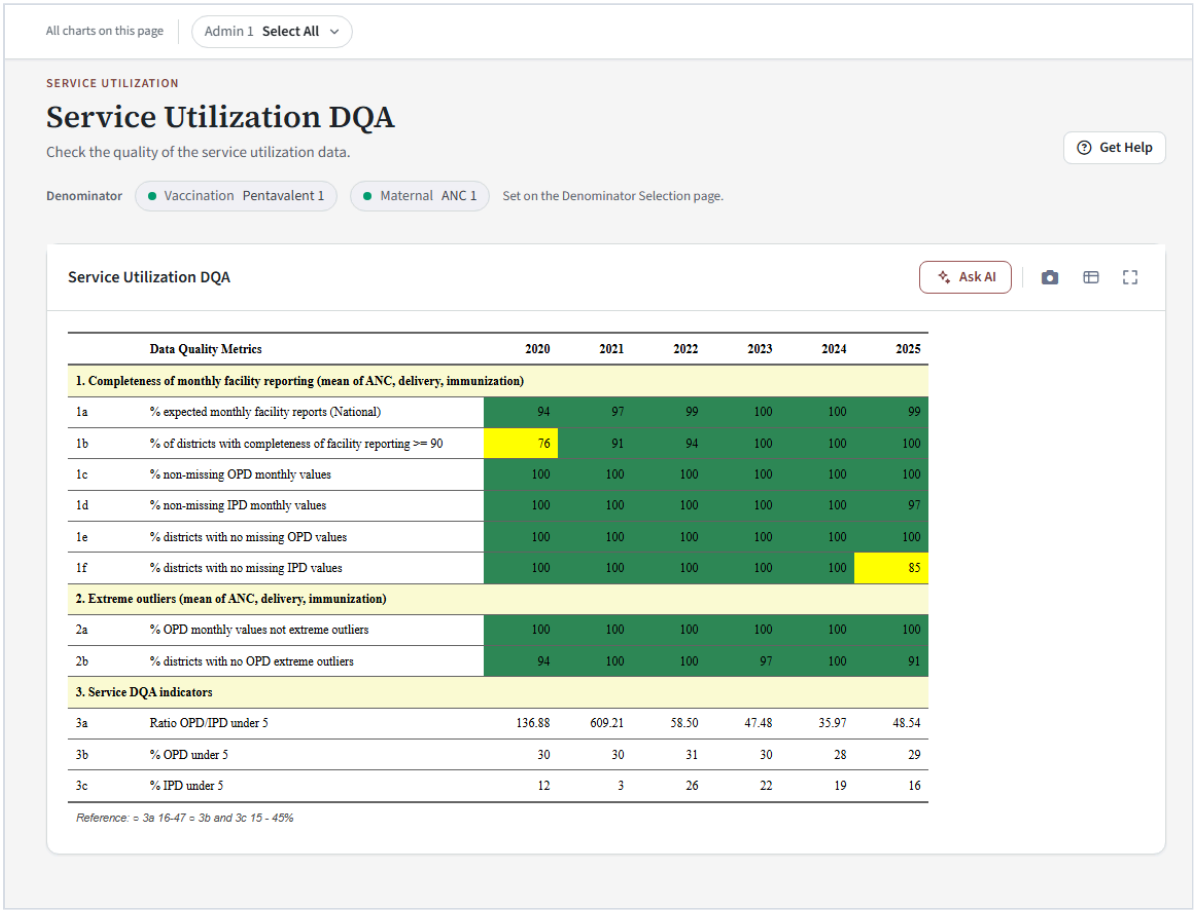
Next: Reports (p. 415).

30.1 Service Utilization DQA

Are the OPD and IPD data good enough to interpret?

This page checks whether the outpatient (OPD (p. 502)) and inpatient (IPD (p. 500)) data are complete and consistent enough to be read, year by year. It is the first page of Service Utilization: open it before National Utilization (p. 395). You leave it knowing which years, and which regions, you can trust when you read the utilization rates. RMNCAH

Before you start: run Data Adjustment (p. 324) and set the threshold on Reporting Rate (p. 309).



	Data Quality Metrics	2020	2021	2022	2023	2024	2025
1. Completeness of monthly facility reporting (mean of ANC, delivery, immunization)							
1a	% expected monthly facility reports (National)	94	97	99	100	100	99
1b	% of districts with completeness of facility reporting >= 90	76	91	94	100	100	100
1c	% non-missing OPD monthly values	100	100	100	100	100	100
1d	% non-missing IPD monthly values	100	100	100	100	100	97
1e	% districts with no missing OPD values	100	100	100	100	100	100
1f	% districts with no missing IPD values	100	100	100	100	100	85
2. Extreme outliers (mean of ANC, delivery, immunization)							
2a	% OPD monthly values not extreme outliers	100	100	100	100	100	100
2b	% districts with no OPD extreme outliers	94	100	100	97	100	91
3. Service DQA indicators							
3a	Ratio OPD/IPD under 5	136.88	609.21	58.50	47.48	35.97	48.54
3b	% OPD under 5	30	30	31	30	28	29
3c	% IPD under 5	12	3	26	22	19	16

Reference: = 3a 16-47 = 3b and 3c 15 - 45%

Figure 30.2 — *The Service Utilization DQA page: the data quality table by year*

Using the page

1 Start with the national table

Keep **Select All** in **Admin 1** to see the whole country.

2 Scan the table

Look for yellow and red cells, and note the years they fall in. Read the three under-5 checks at the bottom against the ranges in the footnote.

3 Check the regions

Choose a region in **Admin 1** to see where a problem comes from. Then go back to **Select All** and go on to National Utilization (p. 395). **Download Data** keeps the table.

Controls

Control	Choices	Default and notes
Admin 1 (All charts on this page)	**Select All** or one region	**Select All**: national figures

Cards

Service Utilization DQA: one table, *Data Quality Metrics*, with one column per year. Its rows cover the completeness of monthly reporting, missing OPD and IPD values, extreme outliers in OPD, and three checks: the ratio of OPD to IPD for children under 5, and the share of OPD visits and of admissions that are for children under 5. Cells are green, yellow or red; the footnote gives the expected ranges. The row on districts meeting the reporting threshold uses the threshold set on Reporting Rate. The toolbar has **Download Plot** (the table as an image) and **Download Data**.

How to read it: green cells in every year mean the data can be read as they are. A yellow or red year needs caution in the utilization pages. A ratio of OPD to IPD far above its range, or an under-5 share of admissions that jumps between years, usually means admissions are under-reported in those years.

Common pitfalls

- Reading the national table only. A region with missing IPD reports can look fine in the national figures.
- Forgetting that the threshold comes from Reporting Rate: change it there, not here.
- Leaving a region selected and then reading the other pages as if the table were national.

Method: Quality of outpatient and inpatient data (p. 207) and its interpretation (p. 208).

30.2 National Utilization

How much outpatient and inpatient care do people, and children under 5, receive each year?

This page shows how much outpatient and inpatient care the population, and children under 5, receive each year, how many admitted patients die, and how use differs on a map of the regions. Open it after the Service Utilization DQA (p. 393) table, which tells you which years to trust. You leave it knowing whether children reach services often enough, and whether that is changing. RMNCAH

Before you start: read the Service Utilization DQA (p. 393) table.

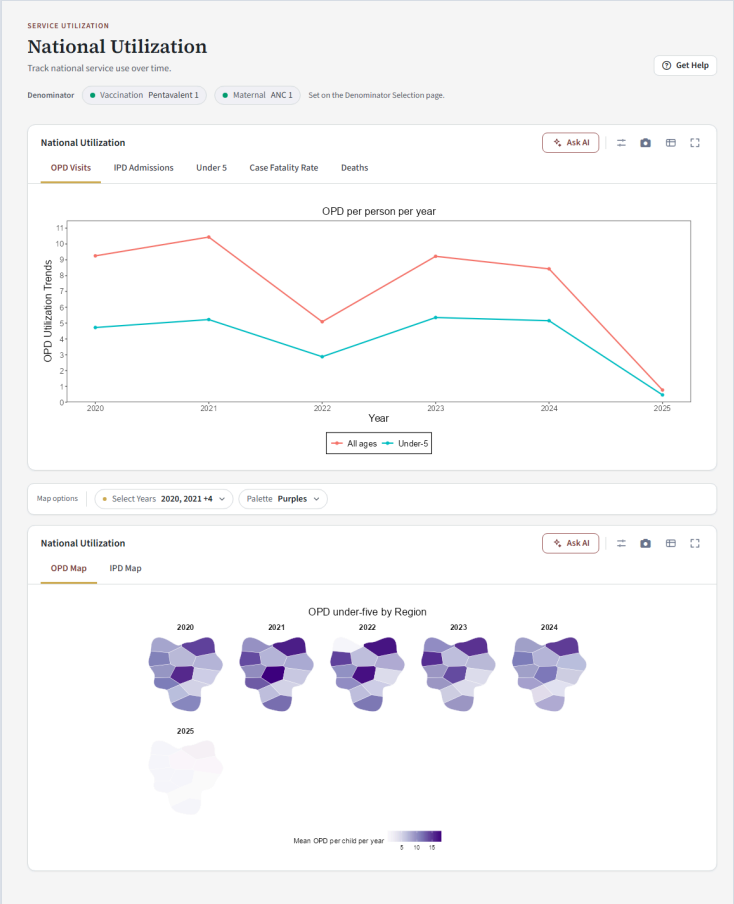


Figure 30.3 — *The National Utilization page: OPD visits by year, then the OPD map by region*

Using the page

1. **Read the trends**

   In the first card, go through the tabs from **OPD Visits** to **Deaths**. Note the level in the latest year and any jump between years.

2. **Map the regions**

   In **Map options**, choose the years in **Select Years**, then open **OPD Map** and **IPD Map** in the second card.

③ Go on

Use **Ask AI** on a card for an explanation, or **Download Data** for the figures. Then open Sub-National Utilization (p. 398) to look at one region.

Controls

Control	Choices	Default and notes
Select Years (Map options)	One or more years	All years; for the map card only
Palette (Map options)	Purples, Greens, Blues, Reds	Purples

Cards

National Utilization (trends): tabs **OPD Visits** (visits per person per year (p. 502), all ages and under 5), **IPD Admissions** (admissions per 100 people per year), **Under 5** (share of OPD visits and admissions that are for children under 5), **Case Fatality Rate** (deaths among admissions, all ages and under 5) and **Deaths** (share of inpatient deaths that are children under 5).

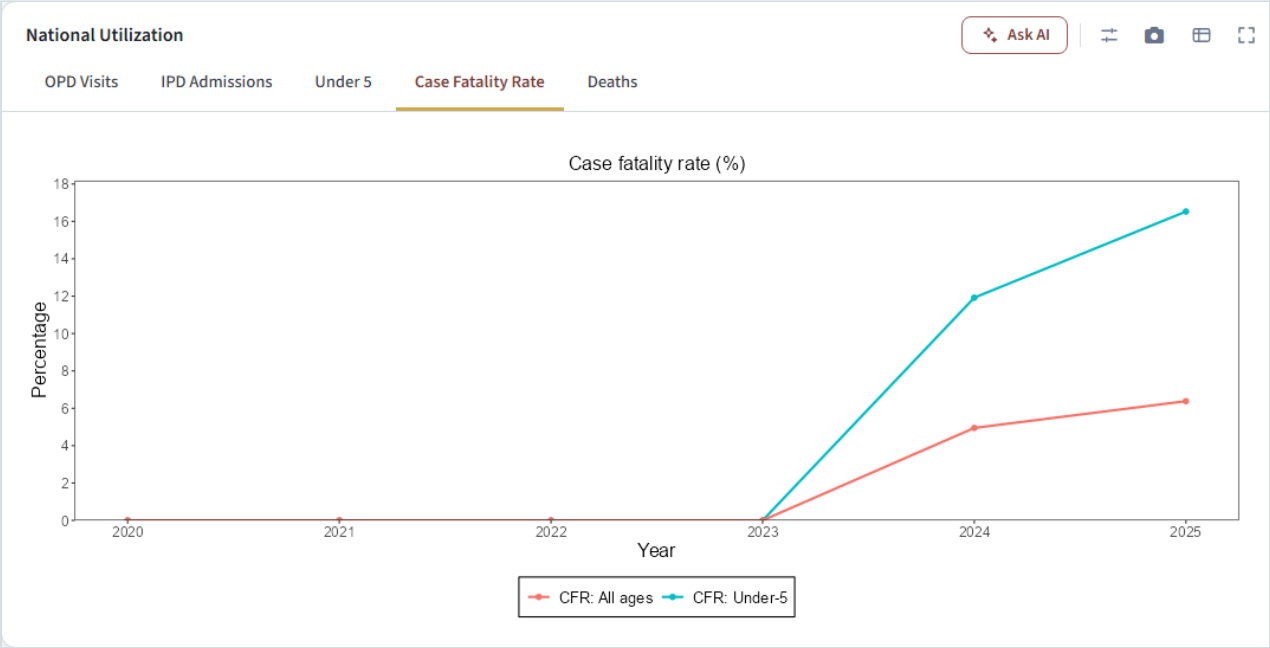


Figure 30.4 — *The Case Fatality Rate tab: all ages and under 5, by year*

How to read it: smooth lines that change little from year to year are what you expect. A big jump up or down in one year usually comes from the data (missing reports, or a change in the population used) rather than from a real change in use; check the DQA table for that year. Under 1 OPD visit per child per year points to barriers to care. A case fatality rate of 0 in a year means no deaths were reported, not that no child died.

National Utilization (maps): tabs **OPD Map** and **IPD Map**, one map per selected year, regions coloured by use per child under 5.

How to read it: darker regions use more care. A region that is pale in every year has low use, or missing reports; compare it with the DQA table for that region.

Common pitfalls

- Reading a trend across years flagged yellow or red in the DQA table.
- Comparing the OPD and IPD rates directly: OPD is per person, IPD per 100 people.
- Taking a rising case fatality rate as worse care before checking that admissions are fully reported.

Method: Outpatient service utilization (p. 210), Inpatient service utilization (p. 215), Pediatric case fatality rate (p. 219); questions to answer: Points to consider (p. 205).

30.3 Sub-National Utilization

How does service use in one region change over time, and how do regions compare?

This page shows the same utilization trends as the national page for one region at a time, and then ranks all regions by how much care children under 5 receive in the latest year. Open it after National Utilization (p. 395). You leave it knowing which regions lag behind and whether a national change comes from a few regions. RMNCAH

Before you start: read National Utilization (p. 395) and the Service Utilization DQA (p. 393) table for the regions you want to look at.

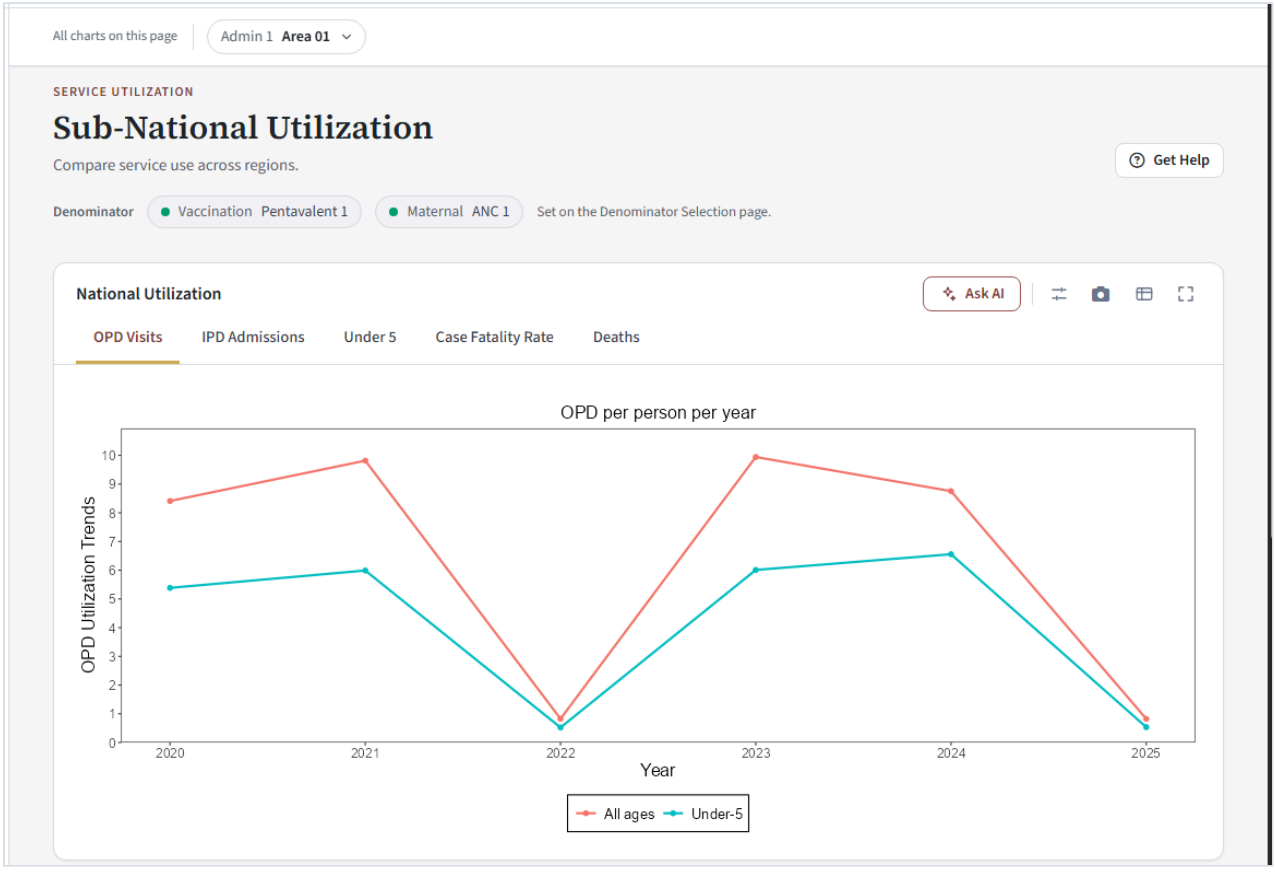


Figure 30.5 — *The Sub-National Utilization page: OPD visits by year for the region chosen in the filter bar*

Using the page

1. **Choose a region**
   Pick a region in **Admin 1** in the filter bar.

2. **Read its trends**
   Go through the five tabs of the first card, as on the national page.

3. **Compare the regions**
   In the second card, open **OPD Visits** and **IPD Admissions** to see where each region stands in the latest year. Then go on to the MCH Preventive vs Curative Index (p. 401).

Controls

Control	Choices	Default and notes
Admin 1 (All charts on this page)	One region	The first region

Cards

National Utilization: the same five tabs as on National Utilization (**OPD Visits**, **IPD Admissions**, **Under 5**, **Case Fatality Rate**, **Deaths**), for the region chosen in **Admin 1**.

How to read it: as on the national page, look for steady lines and question sudden jumps. A region's rates swing more than the national ones because its numbers are smaller.

Sub-National Utilization: one bar per region for the latest year, sorted from the highest rate to the lowest, with the value at the end of each bar. Tabs **OPD Visits** (OPD visits per child under 5 per year, chart *Sub-National OPD Distribution*) and **IPD Admissions** (admissions per 100 children under 5 (p. 496), chart *Sub-National IPD Distribution*). This card does not follow the **Admin 1** choice: it always shows every region.

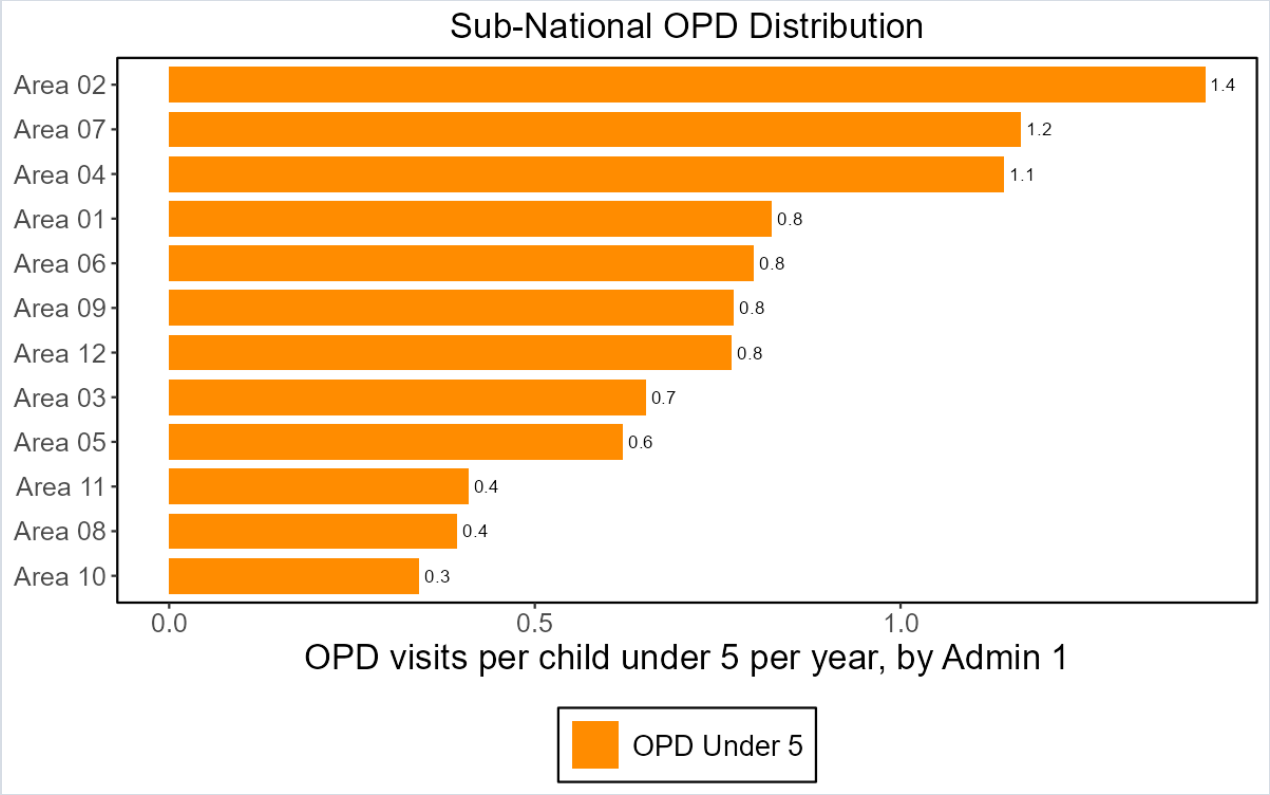


Figure 30.6 — *The chart the Sub-National Utilization card draws for OPD Visits: OPD visits per child under 5 by region, latest year (example dataset)*

How to read it: the gap between the top and bottom bars shows how unequal access is. Regions under 1 OPD visit per child are likely to face barriers to care. A region at or near 0 admissions more likely has missing inpatient reports than no admissions: check it in the DQA table.

Common pitfalls

- Expecting the second card to change with **Admin 1**: it always compares all regions.
- Comparing regions without checking their data quality: choose the region on Service Utilization DQA first.
- Reading the bars as a trend: they show the latest year only.

Method: Outpatient service utilization > Interpretation (p. 211) and Inpatient service utilization > Interpretation (p. 216).

30.4 MCH Preventive vs Curative Index

Do regions that do well on preventive care for mothers and children also provide curative care?

This page places each region on one chart by two scores: how well it delivers preventive care to mothers and children, and how much curative care children use. It is the last page of Service Utilization, after Sub-National Utilization (p. 398). You leave it with the regions that need attention most, and you can create the curative services report from here. RMNCAH

Before you start: read the Service Utilization DQA (p. 393) table; missing OPD or IPD reports move regions on this chart.

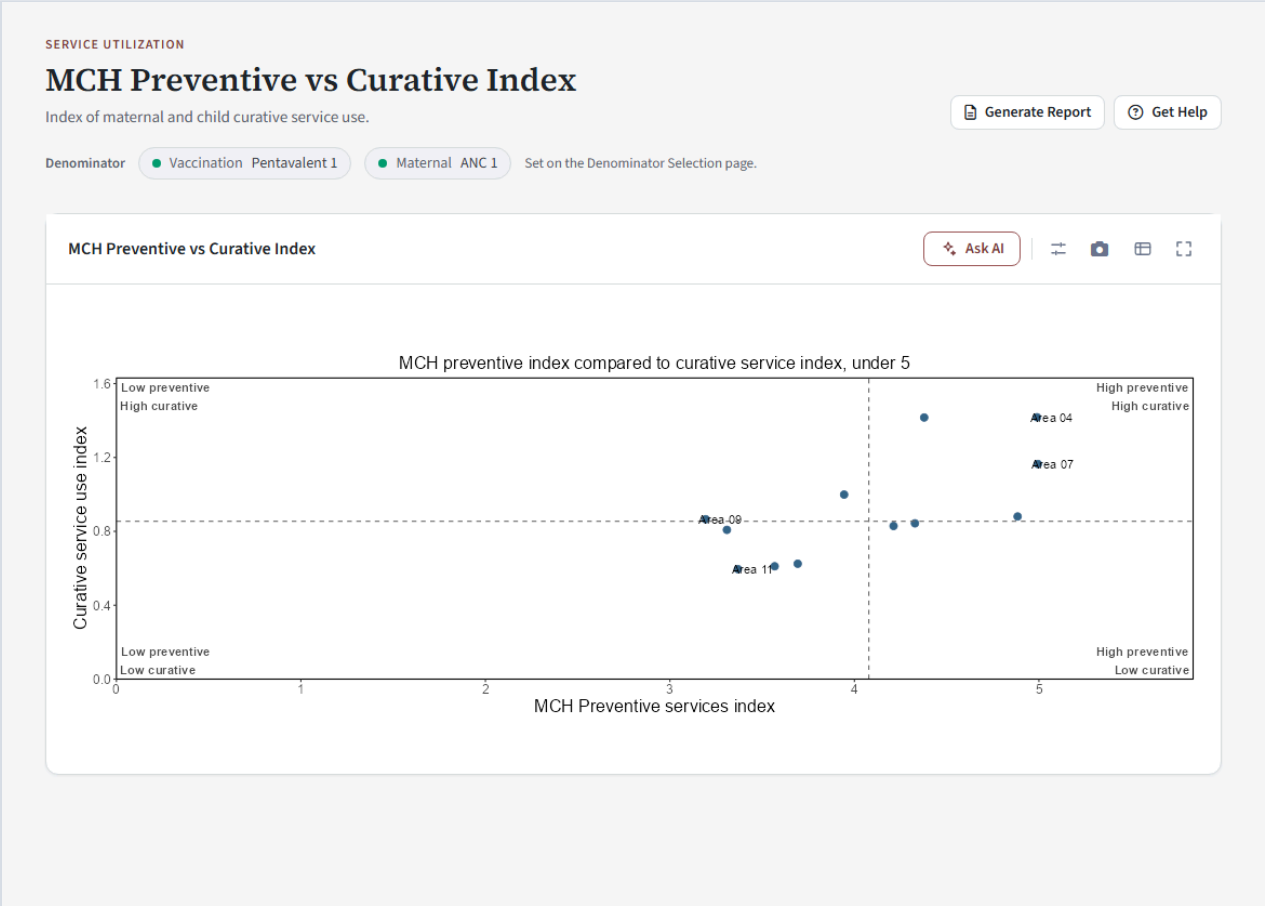


Figure 30.7 — *The MCH Preventive vs Curative Index page: regions in four quadrants*

Using the page

1. **Find the quadrants**
   The dashed lines split the chart into four quadrants, labelled in the corners.

2. **List the regions in each**
   Start with the regions in the low preventive, low curative corner, then look at the regions far from the others.

3 **Create the report**

Press **Generate Report** for the **Curative health services** report, or go on to National Health System (p. 403).

Controls

No filters. **Generate Report** creates the **Curative health services** standard report and opens it on the Reports (p. 416) page.

Cards

MCH Preventive vs Curative Index: each dot is a region. The horizontal axis is the MCH preventive services index (p. 501) and the vertical axis the Curative service use index (p. 498). Dashed median lines split the chart into four quadrants, labelled in the corners (for example *High preventive, Low curative*). Regions far from the others are named.

How to read it: regions in the top right reach families for both prevention and care. Regions in the bottom left reach them for neither and are the priority. High curative with low preventive suggests people come only when ill; high preventive with low curative suggests sick children are not brought in, or OPD and IPD are under-reported.

Common pitfalls

- Reading the quadrants as good or bad performance: the lines are medians, so half of the regions are always below them.
- Trusting a region far from the others before checking its data and its population figures.

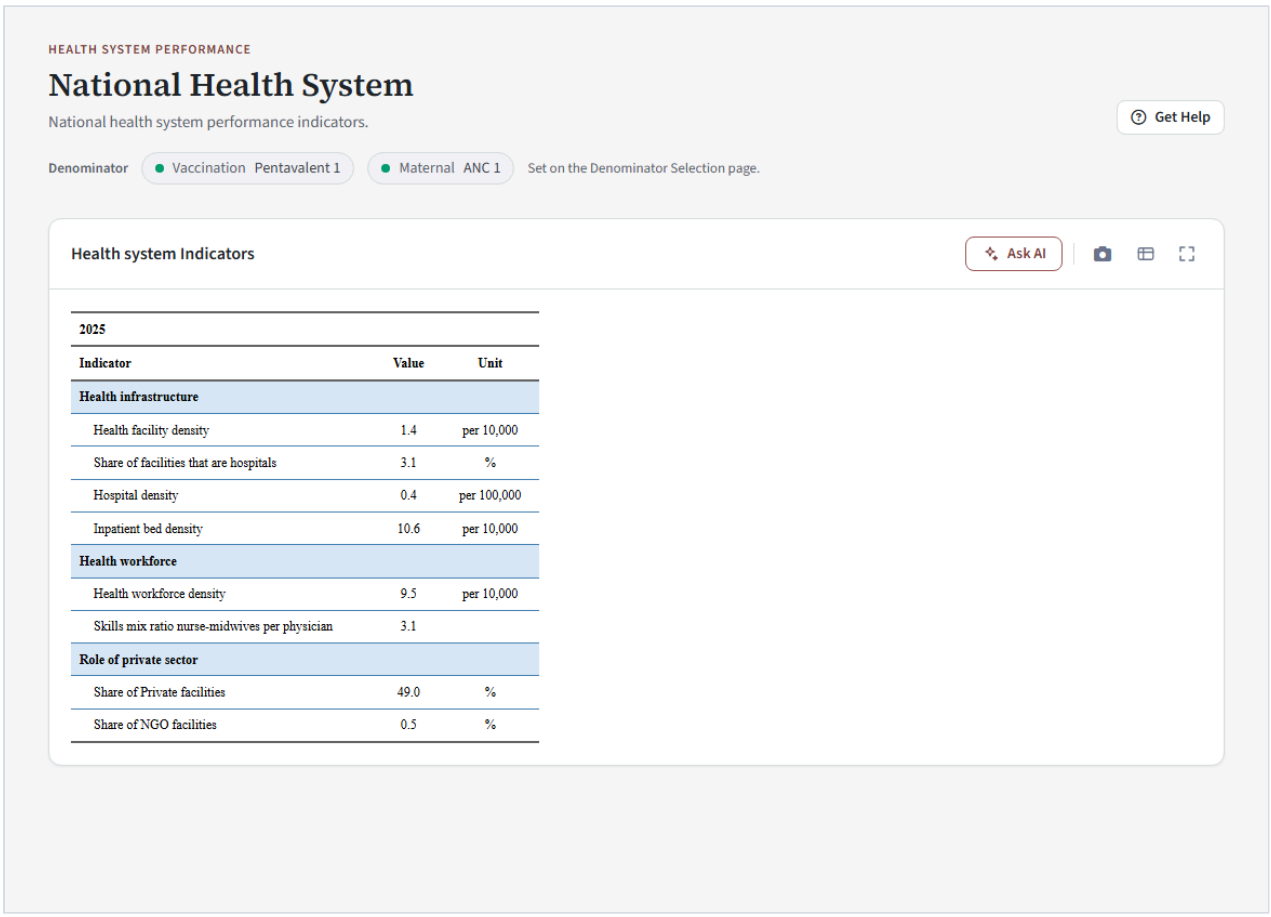
Method: MCH preventive vs curative index (p. 222) and its interpretation (p. 223).

30.5 National Health System

What are the main inputs of the health system, nationally, in the latest year?

This page lists the main inputs of the health system (p. 499) for the whole country in the latest year: facilities, hospitals, beds, health workers and the share of private facilities. It is the first page of Health System Performance. Read it to get the national picture before you compare regions on Sub-National Health System (p. 405). RMNCAH

Before you start: run Data Adjustment (p. 324). The figures come from the facility, bed and workforce data in your dataset.



2025

Indicator	Value	Unit
Health infrastructure		
Health facility density	1.4	per 10,000
Share of facilities that are hospitals	3.1	%
Hospital density	0.4	per 100,000
Inpatient bed density	10.6	per 10,000
Health workforce		
Health workforce density	9.5	per 10,000
Skills mix ratio nurse-midwives per physician	3.1	
Role of private sector		
Share of Private facilities	49.0	%
Share of NGO facilities	0.5	%

Figure 30.8 — *The National Health System page: the health system indicators table*

Using the page

1. **Read the table**
   Go through the three groups of the table and note each value with its unit.

2. **Compare with what you know**
   Compare the values with other national sources on health facilities and the health workforce.

3. **Go on**
   Use **Download Plot** or **Download Data** to keep the table, then open Sub-National Health System (p. 405).

Controls

No filters.

Cards

Health system Indicators: a table for the latest year with the columns **Indicator**, **Value** and **Unit**, in three groups: *Health infrastructure* (facility density (p. 499), share of facilities that are hospitals, hospital density (p. 500), inpatient bed density (p. 500)), *Health workforce* (workforce density and the ratio of nurse-midwives per physician (p. 504)) and *Role of private sector* (share of private and of NGO facilities). The toolbar has **Download Plot** and **Download Data**.

How to read it: values close to other national sources mean the inputs are well reported. A value far from them, or an empty cell, usually means the facility, bed or workforce data are incomplete in your dataset.

Common pitfalls

- Reading the table as a trend: it shows the latest year only.
- Taking a very low share of private facilities at face value: private facilities are often missing from the facility data.

Method: Health systems inputs (p. 229).

30.6 Sub-National Health System

Which regions are below the benchmark for facilities, hospitals, health workers and beds?

This page compares the regions on four inputs of the health system: facilities, hospitals, health workers and beds, each against the national average and, where there is one, a benchmark (p. 497). Open it after National Health System (p. 403). You leave it with the regions where inputs are scarcest, which you then relate to services on PHC Performance (p. 407). RMNCAH

Before you start: read National Health System (p. 403).



Figure 30.9 — *The Sub-National Health System page: facility density by region against the benchmark*

Using the page

1. **Open a tab**

   Start with **Health Facility Density**, then go through the other three tabs.

2. **Find the regions below**

   Note the bars coloured **Below benchmark**, and the regions far below the **National average** line.

3. **Go on**

   Use **Download Data** to keep the figures, then open PHC Performance (p. 407).

Controls

No filters.

Cards

Sub-National Health System: one bar per region, sorted, with the **National average** as a solid line. Tabs **Health Facility Density**, **Hospital Density/10,000 population**, **Workforce Density** and **Bed Density**. Where a benchmark exists, it is drawn as a dashed **Target benchmark** line and the bars are coloured **Meets benchmark** or **Below benchmark**.

How to read it: bars that reach the dashed line meet the benchmark. Regions well below the national average are where inputs are scarcest. A region far above or below all the others on one tab only may have a reporting problem; check it before reading it as a real difference. **Hospital Density/10,000 population** has no benchmark, so its bars are not coloured.

Common pitfalls

- Reading a region as well equipped because it is above the national average: it can still be below the benchmark.
- Comparing tabs with each other: each input has its own unit and benchmark.

Method: Health systems inputs > Settings (p. 230) (the benchmark values) and Interpretation (p. 230).

30.7 PHC Performance

Do regions with more facilities or health workers deliver more primary health care services?

This page asks whether regions with more facilities or more health workers also deliver more primary health care services. Each region is placed by an input and by the PHC service index (p. 502). Open it after Sub-National Health System (p. 405). You leave it with the regions where inputs and services do not match, and you can create the health system performance report from here. RMNCAH

Before you start: read Sub-National Health System (p. 405).

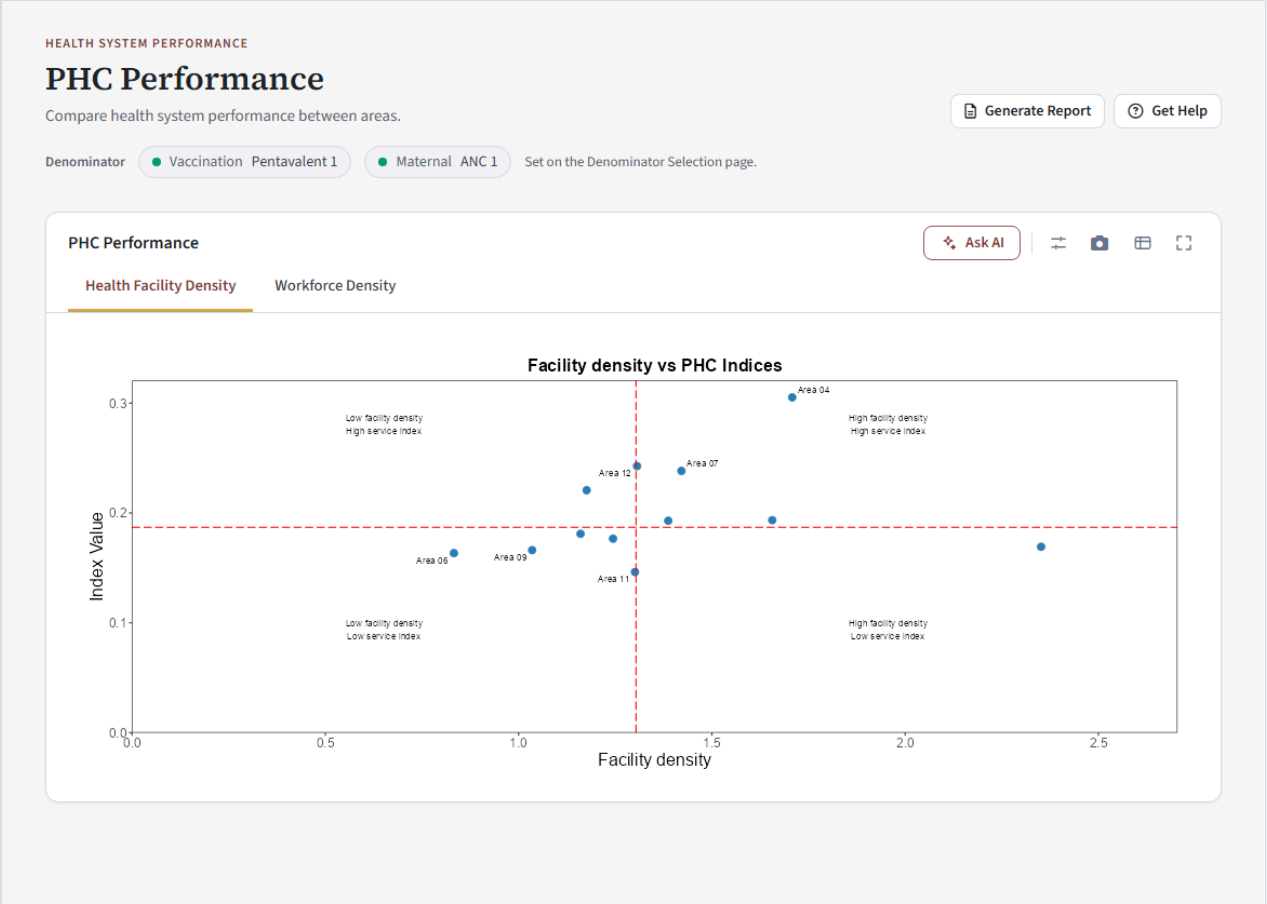


Figure 30.10 — *The PHC Performance page: facility density against the PHC service index, by region*

Using the page

1. **Open a tab**

   Health Facility Density or **Workforce Density** sets the input on the horizontal axis.

2. **Read the quadrants**

   Note the regions in the high input, low service quadrant and in the low input, high service quadrant.

3. **Create the report**

   Press **Generate Report** for the **Health system performance** report, or go on to Public/Private Share (p. 409).

Controls

No filters. **Generate Report** creates the **Health system performance** standard report.

Cards

PHC Performance: each dot is a region, placed by an input (horizontal axis) and the PHC service index (vertical axis). Dashed median lines split the chart into four quadrants, for example *High facility density, Low service index*. Tabs **Health Facility Density** and **Workforce Density**.

How to read it: you expect regions with more inputs to sit higher. A region with many inputs and a low service index may have a bottleneck, or under-reported services. A region with few inputs and a high index may deliver efficiently, or have under-counted inputs. A quadrant describes a region; it does not prove efficiency.

Common pitfalls

- Reading the quadrants as a ranking: the lines are medians, so half of the regions are always below them.
- Drawing a conclusion from one region far from the others: check its data first.

Method: Health systems outputs by inputs > PHC performance (p. 236) and Interpretation (p. 237).

30.8 Public/Private Share

How much of the health system is private, and how much care do private providers give?

This page shows how much of the health system is private: the share of private and NGO facilities in each region, from your facility data, and the share of key services given by private providers, from household surveys. It is the last page of Health System Performance. Read it before you interpret service rates from facility data: where the private sector is large, some care may be missing from those data. RMNCAH

Before you start: run Data Adjustment (p. 324).



Figure 30.11 — *The Public/Private Share page: ownership of health facilities by region*

Using the page

1. **Read the ownership chart**
   Note the regions with a large **Private** or **NGO** share.

2. **Read the survey chart**
   Open **National**, then **Area**, to see the private share of care for sick children, caesarean sections and deliveries.

3. **Create the report**
   Press **Generate Report** for the **Public/private share** report, or go on to Reports (p. 415).

Controls

No filters. **Generate Report** creates the **Public/private share** standard report.

Cards

Public/Private Share (ownership): for each region, the percentage of facilities that are **Private**, **NGO** and **Public**, from the facility data. The minus button in the card header folds the card.

How to read it: regions with a large private or NGO share are where facility data are most likely to miss services. A region with no private facilities at all may simply not report them.

Public/Private Share (services): tabs **National** and **Area** (urban and rural). Bars split care-seeking for child illness, caesarean sections and institutional deliveries into public and private sectors, with the private share written above. These figures come from the household survey estimates built into the app; the chart stays empty when there are none for your country.

How to read it: compare the private share of services with the share of private facilities. A large gap tells you whether private facilities give many services or few. Private shares are usually higher in urban areas.

> The ownership chart title reads *Ownership of health facilities, 2024 (DHIS2)* whatever your data years. Check the years before you quote it.

Common pitfalls

- Quoting the year in the ownership chart title.
- Reading an empty survey chart as no private care: it means there are no survey estimates for your country in the app.

Method: Private sector and RMNCAH-N services (p. 242).

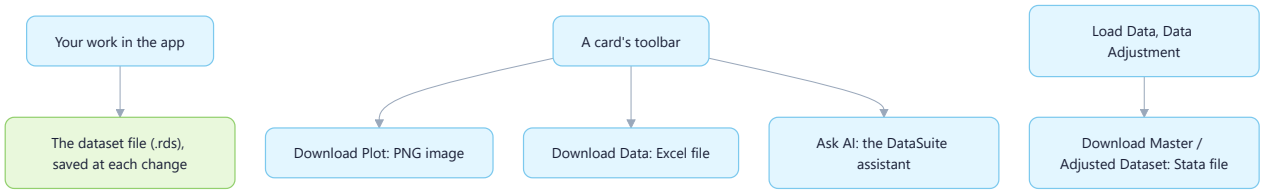
CHAPTER 31

Exports and Ask AI

How to download charts, tables and datasets from the RMNCAH and Vaxx apps, how your work is saved, and how Ask AI opens the DataSuite assistant on a card or a page.

Every chart and table in the apps can be downloaded as an image or as data, and every card can be sent to the DataSuite assistant with **Ask AI**. This page covers the card toolbar, the dataset downloads, how your work is saved, and Ask AI. Reports have their own page: Reports (p. 415). RMNCAH Vaxx

Before you start: finish Load Data (p. 295). Ask AI also needs the app to run inside DataSuite.



Charts and tables

Each card has a toolbar in its header. The buttons are the same in both apps.

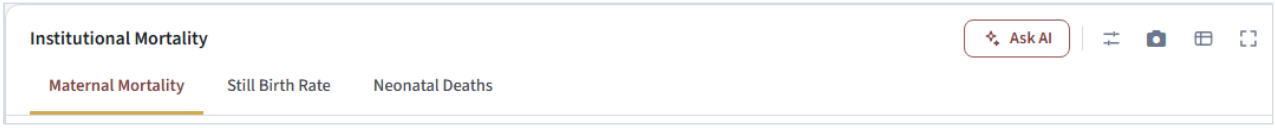


Figure 31.1 — *A card header: the title, the tabs, Ask AI and the toolbar buttons*

Button	What it does
Ask AI	Opens the DataSuite assistant with a question about this card (see Ask AI)
Customize this chart (sliders)	Changes the titles, text, colours, axes and legend of the chart (see below)
Download Plot (camera)	Saves the chart, or the table, as a PNG image 3840 pixels wide
Download Data (table)	Saves the numbers behind the chart as an Excel `.xlsx` file
Expand chart	Enlarges the chart; **Collapse chart** brings it back
More tools (...)	Appears when the window is narrow and holds the buttons that do not fit
Minus sign	On some cards only: folds the card

Downloads follow what you see: the selected tab, the filters of the page and the denominators. The file name says what the chart is and ends with the date and time, for example `rr_national_plot_202610171215.png`. A table with its own year or indicator selector (marked *This table only*) downloads for that choice.

Customize this chart

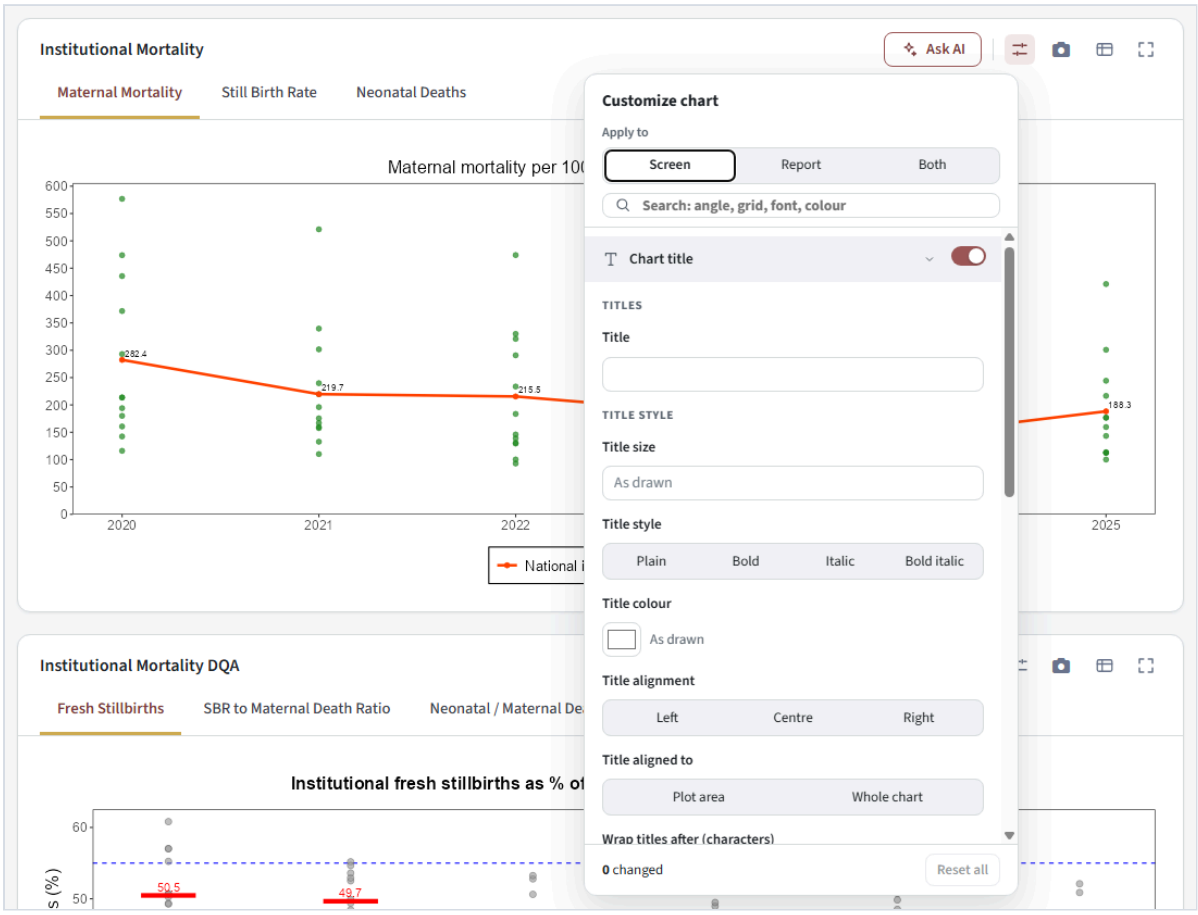


Figure 31.2 — *The Customize chart panel: Apply to, search, and the title settings*

1. Click **Customize this chart** in the card toolbar.
2. Under **Apply to**, choose **Screen** (the chart in the app), **Report** (this chart in your reports only) or **Both**.
3. Open a group (**Chart title**, **Subtitle**, axis titles and labels, **Legend**, **Data labels**, **Gridlines**, **Bars, lines and points**, **Chart area**, **All text**…) or search for a setting, and change it. An empty field keeps the chart as drawn.
4. The footer counts your changes; **Reset all** removes them.

A setting can also be used for the whole card, for every chart of the same type, or for all charts. Settings are saved with the dataset. Text you type here is not translated when you change the language.

Datasets

Button	Page	File
Download Master Dataset	Load Data (p. 295)	The merged data as loaded, as a Stata file (`master_dataset_….dta`)
Download Adjusted Dataset	Data Adjustment (p. 324)	The adjusted data (`master_adj_dataset_….dta`)

Saving your work

There is no Save button. The app keeps everything in the dataset file it opened (the .rds file (p. 503) named in the dataset chip of the top bar): the checks you confirmed, the years removed, the adjustment, the denominators, chart settings and reports. Each change rewrites that file. Keep a copy of the file before you try something you may want to undo, and share the file when a colleague should see your work.

Ask AI

Ask AI opens the DataSuite assistant, which works on the same dataset as the app. There are two buttons:

- **Ask AI** on a card: the chat opens with *Explain "card title":*; finish the question and send it.
- **Ask AI** in the top bar: the chat opens with *Explain the page "page title":*.

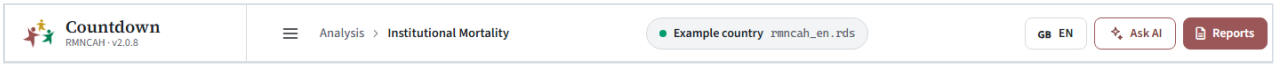


Figure 31.3 — *The top bar: dataset chip, language, Ask AI and Reports*

The assistant sees the page you are on, its cards, tabs and filters. It can read the data behind a chart, switch tabs and open other pages. In a report it can also write text and add charts (see Reports (p. 420)). It ends its answer with its sources. Compare its numbers with the card before you quote them. Good prompts and what the assistant can do: AI assistant (p. 260).

> Ask AI works only when the app runs inside DataSuite. When the app runs on its own (for example started from R), the buttons are greyed out and say *Available when the app runs in DataSuite*.

Limits

- Ask AI and the **Reference documents** of the Reports page need the app to run inside DataSuite.
- A chart's data download holds what the chart shows, not the full dataset: use the dataset downloads for that.
- Text typed in **Customize this chart** stays in the language you wrote it in.

Common pitfalls

- Downloading a chart with the wrong tab or filter selected: the file holds what is on screen.
- Experimenting on your only copy of the dataset: every change is saved in it.
- Quoting the assistant without checking: it reads the same data, so poor data give poor answers, and its output must be checked by an analyst.

Sharing charts and results: Dissemination > Sharing charts and tables (p. 253) and Dissemination > Common pitfalls (p. 247).

Next: Reports (p. 415).

CHAPTER 32

Reports

How to build document and slide-deck reports on the Reports page, generate standard reports, edit them and download them as Word, PowerPoint or PDF files.

The Reports page turns your analysis into documents (Word or PDF) and slide decks (PowerPoint or PDF). Start from a standard Countdown report or a blank page; charts and tables are drawn from the dataset every time, so a report always shows the current state of your analysis. RMNCAH Vaxx

Before you start: finish Load Data (p. 295), run Data Adjustment (p. 324) and choose the denominators on Denominator Selection (p. 338): most charts need them.

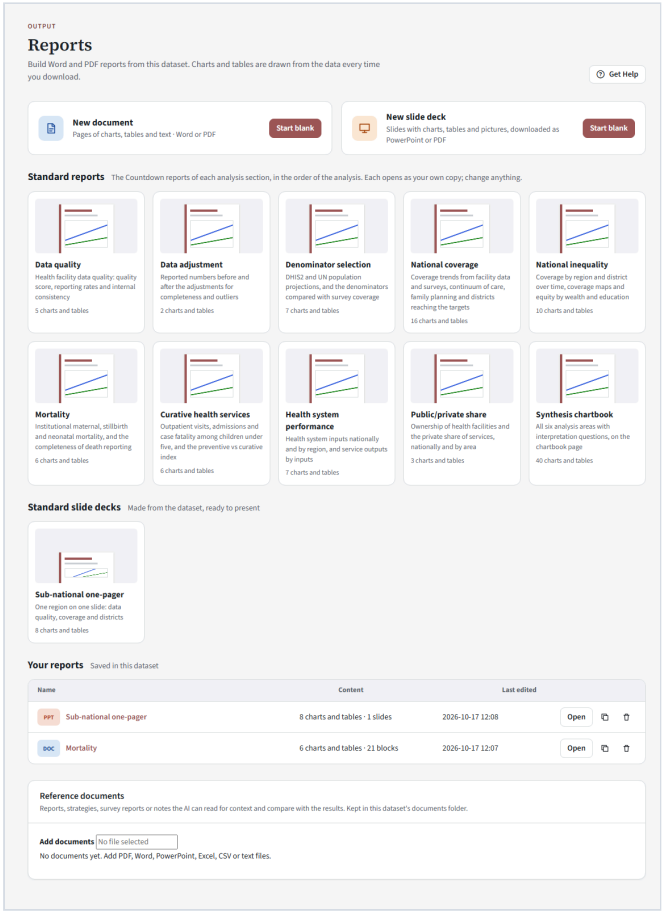


Figure 32.1 — *The Reports page: new document or deck, the standard reports, your reports and the reference documents*

Open the Reports page

Click **Reports** at the right of the top bar; the button appears once a dataset is loaded. On an analysis page, **Generate Report** (next to **Get Help**) creates that page's standard report in the dataset's language and opens it straight away.

The page has five parts:

Part	What it is for
New document / New slide deck	**Start blank**: an empty document or deck
Standard reports	The Countdown report of each analysis section, as an editable document; each card says how many charts and tables it holds
Standard slide decks	The **Sub-national one-pager**, one slide about one region
Your reports	The reports saved in this dataset, with **Open**, duplicate and delete buttons
Reference documents	Reports, strategies or survey reports the AI can read; **Add documents** works only when the app runs inside DataSuite

Standard reports

A standard report is the Countdown report of one analysis section, rebuilt from blocks you can change. It opens as your own copy. Text written as questions, in a **Notes** box, is a prompt: replace it with the interpretation agreed at your review meeting.

Standard report	What it contains	Generate Report on
Data quality	Quality score, reporting rates and internal consistency (5 charts and tables)	Overall Data Quality Score
Data adjustment	Reported numbers before and after the adjustments for completeness and outliers (2)	Data Adjustment Changes
Denominator selection	DHIS2 and UN population projections, and the denominators compared with survey coverage (7)	Denominator Selection
National coverage	Coverage trends from facility data and surveys, continuum of care, family planning and districts reaching the targets (16)	Continuum of Care
National inequality	Coverage by region and district over time, coverage maps and equity by wealth and education (10)	Equity Assessment
Mortality	Institutional maternal, stillbirth and neonatal mortality, and the completeness of death reporting (6)	Mortality Completeness
Curative health services	Outpatient visits, admissions and case fatality among children under five, and the preventive vs curative index (6)	MCH Preventive vs Curative Index
Health system performance	Health system inputs nationally and by region, and service outputs by inputs (7)	PHC Performance
Public/private share	Ownership of facilities and the private share of services, nationally and by area (3)	Public/Private Share
Synthesis chartbook	All six analysis areas with interpretation questions, on the chartbook page (40)	Reports page only
Sub-national one-pager (slide deck)	One region on one slide: data quality, coverage and districts (8)	Reports page only

Vaxx app

The Vaxx app has six standard reports, with vaccine indicators: **Data quality** (6 charts and tables), **Data adjustment** (2), **Denominator selection** (7), **National inequality** (10), the **Synthesis chartbook** (22: data quality, national coverage and equity) and the **Sub-national one-pager** (6). **Generate Report** is on Overall Data Quality Score, Data Adjustment Changes, Denominator Selection and Equity Assessment.

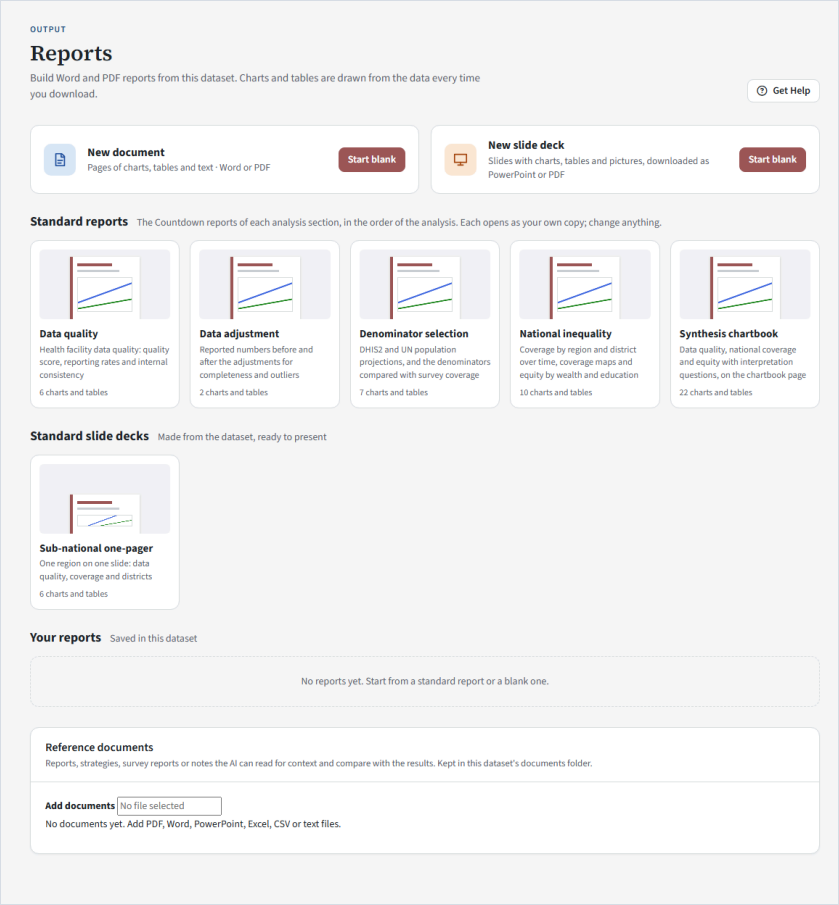
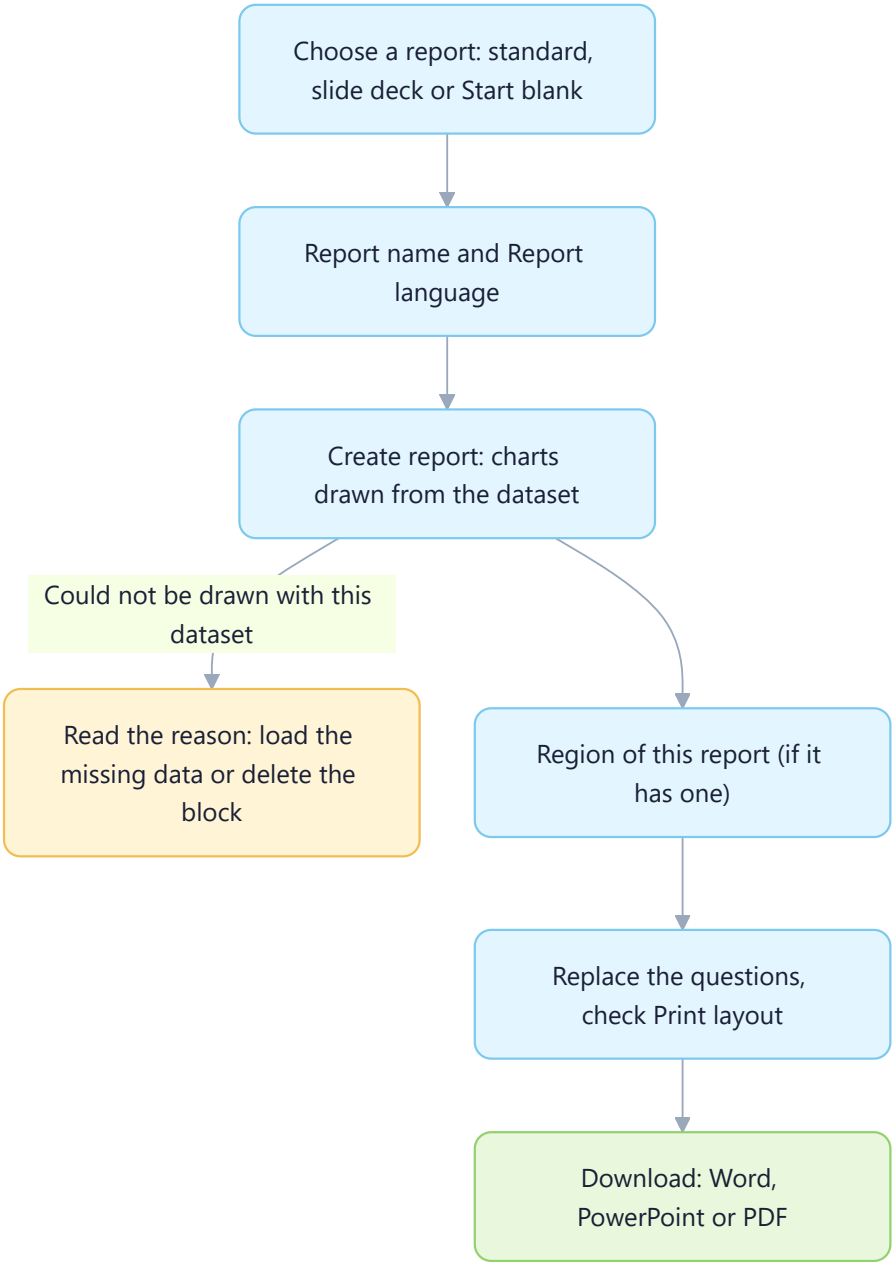


Figure 32.2 — *The Reports page of the Vaxx app*

Which reports stakeholders ask for, and what each is for: Dissemination > Information products (p. 252). The Synthesis chartbook (p. 504) and the Sub-national one-pager (p. 504) are explained in the glossary.

Create a report



① Choose a report

Click a card under **Standard reports** or **Standard slide decks**, or **Start blank**.

② Name it and choose its language

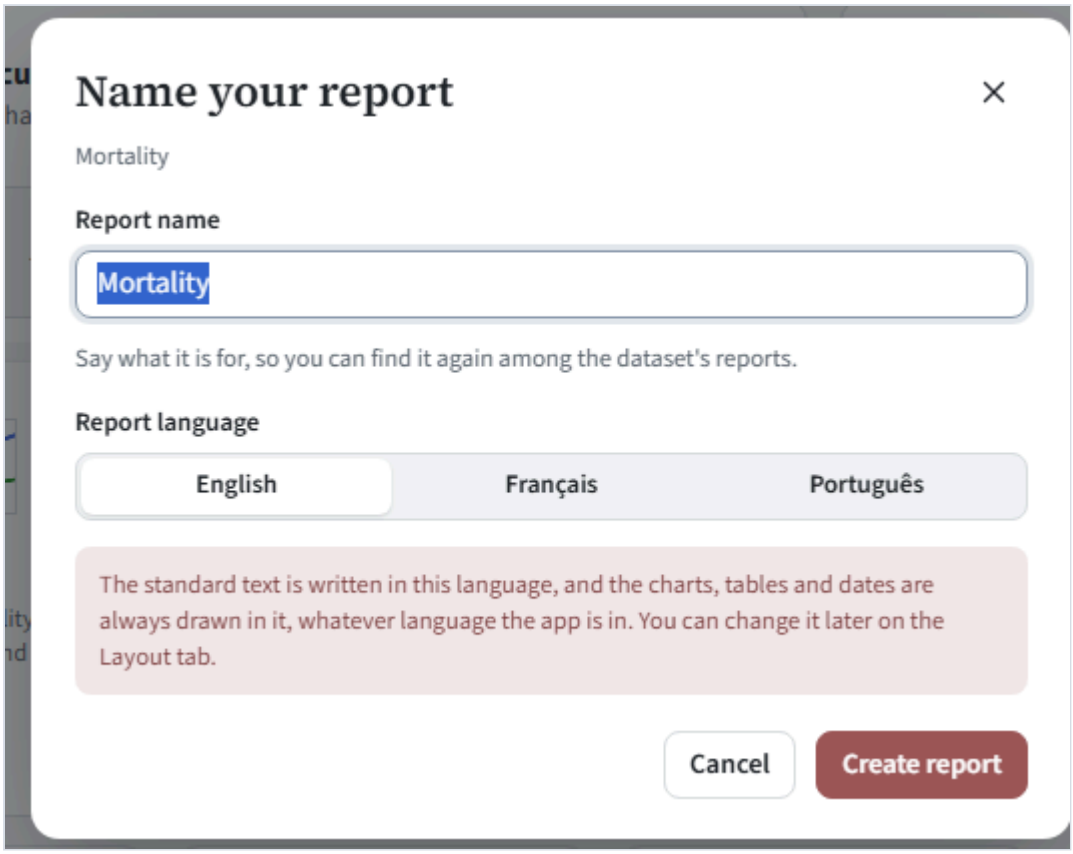


Figure 32.3 — *Name your report: the report name and the Report language*

Type a **Report name** that says what the report is for, and choose the **Report language** (English, Français or Português). The standard text is written in that language and the charts, tables and dates are drawn in it, whatever language the app is in. Click **Create report**.

③ Wait for the charts

The editor opens at once. Each chart shows **Drawing from the dataset…** until it appears. A chart that cannot be drawn shows **Could not be drawn with this dataset** and the reason (see Troubleshooting).

④ Set the region, if the report has one

The one-pager, and any chart drawn for one region, uses **Region of this report** on the **Layout** tab. The first region is chosen for you; **National** gives a national report. For a one-pager per region, duplicate the report in **Your reports** and change the region.

⑤ Edit, then check

Replace the questions, delete what you do not need, add blocks. Then open **Print layout** (in the **File** menu or the **View** tab) to see the pages as the file will have them.

⑥ Download

Click **Download**, choose the format and click the download button. A progress bar shows **Drawing charts from the dataset**, then **Writing the Word file** or **Printing the PDF**. If the file does not arrive, click **Download again**.

Changes are saved in the dataset as you work: the header says **Saved in the dataset**. **All reports** takes you back to the list.

The editor

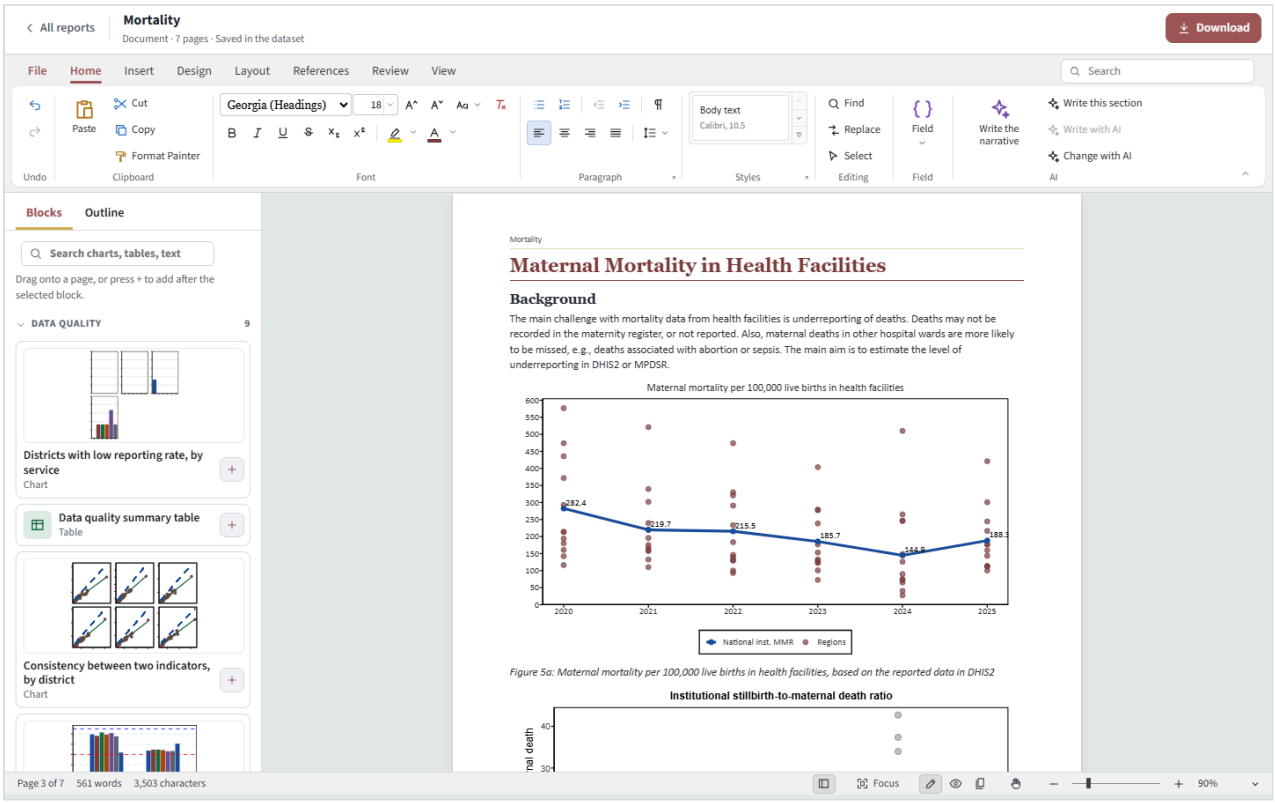


Figure 32.4 — *The document editor: the ribbon, the Blocks panel and a page of the Mortality report*

The editor works like Word. On the left, the **Blocks** panel lists every chart and table the app can draw, grouped by analysis area, with a search box (**Search charts, tables, text**); drag a block onto a page, or press + to add it after the selected block. **Outline** shows the headings of the report. The ribbon has these tabs:

Tab	What it holds
File	**Download**, **Print layout**, **Close**
Home	Undo and redo, clipboard, font, paragraph, styles, find and replace, **Field**, and the AI buttons
Insert	Cover page, blank page, page break, table, picture, chart, blocks, notes box, heading, paragraph, date, link, header and footer, page numbers, fields
Design	Theme (**Countdown**, **Ministry of Health**, **Minimal**, **Formal**), colours, fonts, paragraph spacing, **Customise theme**
Layout	Margins, orientation, page size, page numbers, indents and spacing, **Report language** and **Region of this report**
References	Table of contents
Review	**Report language** and the AI buttons
View	**Edit**, **Read mode**, **Print layout**, navigation, zoom, **Focus**

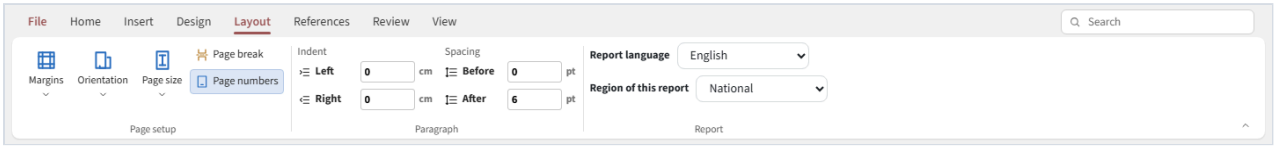


Figure 32.5 — *The Layout tab: page setup, paragraph, Report language and Region of this report*

Select a chart or table to change what it shows (indicator, level, year, region, width, title, caption) or its style with **Customize this chart**. Charts you customised in the app with **Apply to: Report** keep that style. **Fields** such as the country, the region of the report or the latest year are filled in from the data.

When the app runs inside DataSuite, the AI buttons are active: **Write the narrative** (the whole report, section by section), **Write this section**, **Write with AI** (the selected paragraph) and **Change with AI** (the selected block). They open the DataSuite assistant with the request written for you; check what it writes. See Exports and Ask AI (p. 413).

Slide decks

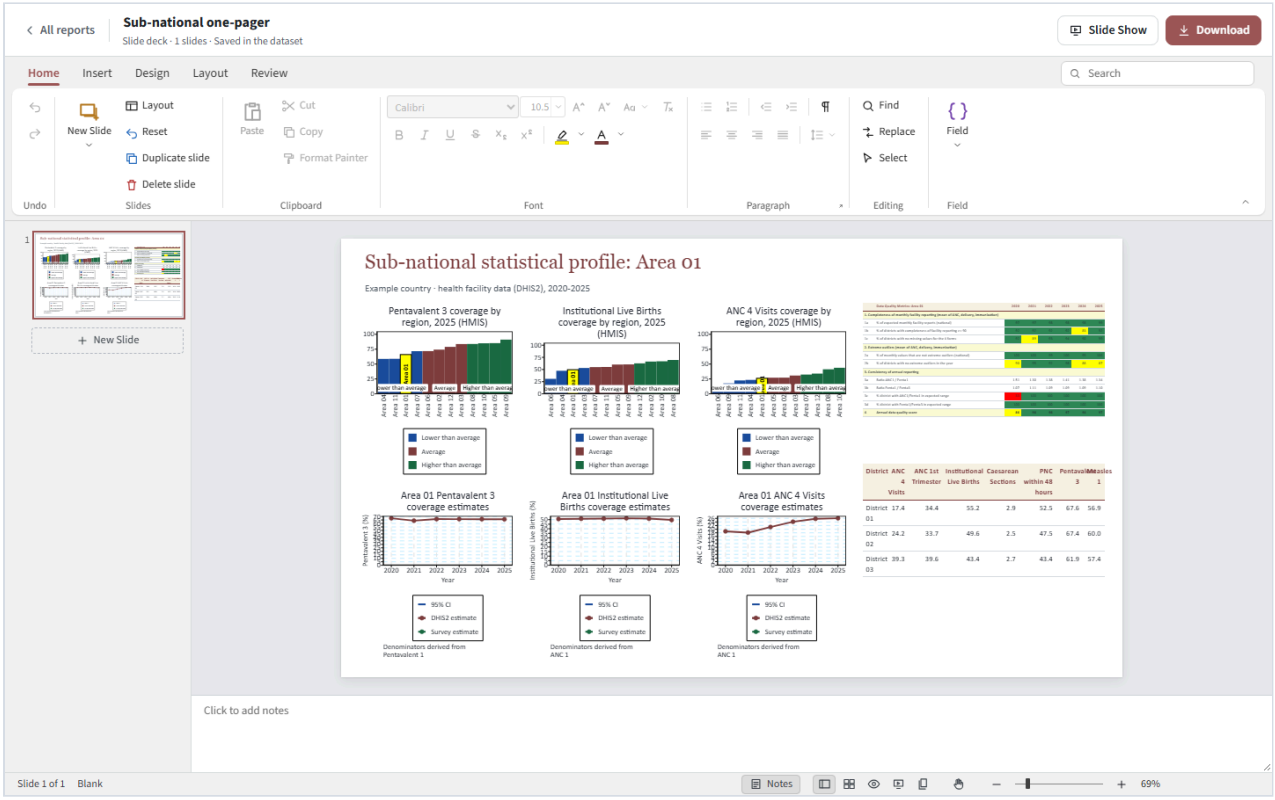


Figure 32.6 — *The slide deck editor with the Sub-national one-pager*

A deck opens in a slide editor: slides on the left, the slide in the middle, speaker notes below. **New Slide**, **Layout**, **Duplicate slide** and **Delete slide** are on the **Home** tab; **Slide Show** presents the deck. Charts and tables are added from the same blocks.

Formats and downloads

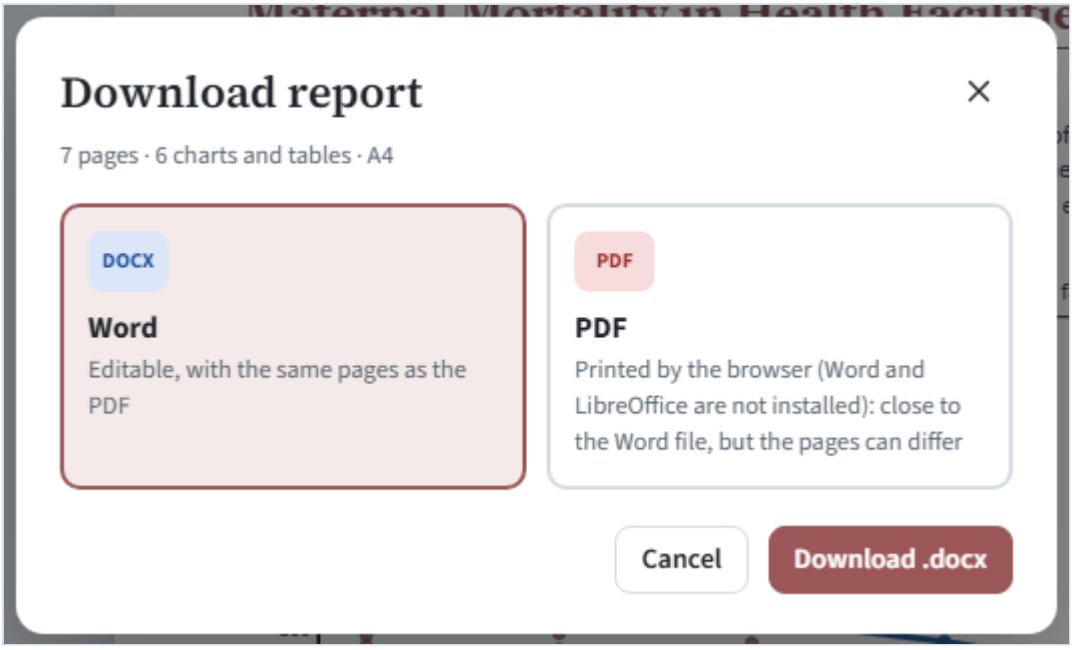


Figure 32.7 — *Download report: Word or PDF*

Format	Made by	Notes
Word `.docx`	The app	Editable, with the same pages as the PDF. On Windows with Word installed, Word also fills in the contents page and embeds the fonts
PDF of a document	Microsoft Word or LibreOffice, from the Word file	Same pages as the Word file. Without either, the browser prints it: close to the Word file, but pages can differ
PowerPoint `.pptx`	The app	Text, charts and tables stay editable
PDF of a deck	Microsoft PowerPoint or LibreOffice	Not possible without one of them

The file is named after the country, the report name and the date, for example `Example_country_Mortality_2026-10-17.docx`, and goes to your downloads folder. The dialog says which program will make the PDF on your computer.

What changes a report

- **Removed years and the adjustment** decide the numbers in every coverage, adjustment and mortality chart.
- **The denominators** chosen on Denominator Selection are used by the coverage charts.
- **Survey files** feed the equity charts and the comparisons of facility data with surveys.
- **Country reference data** (RMNCAH): the family planning (FPET) chart needs estimates for your country.

Change one of these and the charts are redrawn the next time you open or download the report; the text you wrote is not.

Troubleshooting

Message or symptom	What it means and what to do
Could not be drawn with this dataset in a block	The data for that chart are missing. Read the reason under it (below), load the missing data or delete the block. The same text appears in the downloaded file
Choose the denominators first.	Choose the denominators on Denominator Selection, then reopen the report
Choose a region.	A regional chart has no region: set **Region of this report** on the **Layout** tab, or the block's region
No district coverage for … in ….	That region has no district coverage for that year: pick another year or region
No FPET estimates for this country.	The app has no family planning estimates for your country: delete the block
No health system data for this dataset.	The dataset has no facility, bed or workforce data: delete the health system blocks
The PDF option says *Printed by the browser (Word and LibreOffice are not installed)*	Install Microsoft Word or LibreOffice on the computer for a PDF identical to the Word file, or download Word
The final pages need Microsoft Word or LibreOffice on the computer running the app.	**Print layout** cannot show the final pages. Download the Word file, or install Word or LibreOffice
Making a PDF of slides needs Microsoft PowerPoint or LibreOffice.	Download the deck as PowerPoint, or install PowerPoint or LibreOffice
Open this dataset in DataSuite to add documents.	Reference documents are kept in DataSuite's analysis folder: open the dataset from DataSuite
The first download is slow	A long report, such as the synthesis chartbook, draws every chart. Wait for the progress bar

Common pitfalls

- Sending a standard report with its questions still in it: replace every **Notes** box.
- Quoting a report built before you changed the adjustment or the denominators without reopening it: open it again so the charts are redrawn.
- Experimenting on your only copy of the dataset: reports are saved in it.

Preparing reports for stakeholders: Dissemination > Preparing reports for stakeholders (p. 252).

Next: Exports and Ask AI (p. 411).

CHAPTER 33

Pooled App overview

Overview of the Pooled App, which combines several countries' saved datasets into one file for regional and multi-country comparison, and when to use it.

Put several countries side by side

Load the datasets each country saved in the RMNCAH or Vaxx app (p. 289), combine them into one pooled file, then explore, compare and export the tables. Pooled starts where the country apps stop: it does not replace them.

Learn about how it works (p. 425)

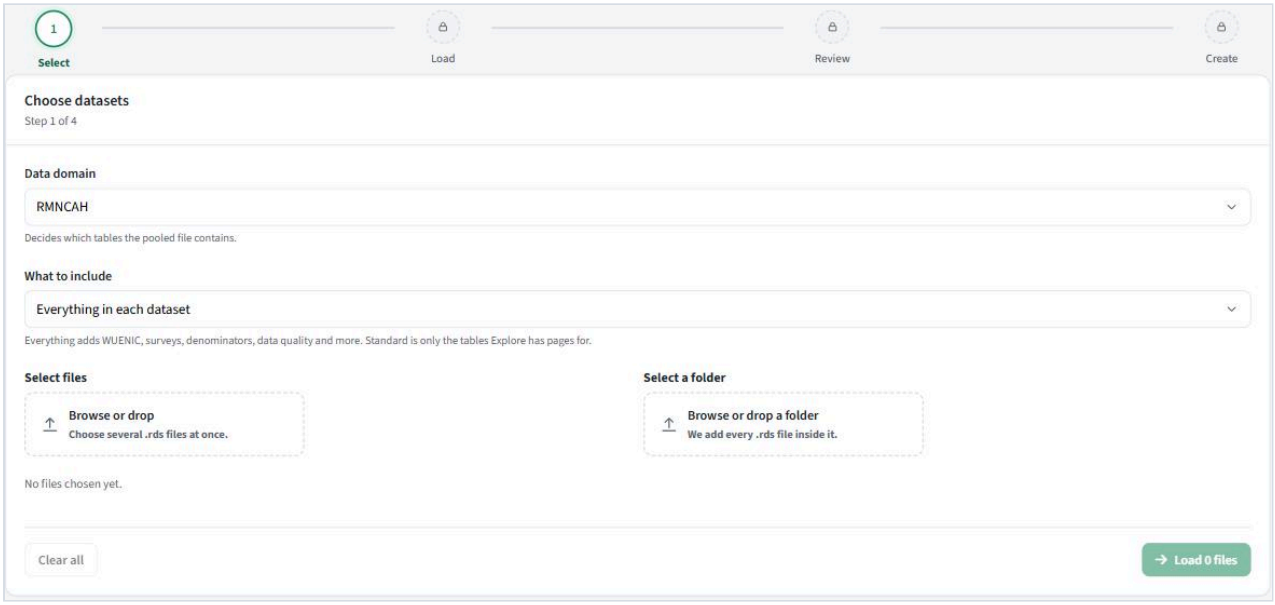
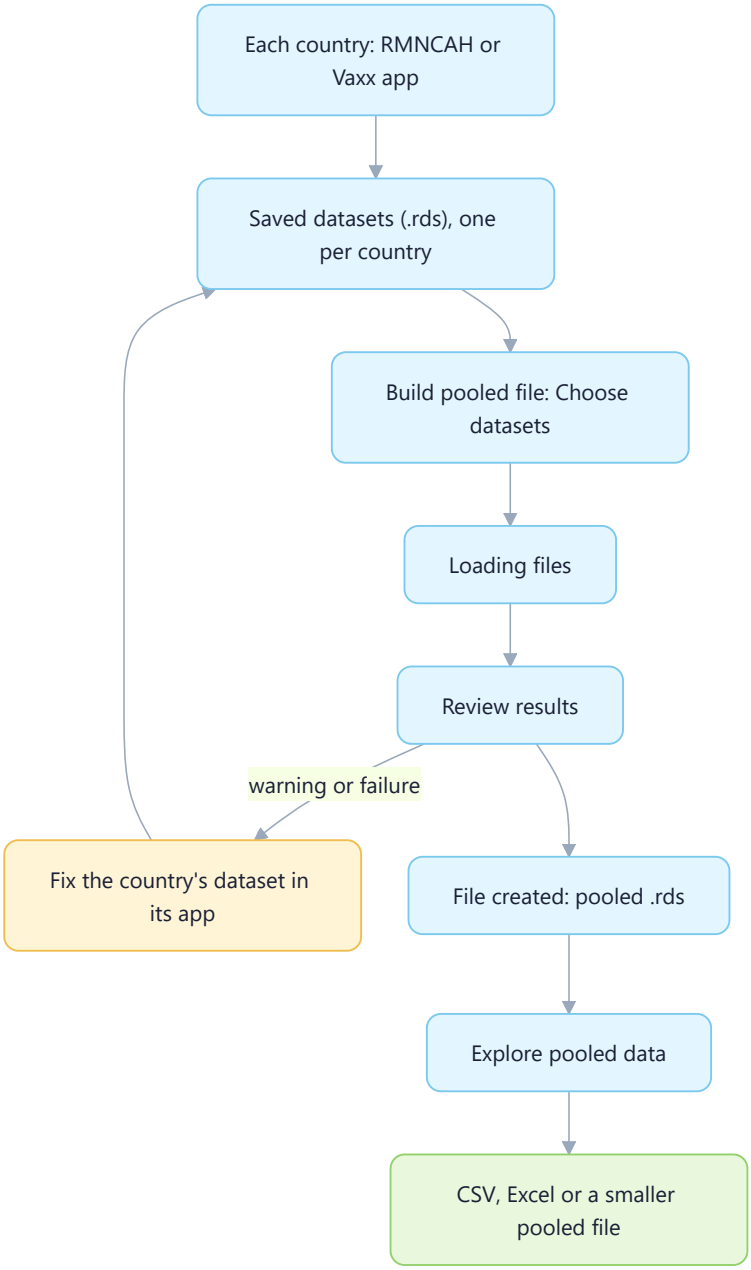


Figure 33.1 — *Build pooled file, step 1: choose the data domain and the datasets*

Pages

Before you start: finish each country's analysis in the RMNCAH or Vaxx app, at least to Denominator Selection (p. 338), so that its saved dataset is adjusted. See Preparing caches (p. 428).

How it works



When to use Pooled

You want to...	Use
Check, adjust and analyse one country's facility data	RMNCAH or Vaxx App
Put several countries side by side (coverage, sources, trends)	Pooled App
Get one table containing all countries for a regional report	Pooled App
Extract only some countries, years or columns for another tool	Pooled App

Requirements

- A set of `.rds` files, one per country, saved by the RMNCAH or Vaxx App. See **Preparing caches** (p. 428).
- All files should belong to the same **data domain**: RMNCAH or Immunization (VAXX). You choose the domain once, before loading.
- Enough memory: the app accepts up to 2 GB of selected files at a time, and reads them one after another.

> Pooled keeps only the analysis tables (coverage, data quality, denominators, surveys and so on). The raw facility-by-month data is left out, which keeps the pooled file small.

How the app is organised

The sidebar has one section, **Pooling**, with two parts.

Build pooled file is a four-step wizard:

1. Choose datasets: the domain, what to include, and the files.
2. Loading files: each file is read and reported as loaded, loaded with a warning, or failed.
3. Review results: decide what to do about warnings and failures.
4. File created: download the pooled `.rds` or go straight to exploring it.

Explore pooled data opens a pooled file and offers these pages:

Page	What it shows
Open a file	Use the file you just built, or upload one made earlier
Parameters	The settings and survey values each country brings
Overall score	The data quality score of each country
Indicator coverage	Coverage by indicator: national, Admin 1, District
Coverage	Coverage with each country's chosen denominators
Mortality	Facility maternal and newborn deaths (RMNCAH domain)
Service utilization	Outpatient and inpatient use, children under 5 (RMNCAH domain)
Compare countries	Continuum of care, source triangulation, level and change, trends
All datasets	Every table in the file, including those without their own page
Extract a piece	Pull chosen datasets, countries, years and columns into a new file

Common pitfalls

- **Mixing domains.** Pool RMNCAH datasets together and Vaxx datasets together; the domain is chosen once, before loading.

- **Unadjusted datasets.** A country that was not adjusted loads with a warning, and its coverage is not comparable with the others.
- **Two files for one country.** Only the newer one is kept; remove old versions from the folder.
- **Not keeping the pooled file.** It is built for the session only: click **Download .rds** to keep it.

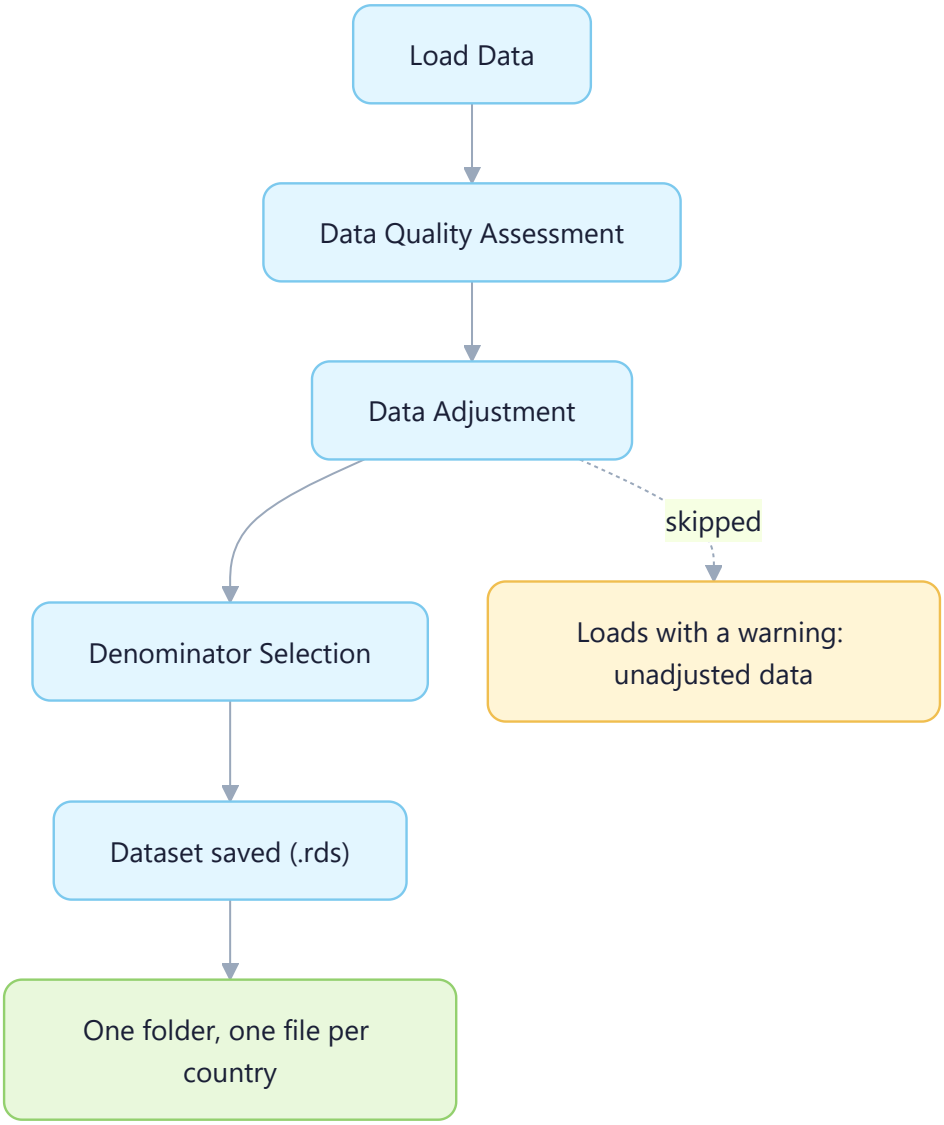
33.1 Preparing caches

How to produce the .rds caches that the Pooled App reads from the RMNCAH or Vaxx app, and what makes a cache loadable and comparable across countries.

Pooled reads the **caches** that the country apps save. A cache is an `.rds` file that holds a country's dataset together with everything computed on it: adjusted data, coverage, data quality results and the choices made in the analysis. Good pooled results depend on well-prepared caches, so it is worth a few minutes to get them right.

This is the step before Pooled: you prepare one cache per country in the RMNCAH or Vaxx app, put them in one folder, then build the pooled file (p. 431).

Before you start: have each country's data loaded in the RMNCAH or Vaxx app (Load Data (p. 295)).



Producing a cache

1. Open the **RMNCAH App** (p. 289) (or the **Vaxx App** (p. 289)) and upload the country's data. The data usually comes from the **Data Extractor** (p. 269).

2. Work through the analysis: data quality checks, data preparation and adjustment, and denominator selection. See the **analytical framework** (p. 42).
3. The app saves the cache as an `.rds` file in the same folder as the file you opened, named after it: `mydata.xlsx` is saved as `mydata_rmncah.rds` by the RMNCAH app and `mydata_vaccine.rds` by the Vaxx app. If that `.rds` already exists, it is opened again instead of a new one being created. See The saved .rds file (p. 47).

Repeat for each country, using the same app and the same domain for all of them.

> Complete the adjustment and denominator steps before pooling. A cache that is not adjusted still loads, but Pooled warns that its tables use unadjusted data, and coverage will not be comparable with countries that were adjusted.

What makes a cache loadable

For each file, Pooled opens it as a Countdown cache for the domain you chose and then checks:

- The file can be read by R. A damaged file, or one saved by a newer version of R, is reported as unreadable.
- The cache contains a **country** name. Files without a country or without adjusted data are reported as "Not a Countdown dataset".
- At least one table can be built. Standard tables that cannot be built produce a warning that lists them.

The country and its ISO3 code come from inside the cache, not from the file name.

Naming and folders

- Only files ending in `.rds` (any letter case) are used. Other files are listed as skipped.
- Put all the countries you want to pool in one folder. Selecting a folder adds every `.rds` file inside it.
- Keep a clear name, ideally with the date at the end, for example `Country_master_adj_dataset_20250918.rds`. Pooled reads the last run of eight digits in the name as the file date (year, month, day). If there is none, the file's modified time is used.
- Keep **one file per country**. If two selected files are for the same country (matched by ISO3 code), Pooled keeps the one with the newer date and leaves the other out, with a reason in the log.
- Remove old versions and unrelated `.rds` files from the folder, or untick them in step 1.

Compatibility notes

- All caches must be loadable by the version of `cd2030.core` installed with DataSuite (the Pooled package requires version 1.1.0 or later). Caches carry an internal data version, so if an old file fails to load, open and re-save it in the country app first.
- Countries can have different tables. Data quality or survey tables, for example, may exist for some countries only. Pooled records what each file lacks and flags it in the review step rather than failing.

- If the columns of one table differ between countries in type, they are combined as text and a note is written to the file log.

Common pitfalls

- **Pooling the Excel workbook.** Pooled reads the `.rds` the app saved, not the workbook from the Data Extractor.
- **Renaming the saved `.rds`.** The app finds a dataset again by its name next to the workbook; copy it to the pooling folder rather than renaming the original.
- **Old and new versions together.** Keep one file per country in the folder, with the date at the end of its name.

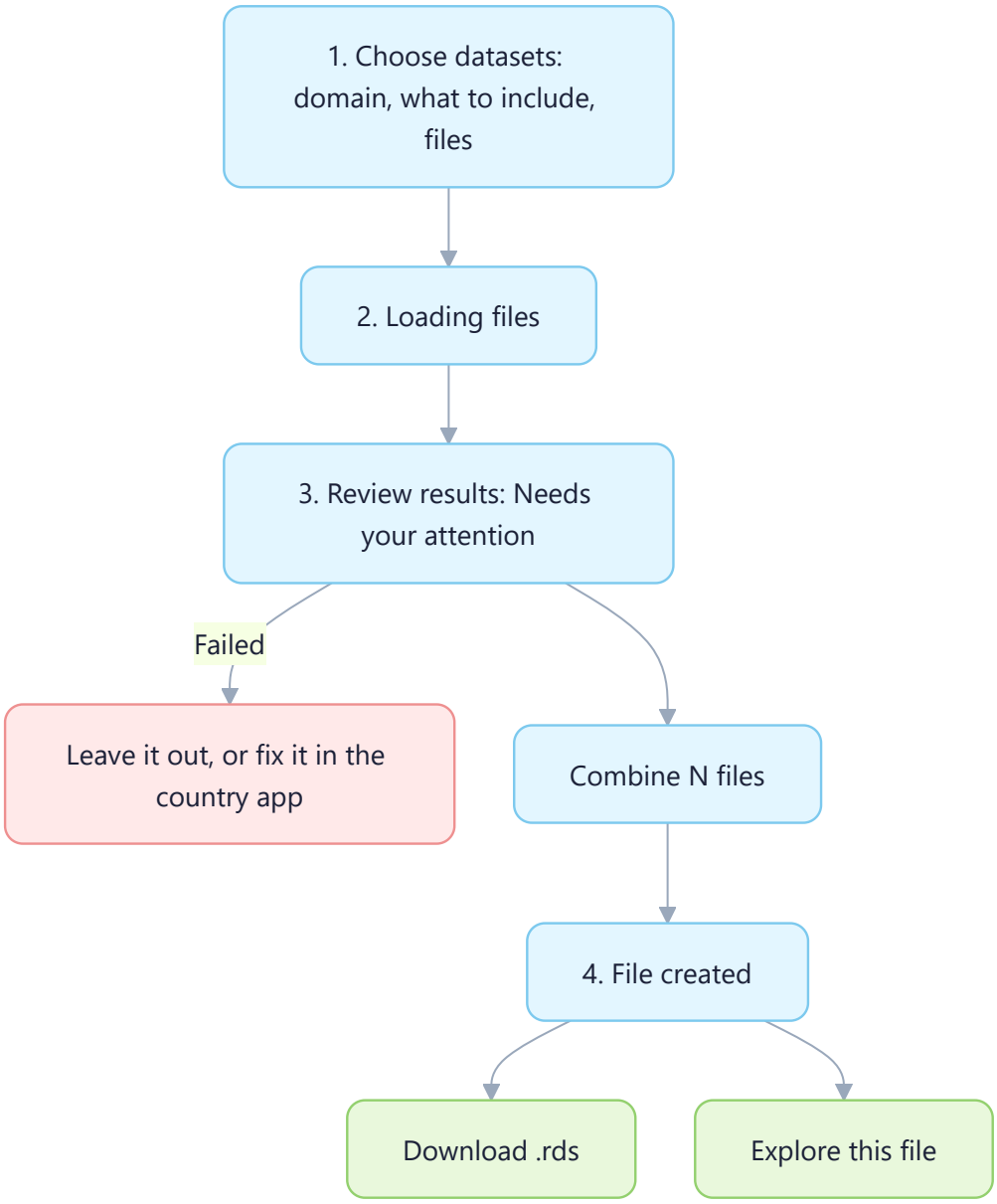
Next: **Using the app** (p. 431).

33.2 Using the app

Step-by-step guide to building a pooled file in the Pooled App from saved country datasets, exploring the pooled tables and reading the output.

You use the Pooled app in two parts: **Build pooled file** reads the countries' saved datasets and combines them into one file, and **Explore pooled data** shows that file's tables and charts, country by country or side by side. You build once, after preparing the caches (p. 428); you explore and export (p. 439) as often as you need.

Before you start: have one `.rds` file per country, saved by the RMNCAH or Vaxx app, in one folder (Preparing caches (p. 428)).



Build a pooled file

① Launch Pooled

Start **Pooled** from the DataSuite apps. When DataSuite starts it with a folder selected, every `.rds` file in that folder is already added to the file list and a banner shows how many files were pre-loaded. The app opens on **Build pooled file**. Most Pooled pages and buttons are currently shown in English whatever the language selected.

② Choose the datasets (step 1 of 4)

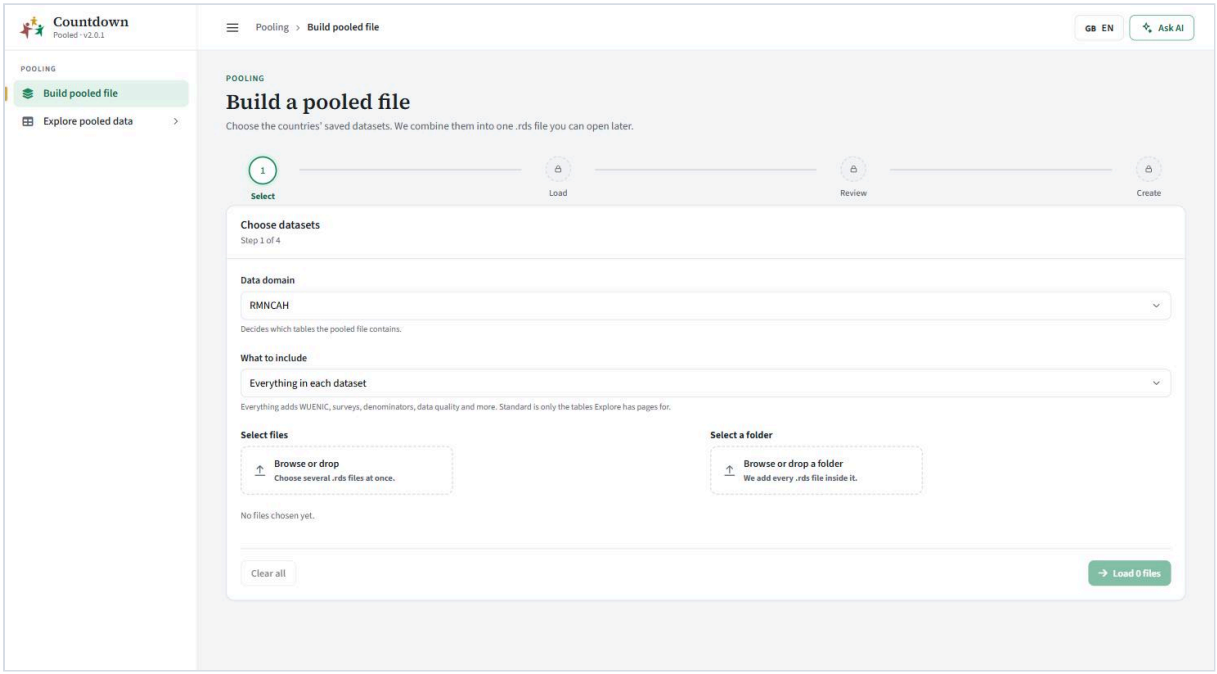


Figure 33.2 — *Choose the countries' saved datasets*

- **Data domain**: RMNCAH or Immunization (VAXX). This decides which tables the pooled file contains. With Immunization, the tables that only exist for RMNCAH (mortality and service utilization) are not built.
- **What to include**: **Everything in each dataset** adds WUENIC, surveys, denominators, data quality and more. **Standard tables only** keeps just the tables that Explore has pages for.
- **Select files** to pick several `.rds` files at once, or **Select a folder** to add every `.rds` file in it. A file with the same name replaces the earlier one in the list.

Each listed file has a tick to include it and a button to remove it. The counter shows the total size against the 2 GB limit. **Clear all** empties the list. Click **Load N files**.

③ Load the files (step 2 of 4)

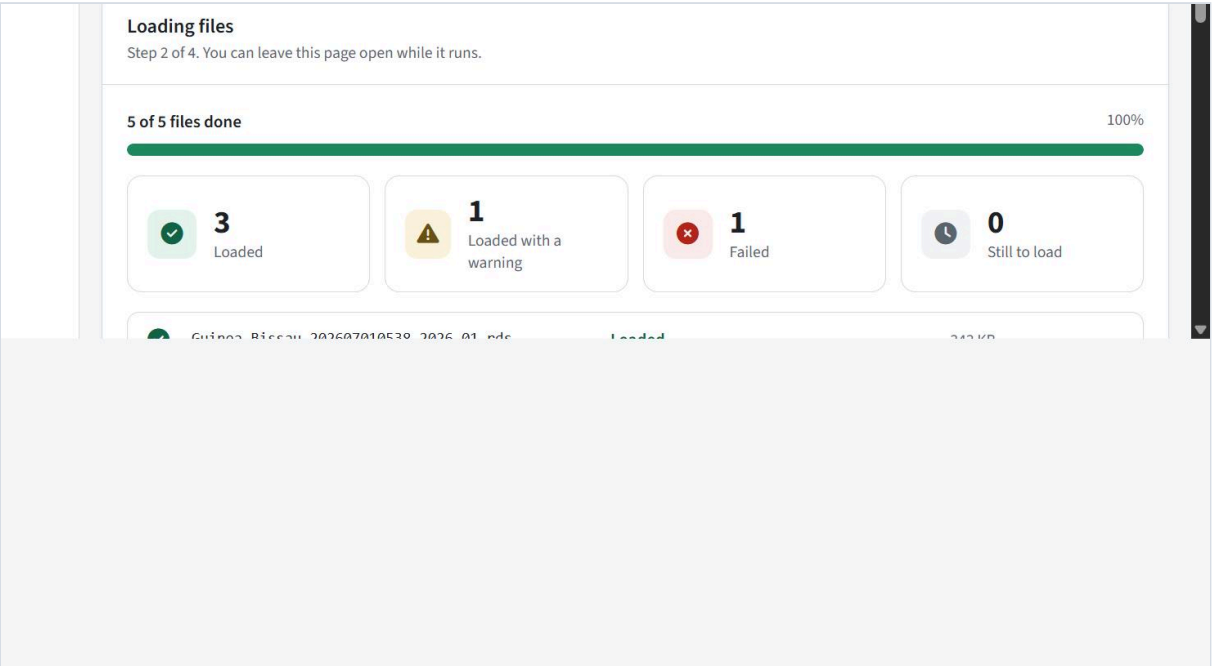


Figure 33.3 — *Loading the selected files*

Files are read one at a time and each shows its result as it finishes, with an overall progress bar and four counters: Loaded, Loaded with a warning, Failed, Still to load. You can leave the page open while it runs. **Cancel loading** returns to step 1; a failed file has a retry option. Then click **Review results**.

④ Review and confirm (step 3 of 4)

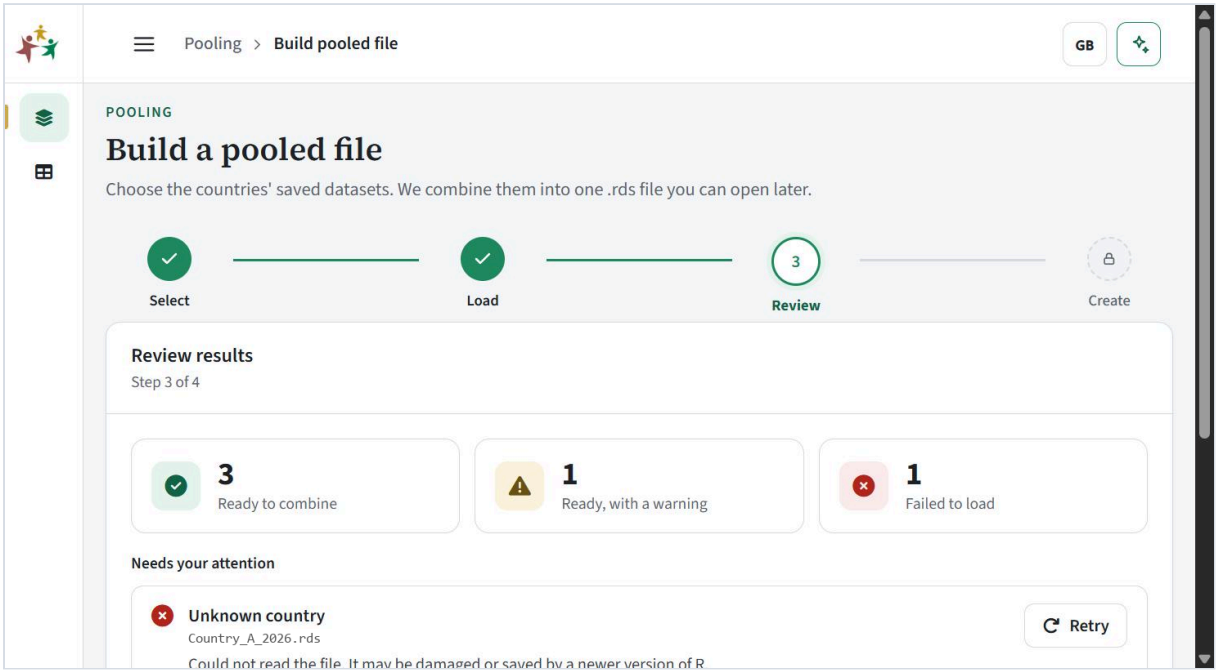


Figure 33.4 — *Review results before combining*

The **Needs your attention** list explains every warning and failure. Typical messages:

Message	Meaning and fix
Not adjusted yet, so its tables use unadjusted data	Finish the adjustment in the country app, save, and reload
N of M standard tables could not be built	The cache lacks some analysis; open it in the country app
Has N of the M tables the loaded files have	Other countries have extra tables this one lacks; usually fine
Another file for the country is newer	Duplicate country; the newer file is used
Not a Countdown dataset	The file has no country or adjusted data; it is not a country cache
Could not read the file	Damaged, or saved by a newer R; re-save it

You can leave a file out. Click **Combine N files**. If some files are left out, a dialog confirms and notes that they are recorded in the file's log with the reason.

⑤ File created (step 4 of 4)

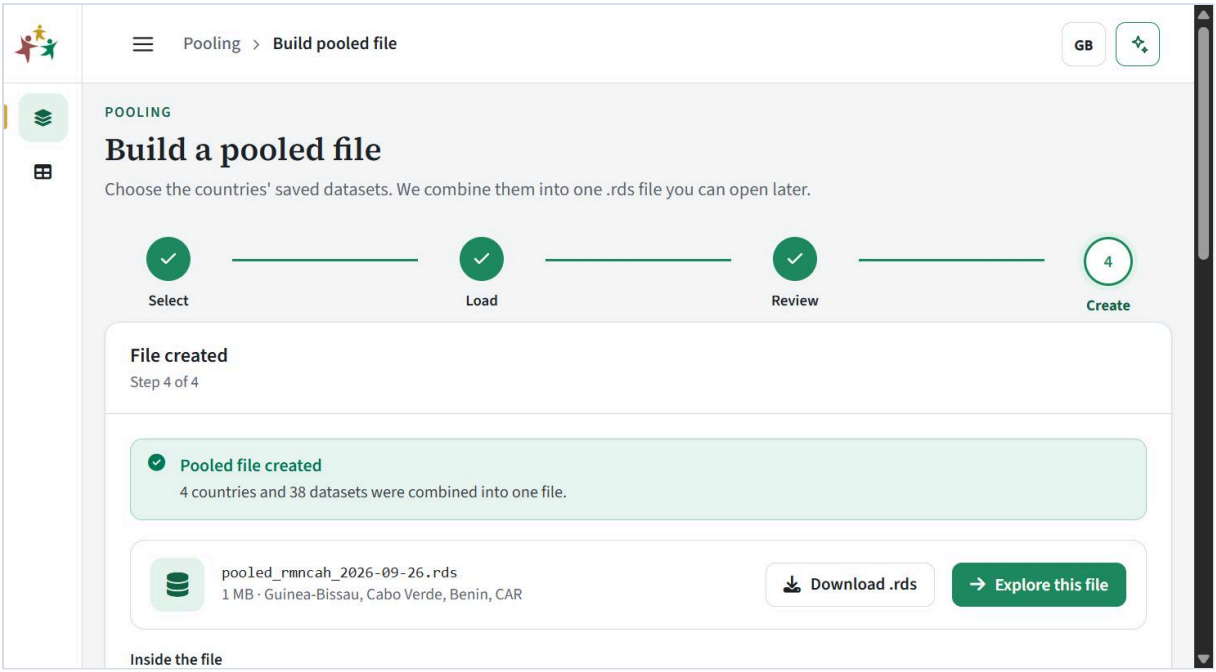


Figure 33.5 — *File created, ready to explore or download*

The app reports the number of countries and datasets and lists each table with its row count. The file is named `pooled_rmncah_` (or `pooled_vaccine_`) followed by the date, ending in `.rds`. Use **Download .rds** to keep it, or **Explore this file**. **Build another file** starts again.

> The pooled file is built in temporary storage for the session. Download the `.rds` if you want to open it again later.

Explore pooled data

Open a file on **Open a file**: use the one just built, or upload a pooled `.rds` made earlier. Only files made by **Build pooled file** are accepted; any other `.rds` is rejected with "This is not a pooled file". After opening, the app jumps to the first page that has data.

Every dataset page has the same layout: a file bar (name, countries, datasets, **Change file**, **Download**), Countries and Years filters, tabs for National, Admin 1 and District where the table has them, a summary strip (Rows, Countries, Years, Columns), graphs, and a searchable **Data** table. The Countries and Years filters are shared by all pages, and choosing none means all.

Pages with numbers also have a **Measure** filter. Graphs use one line or bar per country, with a fixed colour per country. For Admin 1 and District tables the graphs show the median across areas.

- **Parameters**: settings and survey values per country, with survey coverage dots, neonatal mortality and survey year.

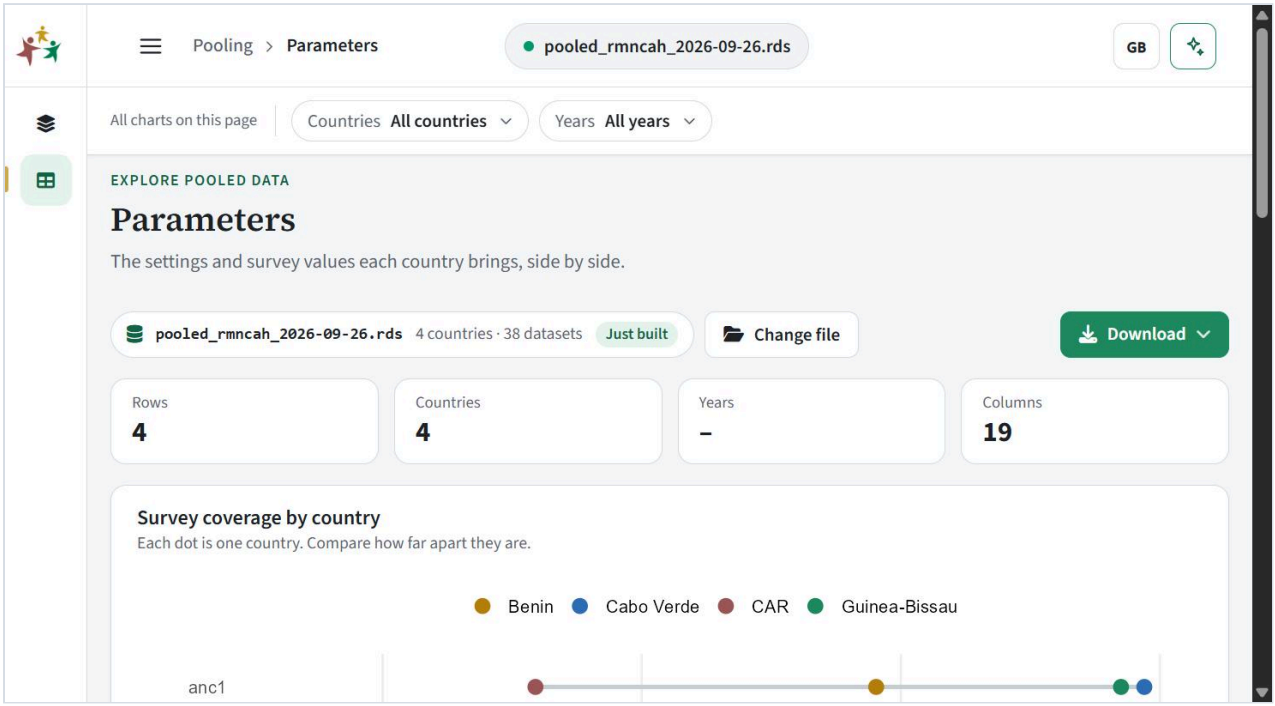


Figure 33.6 — *Parameters page*

- **Overall score**: ranking, and a heatmap of the parts of the score.

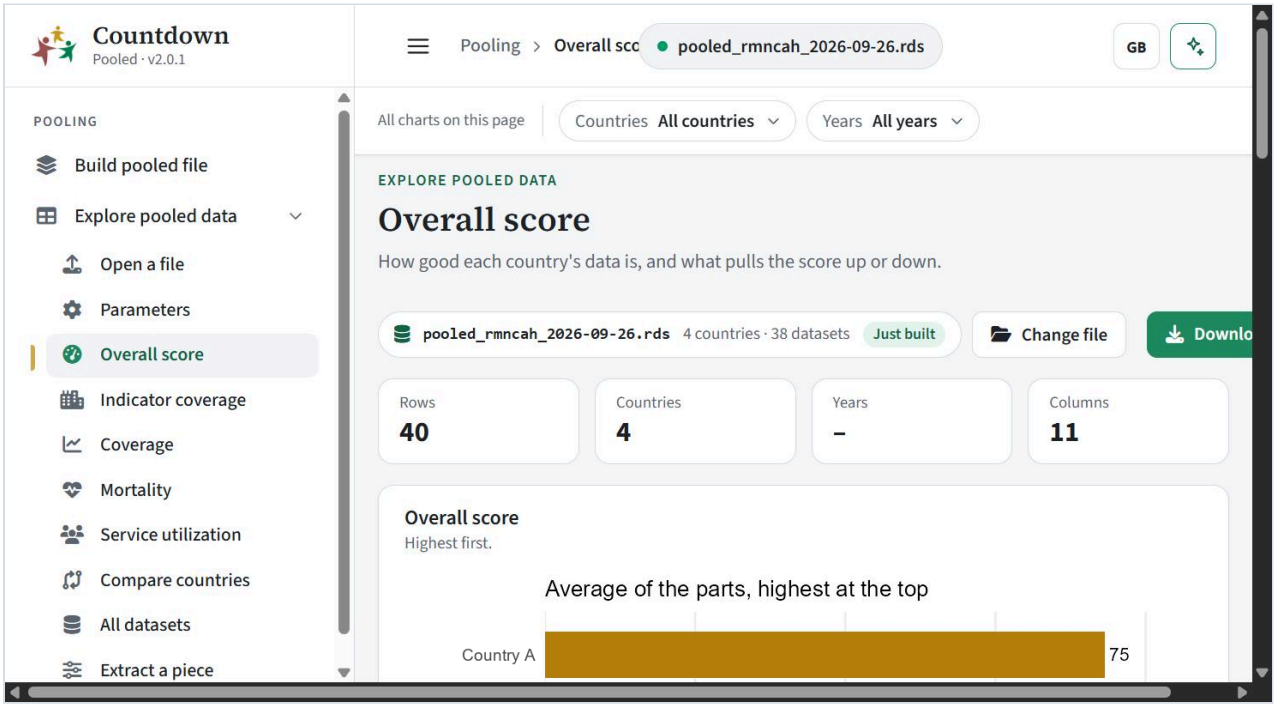


Figure 33.7 — *Overall score page*

- **Indicator coverage** and **Coverage**: trend over time, ranking of the latest year, and either spread across districts or a change-since-first-year table.

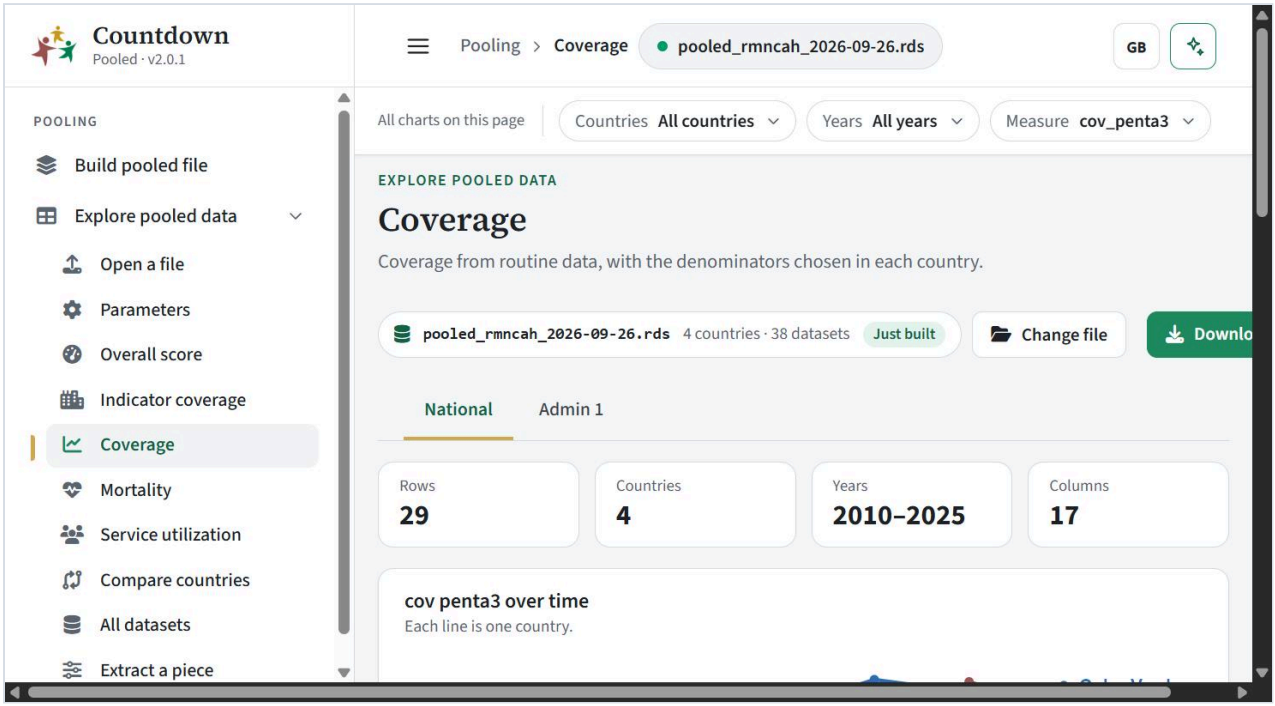


Figure 33.8 — *Coverage page*

- **Mortality** and **Service utilization**: trend, ranking and change.
- **Compare countries**: four tabs. *Continuum of care* is a heatmap of the latest facility coverage from the first antenatal visit to the vaccines. *Source triangulation* shows facility, survey and WUENIC values for DTP3 and MCV1 with the 90% target line. *Level and change* plots coverage now against change over the period. *Trends over time* shows the median and the middle half of countries. It needs the `Coverage - National` table.

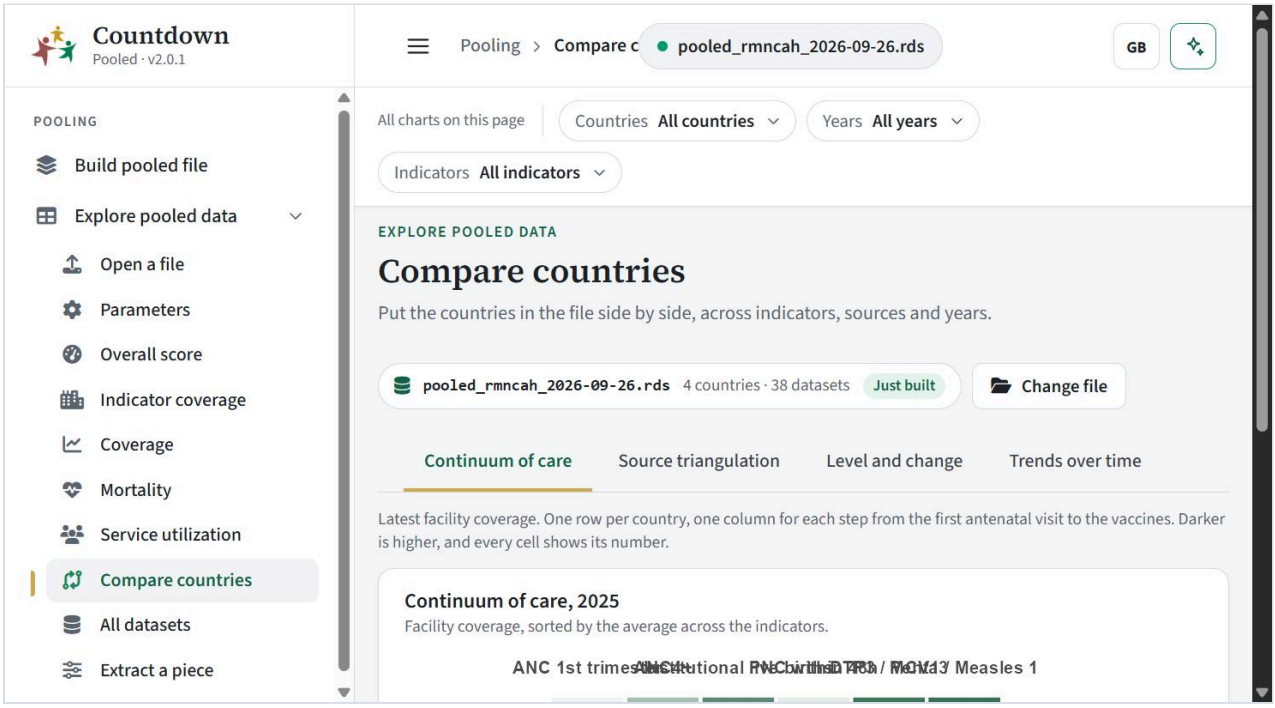


Figure 33.9 — *Compare countries page*

- **All datasets**: pick any table from the grouped list and see its data.

Reading the output

Key columns you will see:

Column	Meaning
`country`, `iso3`	Country name and ISO3 code, always first
`adminlevel_1`, `district`	Area, when the table is subnational
`year`	Reporting year
`cov_penta3`, `cov_measles1`, …	Facility coverage; the same name in every country, whichever denominator the country chose

Coverage indicators are ANC 1st trimester, ANC4+, institutional delivery, C-section, HIV test, institutional live births, PNC within 48h, Penta1, Penta3, Measles 1 and 2, BCG, OPV1-3, syphilis test and IPT2-3 (Admin 1 excludes OPV). The `Parameters` table holds each country's performance threshold, k-factors, national estimates such as `nmr` and `sbr`, survey estimates, `survey_year` and the chosen denominators. For the meaning of scores, adjustments and denominators, see the **framework** **(p. 42)**.

Common pitfalls

- **Choosing the wrong domain.** With **Immunization (VAXX)**, RMNCAH datasets lose their mortality and service utilization tables; choose the domain of the files before loading.
- **Closing the app before downloading.** The pooled file lives in temporary storage: use **Download .rds** first.
- **Opening a country's `.rds` in Explore. Open a file** accepts only files made by **Build pooled file**.
- **Reading a median as a country value.** For Admin 1 and District tables, the graphs show the median across areas.

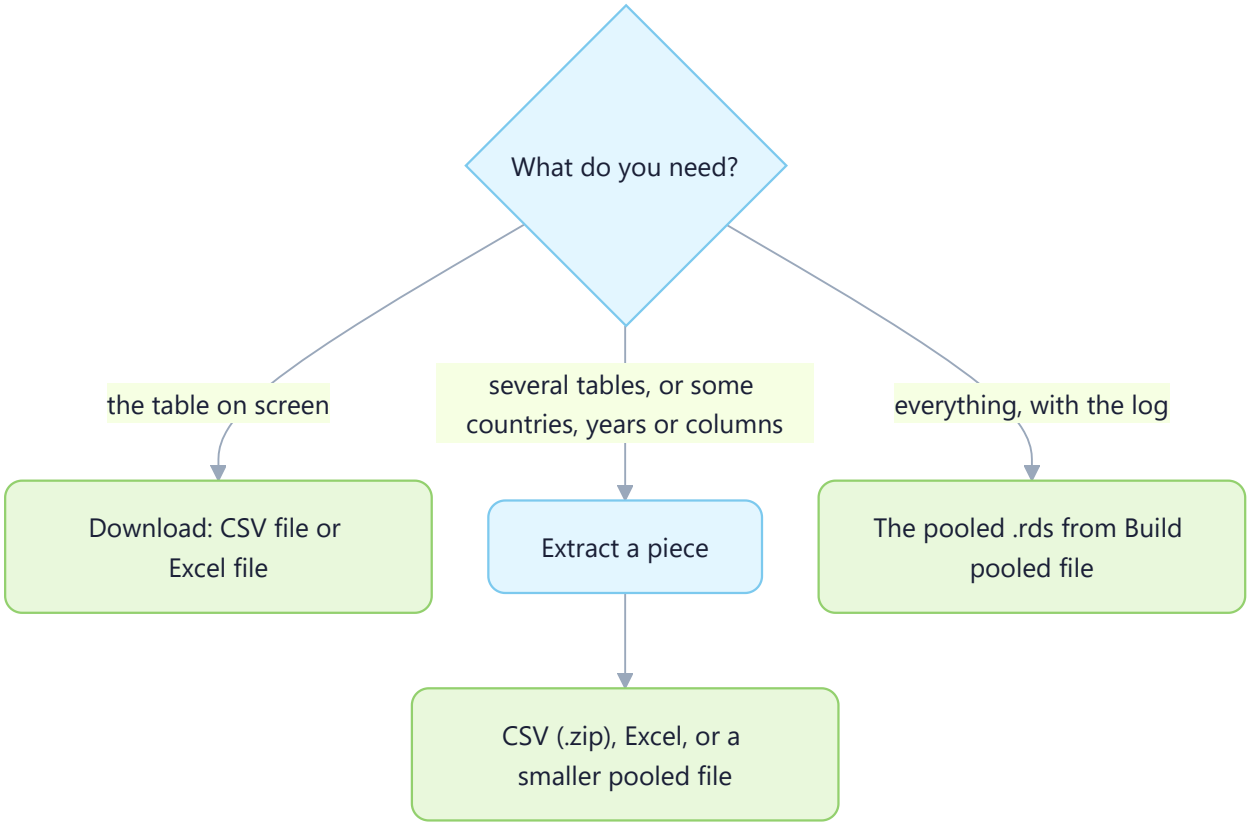
Continue with **Exporting results** **(p. 439)**.

33.3 Exporting results

How to export from the Pooled App, either one table as CSV or Excel, or a selected piece of the pooled file with the datasets, countries and years you choose.

Pooled can export a single table as CSV or Excel, or a selected piece of the pooled file in several formats. Values that are missing are written as empty cells. You export once the pooled file is built and open (p. 431), to take the tables into a regional report or another tool.

Before you start: open a pooled file in **Explore pooled data** (Using the app (p. 435)).



Quick downloads of one table

- On a dataset page, use **Download** and choose **CSV file** or **Excel file**. The export is the table for the tab you are viewing, with the Countries and Years filters applied.
- On **All datasets**, the **CSV** and **Excel** buttons export the table currently selected in the list, also filtered.

Format	File name	Contents
CSV	the table name and date, for example `Coverage_-_National_2026-01-31.csv`	One table, header row, one row per country, year and area
Excel	`pooled_rmncah_2026-01-31.xlsx`	One sheet named after the table

Extract a piece

Use **Extract a piece** when you need several tables, or only part of them. The chips at the top choose **Datasets**, **Countries** and **Years**; choosing nothing means all. Under **Columns**, keep **All columns** or choose the columns per dataset: country, year and area columns are always kept. The **Preview** shows up to six rows per dataset.

The **Get the piece** card counts datasets, countries, years and rows, then offers three formats:

Format	Result
CSV	One CSV for one dataset, or a `.zip` with one CSV per dataset for several
Excel	One workbook with one sheet per dataset
Pooled file	A smaller `.rds` you can reopen in Pooled

Files are named with the domain and date, for example `pooled_rmncah_2026-01-31_piece.zip` or `_piece.xlsx`. Sheet names are limited to 31 characters and the characters not allowed in Excel sheet names are replaced by spaces. Excel sheets have a bold header row, a frozen first row and automatic column widths. **Explore this piece** opens the result in the app. **Start again** resets all filters and columns.

> Excel and CSV files do not contain the build log. To keep the list of countries, the files that were left out and the log, keep the pooled `.rds` from **Build pooled file**.

What is inside the pooled file

The `.rds` is a plain R list, not a country cache. It holds `domain`, the creation time, the app version, a table of `countries` (country, ISO3, source file, warning), a table of `left_out` files with reasons, `datasets` (one table per dataset) and a time-stamped `log`. Tables are ordered as in the app's dataset list: Parameters and score, Indicator coverage, Coverage, Mortality, Service utilization, Data quality, Denominators, Estimates and surveys, Equity, Health systems.

Next steps for analysis

- Read the CSV or Excel file in R, Stata, Python or Excel. `country` and `iso3` identify the country, and coverage columns already share names across countries.
- For a full analysis in R, read the pooled file with `readRDS()` and work on `pooled$datasets`.
- Compare your results with the methods in the **analytical framework** (p. 42), including **national coverage** (p. 115).
- For issues with a single country, return to its cache in the **RMNCAH App** (p. 289), fix it, save, and rebuild the pooled file.

Using the DataSuite chat assistant

The DataSuite chat assistant works on a single country's dataset open in the RMNCAH or Vaxx app; see the **AI assistant** (p. 260) page. The pooled file has a different structure (a list of tables): ask the assistant to read the file and analyse its `datasets` directly, and check the results against the tables you downloaded.

PART IV

Developer resources

Build apps that DataSuite installs and runs, work with Countdown datasets in code, and call the DataSuite API.

CHAPTER 34

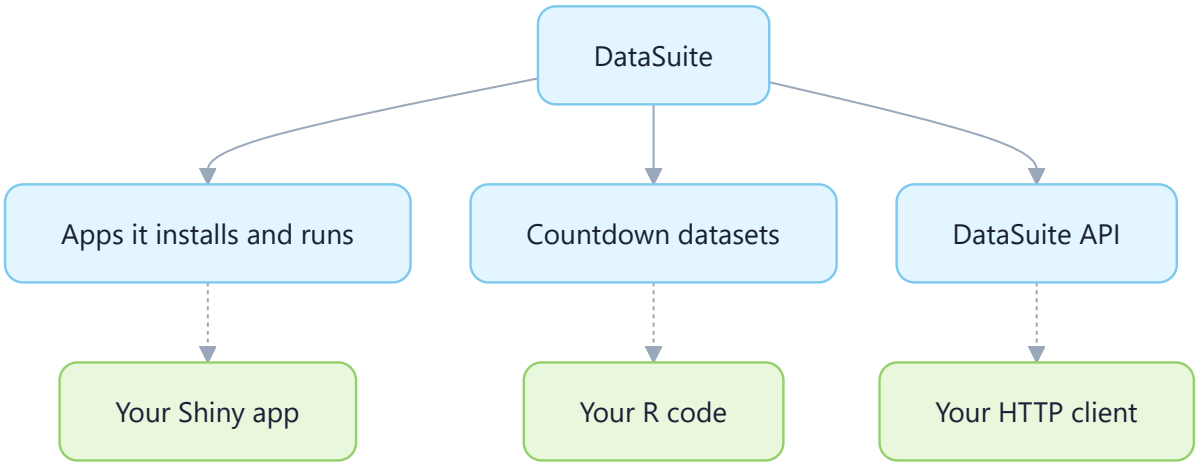
Developer resources

Build apps that DataSuite installs and runs, work with Countdown datasets in code, and call the DataSuite API.

Extend DataSuite with your own code

Package your own Shiny app so DataSuite installs and runs it. Read the Countdown datasets in code, and call the DataSuite API.

Learn about building apps for DataSuite (p. 445)



For AI tools

These docs are also published for AI assistants and coding agents: every page has a Markdown version (add `.md` to its address), and llms.txt <https://datasuite.damurka.com/en/llms.txt> lists them all.

CHAPTER 35

Building apps for DataSuite

How to build a Shiny app that DataSuite installs, updates and runs -- how to package it, declare its dependencies, version it and release it.

DataSuite runs R Shiny apps in its own tabs. An app comes to DataSuite inside an extension: the extension's `package.json` describes the app, and DataSuite installs the app's R code, keeps it up to date and starts it when the user opens a file.

This page is for people who write those apps. It covers what DataSuite expects so that an app installs cleanly on a user's computer and keeps working as R packages change.

> Users install nothing by hand: DataSuite needs R 4.1 or newer on the computer, and installs everything else the app declares into its own R library. Whatever an app does not declare, DataSuite cannot install.

The short version

1. **Make the app an R package** that exports `run_app()`. A plain app folder works too, but a package is easier to install, test and version.
2. **Declare every dependency in DESCRIPTION, each with a minimum version**: `dplyr (>= 1.1.4)`, never a bare `dplyr`.
3. **Never install packages while the app runs**: no `install.packages()`, `pak` or `remotes` in the app's code.
4. **Version with semantic versioning**, and write every change in NEWS.md.
5. **Release through the extension**: a new extension version is what makes DataSuite update users' R packages.

The rest of this page explains each rule.

1. Package the app

As an R package (recommended)

The package exports a `run_app()` function that returns a Shiny app object:

```
#' @export
run_app <- function(...) {
  shiny::shinyApp(ui = app_ui, server = app_server, ...)
}
```

The extension ships a folder with a one-line `app.R`:

```
mypkg::run_app()
```

and names the package in its `package.json`:

```
"contributes": {
  "shinyApps": [{
    "id": "myapp",
    "name": "My App",
    "location": "apps/myapp",
    "accepts": ["xlsx", "rds"],
    "capabilities": { "requiresFile": true },
    "package": {
      "name": "mypkg",
      "version": "1.2.0",
      "repos": ["https://<owner>.r-universe.dev"]
    }
  }]
}
```

- `package.version` is the **oldest** version of the package this release of the extension works with.
- `package.repos` are looked in before CRAN. r-universe <https://r-universe.dev> is the easiest place to publish: it builds Windows, macOS and Linux binaries from your GitHub repository.

As a plain app folder

An `app.R` (or `ui.R` and `server.R`) in the folder named by `location`, with no `package`. DataSuite runs it with `shiny::runApp()`.

Put a DESCRIPTION file next to `app.R` that lists what the app uses, with minimum versions, in the same format a package uses:

```
Depends: R (>= 4.1.0)
Imports:
    dplyr (>= 1.1.4),
    ggplot2 (>= 3.5.0),
    shiny (>= 1.8.0)
```

> DataSuite does not read an app folder's DESCRIPTION yet: today a plain app gets its packages only when it fails with "there is no package called ...", when DataSuite offers to install the missing package. Write the DESCRIPTION now anyway -- it documents the app's needs, renv and rsconnect (Posit Connect) already read it, and DataSuite will install from it.

2. Declare dependencies

In DESCRIPTION:

Field	What goes there
`Depends`	`R (>= 4.1.0)` -- the oldest R the app works with (DataSuite needs 4.1 or newer).

`Imports`	Every package the code needs, **each with a minimum version**.
`Suggests`	Optional features the code checks for with `requireNamespace()` (e.g. PDF export), also with minimum versions. DataSuite installs these for the app's own packages when it can, but never fails because one is missing.
`Remotes`	Packages that are not on CRAN, e.g. `owner/otherpkg`, so installs from GitHub work too.

Why every dependency needs a minimum version

DataSuite updates a CRAN package on a user's computer **only when a DESCRIPTION asks for a newer version than the user has**. It never updates CRAN packages just because a newer one exists -- that would change packages under apps that were never tested with them.

So with a bare `dplyr`, a user who has an old dplyr keeps it, and the first time the app calls a function newer than their copy, it stops with "could not find function" or "unused argument". `R CMD check` does not catch this: it runs against your own, newer library.

How to choose the minimum:

- The version that introduced the newest function or argument your code uses from that package (see the package's NEWS).
- When unsure, the version you develop and test with (`packageVersion("dplyr")`).
- Not a CRAN version released in the last day or two: Windows gets CRAN's prebuilt binaries, which can lag a release.
- Packages that come with R (base, stats, utils, graphics, grDevices, tools, methods) need no minimum: they come with R, so `Depends: R (>= ...)` covers them.

When you start using something newer from a package, raise its minimum **in the same change**.

Call functions explicitly

Use `pkg::fun()` or `@importFrom pkg fun` in package code, never `library()` or `require()`. That way `R CMD check` tells you when something is used but not declared.

3. Don't install at run time

Don't call `install.packages()`, `pak::pkg_install()`, `remotes::install_*()` or `devtools::*` from the app. DataSuite installs what the app declares -- into its own library, in a separate R process, and only when no app tab has the packages loaded (Windows can't replace a loaded package). An app that installs packages itself can break other apps and fail on computers without the right permissions.

4. Version numbers

Use semantic versioning <https://semver.org> for the package:

Change	Bump	Example
Breaks code or data that uses it: a renamed or removed function, changed arguments, a saved-file format that older versions can't read	major	1.4.2 → 2.0.0

Adds something without breaking anything	minor	1.4.2 → 1.5.0
Fixes only	patch	1.4.2 → 1.4.3

Write every release in NEWS.md; for a major version, say what users or other packages must change.

The extension has its own version (in its `package.json`), bumped on the same rules for its own changes. An extension release that only raises `package.version` is a patch.

5. Release

1. Update your development library to current CRAN (`update.packages()`), so you test what a new user would get.
2. Run `devtools::check()`: 0 errors, 0 warnings, 0 notes.
3. Bump `Version:`, write NEWS.md, commit, tag `vX.Y.Z` and push. Wait for your repository (e.g. r-universe) to build it.
4. If you publish several packages that depend on each other, release them in dependency order and wait for each build before releasing the next -- a package published before the version it requires fails to install.
5. In the extension, raise `package.version` if users must have the new version, bump the extension's version and publish it.

How updates reach users

When	What DataSuite does
The extension is installed	Installs the app's package and everything it declares.
The extension is updated (any new version)	Updates the app's package and every dependency from its own repositories to the newest version; installs or updates any CRAN package that is missing or older than a DESCRIPTION requires.
Before every launch	Checks the package is installed at `package.version` or newer; installs only if it isn't. Works offline once set up.
An app tab is open	Waits: updates run once the app's tabs are closed.
Offline	Keeps what is installed if it is new enough, and tries again later.

So publishing a new package version alone reaches **new** installs only (they get the newest); existing users get it with the next extension release. That is deliberate: users only ever run combinations you released together.

Keeping an app working over time

Minimum versions protect against packages that are **too old**. They don't protect against a package's **new** version breaking your app -- and apps outside CRAN get no warning when a CRAN package they use changes.

- Run `R CMD check` on a schedule against current CRAN, e.g. a weekly GitHub Actions workflow (r-lib/actions <https://github.com/r-lib/actions> has ready-made ones). A breaking release then shows up as a failed run, usually before your users meet it.
- Don't pin exact versions or add upper limits (`dplyr (< 1.2)`): DataSuite only understands minimums, and upper limits conflict with other packages that need newer versions.

> Planned in DataSuite: installing CRAN packages as of a date the app declares (a dated CRAN snapshot), so every user gets exactly the versions the author tested with; a separate R library per extension, so one app's updates can't change another's packages; and installing a plain app folder's dependencies from its DESCRIPTION. Following the conventions on this page now means nothing changes for your app when they arrive.

Python apps

DataSuite also runs apps written with Shiny for Python <https://shiny.posit.co/py/>, in the same tabs and with the same restart, crash recovery and environment variables as R apps. Users need **Python 3.10 or newer** from python.org <https://www.python.org/downloads/> (DataSuite finds it through the `py` launcher, the `PATH` or the usual install folders, or the `datasuite.pythonPath` setting); DataSuite installs everything else.

The app's folder holds an `app.py` whose Shiny app object is named `app`:

```
from shiny import App, ui, render

app_ui = ui.page_fluid(ui.output_text("hello"))

def server(input, output, session):
    @render.text
    def hello():
        return "Hello"

app = App(app_ui, server)
```

and the extension marks it with `"runtime": "python"`:

```
"shinyApps": [{
  "id": "myapp", "name": "My App", "location": "apps/myapp", "runtime": "python",
  "accepts": ["csv"],
  "package": { "name": "mypkg", "version": "1.2.0" }
}]
```

Dependencies go in a `requirements.txt` next to `app.py`, each with a minimum version, for the same reasons as above:

```
pandas>=2.1
plotly>=5.18
```

A `pyproject.toml` with a `[project] dependencies` list works too (when both exist, `requirements.txt` is used; reading `pyproject.toml` needs Python 3.11 or newer). Shiny for Python itself is always installed. `package` is optional: a pip package the app comes from, with `version` its oldest supported version and `index` an extra package index to look in besides PyPI; `app.py` then only imports the app, e.g. `from mypkg.app import app`.

DataSuite installs all of this into its own virtual environment, one per Python version (`<user data>/python/venv/<version>`), never into the user's Python, with the same rules as R: when the extension is installed or updated (the package is brought up to its newest version), a quick check before every launch that downloads nothing when every requirement is met, and working offline once set up. If a running app imports a module that isn't installed, DataSuite offers to install the pip package of the same name -- if the name differs (e.g. `sklearn` is `scikit-learn`), list it in `requirements.txt`.

> On Windows, keep paths short: the Python kernel can't open files whose full path is longer than 260 characters, so a deeply nested app folder or package can fail to import.

What DataSuite gives a running app

DataSuite starts the app in a fresh R (or Python) session with these environment variables (`Sys.getenv()` in R, `os.environ` in Python):

Variable	Value
`CDSUITE_SHINY_ID`	The app's `id`.
`CDSUITE_SHINY_NAME`	The app's `name`.
`CDSUITE_SHINY_SELECTED_FILE`	The file the user opened, if any.
`CDSUITE_SHINY_WORKSPACE_DIR`	A folder for the app's work on that file, `<file name>.shiny-workspace` next to it.
`CDSUITE_SHINY_LOCALE`	The display language, e.g. `en`, `fr`, `pt`.

Other `shinyApps` fields:

Field	Purpose
`description`, `icon`	Shown in menus, tooltips and the Start page (`icon` is a codicon id).
`showOnStartPage`	Also list the app on the Start page.
`capabilities`	`requiresFile`, `canSelectFiles`, `canSelectFolders`, `canSelectMany`: what the file picker allows.
`cache`	`{ "file": "<stem>.rds" }`: the file the app saves its work to next to the data file. Opening the data file again opens the saved file.
`aiModelRole`	`fast`, `balanced` or `reasoning`: the least capable AI model the app's analysis can use. While the app's tab is active, the chat's Auto model choice uses at least this.

Checklist before publishing

- ☐ The app runs from `run_app()` (package) or `app.R` (folder), with nothing sourced from outside it.
- ☐ Every package the code uses is in DESCRIPTION, with a minimum version.
- ☐ `devtools::check()` is clean, against an up-to-date library.
- ☐ No `install.packages()` or similar anywhere in the app.
- ☐ Version bumped and NEWS.md written.
- ☐ The package is built in its repository before the extension is published.
- ☐ The extension's `package.version` is raised if users need this version.

CHAPTER 36

CacheConnection reference

Every data point of a Countdown dataset (the cd2030.core CacheConnection) -- what it is, its grain and unit, its default, where it is set and shown in the app, and the methodology section it comes from.

CacheConnection reference

A Countdown dataset (the `.rds` file the RMNCAH and Vaxx apps save) is a `CacheConnection`: the data, the choices made on each page and everything computed from them. This page defines each of its data points. The Countdown AI reads the same definitions.

Generated from cd2030.core 1.4.5. In R, `cd2030.core::cache_definition("adjusted_data")` returns one definition; `cache_manifest()` lists every member with its arguments. Members marked *draft* are not yet fully described.

Data

Tables the analysis produces or starts from. Their columns are explained in the Data dictionary (p. 483).

Member	What it is	Grain / unit	Default	In the app	Computed from
`adequacy_ratios`	The internal consistency ratios (ANC1/Penta1, Penta1/Penta3, and OPV1/OPV3 for immunization) each year, with the share of districts whose ratio is in the adequate range.	one row per year, unit: ratio; % of districts	1, 1.5 ratio (`dq_ratio_adequate_range`); anc1/penta1, penta1/penta3 (`dq_ratio_pairs`)	shown on Internal Consistency, Overall Data Quality Score	`countdown_data`
`adjusted_data`	The adjusted numerators: data_with_excluded_years with each count corrected for incomplete reporting with k (reporting rates below the cutoff replaced first) and extreme outliers replaced; carries each indicator's median, MAD and outlier flag. Every coverage, mortality, utilization and health-system result is computed from it.	one row per district x year x month, unit: counts (adjusted)	75 % (`adj_reporting_rate_cutoff`); 5 x MAD (`adj_outlier_mad_multiplier`)	set on Data Adjustment (Adjust data), shown on Data Adjustment, Data Adjustment Changes	`data_with_exclu` `k_factors`

`adjustment_values`	The yearly numbers of every indicator as reported (after the years and areas removed) and after each step of the dataset's adjustment: completeness, outliers, missing values (the adjusted numbers).	one row per year, unit: counts	shown on Data Adjustment Changes, Reports	`countdown_data`, `adjusted_data`	
`admin1_estimates`	The same rates per first-level region, from the regional survey where it has them, else the national rates.	one row per adminlevel_1, unit: proportion		`regional_survey`, `national_estimates`	analysis setup: survey parameters (p. 54)
`completeness_admin1`	The share of monthly values that are not missing, for each indicator, each year, per first-level region.	one row per adminlevel_1 x year, unit: % non-missing	shown on Data Missingness	`countdown_data`, `performance_threshold`	data quality: missing data (p. 62)
`completeness_district`	The share of monthly values that are not missing, for each indicator, each year, per district.	one row per district x year, unit: % non-missing	shown on Data Missingness	`countdown_data`, `performance_threshold`	data quality: missing data (p. 62)

`completeness_national`	The share of monthly values that are not missing, for each indicator and on average, each year, nationally.	one row per year, unit: % non-missing	shown on Data Missingness, Service Utilization DQA	`countdown_data`, `performance_threshold`	data quality: missing data (p. 62)
`countdown_data`	The HMIS data as loaded: for each district and month, the service counts, the reporting rates (anc_rr, idelv_rr, vacc_rr, opd_rr, ipd_rr) and the population and health-system figures, merged from the uploaded sheets. Nothing is removed or adjusted in it.	one row per district x year x month, unit: counts (as reported)	set on Load Data (Finish merges the uploaded sheets), shown on Reporting Rate, Data Missingness, Internal Consistency, Outlier Detection, Overall Data Quality Score, Data Adjustment		analysis setup: facility data (p. 46)
`data_with_excluded_years`	countdown_data without the excluded years: the data the adjustment starts from.	one row per district x year x month, unit: counts (as reported)	shown on Data Adjustment, Data Adjustment Changes	`countdown_data`, `excluded_years`	data adjustment: remove years (p. 93)
`denominator_metrics`	The population figures compared on the Denominator Assessment page: DHIS2 projections (total population, births, live births, under-1) next to the UN estimates, with the UN growth rate, nationally each year.	one row per year (national), unit: counts; % growth	shown on Denominator Assessment, Reports	`adjusted_data`, `un_estimates`	denominators: population projections (p. 100)

`district_completeness`	The share of districts with no missing values, for each indicator and on average, each year.	one row per year, unit: % of districts	
`district_outliers_summary`	The share of districts with no extreme outlier in the year, for each indicator and on average.	one row per year, unit: % of districts	5 x MAD (`dq_outlier_mad_multi`
`district_reporting_rate`	The share of districts whose reporting rate reaches the threshold, for each service group and on average, each year.	one row per year, unit: % of districts	90 % (`dq_reporting_thresho`
`health_system_comparison`	Health-system inputs next to service coverage (dis_cov_<indicator>_<denominator>) per district and region, for the latest year: the data behind the PHC performance comparison.	one row per district (latest year), unit: coverage %; densities	
`health_system_metrics_admin1`	The same health-system inputs and densities per first-level region.	one row per adminlevel_1 x year, unit: counts; per 10,000 or 100,000 population	
`health_system_metrics_national`	The health-system inputs nationally: facilities, hospitals, beds and health workforce, their densities per population, and service use, for the latest year.	one row per year, unit: counts; per 10,000 or 100,000 population	
`indicator_coverage_admin1`	Coverage of every indicator by every denominator, per first-level region each year (no UN denominator below national).	one row per adminlevel_1 x year, unit: counts; coverage %	

`indicator_coverage_district`	Coverage of every indicator by every denominator, per district each year.	one row per district x year, unit: counts; coverage %	
`indicator_coverage_national`	Coverage of every indicator by every denominator, nationally each year: the adjusted numerators, the population figures, the denominators (tot<population>*<denominator>) and the coverage (cov*<indicator>_<denominator>).	one row per year (national), unit: counts; coverage %	0.025 proportion (`den_nmr`); 0.024 proportion (`den_pnmr`); 0.02 proportion (`den_sbr`); 0.015 proportion (`den_twin`); 0.03 proportion (`den_preg_loss`); 0.98 proportion (`den_anc1survey`); 0.97 proportion (`den_dpt1survey`)
`inequality_admin1`	Coverage by first-level region with the inequality measures against the national value for every indicator and denominator: difference from national, population share, MADM and MRDM (mean absolute and mean relative difference from the national mean), each year.	one row per adminlevel_1 x year, unit: coverage %; differences in percentage points	
`inequality_district`	The same inequality measures per district, against the national value.	one row per district x year, unit: coverage %; differences in percentage points	
`list_outlier_units`	Every district-month with each indicator's value, its median and MAD, and its extreme-outlier flag (1 = more than the bound in MADs from the median).	one row per district x year x month, unit: counts; flags 0/1	5 x MAD (`dq_outlier_mad_multiplier`)
`mortality_ratios`	The latest national institutional MMR and SBR next to the UN estimates (best, lower and upper), for the completeness comparison.	one row per indicator (latest year), unit: per 100,000 live births (MMR); per 1,000 (SBR)	0.5, 1, 1.5, 2 ratio (`mort_completeness_ratios`)

`mortality_summary`	Institutional mortality per first-level region and year: live births, stillbirths (total, fresh), maternal and neonatal deaths, and the institutional rates (iMMR, iSBR, neonatal) and ratios (stillbirths and neonatal deaths per maternal death).	one row per adminlevel_1 x year (and national), unit: counts; per 100,000 live births (iMMR); per 1,000 (iSBR)	25 per 100,000 live births (`mort_mmr_low`); 6 per 1,000 births (`mort_sbr_low`)	shown on Institutional Mortality, Mortality Mapping, Reports	`adjusted_data`	mortali outliers (p. 201)
`outliers_admin1`	The share of monthly values that are not extreme outliers, for each indicator, each year, per first-level region.	one row per adminlevel_1 x year, unit: % non-outlier	5 x MAD (`dq_outlier_mad_multiplier`)	shown on Outlier Detection	`countdown_data`	data quality: outliers (p. 64)
`outliers_district`	The share of monthly values that are not extreme outliers, for each indicator, each year, per district.	one row per district x year, unit: % non-outlier	5 x MAD (`dq_outlier_mad_multiplier`)	shown on Outlier Detection	`countdown_data`	data quality: outliers (p. 64)

`outliers_national`	The share of monthly values that are not extreme outliers (more than the outlier bound in MADs from the district's median), for each indicator, each year, nationally.	one row per year, unit: % non-outlier	5 x MAD (`dq_outlier_mad_multiplier`); all years except the latest (`dq_outlier_baseline`)	shown on Outlier Detection, Overall Data Quality Score	`countdown`
`overall_score`	The overall data quality score table: each metric of the score (1a-1c completeness, 2a-2b outliers, 3a-3d consistency) by year, and the annual score.	one row per score metric (columns are years), unit: % (ratios for 3a-3b)	1a, 1b, 1c, 2a, 2b, 3c, 3d (`dq_score_components_rmncah`); 1a, 1b, 1c, 2a, 2b, 3f, 3g, 3h (`dq_score_components_vaccine`); 90 % (`dq_reporting_threshold`)	shown on Overall Data Quality Score, Reports	`reporting` `district_` `district_` `outliers_` `district_` `adequacy_` `performan`
`ratios_summary`	The consistency ratios each year next to the ratios expected from the survey coverage (the extra row), as the Internal Consistency page compares them.	one row per year, plus the expected ratios, unit: ratio	0.98 proportion (`dq_ratio_expected_coverage_anc1`); 0.97 proportion (`dq_ratio_expected_coverage_penta1`); 0.89 proportion (`dq_ratio_expected_coverage_penta3`)	shown on Internal Consistency, Reports	`adequacy_` `survey_es`
`reporting_rate_admin1`	The average reporting rate of each service group and their mean, each year, per first-level region.	one row per adminlevel_1 x year, unit: %		shown on Reporting Rate	`countdown`

`reporting_rate_district`	The average reporting rate of each service group and their mean, each year, per district.	one row per district x year, unit: %	shown on Reporting Rate	`countdown_data`	data quality: reporting completeness (p. 59)
`reporting_rate_national`	The average reporting rate of each service group (ANC, institutional delivery, vaccination, OPD) and their mean, each year, nationally.	one row per year, unit: %	shown on Reporting Rate, Overall Data Quality Score	`countdown_data`	data quality: reporting completeness (p. 59)
`service_utilization_admin1`	The same service-use measures per first-level region each year.	one row per adminlevel_1 x year, unit: counts; per capita; %	shown on Sub-National Utilization, MCH Preventive vs Curative Index, Reports	`adjusted_data`	service utilisation: inpatient (p. 215)
`service_utilization_national`	Outpatient and inpatient service use nationally each year: visits and admissions (total and under-5), per population, under-5 shares, case fatality and the reporting rates.	one row per year, unit: counts; per capita; %	shown on Service Utilization DQA, National Utilization, Reports	`adjusted_data`	service utilisation: outpatient (p. 210)

`subnational_regions`	The dataset's geography: every district with the first-level region (adminlevel_1) it belongs to.	one row per adminlevel_1 x district	shown on the region and district filters of every sub-national page	`countdown_data`

Calculations

Functions that compute a result for the options given (indicator, level, denominator, years).

Member	What it is	Grain / unit	Default	In the app	Computed from	Methodology
`bayes_estimates`	The estimates of the Bayesian models already fitted at an admin level, as one table: coverage by year from 2010 to 2030, in percent, with the 95% and 80% intervals, for each indicator fitted with the denominator selected now (a row per region at admin level 1). NULL when none is fitted; nothing is fitted here.			shown on Bayesian Analysis	`bayes_model_key`, `get_denominator`	
`bayes_model_cached`	The Bayesian model already fitted for an indicator (by get_bayes_model() or kept with keep_bayes_model()), or NULL: nothing is fitted.			shown on Bayesian Analysis	`bayes_model_key`	
`bayes_model_inputs`	What fitting one indicator's Bayesian model needs (its coverage data, the overall score, the indicator and denominator, the key), so the app fits it in another R process without freezing.			shown on Bayesian Analysis	`calculate_coverage`, `overall_score`, `get_denominator`, `bayes_model_key`	

`bayes_model_key`	The key a Bayesian model is kept under: its admin level, indicator and chosen denominator.		shown on Bayesian Analysis	`get_denominator`	
`bayes_national_routine`	The country's coverage of one of the Bayesian model's indicators from the facility data, by year, as a proportion, with the indicator's denominator: what a sub-national model's charts show beside the regions. NULL when coverage cannot be computed yet.		shown on Bayesian Analysis	`indicator_coverage_national`, `get_denominator`	
`calculate_completeness_summary`	The share of non-missing monthly values at a level: what completeness_* hold, on demand.	national: one row per year; adminlevel_1: per adminlevel_1 x year; district: per district x year, unit: % non-missing	shown on Data Missingness	`countdown_data`	dat qua mis dat 62)
`calculate_coverage`	Coverage of every indicator by every denominator next to the survey estimates and WUENIC, at a level: the table the coverage charts compare.	one row per year and source (national); per region or district x year, unit: coverage %	shown on National Coverage, Reports	`indicator_coverage_national`, `national_survey`, `regional_survey`, `wuenic_estimates`	cov sur and targ (p.

`calculate_derived_coverage`	One indicator's coverage under each of the six denominators next to the survey value (with its bounds), by year: the table behind the Denominator Selection charts.	one row per year, unit: coverage %		sh De Se Re
`calculate_district_completeness_summary`	The share of districts with no missing values, optionally within one region: what district_completeness holds, on demand.	one row per year, unit: % of districts		sh Da Mi
`calculate_district_outlier_summary`	The share of districts with no extreme outlier, optionally within one region: what district_outliers_summary holds, on demand.	one row per year, unit: % of districts	5 x MAD (`dq_outlier_mad_multiplier`)	sh Ou De
`calculate_district_reporting_rate`	The share of districts reaching the reporting threshold, optionally within one region: what district_reporting_rate holds, on demand.	one row per year, unit: % of districts	90 % (`dq_reporting_threshold`)	sh Re Ra Re
`calculate_indicator_coverage`	Coverage of every indicator by every denominator at a level (national, adminlevel_1 or district), optionally for one region: what indicator_coverage_* hold, computed on demand.	national: per year; adminlevel_1: per adminlevel_1 x year; district: per district x year, unit: counts; coverage %		
`calculate_inequality`	The inequality measures for a level against the level above: what inequality_admin1 and inequality_district hold, computed on demand (optionally for one region).	per region or district x year, unit: coverage %; percentage points		

`calculate_outliers_summary`	The share of monthly values that are not extreme outliers at a level: what outliers_* hold, on demand.	national: one row per year; adminlevel_1: per adminlevel_1 x year; district: per district x year, unit: % non-outlier	5 x MAD (`dq_outlier_mad_multiplier`)	shown on Outlier Detection	`count`
`calculate_overall_score`	The overall data quality score table, nationally or for one region: what overall_score holds, on demand.	one row per score metric (columns are years), unit: %	1a, 1b, 1c, 2a, 2b, 3c, 3d (`dq_score_components_rmncah`); 1a, 1b, 1c, 2a, 2b, 3f, 3g, 3h (`dq_score_components_vaccine`)	shown on Overall Data Quality Score, Reports	`repo` `distr` `distr` `outl` `distr` `adeq`
`calculate_ratios_and_adequacy`	The consistency ratios with the share of districts in the adequate range, optionally within one region: what adequacy_ratios holds, on demand.	one row per year, unit: ratio; % of districts	1, 1.5 ratio (`dq_ratio_adequate_range`)	shown on Internal Consistency	`count`
`calculate_reporting_rate`	The average reporting rate of each service group at a level, optionally for one region: what reporting_rate_* hold, on demand.	national: one row per year; adminlevel_1: per adminlevel_1 x year; district: per district x year, unit: %		shown on Reporting Rate	`count`
`calculate_service_dqa_summary`	The data quality table for the service-use data (OPD, IPD): reporting, completeness and outliers by year, nationally or for one region.	one row per quality metric (columns are years), unit: %		shown on Service Utilization DQA, Reports	`repo` `comp` `outl` `serv`

`check_indicator_emptiness`	The indicators with no value anywhere in the dataset.	unit: indicator names	shown on Load Data > Data Quality	`countdown_data`	analysis setup: data checks (p. 48)
`check_population_service_collision`	District-years where a service count looks like population data (a sign the sheets were mixed up).	one row per district x year flagged	shown on Load Data > Data Quality	`countdown_data`	analysis setup: data checks (p. 48)
`check_shapefile_admin_names`	The shapefile region names that do not match a region of the dataset (and are not matched yet), with the closest dataset name and its distance.	one row per unmatched name	shown on Load Data > Shapefile, Load Data > Map Shapefile	`shapefile`, `shapefile_name_field`, `map_mapping`, `subnational_regions`	analysis setup: survey files (p. 50)

`check_survey_admin_names`	The survey region names that do not match a region of the dataset (and are not matched yet), with the closest dataset name and its distance.	one row per unmatched name	shown on Load Data > Survey Files, Load Data > Map Survey	`regional_survey`, `survey_mapping`, `subnational_regions`	analysis setup: survey f (p. 50)
`compute_service_utilization`	The service-use measures at a level (national or adminlevel_1): what service_utilization_* hold, computed on demand.	per year, or per adminlevel_1 x year, unit: counts; per capita; %		`adjusted_data`	service utilisatio data qu (p. 207)
`decompose_change`	Which districts (or regions) drive a change in an indicator's coverage between two years: each unit's counts and denominators in both years, its coverage change and its contribution in percentage points, split into service delivery and denominator, with a reporting-rate flag.	one row per district (or adminlevel_1), unit: counts; coverage %; percentage points		`indicator_coverage_district`, `indicator_coverage_admin1`, `reporting_rate_district`	subnatio coverag 168)

`denominator_comparison`	The Denominator Selection comparison in one tidy table: each indicator's coverage under each denominator option next to the survey estimate of that year and the difference, with the chosen denominator marked. Use it to compare or judge the denominators.	one row per indicator x denominator x year (x region or district), unit: coverage %; percentage points	shown on Denominator Selection	`calculate_derived_coverage`, `denominator`, `maternal_denominator`, `survey_year`
`filter_mortality_summary`	One mortality indicator by region for the chosen years, joined to the map, for the mortality maps.	one row per adminlevel_1 x year, unit: per 100,000 live births (iMMR); per 1,000 (iSBR)	shown on Mortality Mapping, Reports	`mortality_summary`, `mortality_mapping_years`, `map_mapping`
`filter_service_utilization`	The service-use measures for one indicator at a level, optionally for one region.	per year (or per region x year), unit: counts; per capita; %	shown on National Utilization, Reports	`service_utilization_national`, `service_utilization_admin1`
`generate_admin1_mch_curative_index`	The MCH preventive services index and the curative services index per region.	one row per adminlevel_1 x year, unit: index	shown on MCH Preventive vs Curative Index, Reports	`service_utilization_admin1`
`generate_admin1_service_utilization`	Under-5 OPD (or IPD) use per capita per region, for the regional utilization chart.	one row per adminlevel_1 x year, unit: per capita	shown on National Utilization	`service_utilization_admin1`

`generate_coverage_data`	The continuum of care: coverage of the maternal (type maternal) or child (type child) indicators along the continuum, with the chosen denominator, in long form.	one row per year x indicator x source, unit: coverage %	shown on Continuum of Care, Reports	`calculate_coverage`, `denominator`, `maternal_denominator`	cov cor of 139
`generate_health_system_table`	The national health-system table: facility, hospital and bed densities, workforce density and skills mix, and the private and NGO shares of facilities, each with its unit.	one row per health-system indicator, unit: per 10,000 or 100,000 population; %	shown on National Health System, Reports	`health_system_metrics_national`	hea sys inp 228
`generate_phc_scatter_data`	Per region, a health-system density (facilities or workforce) against service coverage, with outliers flagged, for the PHC performance scatter.	one row per adminlevel_1, unit: density; coverage %	shown on PHC Performance, Reports	`health_system_metrics_admin1`	hea sys out inp 235
`generate_private_sector_data`	Per region, the share of facilities that are public, private for profit and NGO.	one row per adminlevel_1 x facility type, unit: %	shown on Public/Private Share, Reports	`health_system_metrics_admin1`	hea sys inp 228

`get_base_indicator_coverage`	The coverage table at a level (indicator_coverage_* or, for one region, computed): the numerators, populations, denominators and coverage before the survey comparison.	per year, region x year or district x year, unit: counts; coverage %		`indicator_coverage_national`, `indicator_coverage_admin1`, `indicator_coverage_district`	denc deriv denc (p. 10
`get_bayes_model` (draft)	The Bayesian model of one indicator's coverage (anc4, anc_1trimester, ideliv, measles1 or penta3) with its chosen denominator, nationally or by region, combining the routine data and the surveys; fitted once and kept in the dataset.		shown on Bayesian Analysis	`calculate_coverage`, `overall_score`, `get_denominator`	baye cove estim 145)
`get_filtered_coverage`	One indicator's coverage with its chosen denominator next to the survey and WUENIC estimates, by year: the table under the National Coverage chart.	one row per estimate (columns are years), unit: coverage %	shown on National Coverage, Reports	`calculate_coverage`, `get_denominator`	cove surve targe 134)
`get_filtered_inequality`	One indicator's coverage per region (or district) with its chosen denominator and the inequality measures: the data of the inequality charts.	one row per region or district x year, unit: coverage %; percentage points	shown on National Inequality, Reports	`inequality_admin1`, `inequality_district`, `get_denominator`	subn ineq 172)
`get_filtered_mapping_data`	One indicator's regional coverage with its chosen denominator for the chosen years, joined to the map.	one row per adminlevel_1 x year, unit: coverage %; geometry	shown on National Inequality, Reports	`get_mapping_data`, `mapping_years`, `get_denominator`	

`get_filtered_threshold`	The share of districts (or regions) reaching the global coverage target for an indicator group (anc4, instlivebirths, vaccine, dropout), each year.	one row per year, unit: % of units	90 % (`cov_target_vaccine_national`); 80 % (`sub_target_vaccine`); 70 % (`cov_target_anc4`); 80 % (`cov_target_instlivebirths`); 10 % (`cov_target_dropout`)	shown on Coverage Target, Reports	`get_base_indic` `get_denominato`
`get_high_performers`	The regions (or districts) whose coverage of an indicator reaches its target, each year.	one row per unit x year reaching the target, unit: coverage %	90 % (`cov_target_vaccine_national`); 80 % (`sub_target_vaccine`); 70 % (`cov_target_anc4`); 80 % (`cov_target_instlivebirths`); 10 % (`cov_target_dropout`); 80 % (`cov_target_other`)	shown on Coverage Target	`get_base_indic` `get_denominato`
`get_mapping_data`	Regional coverage joined to the map regions (through map_mapping) for the coverage maps.	one row per adminlevel_1 x year, unit: coverage %; geometry		shown on National Inequality	`indicator_cove` `map_mapping`, `s`
`get_regional_estimates`	One region's rates from the survey files: its own mortality rates and survey coverage (the latest survey value of each). A rate its surveys do not give keeps the national value; the national rates are returned whole only when the region is not in the survey files.	unit: proportion			`national_estim` `regional_surve` `survey_mapping` `subnational_re` `start_survey_y`

`keep_bayes_model`	Keeps a Bayesian model fitted in another R process under its key, so the page and reports use it without fitting it again.			shown on Bayesian Analysis	`bay`
`list_missing_units`	The district-months where an indicator is missing, optionally within one region.	one row per district x year x month		shown on Data Missingness	`cou`
`prepare_mapping_service_utlization`	OPD or IPD use per region for the chosen years, joined to the map.	one row per adminlevel_1 x year, unit: per capita; geometry		shown on National Utilization, Reports	`ser` `uti` `map`
`summarise_completeness_ratio`	The institutional MMR (or SBR) as a share of the UN estimate for each assumed completeness ratio: how complete death reporting would have to be to match the UN level.	one row per completeness ratio, unit: per 100,000 live births	0.5, 1, 1.5, 2 ratio (`mort_completeness_ratios`)	shown on Mortality Completeness, Reports	`mor`

Settings and values

Choices and numbers the analysis uses, with their defaults (each methodology page lists its settings).

Member	What it is	Grain / unit	Default	In the app	Computed from	Methodology

`adjustment_settings`	What the adjustment does: the years removed everywhere, an area's data removed for some or every year (its services and population), and each indicator's completeness k, outlier and missing-value correction -- everywhere, or with a region's or district's own settings (a district's win over its region's).	unit: settings	set on Data Adjustment (Adjust data), shown on Data Adjustment, Data Adjustment Changes, Reports		data adjustment: completeness adjustment (p. 81)
`country`	The country the dataset is for (its name), from the Admin sheet.	unit: country name	shown on the app header and the reports	`countdown_data`	
`country_iso`	The country's ISO3 code; it selects the built-in reference data (UN estimates, WUENIC, surveys, shapefile) for the country.	unit: ISO3 code		`countdown_data`	
`data_years`	The years present in the loaded data, in order.	unit: years	shown on Overall Data Quality Score, Data Adjustment, Service Utilization DQA, National Utilization, National Health System, Mortality Mapping, Reports	`countdown_data`	

`denominator`	The denominator chosen for the immunization and child indicators (one of the denominator ids: dhis2, anc1, penta1, anc1derived, penta1derived, un).	unit: denominator id	set on Denominator Selection (Select Best Vaccination Denominator), shown on Denominator Selection, National Coverage, Continuum of Care, National Inequality, Coverage Target, Reports		denominators: best denominator (p. 108)
`derivation_population`	The population the growth-projected denominators (ANC1/Penta1 population growth) are built from: one of totbirths_dhis2, totlivebirths_dhis2, totunder1_dhis2, totpop_dhis2, un_population, un_births, un_under1.	unit: column name	set on Denominator Assessment (Select Best Population), shown on Denominator Assessment, Reports		denominators: derived denominators (p. 105)
`excluded_years`	Years removed from the analysis because their data are unreliable; every result after the adjustment leaves them out (the removed years of adjustment_settings).	unit: years	set on Data Adjustment (Data kept for analysis), shown on Data Adjustment		data adjustment: remove years (p. 93)
`get_chart_options`	The chart options that apply to one chart: the dataset's default options with the chart's own on top.			`chart_options`	
`get_denominator`	The chosen denominator for an indicator: maternal_denominator for maternal and newborn indicators, denominator for the others.	unit: denominator id		`denominator`, `maternal_denominator`	denominators: best denominator (p. 108)

`is_default`	Whether a field is still at its built-in value (nothing uploaded or set for it).	unit: logical				
`k_factors`	The completeness adjustment factor k for each service group (anc, idelv, vacc, and opd/ipd for RMNCAH): the assumed ratio of service volume in non-reporting facilities to that in reporting ones.	unit: factor (0 to 1)	0 (`adj_k_start`)	set on Data Adjustment (K-Factors), shown on Data Adjustment, Data Adjustment Changes, Reports		data adjustment: completeness adjustment (p. 81)
`lbr_mean`	The mean national coverage of institutional live births over the years, with the maternal denominator: the completeness assumption of the mortality comparison.	unit: %		shown on Mortality Completeness	`indicator_coverage_national`, `maternal_denominator`	mortality: completeness (p. 194)
`mapping_years`	The years the coverage maps show.	unit: years		set on National Inequality (map years), shown on National Inequality, Reports		

`maternal_denominator`	The denominator chosen for the maternal and newborn indicators (a denominator id, as for denominator).	unit: denominator id		set on Denominator Selection (Select Best Maternal Denominator), shown on Denominator Selection, National Coverage, Continuum of Care, PHC Performance, Reports	
`mortality_mapping_years`	The years the mortality maps show.	unit: years		set on Mortality Mapping, shown on Mortality Mapping	
`national_estimates`	The national rates used to compute the denominators: neonatal (nmr) and post-neonatal (pnmr) mortality, stillbirth rate (sbr), twin rate, pregnancy loss, and the ANC1 and Penta1 survey coverage (from survey_estimates).	unit: proportion	0.015, 0.03 proportion (`den_app_twin_preg_loss`)	set on Load Data > National Rates, shown on Load Data > National Rates	`survey_estima`
`performance_threshold`	The reporting completeness threshold: a district counts as reporting well in a year when its reporting rate is at least this.	unit: %	90 % (`dq_reporting_threshold`)	set on Reporting Rate (Performance Threshold), shown on Reporting Rate, Overall Data Quality Score, Service Utilization DQA, Reports	

`shapefile_name_field`	Which column of the shapefile holds the region names (NAME_1 for the built-in shapefile).	unit: column name	set on Load Data > Shapefile, shown on Load Data > Shapefile, Load Data > Map Shapefile	
`start_survey_year`	The first survey year used: survey data before it are left out of the comparisons.	unit: year	set on Load Data > National Rates (Survey Data Start Year), shown on Load Data > National Rates	
`survey_estimates`	The national survey coverage of the indicators the denominators and consistency checks need (anc1, penta1, penta3, bcg, measles1, instlivebirths, and anc4, csection, low_bweight or opv1, opv3 by indicator group), from the most recent survey.	unit: % (low_bweight: proportion)	set on Load Data > National Rates (pre-filled from the national survey), shown on Load Data > National Rates, Internal Consistency	analysis setup: survey parameters (p. 54)

`survey_year`	The year of the survey the ANC1/Penta1 coverage comes from: the base year the derived denominators start from.	unit: year	2019 year (`den_survey_year`)	set on Load Data > National Rates (Recent Survey Year), shown on Load Data > National Rates, Denominator Selection, Reports		denominators derived denominators (p. 105)
`survey_years`	The years that have national survey data.	unit: years		shown on Load Data > National Rates	`national_survey`	
`utilization_mapping_years`	The years the service utilization maps show.	unit: years		set on National Utilization, shown on National Utilization		

Reference data and mappings

Estimates, surveys, shapes and the name matching that connect them to the dataset.

Member	What it is	Grain / unit	Default	In the app	Computed from	Methodology
`area_private_share`	The survey shares of care received in the private and public sectors, by area of residence (built in).	one row per indicator x area, unit: %		shown on Public/Private Share, Reports		health system: inputs (p. 228)
`area_survey`	The survey estimates by area of residence (urban, rural): built in, or uploaded.	one row per area x survey year, unit: %		set on Load Data > Survey Files, shown on Equity Assessment, Reports		equity: wealth quintile (p. 156)

`csection_area_estimates` (draft)	C-section estimates by area, when given to the dataset.			
`csection_national_estimates` (draft)	C-section estimates, nationally, when given to the dataset.			
`education_survey`	The survey estimates by the mother's education (none, primary, secondary+): built in, or uploaded.	one row per education level x survey year, unit: %	set on Load Data > Survey Files, shown on Equity Assessment, Reports	equity: wealth quintile (p. 156)
`fpet_data`	The Family Planning Estimation Tool (FPET) projections for the country: each indicator's median and uncertainty bounds, each year.	one row per indicator x year, unit: %	shown on National Coverage, Reports	coverage: family planning (p. 133)
`map_mapping`	How the shapefile's region names match the dataset's first-level regions; NULL until the names are matched.	one row per map region	set on Load Data > Map Shapefile, shown on Load Data > Map Shapefile	analysis setup: survey files (p. 50)
`national_private_share`	The survey shares of care received in the private and public sectors, nationally (built in).	one row per indicator, unit: %	shown on Public/Private Share, Reports	health system: inputs (p. 228)
`national_survey`	The national household survey estimates (coverage r_<indicator> with standard error se_, bounds ll_/ul_) for each survey year: built in, or uploaded.	one row per survey year, unit: %	set on Load Data > Survey Files, shown on Load Data > National Rates, National Coverage, Denominator Selection	analysis setup: survey files (p. 50)
`regional_survey`	The survey estimates per first-level region (same columns as national_survey): built in, or uploaded.	one row per survey region x survey year, unit: %	set on Load Data > Survey Files, shown on Load Data > Map Survey, Denominator Selection, Sub-National Coverage	analysis setup: survey files (p. 50)
`sector_area_estimates` (draft)	Public/private sector estimates by area, when given to the dataset.			
`sector_national_estimates` (draft)	Public/private sector estimates, nationally, when given to the dataset.			

`shapefile`	The map of the country's first-level regions: the built-in one, or one uploaded to replace it.	one row per map region, unit: geometry	set on Load Data > Shapefile, shown on Load Data > Map Shapefile, National Inequality, Mortality Mapping, National Utilization	analysis setup: survey files (p. 50)
`survey_mapping`	How the survey's region names match the dataset's first-level regions (adminlevel_1); NULL until the names are matched.	one row per survey region	set on Load Data > Map Survey, shown on Load Data > Map Survey	analysis setup: survey files (p. 50)
`un_estimates`	The UN population estimates for the country (population, births, under-1 and under-5 population, women 15-49, growth, crude birth and death rates, total fertility), each year: built in, or uploaded to replace them.	one row per year, unit: counts; rates	set on Load Data > Upload Data (Additional reference data), shown on Denominator Assessment	analysis setup: reference datasets (p. 52)
`un_mortality_estimates`	The UN estimates of maternal mortality (MMR), neonatal mortality and stillbirth rates with their bounds, each year: built in, or uploaded.	one row per year, unit: per 100,000 live births (MMR); per 1,000 (NMR, SBR)	set on Load Data > Upload Data (Additional reference data), shown on Institutional Mortality, Mortality Completeness	analysis setup: reference datasets (p. 52)
`wiq_survey`	The survey estimates by household wealth quintile (Q1 poorest to Q5 richest): built in, or uploaded.	one row per quintile x survey year, unit: %	set on Load Data > Survey Files, shown on Equity Assessment, Reports	equity: wealth quintile (p. 156)
`wuenic_estimates`	The WHO/UNICEF estimates of national immunization coverage (WUENIC) for each vaccine, each year: built in, or uploaded.	one row per year, unit: %	set on Load Data > Upload Data (Additional reference data), shown on National Coverage	analysis setup: reference datasets (p. 52)

State, stores and checks

What the dataset records about itself: saved reports and graphs, revision, checks.

Member	What it is	Grain / unit	Default	In the app	Computed from	Methodology
`adjusted_flag`	Whether the data have been adjusted (TRUE once Adjust data has run with the current years and k-factors).	unit: logical		set on Data Adjustment (Adjust data), shown on Data Adjustment, Data Adjustment Changes		

`cache_path`	Where this dataset is saved: the path of its .rds file (the saved copy the app works on).	unit: file path	set on Load Data (Finish writes the saved copy)	
`chart_options`	What the user changed about how charts look: the dataset's default options and each chart's own, by chart id.		set on the chart options menu of every chart, shown on every chart, and the reports	
`check_coverage_params`	Whether everything the survey comparisons need is present: the inequality inputs plus WUENIC and the national and regional surveys.	unit: logical		`check_inequality_params`, `wuenic_estimates`, `national_survey`, `regional_survey`
`check_inequality_params`	Whether everything coverage by denominator and inequality need is present: adjusted data, a survey year, UN estimates and the national rates.	unit: logical		`adjusted_data`, `survey_year`, `un_estimates`, `national_estimates`
`check_mortality_params`	Whether the mortality analysis can run: adjusted data and the UN mortality estimates are present.	unit: logical		`adjusted_data`, `un_mortality_estimates`
`check_sector_params`	Whether the four public/private and c-section estimate tables are present.	unit: logical		`sector_national_estimates`, `sector_area_estimates`, `csection_national_estimates`, `csection_area_estimates`

`graphs`	The saved custom charts, by id: where each draws its data from, its transforms and its plot (a description, never code); they redraw with the data and can go in any report.		set on the AI (countdown_graph with save) or the report builder, shown on Reports	
`language`	The language of the app's labels for this dataset (two-letter code: en, fr, pt).	unit: language code	set on the language switcher in the app header	
`quality_confirmed`	Whether this dataset finished the Load Data wizard with its quality checks passed (set at Finish). Older datasets and ones built outside the wizard have NULL or FALSE.	unit: logical	set on Load Data (Finish)	
`read_only`	Whether this copy of the dataset never writes its .rds (the AI's copy is read-only).	unit: logical		
`report_assets`	The pictures the reports use, by id (a block's src is asset:<id>).		set on Reports, shown on Reports	
`report_projects`	The reports built in the report builder (or saved by the AI), by id: name, language, design, cover and blocks.		set on Reports, shown on Reports	dissemination: information products (p. 252)
`report_themes`	The report themes made from Office files, by id.		set on Reports, shown on Reports	
`revision`	A number that goes up each time a change to the dataset is saved, so another reader of the same .rds (the AI's copy) knows to reload.	unit: count		
`selected_admin_level_1` (draft)	A first-level region kept as a view selection.	unit: region name		
`selected_district` (draft)	A district kept as a view selection.	unit: district name		

`wizard_parts`	The uploaded sheets while the Load Data wizard is in progress, before Finish merges them into countdown_data; NULL afterwards.	set on Load Data > Upload Data, shown on Load Data > Data Quality	
`wizard_quality_results`	The quality checks as they were at Finish, on the uploaded sheets, so the Data Quality step shows the same results when revisited.	set on Load Data (Finish), shown on Load Data > Data Quality	analysis setup: data checks (p. 48)

Changing a value

Each of these sets or clears the member it names (in the app, the page's control does it):

- `adjust_data` -- changes `data_with_excluded_years`, `k_factors`
- `clear_area_survey` -- changes `area_survey`
- `clear_education_survey` -- changes `education_survey`
- `clear_map_mapping` -- changes `map_mapping`
- `clear_national_survey` -- changes `national_survey`
- `clear_regional_survey` -- changes `regional_survey`
- `clear_shapefile` -- changes `shapefile`
- `clear_survey_mapping` -- changes `survey_mapping`
- `clear_survey_year` -- changes `survey_year`
- `clear_un_estimates` -- changes `un_estimates`
- `clear_un_mortality_estimates` -- changes `un_mortality_estimates`
- `clear_wiq_survey` -- changes `wiq_survey`
- `clear_wizard_parts` -- changes `wizard_parts`
- `clear_wuenic_estimates` -- changes `wuenic_estimates`
- `load_from_disk`
- `reactive`
- `reset_chart_options` -- changes `chart_options`
- `save_to_disk`
- `set_adjusted_data` -- changes `adjusted_data`
- `set_adjusted_flag` -- changes `adjusted_flag`
- `set_adjustment_settings` -- changes `adjustment_settings`
- `set_area_survey` -- changes `area_survey`
- `set_cache_path` -- changes `cache_path`
- `set_chart_options` -- changes `chart_options`
- `set_countdown_data` -- changes `countdown_data`
- `set_csection_area_estimates` -- changes `csection_area_estimates`
- `set_csection_national_estimates` -- changes `csection_national_estimates`
- `set_denominator` -- changes `denominator`

- `set_derivation_population` -- changes `derivation_population`
- `set_education_survey` -- changes `education_survey`
- `set_excluded_years` -- changes `excluded_years`
- `set_fpet_data` -- changes `fpet_data`
- `set_graph` -- changes `graphs`
- `set_k_factors` -- changes `k_factors`
- `set_language` -- changes `language`
- `set_map_mapping` -- changes `map_mapping`
- `set_mapping_years` -- changes `mapping_years`
- `set_maternal_denominator` -- changes `maternal_denominator`
- `set_mortality_mapping_years` -- changes `mortality_mapping_years`
- `set_national_estimates` -- changes `national_estimates`
- `set_national_survey` -- changes `national_survey`
- `set_performance_threshold` -- changes `performance_threshold`
- `set_quality_confirmed` -- changes `quality_confirmed`
- `set_regional_survey` -- changes `regional_survey`
- `set_report_asset` -- changes `report_assets`
- `set_report_project` -- changes `report_projects`
- `set_report_theme` -- changes `report_themes`
- `set_sector_area_estimates` -- changes `sector_area_estimates`
- `set_sector_national_estimates` -- changes `sector_national_estimates`
- `set_selected_admin_level_1` -- changes `selected_admin_level_1`
- `set_selected_district` -- changes `selected_district`
- `set_shapefile` -- changes `shapefile`
- `set_shapefile_name_field` -- changes `shapefile_name_field`
- `set_start_survey_year` -- changes `start_survey_year`
- `set_survey_estimates` -- changes `survey_estimates`
- `set_survey_mapping` -- changes `survey_mapping`
- `set_survey_year` -- changes `survey_year`
- `set_un_estimates` -- changes `un_estimates`
- `set_un_mortality_estimates` -- changes `un_mortality_estimates`
- `set_utilization_mapping_years` -- changes `utilization_mapping_years`
- `set_wiq_survey` -- changes `wiq_survey`
- `set_wizard_country` -- changes `country`
- `set_wizard_parts` -- changes `wizard_parts`
- `set_wizard_quality_results` -- changes `wizard_quality_results`
- `set_wuenic_estimates` -- changes `wuenic_estimates`

CHAPTER 37

Data dictionary

What the ids and column names in the Countdown apps' data mean -- the six denominator options, how column names are built, the target populations, indicators and reporting rates.

Data dictionary

The apps' data names things with short ids -- an indicator (`penta1`), a denominator option (`anc1derived`), a target population (`totinftpenta`) -- combined into column names such as `cov_penta3_penta1derived`. This page says what each part means. Read the meaning here rather than from the spelling: **`anc1` and `penta1` are the ANC1-derived and Penta1-derived denominators, while `anc1derived` and `penta1derived` are the population-growth options.**

Generated from cd2030.core 1.4.5. In R, `cd2030.core::cd_dictionary()` returns these tables and `cd2030.core::cd_describe_columns()` reads column names with them.

Denominators

The six denominator options (Denominator assessment and selection (p. 95)).

Id	Label	What it is	Levels
`un`	UN projections	The UN estimates of live births and population (UN projections). National only.	national
`dhis2`	DHIS2 projections	The population projections in DHIS2 (live births, total population, under-1s).	national, adminlevel_1, district
`anc1`	ANC1-derived	Derived each year from that year's reported ANC1 visits and the survey's ANC1 coverage: pregnancies = ANC1 / ANC1 survey coverage, then deliveries, births and infants from them.	national, adminlevel_1, district
`penta1`	Penta1-derived	Derived each year from that year's reported Penta1 doses and the survey's Penta1 coverage: infants = Penta1 / Penta1 survey coverage, then births, deliveries and pregnancies from them.	national, adminlevel_1, district
`anc1derived`	ANC1 population growth	The ANC1-derived value of the survey year, carried to the other years with the population growth rate of the chosen population (not re-derived from each year's ANC1).	national, adminlevel_1, district
`penta1derived`	Penta1 population growth	The Penta1-derived value of the survey year, carried to the other years with the population growth rate of the chosen population (not re-derived from each year's Penta1).	national, adminlevel_1, district

How column names are built

Pattern	Meaning	Example
`<indicator>`	The number of services or events reported for the indicator (a count).	`penta1`
`cov_<indicator>_<denominator>`	Coverage of the indicator (%) with that denominator option.	`cov_penta3_penta1derived`
`<population>_<denominator>`	The target population (thousands) as estimated by that denominator option.	`totinftpenta_penta1`
`<population>_<denominator>derived`	The same population carried from the survey year with population growth (the population-growth options).	`totinftpenta_penta1derived`
`r_<indicator>`	The indicator's coverage in the survey (%).	`r_penta1`
`ul_<indicator>`	The upper bound of the survey estimate's confidence interval (%).	`ul_penta1`
`ll_<indicator>`	The lower bound of the survey estimate's confidence interval (%).	`ll_penta1`
`<group>_rr`	Reporting rate of that service group (% of expected reports received).	`vacc_rr`
`nat_<name>`	The national value of <name> (intermediate).	`nat_pop`
`admin1_<name>`	The first-administrative-level value of <name> (intermediate).	`admin1_pop`
`<name>_survey`	<name> in the survey year.	`penta1_survey`

Target populations

In thousands: coverage = 100 x count / (population x 1000).

Id	Label	What it is
`totpreg`	Pregnancies	Expected pregnancies: the denominator of the antenatal care indicators.
`totdeliv`	Deliveries	Expected deliveries: the denominator of institutional deliveries and caesarean sections.
`totbirths`	Total births	Expected births, live and stillborn (DHIS2 denominator of BCG, live births, low birth weight, PNC).
`totlbirths`	Live births	Expected live births (ANC1- and Penta1-derived denominator of BCG, live births, low birth weight, PNC).
`totlivebirths`	Live births	Live births in the DHIS2 projections (the base of the DHIS2 denominators).
`totinftpenta`	Infants eligible for Penta	Infants surviving the neonatal period: the denominator of Penta, PCV, rotavirus, IPV, OPV, zero-dose, under-vaccinated and the Penta dropouts.

`totinftmeasles`	Infants eligible for Measles 1	Infants surviving to the age of Measles 1: the denominator of Measles 1 and its dropouts.
`totmeasles2`	Children eligible for Measles 2	Children surviving to the age of Measles 2: the denominator of Measles 2.
`totpop`	Total population	Total population in the DHIS2 projections.
`totunder1`	Children under 1	Children under 1 in the DHIS2 projections.

Indicators

An indicator's id is also the column of its monthly count; `computed` indicators are calculated from the counts.

Id	Label	Group	Reported or computed
`anc1`	ANC 1	anc	reported
`anc_1trimester`	ANC 1st Trimester	anc	reported
`anc4`	ANC 4 Visits	anc	reported
`ipt2`	IPT 2nd Dose	anc	reported
`ipt3`	IPT 3rd Dose	anc	reported
`syphilis_test`	Syphilis Test	anc	reported
`ifa90`	IFA 90+ Days	anc	reported
`hiv_test`	HIV Test	anc	reported
`sba`	Skilled Birth Attendance	idelv	reported
`ideliv`	Institutional Deliveries	idelv	reported
`instlivebirths`	Institutional Live Births	idelv	reported
`csection`	Caesarean Sections	idelv	reported
`low_bweight`	Low Birth Weight	idelv	reported
`pnc48h`	PNC within 48 hours	idelv	reported
`total_stillbirth`	Total Stillbirths	idelv	reported
`stillbirth_f`	Fresh Stillbirths	idelv	reported
`stillbirth_m`	Macerated Stillbirths	idelv	reported
`maternal_deaths`	Maternal Deaths	idelv	reported
`neonatal_deaths`	Neonatal Deaths	idelv	reported
`penta1`	Pentavalent 1	vacc	reported
`penta2`	Pentavalent 2	vacc	reported
`penta3`	Pentavalent 3	vacc	reported
`measles1`	Measles 1	vacc	reported
`measles2`	Measles 2	vacc	reported
`measles3`	MCV 3		reported

`bcg`	BCG	vacc	reported
`ipv1`	IPV 1	vacc	reported
`ipv2`	IPV 2	vacc	reported
`opv1`	OPV 1	vacc	reported
`opv2`	OPV 2	vacc	reported
`opv3`	OPV 3	vacc	reported
`pcv1`	PCV 1	vacc	reported
`pcv2`	PCV 2	vacc	reported
`pcv3`	PCV 3	vacc	reported
`rota1`	Rotavirus 1	vacc	reported
`rota2`	Rotavirus 2	vacc	reported
`opd_total`	OPD Total (Outpatient)	opd	reported
`opd_under5`	OPD Under 5 Years	opd	reported
`ipd_total`	IPD Total (Inpatient)	ipd	reported
`ipd_under5`	IPD Under 5 Years	ipd	reported
`zerodose`	Zero-dose	vacc	computed
`undervax`	Under-vaccinated	vacc	computed
`dropout_penta13`	Penta 1 to Penta 3 Dropout	vacc	computed
`dropout_penta1mcv1`	Penta 1 to Measles 1 Dropout	vacc	computed
`dropout_penta3mcv1`	Penta 3 to Measles 1 Dropout	vacc	computed
`dropout_measles12`	Measles 1 to Measles 2 Dropout	vacc	computed
`instdeliveries`	Institutional Live Births per Delivery		reported

Reporting rates

Id	Meaning
`anc_rr`	Reporting rate, antenatal care (% of expected reports received).
`idelv_rr`	Reporting rate, institutional deliveries (% of expected reports received).
`vacc_rr`	Reporting rate, vaccination (% of expected reports received).
`opd_rr`	Reporting rate, outpatient (% of expected reports received).
`ipd_rr`	Reporting rate, inpatient (% of expected reports received).
`mean_rr`	Reporting rate, all services (mean) (% of expected reports received).

Other columns

Column	Meaning
`country`	The country.

`iso3`	The country's ISO3 code.
`adminlevel_1`	The first administrative level (region, province).
`district`	The district.
`year`	The year.
`month`	The month.
`denominator`	A denominator option id: un = UN projections, dhis2 = DHIS2 projections, anc1 = ANC1-derived, penta1 = Penta1-derived, anc1derived = ANC1 population growth, penta1derived = Penta1 population growth.
`indicator`	An indicator id (see the indicators).
`un_births`	Live births in the UN projections (thousands).
`un_population`	Total population in the UN projections (thousands).
`un_under1`	Children under 1 in the UN projections (thousands).
`un_popgrowth`	Annual population growth rate in the UN projections.
`population_growth_change`	Population change from the survey year to this year (percent, annualised log2), used to carry the population-growth denominators.
`population_proportion`	This unit's share of the national population in the survey year.
`denominator_label`	The denominator option's label (see denominator).
`selected`	Whether this is the denominator chosen for the indicator (denominator or maternal_denominator).
`coverage`	Coverage of the indicator (%) with this row's denominator option.
`survey`	The indicator's survey coverage estimate for this year (%), when a survey has one.
`survey_lower`	The lower bound of the survey estimate's confidence interval (%).
`survey_upper`	The upper bound of the survey estimate's confidence interval (%).
`difference`	Coverage minus the survey estimate (percentage points); positive when the facility-based coverage is higher.
`survey_year`	The survey year the derived denominators start from (the base year of the ANC1/Penta1-derived and population-growth options).

PART V

FAQs

Frequently asked questions about DataSuite, grouped by topic: installing and updating, the Data Extractor, loading data, the analysis and reports.

CHAPTER 38

FAQs

Frequently asked questions about DataSuite, grouped by topic: installing and updating, the Data Extractor, loading data, the analysis and reports.

Short answers to the questions users ask most, each with the page that covers it.

Installing and updating

Q. What do I need to run DataSuite?

Windows 10 or 11 (64-bit), or a recent 64-bit Linux, and R 4.1 or newer: DataSuite runs the analyses in R and does not include it. Python 3.10 or newer and a licensed Stata 17 or newer are needed only for notebooks in those languages. See Requirements (p. 16) and, for Linux, Linux <https://datasuite.damurka.com/en/downloads#linux>.

Q. DataSuite says it needs R. What do I do?

The message "DataSuite needs R 4.1.0 or newer to run its apps and analyses, and none was found" means R is not installed, or DataSuite cannot find it. Choose **Download R** and install it, or **Choose R Folder…** if R is already installed somewhere else, then **Check Again**. If R is older than 4.1, install a newer one.

Q. Should I use the User or the System installer?

The **User** installer, unless several people use the same computer: it needs no administrator rights and updates itself in the background. The **System** installer installs for every user and needs administrator rights to install and update. See Installer options <https://datasuite.damurka.com/en/downloads#installer-options>.

Q. Which version of DataSuite am I running?

Click **Help** and choose **About**.

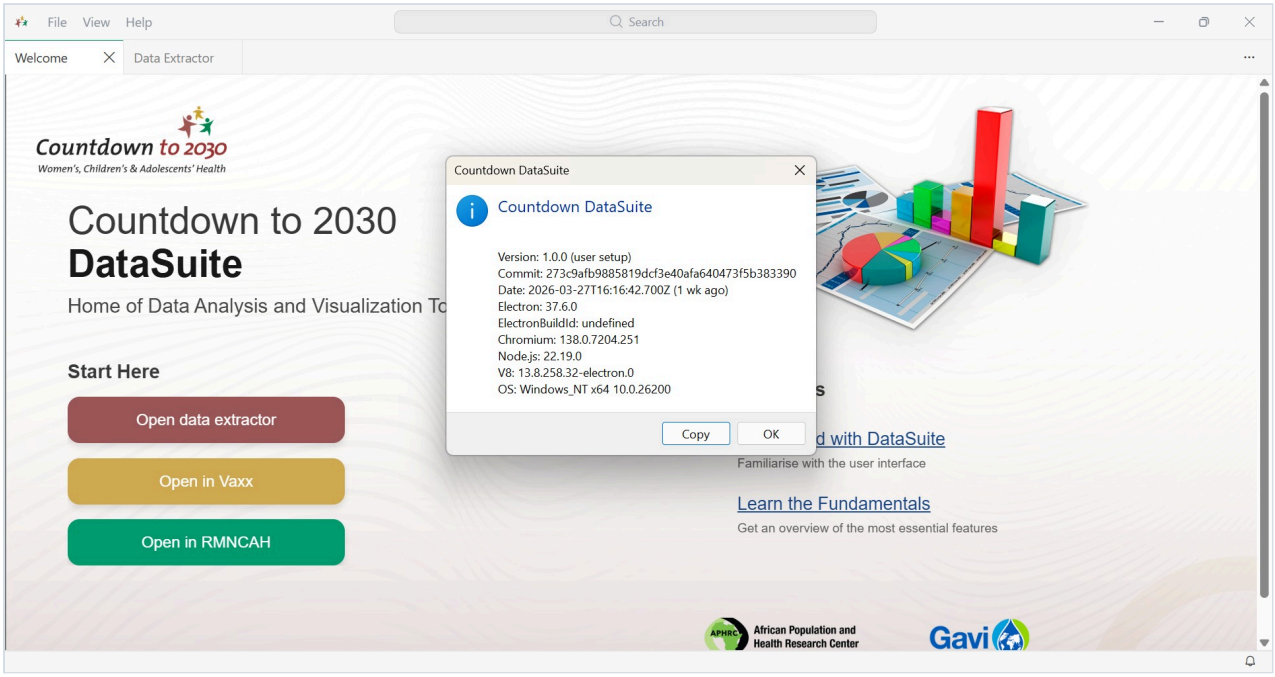


Figure 38.1 — *DataSuite About dialog showing the version number*

Q. How do I update DataSuite?

DataSuite tells you when a new version is out, when you open it. To check yourself, click **Help** and choose **Check for Updates**. The `.zip` (portable) version does not update itself: download the new version and copy your `data` folder into it (see Portable mode (p. 19)). What changed: Release notes <https://datasuite.damurka.com/en/release-notes>.

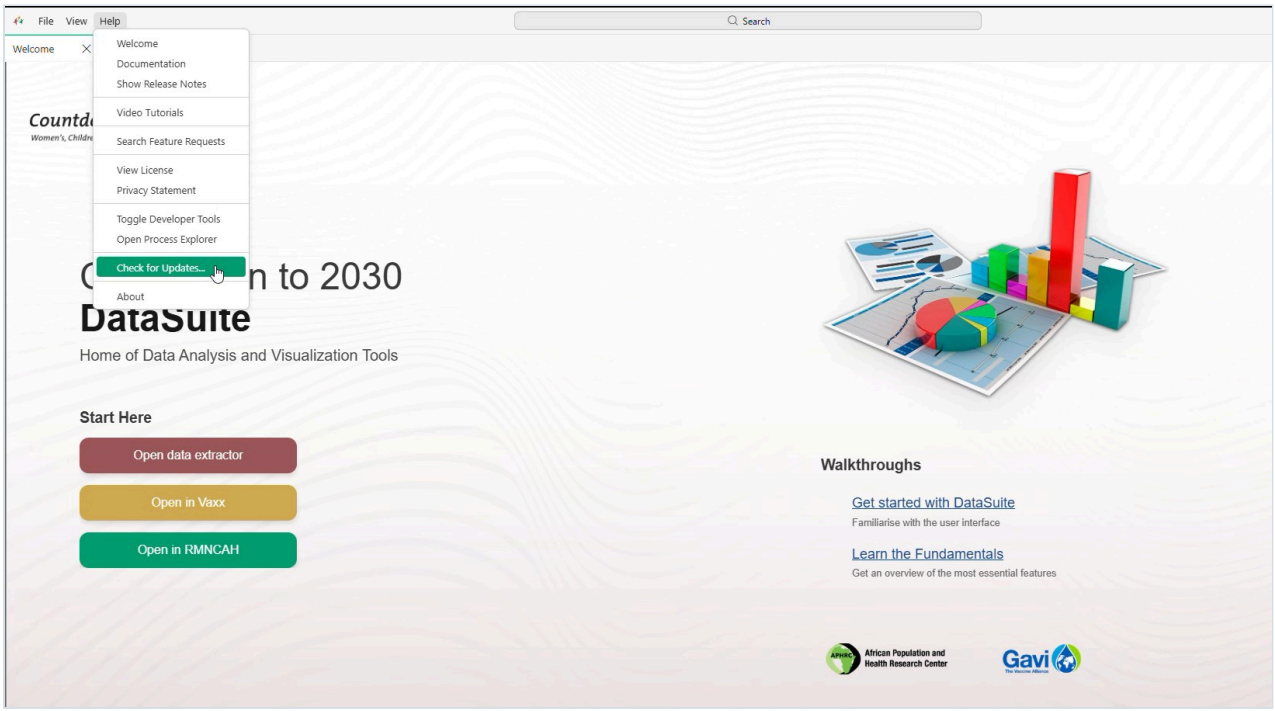


Figure 38.2 — *Help menu with the Check for Updates option highlighted*

Q. How do I remove DataSuite completely, or reset its settings?

Uninstall it <https://datasuite.damurka.com/en/get-started/install/uninstall>, then delete its two user data folders, `%APPDATA%\DataSuite` and `%USERPROFILE%\.datasuite`. Deleting them without uninstalling resets DataSuite to its first start: settings, DHIS2 sign-ins, mappings and the download history go. Your data files (workbooks, saved datasets, downloads) are not in these folders and are kept. See Clean uninstall <https://datasuite.damurka.com/en/get-started/install/uninstall#clean-uninstall>.

Data Extractor

Q. I cannot sign in to DHIS2.

The message says why:

- **The server's address starts with https://** -- type the full address of the DHIS2 server, for example `https://dhis2.example.org`, not the address of a page inside it.
- **Invalid username or password.** -- check them by signing in to DHIS2 in your browser. You can sign in with a personal access token instead (DHIS2 2.39 or later: **Profile** > **Edit profile** > **Personal access tokens**).
- **Unable to reach the DHIS2 server.** -- check your internet connection and the address.

See DHIS2 login (p. 273).

Q. I signed in, but a data element, an indicator or an area is missing.

The extractor sees only what your DHIS2 user may see: access is set in DHIS2, not in DataSuite. Ask your DHIS2 administrator for access. If the item was added to DHIS2 recently, refresh the extractor's copy of the metadata (**Settings** > **Metadata** > **Refresh now**).

Q. A download fails or is very slow.

A failed download says why in **Downloads** > **History**; press **Retry** to run it again. Large downloads (many years, a low level) take time. If the server is slow or shared, lower **Downloads at the same time** and **Values per request** in **Settings** > **Downloads**, or download fewer years at a time. See Downloading data (p. 283).

Q. Where is the file I downloaded?

In `Documents\DataSuite\Downloads`, unless you chose another folder in **Settings** > **Files** > **Save to**. In **Downloads** > **History**, open a finished download as Excel or JSON, or choose **Show in folder**. A download made with a Countdown mapping can also open straight in the RMNCAH or Vaccination app: its workbook is then written to `Documents\Countdown data\<server>`, and the app saves its dataset next to it.

Loading data

Q. Do I load the Excel workbook or the .rds file?

The Excel workbook from the Data Extractor the first time: the app checks it and walks you through the setup (survey values, survey files, shapefile). When you press **Finish Loading Data**, the app saves the dataset as an `.rds` file. Later, open the same workbook again, or the `.rds`: either way the

saved dataset opens, with your choices. See Accepted files (p. 295).

Q. Where is the saved dataset, and what is it called?

In the same folder as the workbook you opened, named after it: `mydata.xlsx` is saved as `mydata_rmncah.rds` by the RMNCAH app and `mydata_vaccine.rds` by the Vaxx app. Keep the workbook and its `.rds` together and do not rename the `.rds`. See The saved .rds file (p. 47).

Q. I corrected the workbook, but the app still shows the old data.

When a saved dataset exists next to the workbook, the app opens it instead of reading the workbook again. Delete (or move) the `.rds`, then open the workbook.

Q. The upload fails.

- **Upload failed: Unsupported file format.** -- the app reads `.xlsx`, `.xls`, `.dta` and `.rds` files only.
- **This dataset was built for the 'vaccine' indicator group, but this app is for 'rmncah'** (or the reverse) -- the `.rds` was saved by the other app. Open it in that app, or load the workbook.
- **Some checks found issues that need fixing before you can continue** -- a blocking check failed. Correct the workbook and load it again.

All the messages and what to do: Troubleshooting (p. 304).

Q. Where do the survey values on the National Rates step come from?

From the country's survey data built into the app, or from the national survey file if you upload one. Check them against the latest survey report (DHS, MICS) and correct them if you have more recent, validated values. What each value means: National survey parameters (p. 54).

Q. Which other data does the analysis use?

Global estimates built into the app for the country: UN population projections, WUENIC immunization estimates, UN mortality estimates, family planning (FPET) estimates and survey results. You can replace the UN, WUENIC and UN mortality estimates with your own `.dta` files on the **Upload Data** step. See Global estimates (p. 52).

Analysis

Q. The analysis pages are greyed out, or show nothing.

Pages unlock in order. Finish loading the dataset first ("Load and finish a dataset first to unlock this page"), then run the adjustment on the **Data Adjustment** page: the denominator, analysis and report pages open after it, even with every adjustment factor at 0. The coverage charts also need the denominators chosen on **Denominator Selection**. See Data preparation (p. 324).

Q. I changed the adjustment factors or removed a year, and nothing changed.

The analysis pages keep the last adjustment until you run it again on the **Data Adjustment** page. Then check the denominators again. See Data preparation (p. 322).

Q. Why does the Vaxx app have no mortality, service utilization or health system pages?

A vaccine analysis covers immunization only. What the two analyses share and where they differ: RMNCAH and vaccine analyses (p. 258).

Reports

Q. Which reports can I make?

The standard Countdown reports: one per analysis (data quality, data adjustment, denominators and inequality; in the RMNCAH app also national coverage, mortality, curative services, the health system and the public/private share), a synthesis chartbook and a sub-national one-pager. You can also start your own document or slide deck. Each opens as an editable copy. See Standard reports (p. 416).

Q. Where do I make and download them?

On the **Reports** page (under **Output**), which opens once the data are adjusted. A **Generate Report** button on an analysis page starts its standard report. Choose **Download**: a document as Word or PDF, a slide deck as PowerPoint or PDF. See Reports (p. 415).

Q. The PDF cannot be made.

A document's PDF is made with Microsoft Word or LibreOffice, a slide deck's with PowerPoint or LibreOffice: install one of them, or download the Word or PowerPoint file. See Formats and downloads (p. 422).

Q. How do I share my reports with a colleague?

Reports are saved inside the dataset (`.rds`): share that file. The colleague opens it in the same app and finds the reports under **Your reports**.

Appendix

APPENDIX A

Glossary

Alphabetical glossary of terms as the DataSuite documentation defines them, from adjusted numerator to zero-dose, each linked to the section it comes from.

Terms as the DataSuite documentation defines them, each linked to the section it comes from. Where a page and the methodology word a term differently, the methodology page applies. For abbreviations and acronyms, see Abbreviations (p. 506).

Term	Definition
Adjusted numerator (N_adj)	N_adj = N_rep + N_rep x (1/C - 1) x k, where N_rep is the raw reported number and C the reporting completeness (proportion of expected reports received) (source (p. 80))
Admin 1	The first level below the national level, such as a region, province or department. Districts (Admin 2) are the level below it.
Admin 2	The administrative level below Admin 1, usually the district. Subnational coverage, inequality and targets are computed for Admin 1 regions and for Admin 2 districts (source (p. 168)).
Admin data	District, region, country; hospitals, health centres, all facilities, private for-profit and non-profit facilities; physicians, nurses/midwives, clinical officers, non-clinician physicians; hospital beds (source (p. 509))
Admissions per 100 children under 5 per year	Under-5 hospital admissions in a year × 100 ÷ the projected under-5 population. Fewer than 2 means very low access to inpatient care; the median across sub-Saharan African countries for 2018-2022 was 4.5 (source (p. 215)).
ANC1	The first antenatal care visit of a pregnancy, usually at about 20 weeks (4-5 months) of pregnancy (source (p. 67)).
ANC4	Four or more antenatal care visits during a pregnancy. Its survey coverage is the percentage of women aged 15–49 with a live birth who had four or more visits during their last pregnancy.
Antenatal care (ANC)	Care during pregnancy at a health facility. ANC1 is the first visit; ANC4 the fourth.
Bayesian model	A statistical model that starts from prior knowledge (here, that coverage usually rises along an S-shaped curve) and updates it with the data, giving estimates with their uncertainty. The coverage model takes its levels from surveys and its year-to-year change from the routine data (source (p. 143)).
BCG	The vaccine against tuberculosis, given at birth.

Benchmark	A reference value an indicator is compared with, such as 2 health facilities, 25 inpatient beds or 44.5 core health professionals per 10,000 population. A reference, not a target for every region (source (p. 230)).
Blocking check	A check on the uploaded workbook (sheets, key columns, country, districts, months) that must pass before the analysis can continue; informational checks only warn (source (p. 48)).
Bottom inequality (marginalization)	The poorest are left behind (source (p. 156))
Caesarean section rate	Caesarean sections as a percentage of deliveries. WHO's rule of thumb puts the population need at 10-15%: below 10% indicates unmet need for surgical delivery care, above 15% likely overuse (source (p. 124)).
Cascade training	Training in which the people trained then train others, from national to regional and district level, here on the concepts, interpretation and use of the analyses (source (p. 251)).
Community maternal mortality ratio (cMMR)	Maternal deaths outside health facilities per 100,000 live births outside health facilities. It is not reported, but can be back-calculated from the population and institutional ratios (source (p. 194)).
Community-to-institutional ratio	The assumed ratio of community to institutional mortality (M_c / M_i), used to compute the expected institutional rate; the apps compare 0.5, 1, 1.5 and 2 (source (p. 193)).
Completeness of facility death reporting	Completeness of facility reporting = reported iMMR / expected iMMR (source (p. 194))
Composite coverage index	The average coverage of seven mother and child health indicators, with equal weight: four or more antenatal care visits, institutional live births, skilled attendance at birth, the second dose of intermittent preventive treatment of malaria in pregnancy, postnatal care, the third dose of pentavalent vaccine and the first dose of measles vaccine (source (p. 169)).
Confidence interval	The range around a survey estimate that reflects its sampling uncertainty; a 95% confidence interval would contain the true value in 95 of 100 repeated surveys. A wide interval means the survey point should be read with caution (source (p. 135)).
Consistent district	Its yearly ratio lies between 1 and 1.5 (source (p. 317))
Continuum of care	The coverage indicators put side by side in the order a woman or child meets them: first-trimester antenatal care, ANC1, ANC4, institutional live births and postnatal care (maternal and newborn); BCG, penta1, penta3 and measles 1 (child). A step much lower than the one before shows where women and children are lost (source (p. 137)).
Core health professionals per 10,000 (physicians, non-physician clinicians, nurses, midwives)	Thresholds 23 (2006 WHO) and 44.5 (UHC) (source (p. 229))

Country platform	The national forum, usually led by the Ministry of Health, where government and partners review RMNCAH-N data and plans together (for example the RMNCAH platform and its technical working groups). The main audience for the workshop's key findings (source (p. 248)).
Crude birth rate (CBR)	The number of live births per 1,000 population in a year (source (p. 99)).
Crude death rate (CDR)	The number of deaths per 1,000 population in a year. Without major migration it is the crude birth rate minus the population growth rate (both per 1,000) (source (p. 99)).
Curative service use index	Curative services per child under 5 in a region: (under-5 OPD visits + 10 × under-5 admissions) ÷ under-5 population, an admission counted ten times a visit, for the latest year (source (p. 222)).
Data checks	Per district, whether it is missing from, and how often it occurs in, each sheet (source (p. 513))
Data quality indicators 1a, 1b, 1c	• 1a: monthly facility reports received / total expected reports (12 x number of facilities) (source (p. 59)) - 1b: districts with at least 90% reporting completeness / total districts (source (p. 59)) - 1c: districts with no missing values across all forms / total districts (source (p. 59))
Data quality indicators 2a, 2b	• 2a: monthly values not classified as outliers / total monthly values; 2b: districts with no outliers / total districts (source (p. 64))
Data quality indicators 3a-3d	• 3a: ANC1 reported / Penta1 reported (expected ratio range); 3b: Penta1 reported / Penta3 reported (based on survey data); 3c and 3d: districts within the expected ratio or range / total districts (source (p. 67))
Demand satisfied (family planning)	Family planning demand satisfied = use of modern contraceptives / (unmet need + use of modern contraceptives) (source (p. 132))
DHIS2	District Health Information Software 2, the software most countries use as their health management information system (HMIS); health facilities' monthly reports are entered in it, and the facility data workbook is exported from it (source (p. 46)).
Discharge	A patient leaving the hospital alive after an admission. Discharges = admissions − deaths; where available, discharge data are the preferred input for inpatient analyses (source (p. 215)).
District reporting threshold	The reporting rate a district must reach to count as reporting well: 90% by default; it can be changed on the Reporting Rate page (source (p. 59)).
Dropout	Children who received the first dose minus those who received the second, divided by the first, as a percentage (source (p. 345))
ENAP and EPMM	The Every Newborn Action Plan (ENAP) and the strategies for Ending Preventable Maternal Mortality (EPMM), the global plans for newborn and maternal health. Their coverage targets include 80% of districts with institutional delivery coverage of at least 80% (source (p. 175)).

Equiplot	A chart that shows the coverage of each population group (wealth quintiles, mother's education, urban and rural) as dots on one line per survey year. The distance between the dots is the absolute inequality between the groups; comparing the lines shows whether the gap narrows over time (source (p. 154)).
Expected institutional maternal mortality ratio (expected iMMR)	Expected iMMR = population MMR / (P_i + (M_c/M_i) x (1 - P_i)), with P_i the share of live births in facilities (source (p. 194))
Expected ratio	The ratio of two reported numbers (such as ANC1 to penta1) implied by the survey coverage and, for ANC1 to penta1, the mortality between the two; a national ratio within plus or minus 0.05 of it is acceptable (source (p. 67)).
Extreme outlier	A monthly value more than 5 times the median absolute deviation (MAD) from the median monthly value of the same series (district, or the national or admin-1 series being assessed), the median and MAD being computed from all years except the most recent one (source (p. 64)). In data adjustment it is replaced by the median of the district's non-outlier months in the same year (source (p. 84)).
Facility data derived denominator	A denominator computed from the health facility data: the reported number of ANC1 visits or Penta1 doses divided by their survey coverage, then converted to the other target populations (deliveries, live births, infants) with assumptions on pregnancy loss, twins, stillbirths and mortality (source (p. 103)).
Facility-based coverage	Coverage computed from the health facility data: the adjusted number of services reported divided by the target population estimated with a denominator option, × 100 (source (p. 109)).
First-trimester ANC	A first antenatal care visit made in the first three months of pregnancy; its coverage shows how early women enter pregnancy care (source (p. 118)).
FPET	Family Planning Estimation Tool: modelled national estimates of family planning indicators, such as contraceptive use and demand satisfied, built into the app (source (p. 52)).
Fresh stillbirth	An intra-partum stillbirth: the baby died during labour or delivery. Fresh stillbirths are expected to be 30% to 55% of all stillbirths (source (p. 185)).
Health facilities per 10,000 (all tiers, public and private)	Below 2 indicates low infrastructure access (source (p. 229))
Health system inputs	The resources the health system works with: infrastructure (health facilities, hospitals, inpatient beds), the health workforce, and financing, expressed per population as densities (source (p. 228)).
Health system outputs	What the health system delivers: service utilization (such as OPD visits and admissions per child) and coverage of interventions (such as institutional deliveries), compared with the inputs region by region (source (p. 235)).
HMIS	Health management information system: the national system that collects routine data from health facilities, such as their monthly service reports. In most countries it runs on DHIS2.

Hospitals per 100,000 population	Number of hospitals × 100,000 ÷ total population: the infrastructure available for acute inpatient services; no benchmark is drawn (source (p. 229))
Household survey	A national survey, such as a DHS or MICS, that measures coverage by interviewing a sample of households. Its coverage figures are used as a benchmark for the facility data.
IA2030	The Immunization Agenda 2030, the global immunization strategy for 2021-2030. Its coverage targets include 90% coverage for all routine vaccines and geographic equity of DTP3 (Penta3) coverage in the 20% of districts with the lowest coverage (source (p. 175)).
Immunization denominator (BCG, Penta)	Penta and the other infant vaccines: live births − neonatal deaths (surviving infants); BCG uses the live births (source (p. 127))
Implementation research	Research on how to deliver interventions that are known to work: why coverage is stuck (bottlenecks), and how successes can be scaled up or replicated (source (p. 251)).
Imputation	Replacing a missing or wrong value with an estimated one. In data adjustment, a missing monthly value, or an extreme outlier, is given the median of the district's other months in the same year (source (p. 90)).
Inpatient beds per 10,000 population	Number of official inpatient beds in all functioning facilities × 10,000 ÷ total population; the benchmark drawn is 25 (source (p. 229))
Institutional delivery	A delivery that takes place in a health facility. Its survey coverage is the percentage of live births that occurred in a health facility.
Institutional live births	Live births that take place in a health facility. Their coverage (institutional live births ÷ live births in the population) is one of the two indicators used to select the denominator (source (p. 108)).
Institutional mortality (iMMR, iSBR, neonatal deaths before discharge)	IMMR: maternal deaths in facilities per 100,000 facility live births; iSBR: stillbirths per 1,000 facility births; neonatal deaths before discharge per 1,000 facility live births (source (p. 182))
Intermittent preventive treatment (IPT)	Doses of an antimalarial drug given to pregnant women during antenatal care where there is malaria risk; IPT2 and IPT3 are two and three doses (source (p. 118)).
IPD	Inpatient department: admissions of patients who stay at least one night in a hospital or a facility with beds. Admissions include new admissions and re-admissions (source (p. 215)).
Iron-folic acid (IFA) supplementation	Iron and folic acid tablets given to pregnant women during antenatal care; coverage counts women who received at least 90 tablets (source (p. 118)).
k (k-factor)	The ratio of the volume of services provided by non-reporting facilities to that of reporting facilities (source (p. 80))

Linear inequality	Equidistant differences; coverage rises linearly with household wealth (source (p. 156))
Live birth	A baby born showing any sign of life, however briefly. Live births are the denominator of the maternal mortality ratio and of neonatal mortality; births are live births plus stillbirths (source (p. 183)).
Low birth weight	A birth weight of less than 2,500 grams.
Macerated stillbirth	An ante-partum stillbirth: the baby died before the onset of labour and shows signs of tissue deterioration (source (p. 185)).
MADM	The guidebook's summary measure of inequality between subnational units: the average distance, in percentage points, between each region's or district's coverage and the national coverage (the region's coverage when the units are the districts of one region). A falling MADM means that inequalities between units are being reduced (source (p. 172)).
Master adjusted dataset	The facility dataset after data adjustment (removed years left out, numbers adjusted for completeness of reporting, extreme outliers and missing values corrected). Every later step uses it; the app can download it as a Stata file, `master_adj_dataset.dta` (source (p. 75)).
Maternal death	The death of a woman from any cause related to or aggravated by pregnancy or its management (excluding accidental or incidental causes) during pregnancy and childbirth or within 42 days of termination of pregnancy, irrespective of the duration and site of the pregnancy (source (p. 182)).
Maternal denominator	In an RMNCAH analysis, the denominator option selected for the antenatal care and delivery indicators (ANC1-derived by default); the immunization indicators use the vaccination denominator (source (p. 109)).
MCH preventive services index	Preventive maternal and child health services per child under 5 in a region, weighted by importance: (ANC1 + 3 × ANC4 + postnatal care within 48 hours + BCG + (penta1 + penta3) ÷ 2 + penta3 + measles 1 + measles 2 + 10 × institutional deliveries) ÷ under-5 population, for the latest year (source (p. 222)).
Measles 1	The first dose of measles (or measles-rubella) vaccine, usually given at 9 months of age. Its survey coverage is the percentage of children aged 12–23 months who received it.
Measles denominators	Measles 1 denominator: live births - neonatal deaths - postneonatal deaths; Measles 2: live births - neonatal deaths - 2 x postneonatal deaths (source (p. 126))
Median	The middle value of a series once sorted: half the values are above it, half below. A few extreme values barely move it, unlike the mean (source (p. 64)).
Median absolute deviation (MAD)	How much the values of a series usually differ from their median: the median of the distances between each value and the median. A monthly value more than 5 MAD from the median is an extreme outlier (source (p. 64)).

Missing value	A month with no number reported for an indicator in a district. Measured by the share of districts with no missing monthly values, and replaced in data adjustment by the district's median monthly value for the year (source (p. 62)).
Modern contraceptive prevalence (mCPR)	The percentage of women (in the apps, currently married women aged 15-49) who use a modern method of contraception; estimated by FPET (source (p. 132)).
MPDSR	Maternal and Perinatal Death Surveillance and Response: the review of each maternal and perinatal death to learn why it happened and act on it. Sources that rely only on MPDSR or maternity ward logs miss deaths outside the labour ward (source (p. 183)).
Neonatal mortality	Deaths of babies in their first 28 days of life.
Non-physician clinician	A health worker who diagnoses and treats patients without a full medical degree, for example a clinical officer or health officer with surgical training and several years of certification. Counted among the core health professionals (source (p. 229)).
OPD	Outpatient department: visits to a facility that do not need an overnight stay.
OPD visits per child per year	Adjusted under-5 OPD visits (new and re-visits) / DHIS2 projected under-5 population (source (p. 210))
OPV1, OPV3	The first and third doses of oral polio vaccine (source (p. 67)).
Overall data quality score	The summary of performance for a year: the mean of the completeness, missing value, extreme outlier and district consistency indicators, computed automatically (source (p. 72)).
Pediatric case fatality rate	Children who die in hospital / total admissions (discharges + deaths) (source (p. 219))
Penta1, Penta3	The first and third doses of pentavalent (DTP-containing) vaccine, recommended at 6 and 14 weeks of age (source (p. 67)).
Percent of subnational units above a target or threshold	How far universal coverage is reached subnationally; the goal is all units meeting the target, and higher means less inequality (source (p. 172))
PHC service index	The service index of the PHC Performance charts: the mean of a region's MCH preventive services index and curative service use index (source (p. 236)).
Population data	By district and year: growth rate, total population, under-5, under-1, projected live births and total births (census-based), women 15-49 (source (p. 510))
Population growth denominator	The ANC1 or Penta1 population growth option: the ANC1- or Penta1-derived denominator of the reference year (the year before the survey year), carried to the other years with the growth of the DHIS2 population chosen in Select Best Population. Its ids end in `derived` (source (p. 106)).
Population growth rate	The annual change in a population, relative to its size: (population in a year − population the year before) ÷ population the year before, or ln(population in a year ÷ population the year before) (source (p. 99)).

Population maternal mortality ratio (MMR)	Maternal deaths in the whole population, in health facilities and in the community, per 100,000 live births in the population. It comes from the UN estimates or a recent national survey or census (source (p. 194)).
Post-neonatal mortality	Deaths of babies between 28 days and 11 months of life, as a proportion of live births.
Postnatal care within 48 hours (PNC48h)	A postnatal care contact for the mother and newborn within 48 hours of birth. How it is recorded varies between countries (source (p. 122)).
Pregnancy loss	A pregnancy that ends before birth (miscarriage or abortion).
R-squared (R^2)	How closely the points of a scatter plot follow the linear fit, from 0 (not at all) to 1 (exactly on the line) (source (p. 236)).
RDS file	The copy of a loaded dataset that the app saves next to the source file (`.rds` , an R data file). It opens much faster than the Excel workbook, without the checks; delete it to reload a changed workbook (source (p. 47)).
Relative change	The change the data adjustment made to a reported number, as a percentage of it: the absolute value of 100 × (reported − adjusted) ÷ reported. Reported nationally, for each year and on average over the years (source (p. 76)).
Reporting completeness data	Monthly expected, received and rate for ANC, institutional delivery, vaccinations, OPD and IPD (source (p. 510))
Reporting form	The monthly form a health facility sends for one service: antenatal care, delivery, immunization or outpatient care.
Reporting rate	The percentage of expected monthly facility reports that were received: the completeness of monthly facility reporting. Reporting rates over 90% are good (source (p. 59)).
RMNCAH	Reproductive, maternal, newborn, child and adolescent health: the services the RMNCAH analysis covers, from antenatal care to child immunization.
Service coverage	The population who received the service divided by the population who need it (the target population) (source (p. 95))
Service data 1	BCG, Penta1, Penta3, Measles1, Measles2; OPD total and under-5 (new and re-visits); IPD total and under-5; under-5 and all-age facility deaths (source (p. 511))
Service data 2	ANC1, first-trimester ANC, ANC4, IPT2, IPT3, IFA90, syphilis and HIV tests, SBA, institutional deliveries and live births, C-sections, low birth weight (<2500 g), PNC within 48 hours, stillbirths (total, fresh, macerated), maternal and neonatal deaths (source (p. 511))
Service data 3	OPV1-3, Penta2, PCV1-3, rotavirus 1-2, IPV1-2, HPV1-2, measles cases (source (p. 512))
Shapefile	A set of map files (`.shp` , `.dbf` , `.shx` , `.prj`) holding the boundaries of a country's regions, used to draw maps. A standard one is built into the app for each country (source (p. 50)).

Skilled birth attendance (SBA)	A birth attended by skilled health personnel (a doctor, nurse or midwife). Closely related to institutional births, but national definitions of a skilled attendant vary (source (p. 122)).
Skills mix ratio	The number of nurses and midwives per physician, shown in the national health system table (source (p. 229)).
Standard deviation	How far the values of a series usually are from their mean. An annual number of deaths more than 3 standard deviations from the long-term annual mean is inspected as a possible data entry error (source (p. 201)).
Stillbirth	A baby born with no signs of life after 28 weeks of pregnancy.
Sub-national one-pager	The standard report that gives one region its own statistical profile on a single page, duplicated for each region and shared with the regional teams (source (p. 252)).
Synthesis chartbook	The standard report that brings the complete analysis together in one document for the review meeting, with interpretation questions and a cover for the country team (source (p. 252)).
Target population	The population who need a service, the denominator of its coverage: for example pregnancies for antenatal care, live births for institutional live births, infants eligible for vaccination for penta (source (p. 95)).
Top inequality (mass deprivation)	The richest are well ahead of all other wealth quintiles (source (p. 156))
Total fertility rate (TFR)	The average number of children a woman would have over her lifetime at current age-specific fertility rates. Where it is above 4, children under 5 usually make more than 30% of outpatient visits (source (p. 210)).
Tracer vaccines	The vaccines used to follow an immunization programme: BCG, penta1 to penta3, OPV1 to OPV3 and measles 1.
Trajectory	One possible path of coverage over time; each year's estimate is the middle value across trajectories, with a range leaving out the most extreme ones (source (p. 145))
Twinning rate	The share of births that are twins.
UN estimates	The United Nations population projections for a country (total population, births and target groups by year), built into the app; used to assess the population projections of the facility data and as a denominator option (source (p. 52)).
UN mortality estimates	The UN inter-agency estimates of maternal mortality, stillbirth and neonatal mortality for each country, with a lower bound, best estimate and upper bound (for stillbirths, 90% uncertainty intervals). The RMNCAH app has them built in, or you upload your own (source (p. 195)).
Uncertainty interval	The range around a modelled estimate that shows how certain it is: for the Bayesian coverage model, the range covered by 95% of the trajectories in a year (the most extreme 2.5% on each side left out). A narrow interval means the data agree (source (p. 145)).

Universal health coverage (UHC)	All people receiving the health services they need without financial hardship. Its workforce threshold is at least 44.5 core health professionals per 10,000 population (source (p. 229)).
Unmet need for family planning	The percentage of women who want to stop or delay childbearing but use no modern method of contraception. Demand satisfied = use ÷ (unmet need + use) (source (p. 133)).
Wealth quintile	One of five groups of the same size into which a household survey splits households after ranking them by wealth, from the poorest fifth (Q1) to the richest fifth (Q5). Because the groups are the same size, the pattern of inequality (top, linear or bottom) can be read from them (source (p. 156)).
WUENIC	WHO and UNICEF estimates of national immunization coverage, produced each year with countries from all data sources; compared with the coverage estimated from facility data (source (p. 52)).
Zero-dose and under-vaccinated	Zero-dose and Under-vaccinated are derived from Penta1 and Penta3 (source (p. 345))

APPENDIX B

Abbreviations

Alphabetical list of the abbreviations and acronyms used in the DataSuite documentation, each linked to the section that defines or best uses the term.

The abbreviations and acronyms used in the DataSuite documentation, in alphabetical order. **See** links to the section that defines or uses the term best; for defined terms, see also the Glossary (p. 496) and the Indicator definitions (p. 509).

Abbreviation	Meaning	See
Admin1, Admin2	First and second subnational administrative levels (region, district)	Subnational analysis (p. 168)
AI	Artificial intelligence	AI assistant (p. 260)
ANC	Antenatal care	Coverage: antenatal care (p. 118)
ANC1	First antenatal care visit	Indicator definitions: service data 2 (p. 511)
ANC4	Four or more antenatal care visits	Coverage: antenatal care (p. 118)
BCG	Bacille Calmette-Guérin (tuberculosis vaccine, given at birth)	Coverage: immunization (p. 126)
CBR	Crude birth rate	Denominators: population trends (p. 100)
CCI	Composite coverage index	Formulas (p. 169)
CD2030	Countdown to 2030	Methodology (p. 42)
CDR	Crude death rate	Formulas (p. 99)
CI	Confidence interval (95% CI)	Coverage: read the chart (p. 117)
cMMR	Community maternal mortality ratio	Mortality (p. 194)
cSBR	Community stillbirth rate	Mortality (p. 194)
CSV	Comma-separated values (text file format)	Pooled App: exporting results (p. 439)
DHIS2	District Health Information Software 2	Analysis setup: HMIS data (p. 46)
DHS	Demographic and Health Surveys	Analysis setup: survey parameters (p. 54)
DQA	Data quality assessment	Data quality (p. 57)
DTP, DTP1, DTP3	Diphtheria-tetanus-pertussis vaccine; first and third doses	Coverage: immunization (p. 126)
ENAP	Every Newborn Action Plan	Coverage: antenatal care (p. 118)
EPI	Expanded Programme on Immunization	RMNCAH and Vaxx apps (p. 289)
EPMM	Ending Preventable Maternal Mortality	Coverage: antenatal care (p. 118)

EULA	End-User License Agreement	Installing on Windows (p. 13)
FP	Family planning	Coverage: family planning (p. 132)
FPET	Family Planning Estimation Tool (Track20)	Coverage: family planning (p. 132)
HFD	Health facility data	Getting started (apps) (p. 295)
HIV	Human immunodeficiency virus	Indicator definitions: service data 2 (p. 511)
HMIS	Health management information system	Analysis setup: HMIS data (p. 46)
HPV	Human papillomavirus (vaccine)	Indicator definitions: service data 3 (p. 512)
IA2030	Immunization Agenda 2030	Subnational: target values (p. 175)
IFA, IFA90	Iron-folic acid supplementation; IFA90: at least 90 tablets during pregnancy	Coverage: antenatal care (p. 118)
iMMR	Institutional maternal mortality ratio	Mortality (p. 201)
IPD	Inpatient department	Service utilization: inpatient (p. 215)
IPT, IPT2, IPT3	Intermittent preventive treatment of malaria in pregnancy; two and three doses	Coverage: antenatal care (p. 118)
IPV	Inactivated polio vaccine	Indicator definitions: service data 3 (p. 512)
iSBR	Institutional stillbirth rate	Mortality (p. 201)
ISO3	Three-letter ISO 3166 country code	Pooled: preparing caches (p. 428)
M&E	Monitoring and evaluation	Dissemination: building capacity (p. 251)
MAD	Median absolute deviation	Data quality: outlier detection (p. 64)
MADM	Median absolute deviation from the median (subnational inequality)	Glossary: MADM (p. 501)
MCH	Maternal and child health	Service utilization and health system (apps) (p. 391)
mCPR	Modern contraceptive prevalence rate (the app's "Prevalence of Modern Methods")	Advanced analysis (apps) (p. 371)
MCV, MCV1, MCV2	Measles-containing vaccine; first and second doses	Coverage: immunization (p. 126)
MICS	Multiple Indicator Cluster Surveys	Analysis setup: survey parameters (p. 54)
MMR	Maternal mortality ratio	Mortality (p. 201)
MNCH	Maternal, newborn and child health	Data Extractor: mapping (p. 277)
MPDSR	Maternal and Perinatal Death Surveillance and Response	Mortality (p. 183)
NGO	Non-governmental organization	Health system: private sector (p. 226)
NMR	Neonatal mortality rate	Denominators (p. 105)
OPD	Outpatient department	Service utilization: outpatient (p. 210)

OPV, OPV1, OPV3	Oral polio vaccine; first and third doses	Data quality: internal consistency (p. 67)
ORS	Oral rehydration salts	Coverage: child continuum of care (p. 140)
PCV	Pneumococcal conjugate vaccine	Indicator definitions: service data 3 (p. 512)
PDF	Portable Document Format	Reports (apps) (p. 415)
Penta, Penta1, Penta3	Pentavalent vaccine (contains DTP); first and third doses	Glossary: immunization denominator (p. 500)
PHC	Primary health care	Service utilization and health system (apps) (p. 391)
PNC, PNC48h	Postnatal care; within 48 hours of birth	Coverage: delivery care (p. 122)
PNG	Portable Network Graphics (image format)	Exports and AI (apps) (p. 411)
PNMR	Post-neonatal mortality rate	Denominators (p. 105)
RAM	Random-access memory	Requirements (p. 16)
RDS	R data file (`.rds`), the apps' saved dataset	Analysis setup: automatic RDS creation (p. 47)
RMNCAH, RMNCAH-N	Reproductive, maternal, newborn, child and adolescent health (and nutrition)	RMNCAH and Vaxx apps (p. 289)
SBA	Skilled birth attendance	Coverage: delivery care (p. 122)
SBR	Stillbirth rate	Mortality (p. 201)
SD	Standard deviation	Mortality: outlier thresholds (p. 201)
SDGs	Sustainable Development Goals	Subnational: inequality (p. 165)
TFR	Total fertility rate	Service utilization (p. 204)
UAC	User Account Control (Windows)	Installing on Windows (p. 12)
UHC	Universal health coverage	Subnational coverage (p. 168)
UI	User interface	Getting started (p. 21)
UN	United Nations	Denominators: options (p. 103)
UNICEF	United Nations Children's Fund	Coverage: immunization (p. 126)
URL	Web address (uniform resource locator)	Data Extractor: DHIS2 login (p. 273)
WHO	World Health Organization	Coverage: delivery care (p. 122)
WUENIC	WHO/UNICEF Estimates of National Immunization Coverage	Coverage: immunization (p. 126)
ZIP	Compressed archive file format (`.zip`)	Portable version (p. 19)

APPENDIX C

Indicator definitions

Standard definitions mapping HMIS variables to RMNCAH indicators, covering the administrative, population and service data the DataSuite analysis requires.

DataSuite standardizes indicator definitions across health management information systems (HMIS) to enable consistent, cross-country analysis. This section specifies exactly which HMIS variables map to reproductive, maternal, newborn, child, and adolescent health (RMNCAH) indicators—preventing misinterpretation and ensuring your facility data feed the correct numerators and denominators. Standardized definitions also enable comparison across countries and over time, making global progress toward universal health coverage measurable.

Three data categories organize the requirements: administrative data (facility inventory, workforce), population data (demographic estimates, survival rates), and service data (events reported by facilities, stratified by indicator and administrative level). Because country HMIS structures vary widely, this reference enables you to systematically extract or map your data to the standardized variable names DataSuite expects. Precise adherence to these definitions is foundational to reliable analysis.

Data categories

For cross-country standardization of the analysis, this section describes the indicators/variables extracted from the countries' HMIS, that are required for the DataSuite analysis.

The data is grouped into the following main categories:

- **Admin Data**
- **Population Data**
- **Reporting Completeness**
- **Service Data (1, 2, and 3)**
- **Data Checks**

Admin data

The following variables describe health system infrastructure and workforce.

Variable Name	Description
`Num_order`	N°
`District_name`	District
`First_admin_level`	Region
`Country`	Your country name
`Number_hospitals`	Total number of hospitals (based on country definition)

`Number_hcenters`	Total number of health centers (non-hospitals; based on country definition)
`Total_number_health_facilities`	Total number of health facilities (all types/categories)
`Number_pfacilities_profit`	Total number of health facilities that are private-for-profit
`Number_pfacilities_nonprofit`	Total number of health facilities that are private-non-profit
`Total_physicians`	Total number of physicians / Medical officers
`Total_nurses_midwives`	Total number of nurses / midwives
`Total_clinical_officers`	Total number of clinical officers / Assistant medical officers
`Total_nonclinique_physicians`	Total number of non-clinician physicians
`Number_hospital_beds`	Number of hospital beds

Population data

The following variables describe demographic and population estimates.

Variable Name	Description
`District_name`	District name
`Year`	Year
`Pop_growth_rate`	Annual Population growth rate
`Total_Population`	Total Population
`Population_ under_5years`	Population under 5 years
`Population_under_1year`	Population under 1 year
`Live_births`	Projected total Live births (Estimated live births from census)
`Total_births`	Projected total births (Estimated births from census)
`Women_15_49_years`	Total women 15-49 years

Reporting completeness

The following variables track facility reporting rates and expected submissions for various health service categories.

Variable Name	Description
`District`	District name
`Year`	Year
`Month`	Month
`ANC_reporting_expected`	ANC reporting: Expected number (#)
`ANC_reporting_received`	ANC reporting: Received number (#)
`ANC_reporting_rate`	ANC reporting: Reporting completeness rate (%)
`Instdelivey_reporting_expected`	Institutional delivery reporting: Expected number (#)
`Instdelivey_reporting_received`	Institutional delivery reporting: Received number (#)
`Instdelivey_reporting_rate`	Institutional delivery reporting: Reporting completeness rate (%)

`Vacc_reporting_expected`	Vaccinations reporting: Expected number (#)
`Vacc_reporting_received`	Vaccinations reporting: Received number (#)
`Vacc_reporting_rate`	Vaccinations reporting: Reporting completeness rate (%)
`OPD_reporting_expected`	OPD visits reporting: Expected number (#)
`OPD_reporting_received`	OPD visits reporting: Received number (#)
`OPD_reporting_rate`	OPD visits reporting: Reporting completeness rate (%)
`IPD_reporting_expected`	IPD admissions reporting: Expected number (#)
`IPD_reporting_received`	IPD admissions reporting: Received number (#)
`IPD_reporting_rate`	IPD admissions reporting: Reporting completeness rate (%)

Service data 1

General admissions, deaths, and initial child vaccinations.

Variable Name	Description
`District`	District name
`Year`	Year
`Month`	Month
`BCG`	Total number of BCG doses administered <1 year
`Penta1`	Total number of Pentavalent vaccination 1st dose (infants)
`Penta3`	Total number of Pentavalent vaccination 3rd dose (infants)
`Measles1`	Total number of Measles vaccination 1st dose (infants)
`Measles2`	Total number of Measles vaccination 2nd dose
`OPD_total`	Total number of OPD visits all ages (new and re-visit)
`OPD_under5`	Number of OPD visits for children under-5 years
`IPD_total`	Total number of IPD admissions, all ages
`IPD_under5`	Number of IPD admissions of children under-5 years
`under5_deaths`	Total number of under-5 deaths in health facilities
`total_deaths`	Total number of all age deaths in health facilities

Service data 2

Maternal, newborn, and reproductive health indicators.

Variable Name	Description
`District`	District name
`Year`	Year
`Month`	Month
`ANC1`	Total number of ANC 1st visit
`ANC_1trimester`	Total number of ANC during the 1st trimester of pregnancy

`ANC4`	Total number of Pregnant women completing 4 ANC visits
`IPT2`	Total number of pregnant women receiving IPT 2nd dose
`IPT3`	Total number of pregnant women receiving IPT 3rd dose
`IFA90`	Total number of pregnant women who received iron folic acid (IFA) supplementation for 90 or more days during the pregnancy
`Syphilis_test`	Total number of pregnant women tested for syphilis during pregnancy
`HIV_test`	Total number of pregnant women tested for HIV during pregnancy
`SBA`	Total number of skilled birth attendance in health facilities
`Instdelivery`	Total number of deliveries in health facilities
`Instlivebirths`	Total number of live births in health facilities
`Csection`	Total number of Cesarean sections
`Low_BWeight`	Total number of newborns with low birth weight (<2500g)
`PNC_48h`	Total number of mothers attending PNC visit within 48 hours after delivery
`Stillbirth_total`	Total number of stillbirths in health facilities (all)
`Stillbirth_fresh`	Total number of fresh stillbirths in health facilities
`Stillbirth_macerated`	Total number of macerated stillbirths in health facilities
`maternal_deaths`	Total number of maternal deaths reported
`neonatal_deaths`	Total number of neonatal deaths reported

Service data 3

Additional immunizations and disease cases.

Variable Name	Description
`District`	District name
`Year`	Year
`Month`	Month
`opv1`	Total number of <1 year who received Polio 1 vaccine
`opv2`	Total number of <1 year who received Polio 2 vaccine
`opv3`	Total number of <1 year who received Polio 3 vaccine
`penta2`	Total number of Pentavalent vaccination 2nd dose (infants)
`pcv1`	Total number <1 year who received Pneumococcal Conjugate 1 vaccine
`pcv2`	Total number of <1 year who received Pneumococcal Conjugate 2 vaccine
`pcv3`	Total number of <1 year who received Pneumococcal Conjugate 3 vaccine
`rota1`	Total number of <1 year who received Rotavirus 1 vaccine
`rota2`	Total number of <1 year who received Rotavirus 2 vaccine
`ipv1`	Total number of children who received inactivated poliovirus 1 vaccine
`ipv2`	Total number of children who received inactivated poliovirus 2 vaccine

`hpv1`	Total number of children who received human papillomavirus vaccine 1
`hpv2`	Total number of children who received human papillomavirus vaccine 2
`measles_cases`	Total number of Measles cases reported

Data checks

Automated diagnostic variables used for validation of completeness and consistency across sheets.

Variable Name	Description
`District`	District Name
`# of occurrences in 'Admin data'`	Verification count of district presence in the Admin Data sheet
`Is the district missing in 'Population_data'?`	Status flag checking if the district is missing in Population Data
`# of occurrences in 'Population_data'?`	Verification count of district presence in Population Data
`Is the district missing in 'Reporting_completeness'?`	Status flag checking if the district is missing in Reporting Completeness
`# of occurrences in 'Reporting completeness'`	Verification count of district presence in Reporting Completeness
`Is the district missing in 'Service_data_1'?`	Status flag checking if the district is missing in Service Data 1
`# of occurrences in 'Service_data_1'`	Verification count of district presence in Service Data 1
`Is the district missing in 'Service_data_2'?`	Status flag checking if the district is missing in Service Data 2
`# of occurrences in 'Service_data_2'`	Verification count of district presence in Service Data 2