Harmonized Health Facility Assessment (HHFA)

Comprehensive guide

World Health Organization

Harmonized Health Facility Assessment (HHFA)

Comprehensive guide

World Health Organization

Harmonized health facility assessment (HHFA): comprehensive guide

ISBN 978-92-4-006030-2 (electronic version)
ISBN 978-92-4-006031-9 (print version)

Contents

Note: The HHFA tools are periodically updated. The screenshots appearing in this document may therefore not reflect the most recent versions.

Acknowledgements

The Harmonized Health Facility Assessment (HHFA) modules and resource package are a key deliverable of the Health Data Collaborative Facility Surveys Working Group. The modules provide a harmonized approach to health facility assessments/surveys, building on existing internationally tested tools, such as the World Health Organization (WHO) Service Availability and Readiness Assessment (SARA), the United States Agency for International Development Service Provision Assessment (SPA) and the World Bank Service Delivery Indicators (SDI), as well as consolidating best practices and lessons learned through implementation in many countries.

Overall guidance for the development of the initial version HHFA modules was provided by the Health Data Collaborative Facility Surveys Working Group. Kathryn O'Neill, Amani Siyam and Kavitha Viswanathan coordinated the development of the initial version. Wendy Venter coordinated the revisions of the modules, and the development of the HHFA resource package with technical support from the Johns Hopkins Bloomberg School of Public Health. Substantial technical contributions to the resource package were made by Eman Aly, Yolanda Barbera, Sandro Colombo, Benson Droti, Nancy Fronczak, Sherrell Goggin, Fern Greenwell, Geoff Greenwell, Jaya Gupta, Shannon King, Hillary Kipruto, Benito Koubemba, Davy Audrey Liboko Gnekabassa, Geofrey Lutwama, Boniface Muganda, Timothy Roberton, Ashley Sheffel and Moussa Traore. Technical inputs concerning guidelines, service standards, measurement methods and indicators were provided by multiple WHO technical programmes and regional offices as well as other agencies within the health sector.

The ministries of health of Burkina Faso, Kenya, Liberia, Malawi and Zambia are gratefully acknowledged for assistance in testing the implementation of the HHFA modules and resource package.

The HHFA modules and resource package were produced with the support of grants from Bloomberg Philanthropies Data for Health Initiative; Gavi, the Vaccine Alliance; The Global Fund to Fight AIDS, Tuberculosis and Malaria; the Howard G. Buffett Foundation; the Kingdom of Saudi Arabia; the Norwegian Agency for Development Cooperation (Norad); and the Canada Department of Foreign Affairs, Trade and Development (DFTAD).

Abbreviations

CI	confidence interval
CSPro	Census and Survey Processing System
DCMI	Dublin Core Metadata Initiative
DDI	Data Documentation Initiative
deff	design effect
DQR	data quality review
FBO	faith-based organization
GIS	geographic information system
GPS	global positioning system
HHFA	Harmonized Health Facility Assessment
HIV	human immunodeficiency virus
ID	identification
IHSN	International Household Survey Network
MFL	master facility list
NGO	nongovernmental organization
PHC	primary health care
PMTCT	prevention of mother-to-child transmission (of HIV)
PSU	primary sampling unit
RHIS	routine health information system
SARA	Service Availability and Readiness Assessment
SAS	Statistical Analysis System
SDG	Sustainable Development Goal
SDI	Service Delivery Indicators
SPA	Service Provision Assessment
SPSS	Statistical Package for Social Services
TOT	training of trainers
UHC	universal health coverage
WHO	World Health Organization
XML	Extensible Markup Language

How to use the *HHFA Comprehensive guide*

What is the purpose of this guide?

The *HHFA Comprehensive guide* serves as the main reference document for planning and implementing a country HHFA. This guide will promote understanding of:

- What the HHFA is and the information it can and cannot provide.
- The HHFA modules, questionnaires and CSPro electronic data collection tool.
- The HHFA indicators, indices and their organization within the HHFA indicator inventory platform.
- The HHFA data analysis platform.
- The HHFA sampling and data collection methodologies.
- The detailed steps involved in planning and implementing an HHFA.
- Key concepts in review, interpretation and communication of HHFA findings.

An abridged version of this guide, the *HHFA Quick guide*, provides a rapid overview of the key aspects of the HHFA.

Who is this guide for?

The *HHFA Comprehensive guide* is intended to help anyone involved in the detail of planning and implementing an HHFA and using HHFA data to strengthen country health services. This can include staff of the ministry of health, donor agencies and implementing partners, as well as academic institutions supporting the survey. The guidance will be of particular use to the multipartner country HHFA coordination group, the HHFA technical committee, the survey manager and the data analysts and report writers.

How is this guide structured?

Chapter 1 presents a summary of the HHFA: its uses, thematic scope and content, implementation approaches, and available tools and resources. Chapter 2 provides the HHFA's background: its rationale, content concepts, development and role in a country health information system. Chapter 3 examines the HHFA objectives and uses, and then describes its components: modules, questionnaires, indicators and data collection and analysis tools. Chapter 4 examines methodological aspects of the survey and its required adaptation to the country context and information needs. Chapter 5 provides a summary table of the HHFA planning and implementation steps.

Chapter 6 delves into the strategic and operational planning and preparation of the survey: a critical phase of the HHFA, which requires strong engagement by the ministry of health and its partners. Chapter 7 focuses on data collection, which involves the deployment of multiple teams to multiple locations and requires careful planning, preparation and monitoring. Chapter 8 describes data processing, including checking and cleaning of the dataset, and data analysis which produces outputs in the form of tables, graphs, maps and a report outline format. Chapter 9 describes how to carry out a descriptive analysis of the analysis outputs. Chapter 10 provides guidance on how to interpret the survey results, identify key findings, and communicate main conclusions and recommendations to diverse audiences. Chapter 11 highlights the importance of documenting and archiving the survey, for further analyses, research and future reference. Finally, the annexes provide details on index calculation and sampling methodology, offer examples of HHFA job descriptions, provide guidance on conducting a training workshop for data collectors and procedures for data collection, and present a glossary of key terms.

This guide should be used along with other documents, tools and training materials in the HHFA resource package.

1. HHFA summary

1.1 What is the Harmonized Health Facility Assessment?

The HHFA is a comprehensive, standardized health facility survey that provides reliable, objective information on the availability of health facility services and the capacities of facilities to provide the services at required standards of quality.

Availability and quality of health services are integral to achieving universal health coverage (UHC) and the health-related Sustainable Development Goals (SDGs). HHFA data can support health sector reviews and evidence-based decision-making for strengthening country health services. Developed through multistakeholder collaboration, the HHFA builds on previous and existing global facility survey instruments, is based on global service standards, and uses standardized indicators, questionnaires, data collection methodologies and data analysis tools.

HHFA content

The HHFA covers all key facility services and facility-level management systems. The HHFA content is organized into four modules: service availability; service readiness; quality of care; and management and finance.

Fig.1. HHFA modules

A module represents a set of questions (in questionnaire format) for a main topic area. Countries may choose to implement any single module or a combination of modules. Core questions represent the recommended minimum information, while optional additional questions provide further details. All questions must be linked to defined indicators. Various questionnaire options are available (refer to Fig. 2). The questionnaires can also be adapted to country needs. The HHFA questionnaires are programmed into the HHFA Census and Survey Processing System (CSPro) electronic data collection tool. HHFA data are analysed to produce indicators within five service dimensions: general service availability; general service readiness; service-specific availability and readiness; management and finance support systems; and quality of care.

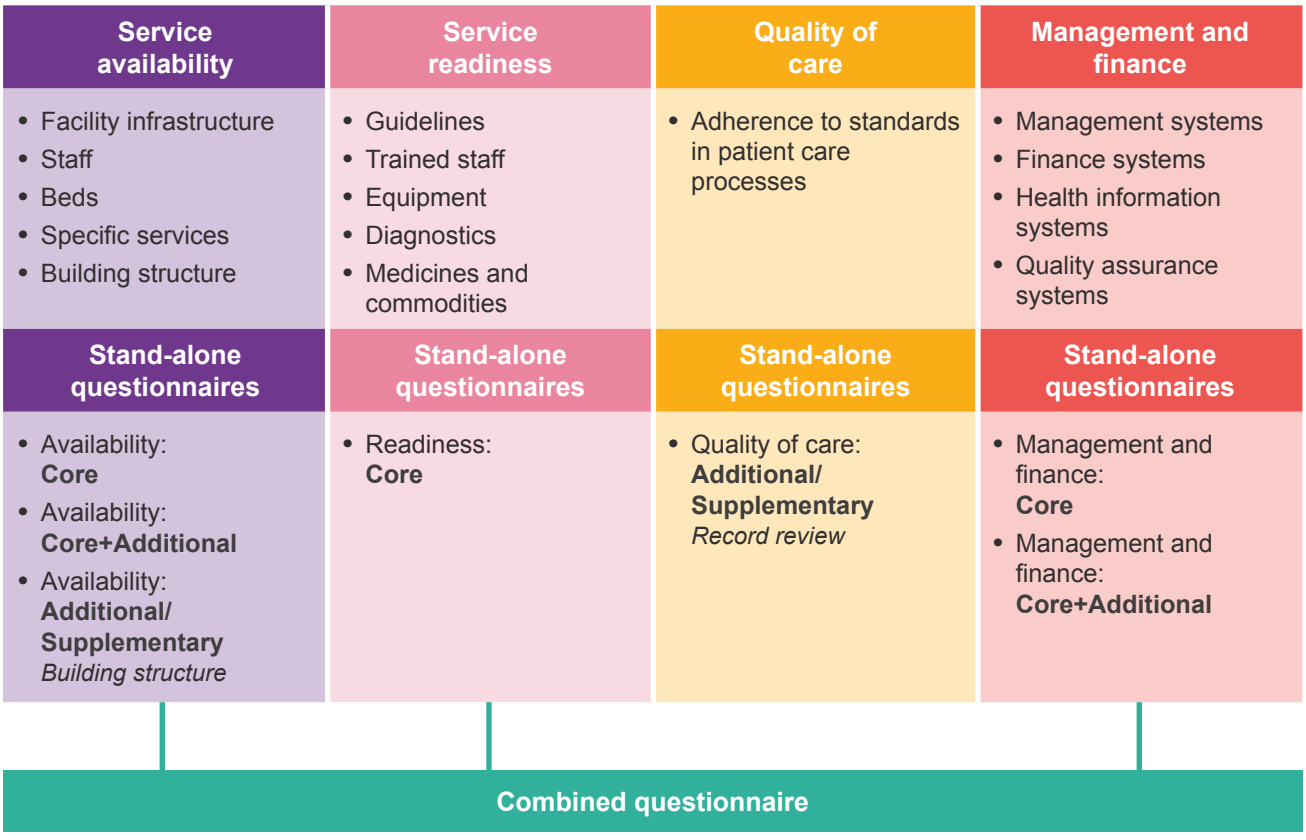
HHFA implementation

The HHFA can be conducted on a representative sample of facilities or as a census of all facilities in the country. An updated master facility list (MFL) including all public and private facilities serves as the foundation of the survey sampling frame.

Trained data collectors visit the facilities to collect data on electronic devices (tablets or mobile phones) using the HHFA CSPro tool. Once data collection is completed, the data are transferred to the HHFA data analysis platform. The analysis platform enables automated production of the HHFA indicators in tables, graphs and maps in a standard report outline format. Countries can use this outline as the basis for a comprehensive survey report, with interpretation of the findings within the country context and recommendations for action.

Countries should establish a plan of regular HHFAs (e.g. 1 to 5 years) as part of the national health sector monitoring and evaluation framework. Selected modules could be implemented as sample surveys in alternating years for monitoring purposes. The HHFAs should be synchronized with the country's schedule of routine analytical reviews and planning processes, so that the results can feed into these processes.

Fig. 2. **HHFA modules and questionnaires**



1.2 The HHFA resource package

The HHFA resource package is a comprehensive set of downloadable tools and guidance to support countries in planning and implementing an HHFA (refer to Fig. 3).

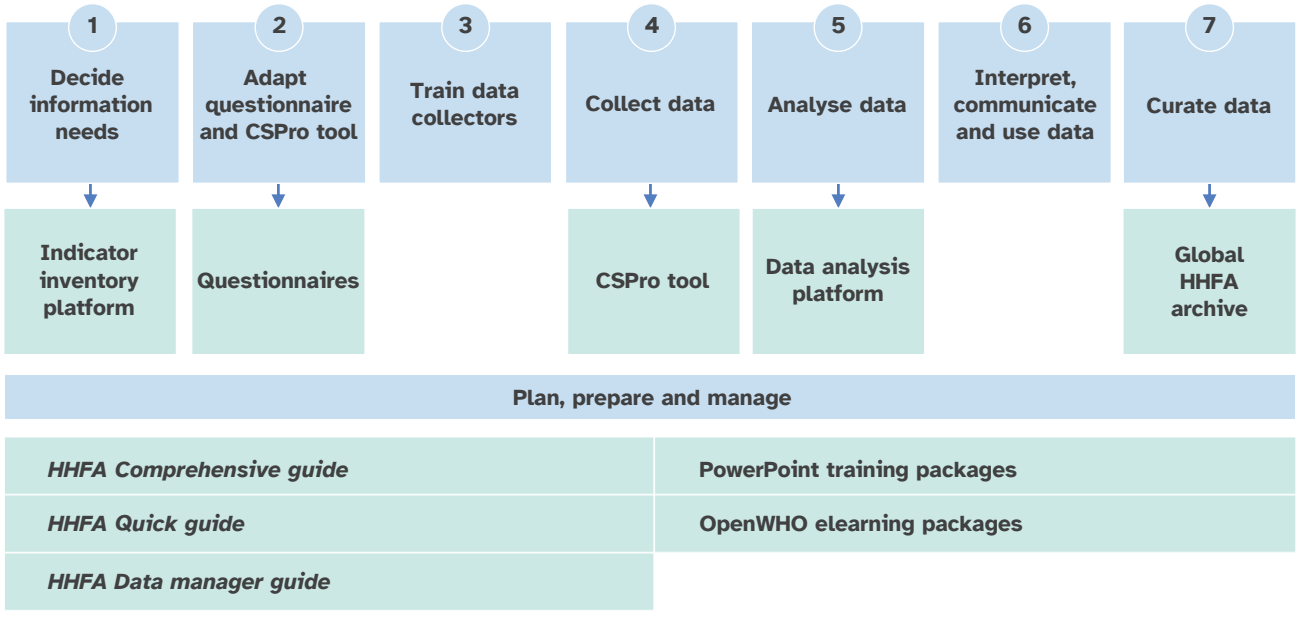
The resource package is available at:
https://www.who.int/data/data-collection-tools/harmonized-health-facility-assessment/introduction

Additional resources, including training materials, may also be accessed upon request to hhfa@who.int

HHFA OpenWHO e-learning courses are available at: https://openwho.org/channels/hhfa

- ***HHFA Quick guide*:** This provides a rapid overview of key HHFA concepts and summarizes the steps of the survey planning and implementation.
- ***HHFA Comprehensive guide*:** This guide provides an expanded description of the HHFA background, concepts and tools, as well as detailed step-by-step guidance for survey planning, preparation, implementation, data analysis, interpretation and dissemination of results.
- **Indicator inventory:** An online platform displays all the HHFA indicators, including the survey questions and code needed to calculate each indicator. The inventory can also be downloaded as an Excel document.
- **Questionnaires:** Questionnaires are available in "combined" and "stand-alone" formats. The "combined" questionnaire includes core questions from multiple HHFA modules, integrated to facilitate data collection. "Stand-alone" questionnaires are also available for each module. The stand-alone questionnaires are further categorized as Core, Core+Additional and Supplementary, based on the types of questions they contain and the data collection methodology used.
- **CSPro electronic data collection tool:** This tool is a CSPro application containing all the HHFA questions. The tool is flexible, enabling countries to select the questionnaires and questions they want to implement and to adapt the questionnaires to the country context.
- ***HHFA Data manager guide:*** The guide defines the data manager's responsibilities in an HHFA and provides detailed explanations on how to adapt and use the CSPro tool.
- **Data analysis platform:** After export from the CSPro tool, HHFA data are uploaded to the HHFA data analysis platform (or other analysis software). The platform automatically calculates the standard HHFA indicators and produces tables, graphs and maps in a standard format. The data analysis platform can also be adapted to country needs.
- **Training resources:** Various training resources are available to support countries in preparing for and implementing an HHFA, including a set of OpenWHO e-learning courses. Training topics include: an introduction to the HHFA; questionnaires, indicators and country adaptation; data collector training; and HHFA data review, interpretation and communication.
- **Global archive:** WHO has developed a central data catalogue where countries may choose to securely store their HHFA data and reports; the archive content can also be made publicly available, based on country authorization.

Fig. 3. **HHFA resource package – tools for every HHFA step**



2. HHFA background

2.1 Why assess health facility services?

All countries are working toward achieving UHC. The goal of UHC is based on country capacity to provide quality health services to all people needing care, while protecting the vulnerable from financial hardship. The COVID-19 pandemic, recent Ebola epidemics, and the global increase in natural disasters and conflicts have exposed the vulnerability of health systems to external shocks. They have highlighted the importance of building health systems that have the resilience to respond to and recover from health crises. Strengthening the availability and quality of health facility services, and improving their resilience, are key steps in the path toward achieving national health goals, UHC and the health-related SDGs.

Sound information on the supply and quality of health services is necessary for health systems policy-making, planning and management. Health facility data are needed for a comprehensive understanding of the functioning of health service delivery systems and for monitoring changes in these systems over time. However, despite decades of investments in health information systems, few countries have accurate, up-to-date information on the availability of health services in both public and private facilities, or their capacity, or "readiness", to provide quality services. The WHO 2020 SCORE Assessment *[1]* revealed that almost 50% of countries assessed had limited capacity for systematic assessment of quality of care, with most of these being low- and middle-income countries.

2.2 Concepts for assessing health facility services

Ensuring access to quality health services is a key function of a health system. Service access includes multiple components: **service availability**, referring to the physical presence or reach of health facilities; affordability, referring to the ability of a client to pay for services; and acceptability, referring to sociocultural aspects. The latter two aspects are not measured by the HHFA.

Availability of services is not enough: facilities must have the capacities to provide the services at required standards of quality. **Service readiness** refers to the availability and functionality of key resources (infrastructure, trained staff, guidelines, equipment, diagnostic tests, medicines and commodities) needed for providing the services. Furthermore, appropriate facility-level **management systems** must be in place to plan, organize, support and monitor the delivery of the services.

Service availability, readiness and management systems are all prerequisites for service quality. However, they do not guarantee the delivery of a high-quality care process. **Quality of care** is a complex concept that includes multiple dimensions. It requires a health system that is able to ensure service availability, readiness and management, and includes technical quality of care, as well as the attitudes and behaviours of service providers, and patient trust in the providers.

Optimal functioning of all these elements contributes to the achievement of key health service outcomes: high coverage of key, effective interventions, people-centred care (care which has considered the preferences and aspirations of individual service users and the cultures of their communities), financial protection of vulnerable families, and, ultimately, improved health outcomes.

2.3 Development of the HHFA

Over the years, various health facility survey tools have been developed, including the WHO Service Availability and Readiness Assessment (SARA),[1] the United States Agency for International Development Service Provision Assessment (SPA),[2] and the World Bank Service Delivery Indicators (SDI)[3] survey. Sometimes multiple, uncoordinated facility surveys have been conducted in a single country, at high cost and often producing non-comparable results. Furthermore, facility surveys have often emphasized specific topics or programmes, rather than providing an integrated assessment across all services.

In an effort to address these issues and to ensure a facility survey tool to meet the needs of the UHC and SDG era, the HHFA was developed. The HHFA is a comprehensive assessment of health facility services, based on global service standards, and using standardized indicators, questionnaires, data collection methodologies and data analysis tools. It represents a consolidated approach, building on previous and current survey instruments and experiences, and reflecting global indicator lists. Key aims of the HHFA development process are to promote support for alignment across facility surveys, to reduce redundancy and costs of multiple surveys in the same country, and to facilitate comparability of results among surveys.

The HHFA incorporates, updates and expands upon the SARA, and also provides an updated and more extensive set of tools and resources than the SARA. Table 1 summarizes key differences between the SARA and the HHFA.

Table 1. Differences between the SARA and the HHFA

	SARA	HHFA
Modules	■ Availability ■ Readiness	■ Availability ■ Readiness ■ Quality of care ■ Management and finance
Indicators	■ Core only ■ Printed indicator inventory	■ Core (expanded) and Additional ■ Indicator platform and Excel download
Questions	■ Core only	■ Core (expanded) and Additional
Questionnaires	■ Single core questionnaire	■ Stand-alone and combined questionnaires ■ Core; Core+Additional; Supplementary
Data collection methods	■ Facility audit (observation and interviews)	■ Facility audit and other methods, e.g. record reviews
Electronic data collection tool	■ CSPro application	■ CSPro application improved with additional features
Data analysis	■ Excel chartbook	■ Data analysis platform: automated analysis and report outline production
Global archive	■ No central data repository	■ Global archive with capacity for countries to store reports and survey data in a secure, central location

[1] https://www.who.int/data/data-collection-tools/service-availability-and-readiness-assessment-(sara)?ua=1

[2] https://dhsprogram.com/methodology/Survey-Types/SPA.cfm

[3] https://www.sdindicators.org/

The HHFA was developed by WHO with inputs from the Health Data Collaborative. The development process began in 2014 with a multipartner technical consultation. Following the consultation, a multipartner working group was established to develop the initial set of HHFA resources, including standard indicators, measurement methods, draft questionnaires and analysis plans. Early versions of HHFA resources were tested in Burkino Faso, Kenya and Malawi. Subsequent versions and additional tools were tested in Liberia and Zambia.

The HHFA is based on global service standards that are continuously evolving. Furthermore, lessons learned from HHFA implementation in multiple countries over time, along with feedback from programmes and partners, contribute to strengthening the tools. The HHFA resource package will therefore require regular updates. A feedback form is available on the HHFA website for submission of ongoing feedback.

2.4 Role of the HHFA in country health information systems

The **HHFA is designed to provide periodic, aggregate information across multiple facilities on overall service status,** which should show whether systems and services are functioning as expected. It does not aim to provide frequent, regular information about individual health facilities for ongoing supervision or management purposes.

The HHFA collects information that is usually not systematically collected by other information systems. It intends to complement other data sources such as the routine health information system (RHIS) and supervision systems, by filling information gaps. Furthermore, as HHFA data are collected by external data collectors (rather than being self-reported by facility staff), the HHFA is able to provide an objective assessment. In countries where health facility accreditation or certification systems are not yet well-established, the HHFA's external assessment of facility adherence to standards can serve as a precursor to such systems.

Health facility surveys form an integral component of a country's health information system and health sector monitoring and evaluation plan. The timing of HHFAs should be synchronized with national planning cycles and review processes, so that the results are available in time to feed into these processes. The time needed to complete an HHFA depends on several factors including the size of the country, the number of health facilities, the sample size, the modules selected, the available resources and the number of data collection teams. The entire process generally requires 3 to 6 months from the time of country adaptation of the questionnaires to the production of the country report. Additional time may be needed if the country is conducting a facility survey for the first time or implementing the survey as a census of all facilities.

The frequency of conducting HHFAs depends on country needs and available resources. Ideally, a country would implement only the core availability module as a census of all facilities every 5 years. The implementation of all modules in all facilities is an enormous and costly undertaking and is not recommended. Implementing the availability module as a census would ensure that the national master facility list (MFL) is updated and that minimum information on service availability is produced for the entire country. A representative sample survey using the other modules could then be conducted at intervals of 1 to 3 years. A further option could be to alternate a sample survey of readiness (the most extensive module) with a sample survey of management and finance, and quality of care. The modules selected should be those required to fill specific data gaps for country planning processes or to review specific service aspects that have previously shown weaknesses and for which corrective measures have been implemented. However, countries should adapt this cycle of assessments according to needs and feasibility.

3. HHFA overview

3.1 HHFA objectives

The HHFA is designed to provide reliable, objective information on the status of health facility services that can be used to measure progress in health system strengthening over time. It generates a set of indicators on key service inputs and processes that measure whether or not facilities meet the required conditions to support provision of basic or specific services at accepted standards of quality. The indicators can monitor changes over time using information that is reliable and comparable because it is collected using consistent definitions and methodologies.

HHFA indicators can be used to:

- provide information on the status of facility services as assessed against agreed-upon standards;
- detect changes and measure progress in facility services over time and among administrative/geographical areas;
- generate evidence for health sector reviews to inform the development of strategies and plans and to guide investments;
- support planning and management of facility services, e.g. to address gaps and to promote equitable distribution of services and resources;
- plan and monitor the scale-up of key interventions to address priority health challenges and achieve UHC and the SDGs;
- benchmark facility performance to support the development of quality improvement plans; and
- provide evidence to motivate political and financial support for improving service quality.

Box 1 provides examples of questions that can be answered through an HHFA.

Box 1. Examples of questions

- What percentage of health facilities offers a specific service, e.g. antiretroviral therapy, family planning?
- What percentage of facilities offers the country's basic package of essential health services?
- Are basic resources and systems for general health service delivery established, adequate and functional?
- Are key resources (trained staff, guidelines, equipment, diagnostics, medicines and commodities) in place to provide a specific service at a required level of quality, e.g. for family planning, malaria diagnosis and treatment, hypertension management?
- What is the distribution of health workers (occupational category) across facilities?
- Are appropriate facility management systems established, adequate and functioning?
- Are quality assurance systems established, adequate and functioning?
- Is there evidence that providers followed required standards in the care of individual patients/clients?
- What are the strengths and weaknesses identified across all services and for specific services?
- Are there differences between urban and rural areas, and among different facility levels?

3.2 HHFA modules

The HHFA covers all key facility services and facility-level management systems. It includes both outpatient and inpatient services and has a strong focus on primary health care (PHC). However, the HHFA is not an appropriate tool for assessing complex inpatient or tertiary level services.

The HHFA content is organized into four modules that may be viewed as "lenses" through which facility services are assessed: **1) service availability; 2) service readiness; 3) quality of care;** and **4) management and finance.**

A **module** is defined as a set of questions that provide information about a main topical area. Countries may choose to implement a single HHFA module or a combination of modules. Refer to Fig. 4 for a summary of the types of information collected within each module.

Fig. 4. HHFA modules

Service availability	Service readiness	Quality of care	Management and finance
• Facility infrastructure • Staff • Beds • Specific services • Building structure	• Guidelines • Trained staff • Equipment • Diagnostics • Medicines and commodities	• Adherence to standards in patient care processes	• Management systems • Finance systems • Health information systems • Quality assurance systems

MODULE 1: Service availability – *Are basic infrastructure and services available?*

Module 1 refers to the physical presence of services. It encompasses key facility resources (infrastructure, staff, beds) as well as the availability of specific services in facilities. (Note that this does not include other aspects of service access, such as geographical and social barriers, travel time and user behaviour, which require data collection methods other than the HHFA.)

MODULE 2: Service readiness – *Are the key prerequisites for providing quality services in place?*

Module 2 measures the extent to which the resources and conditions are in place to provide services according to defined minimum standards, including the presence and functionality of basic amenities, trained staff, guidelines, equipment, diagnostic capacity, and medicines and commodities. Facility-level systems to support quality and safety are also assessed. Readiness is assessed for both the overall capacity of the facility to provide basic services and for specific services.

MODULE 3: Quality of care – *Are services delivered to patients according to required standards?*

Module 3 assesses whether individual patients have received appropriate care, measured by assessing provider adherence to standards in the care process, as documented in individual patient records. (Note that the resources and systems considered prerequisites for quality services are integrated into the service readiness, and management and finance modules.)

MODULE 4: Management and finance – *Are facility management structures in place to support continuous availability and quality of services?*

Module 4 assesses the various management systems and practices implemented in the facility, including facility governance, financial practices, management and support of staff, management of medicines and other commodities, health information systems, and quality assurance systems.

Data quality review for routine health information systems (RHIS) data

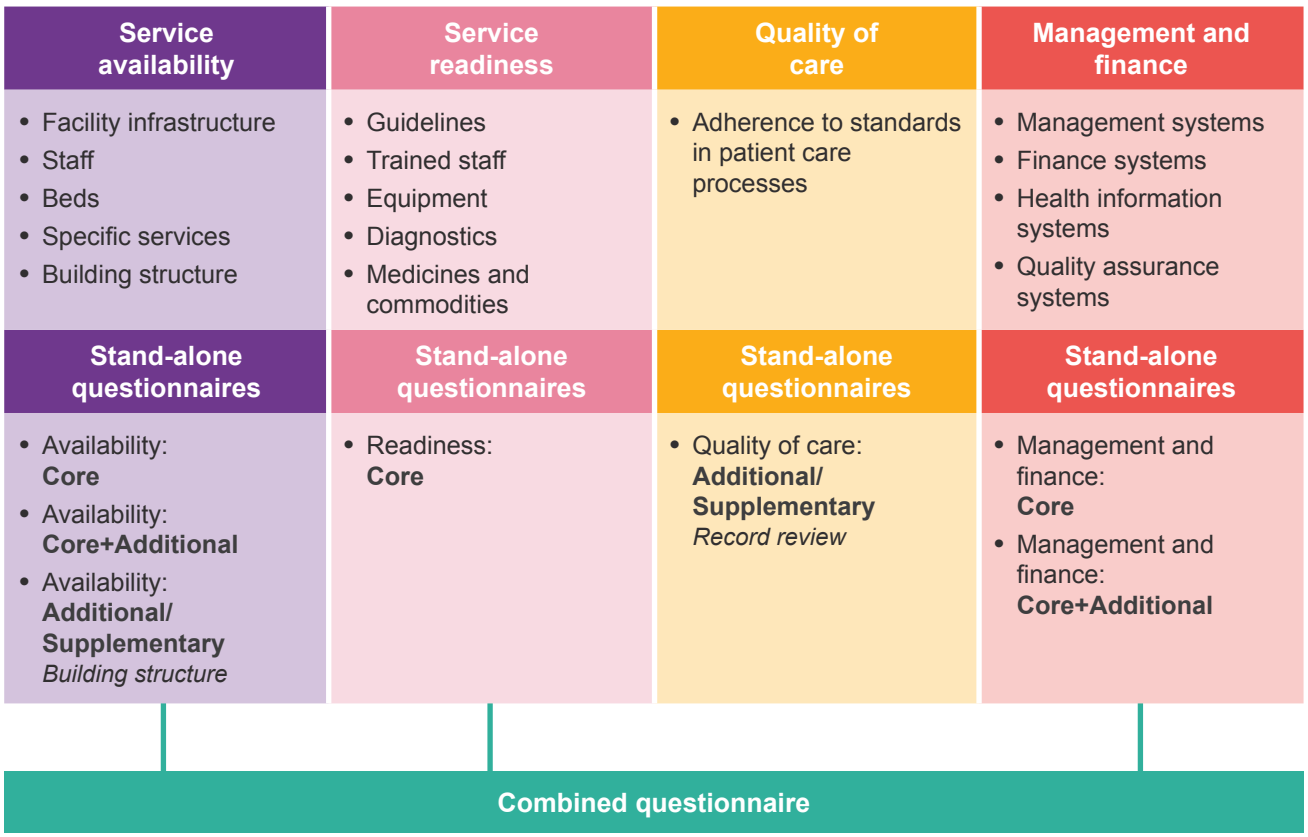
In the past, when implementing the SARA, it was recommended to conduct in parallel a data quality review (DQR) involving a data verification exercise to ascertain the quality of self-reported routine data submitted by facilities through the RHIS. However, given the length and complexity of the HHFA, it is now recommended to conduct the DQR as a separate exercise. The WHO Data Quality Assurance (DQA) toolkit *[2]* provides tools to verify the quality of routinely reported data for selected key indicators and quantifies problems of data completeness, timeliness, accuracy and consistency.

3.3 HHFA questions and questionnaires

The HHFA aims to collect data that are comparable over time, within countries (across regions and/or districts) and across countries. To achieve this, a standardized set of HHFA questions was developed, consisting of core and additional questions. Core questions represent the recommended minimum information, while optional additional questions can provide further details. Typically, a country adopts the set of core questions, with some adaptations related to certain country-specific elements (e.g. types of facilities, managing authorities, staffing categories, national guidelines). Additional questions on selected topics may be included based on country priorities. Questions not relevant to the context can be removed.

The questions are organized into various questionnaires, based on the module, the use of core and additional questions, and the data collection method required. Fig. 5 shows the questionnaires available within each module, as well as the combined questionnaire that includes the facility audit questions of the service availability, service readiness, and management and finance modules.

Fig. 5. **HHFA modules and questionnaires**



The HHFA questionnaires are provided in two formats: stand-alone and combined. Each questionnaire can be downloaded separately as a pdf document or as a Word or an Excel file.

- **Stand-alone questionnaires:**

  Each of the four HHFA modules contains a set of stand-alone questionnaires that are called “Core”, “Core+Additional” and/or “Supplementary”.

  - A **Core** questionnaire contains only core questions.
  - A **Core+Additional** questionnaire contains both core and additional questions.
  - A **Supplementary** questionnaire contains additional questions that are not included in the Core or the Core+Additional questionnaires.

- **Combined questionnaire:**

  - The **Core Combined** questionnaire contains ALL core questions from the three modules that require a facility audit methodology (service availability, service readiness, and management and finance). Questions from the Supplementary questionnaires are not included. Versions of the Core Combined questionnaire that contain different combinations of the facility audit modules are also available.

Question labels

Within a questionnaire, each question has a **unique number** assigned to it. Each question also has a **label** specifying the module to which it belongs (A, R, Q or M) and its designation as Core or Additional (C or A). A further label (“S”) is included for questions in Supplementary questionnaires. A unique three-letter identification (ID) code of the HHFA indicator to which the question is linked, is provided alongside each question in the Excel format of the questionnaire. A single question may be linked to multiple indicators, and a single indicator may require data from multiple questions. Table 2 shows the various question labels related to the module and the Core, Additional and Supplementary designations.

Table 2. HHFA question labels showing module and questionnaire type

Label	Label meaning
A_C	Availability Core
A_A	Availability Additional
A_AS	Availability Additional/Supplementary
R_C	Readiness Core
Q_AS	Quality Additional/Supplementary
M_C	Management Core
M_A	Management Additional
ALL	Question used in all facility audit modules
A_C, R_C	Availability Core, Readiness Core (question used in both Availability and Readiness modules)
R_C, M_C	Readiness Core, Management Core (question used in both Readiness and Management modules)

HHFA questionnaire structure

An HHFA questionnaire is organized into sections and subsections that contain questions related to a specific service aspect or programme. Fig. 6 shows an extract from a questionnaire.

Fig. 6. Extract from an HHFA questionnaire

Mod	No.	Question	Response		Skip
		17. SERVICES FOR SPECIAL NEEDS			
		17.1. PALLIATIVE CARE			
		17.1.1. SERVICE AVAILABILITY			
R_C	1700	Does this facility offer any palliative care services?	YES ..... 1 NO 2		➔Q1706
R_C	1701	Which of the following palliative health services are offered in this facility:	YES	NO	
R_C	01	Inpatient palliative care	1	2	
R_C	02	Outpatient palliative care	1	2	
R_C	03	Home care for palliative care	1	2	
R_C	04	Linkages with other organizations providing home-based palliative care	1	2	
	1702	ASK TO BE SHOWN THE LOCATION IN THE FACILITY WHERE PALLIATIVE CARE SERVICES ARE PROVIDED. FIND THE PERSON MOST KNOWLEDGEABLE ABOUT PALLIATIVE CARE SERVICES IN THE FACILITY. INTRODUCE YOURSELF, EXPLAIN THE PURPOSE OF THE SURVEY AND ASK THE FOLLOWING QUESTIONS.			

The paper questionnaire is structured into five columns:

- Column 1 (Mod) contains the question label.
- Column 2 (No) contains the number of the HHFA question. There may be a single number per question, e.g. 1700 (main question), or a main number with sub-questions below it, e.g. 1701_01 (sub-question). Note that some numbers are not associated with questions, but with headings, clarifying text or instructions, e.g. 1702.
- Column 3 (QUESTION) contains the question that the data collector reads to the respondent. This column may also contain additional clarifying information (in lower case font) that the data collector reads to the respondent, or instructions to the data collector (in upper case font) that are not read to the respondent.
- Column 4 (RESPONSE) contains the response options. Different types of response options are used for different types of questions, e.g. pre-coded responses where one or more options are selected, or fields requiring entry of a number or text, or combinations of these.
- Column 5 (SKIP): This column contains arrows that instruct the data collector to skip to a specific question, to the end of a section, or to other instructions.

Stand-alone questionnaires

MODULE 1: Service availability

Module 1 includes a Core availability questionnaire and a Core+Additional service availability questionnaire. An Additional/Supplementary questionnaire is also available for detailed assessment of the condition of facility building structure

MODULE 2: Service readiness

Module 2 consists of an extensive Core service readiness questionnaire that assesses the basic prerequisites for service delivery in the facility as a whole, as well as the specific prerequisites for specific services. (An Additional/Supplementary questionnaire to assess provider competency may be developed in the future.)

MODULE 3: Quality of care

An Additional/Supplementary questionnaire uses a record review methodology to assess patient care processes in a sample of individual patient records. (A second Additional/Supplementary questionnaire may be developed in the future, using client interviews to assess experiences of care.) Note that Module 3 does not contain any facility audit questions. Questions related to systems and commodities that support quality of care are considered prerequisites for quality and are integrated within the service readiness and the management and finance modules.

MODULE 4: Management and finance

Module 4 includes both a Core questionnaire as well as a Core+Additional questionnaire. (An Additional/Supplementary questionnaire may be made available in the future to assess health worker absenteeism.)

Combined Core questionnaire

The Combined Core questionnaire contains all the questions from all three core questionnaires that use a facility audit data collection methodology (service availability, service readiness, and management and finance). This is the “master” questionnaire. The combined questionnaire is designed to ensure that the data from across the various modules are collected in a coherent, non-repetitive way. From this “master” version, combined questionnaires with various module combinations have been generated. Countries can select the combination they want to implement. For example, a country may want to implement the Core questions of only the service availability and service readiness modules, or only the service availability and management and finance modules. Note that the quality of care questionnaire cannot be integrated into the combined questionnaire, as it requires a different data collection methodology.

Refer to Section 4.4 for further details on the selection of questionnaires for implementation.

3.4 HHFA CSPro electronic data collection tool

The use of electronic data collection devices for field surveys has increased in popularity as a result of decreasing costs, increasing computational and functional capacity, and user-friendliness. Electronic data collection has many advantages. Data validation procedures, including skip patterns, range controls, standardized responses and mandatory question responses, can be programmed into an electronic device to facilitate the collection of accurate and reliable data. Furthermore, automatic progression of the questionnaire and standardized responses make it easy and relatively quick for interviewers to administer the survey. As the size and scope of a survey increase, so also do the benefits of electronic data collection. Large volumes of data are subject to the risk of increased data collection errors. Such errors can be

minimized through the use of electronic data collection devices, in addition to the substantial time saved in data collection, data entry, data cleaning, data integration and data dissemination.

For the HHFA, electronic data collection is conducted using CSPro. CSPro is a software package for entry, editing, tabulation and dissemination of census and survey data. It was developed and is maintained by the United States Census Bureau and partners. CSPro is available at no cost, may be distributed freely, and is available for download at: https://www.census.gov/data/software/cspro.Download.html

The HHFA CSPro data collection tool is a **standard CSPro application** that has been customized for the HHFA and contains all the HHFA questionnaires. The HHFA application for data collection runs on any Android device (recommended), on a Windows 8 touch screen tablet, or on a Windows laptop computer. The tool is available for download at: https://cspro.hhfa.online/

HHFA CSPro electronic data collection tool – key features

The HHFA CSPro electronic data collection system has been designed to provide a robust mobile electronic data entry application with flexibility for country adaptation and mechanisms that facilitate data collection in the field. Key features include:

- the ability to establish survey parameters that enable selection of modules and questionnaires for implementation, with automated turning on or off of questions accordingly;
- simplified processes for loading facility sample, staff listing, and administrative levels;
- a QR code for set-up of tablets;
- assignment of questionnaire sections to data collectors: this allows multiple data collectors to collect data for the same facility at the same time and gives them additional flexibility in selecting the sequence of completing the sections assigned to them;
- merging of facility data from multiple data collectors at the team leader level and report generation to determine facility data completeness;
- data synchronization via Dropbox or CSWeb;
- when data collection is complete, the data are exported from the HHFA CSPro application and can then be uploaded to the HHFA data analysis platform for indicator generation and analysis, or to a statistical programme of choice.

The *HHFA Data manager guide* provides detailed information on adaptation of the HHFA CSPro tool to the country context and its use for survey implementation, including troubleshooting while data collection remains ongoing.

3.5 HHFA indicators and the indicator inventory platform

Any data collected through the HHFA must be indicator driven, with all questions linked to clearly defined indicators. HHFA indicators are designated as either core or additional and are derived from core or additional questions respectively. Core indicators represent the recommended minimum information on key services, items and/or attributes that are needed to deliver services at agreed-upon standards. Additional indicators provide in-depth information on specific topics, based on country needs.

This section describes the various types of HHFA indicators, the grouping of indicators into tables, and the organization of the tables in the HHFA indicator inventory platform.

HHFA indicator types

The HHFA uses six types of indicators: proportion (percentage), mean, count, median, ratio and index. These are described in Table 3. Most HHFA indicators are percentages. Indices (also called summary or composite indicators) are used to summarize and communicate information about multiple indicators. The calculation of HHFA indices is described in Annex 1.

All the HHFA indicators are grouped into the tables provided in the **HHFA indicator inventory platform.** All HHFA indicators are calculated automatically when the data are uploaded into the **HHFA data analysis platform**, as described in Section 3.6.

Table 3. **HHFA indicator types**

Indicator type	HHFA indicator definition and description
Percentage (proportion)	A proportion describes the relationship of the numerator to the denominator, where the numerator is included in the denominator**. A percentage is a type of proportion**, where the denominator is 100. Most HHFA indicators are percentages and are mostly one of two kinds: ■ The percentage of facilities with an available "condition", where the denominator is all facilities, e.g. ▶ Percentage of facilities offering facility planning services ■ The percentage of facilities with an available "condition", where the denominator is limited to those facilities offering the specific service, e.g. ▶ Percentage of facilities offering family planning services with a blood pressure apparatus A "condition" may be a service, an item, a system, or evidence of the implementation of a requirement.
Mean (average)	An arithmetic mean is the **average of a set of values** (the sum of the set of values divided by the number of values in the set). The HHFA mostly uses a mean to express **the average of a set of indicators**, e.g. ▶ Mean percentage of family planning items available among facilities offering family planning services Most means in the HHFA are used to express the average of a set of indicators in a table. A mean used in this way provides **the index** for either a subset of indicators in the table (a "domain" index), or all the indicator in the table. In a very small number of indicators, the HHFA also uses a mean to express the **average of a set of values within a single indicator**, e.g. ▶ Among facilities offering outpatient services, availability of services by mean number of days per week
Count	**Total number of items** The HHFA includes a few indicators that are simple counts, e.g. ▶ Number of maternity beds

Median	**Median number of items** In a set of values ordered from lowest to highest, the median is the value located at the midpoint of the set, with an equal number of values both above it and below it, e.g. ▶ Median number of available inpatient beds among facilities offering inpatient services
Ratio	**Number of items per population** A ratio describes the relationship of the numerator to the denominator, where the numerator is not included in the denominator. The HHFA uses ratios for indicators of **density** of facility infrastructure, inpatient beds and health workforce, e.g. ▶ Number of health facilities per 10 000 population ▶ Number of maternity beds per 1000 pregnant women ▶ Number of midwives per 10 000 population Density indicators should only be calculated using HHFA data if it has involved a census of all facilities in the country and reliable population estimates are available.
Index	An index is **a summary or composite indicator** that is used to communicate information about a selected group of indicators. In the HHFA, indices are useful for summarizing multiple pieces of information, assessing changes over time, or comparing subnational areas. However, indices also have limitations. If only presented with an index, it is difficult to understand the individual factors contributing to the index value; therefore, it is important to present information on the individual indicators within the index, along with the index. The HHFA includes the following types of indices: ■ **An index that is a mean:** ▶ **The mean of all the indicators in a table – the table index or score, e.g.** Basic equipment in the main service area of the facility: mean proportion (percentage) of all items at facilities, e.g. Family planning service readiness: mean percentage of all items at facilities ▶ **The mean of a subset (or domain) of indicators** within a service-specific readiness table – the **domain index or score**, e.g. Family planning service readiness: mean percentage of medicines and commodities items at facilities ▶ **The mean of the scores of a number of tables** – these are called “special tables” and require a complex calculation, e.g. The general service readiness index ▶ **Means within the clinical quality of care dimension** of the indicator inventory (refer to Section 3.5 Dimension 5. Clinical quality of care). ■ **An index that expresses “with all items”:** Each service readiness table (and a few other tables) includes a “with all items” indicator, which describes the percentage of facilities that have available all the items in the table, e.g. ▶ Family planning service readiness: percentage of facilities with all items

HHFA indicator tables

All the indicators in a single HHFA table have the same denominator. The HHFA contains different types of tables that can be described according to the indicator inventory dimension (see below), the presence of indices, and the denominator used.

- **General service availability tables** do not contain indices. The denominator is "all facilities".
- **General service readiness tables** all contain indices: the mean percentage of all items available in facilities, and the percentage of facilities with all items. The denominator is "all facilities".
- **Service-specific readiness tables** all contain three types of indices: domain indices, the mean percentage of all items available in facilities, and the percentage of facilities with all items. The denominator is "facilities offering the specific service".
- **Auxiliary indicator tables** do not contain indices, with a few exceptions. The denominator may be "all facilities" or "facilities offering the specific service".
- **"Special" tables** contain indices based on more complex calculations than the other HHFA indices. These tables are mainly used in the general services readiness dimension. The HHFA contains only a few such tables.
- **Other tables**, mainly used in the management and finance dimension, do not contain any indices and the denominator is almost always all facilities.

The types of HHFA indicator tables, indices, and denominators are summarized in Table 4.

Table 4. Types of HHFA indicator tables

Type of table	Indices	Denominator
1. General service availability	No indices	Mostly all facilities
2. General service readiness	All have indices: mean % of all items, % facilities with all items	All facilities
3. Service-specific availability	No indices (except a few maternal and newborn health services tables)	All facilities
4. Service-specific readiness	All have indices: domain indices; mean % of all items, % facilities with all items	Facilities offering the specific service
5. Auxiliary indicators	Most have no indices; a few have indices similar to readiness tables	All facilities or facilities offering the specific service
6. "Special table"	Have indices with complex calculations; used in general availability and general readiness; indices may be calculated across several tables	
7. Other tables	No indices	All facilities

HHFA indicator inventory platform

The HHFA indicators can be viewed in an online indicator inventory platform (https://indicator-inventory.hhfa.online/). An "instructions" tab in the platform provides explanations and short videos on how to use the platform.

The indicator inventory platform provides a quick way of gaining an overview of the HHFA content and the way indicators are organized, as well as providing detailed information on each indicator. Countries planning an HHFA should firstly review the indicator inventory to identify the indicators they need, and then select the relevant questionnaires and questions.

The complete indicator inventory can be downloaded from the platform as an Excel document. Users can also download a selection of indicators or a selection of data fields. An indicator tabulation plan can be generated from the indicator platform in pdf format, showing how the indicators are tabulated and organized in the report outline produced by the HHFA data analysis platform.

The platform contains all the HHFA indicators. Filter options are available to select the indicators related to specific modules and questionnaires. A search function enables the user to find a specific indicator, or a group of indicators related to a specific topic.

The platform provides two views of the indicators: a "listing" view and a "tabulation" view. The "listing" view displays a simple list of the indicators. The "tabulation" view shows a "nested" list that groups the indicators according to the following hierarchy:

Service dimension → Service area → Service subarea → Table → Indicators

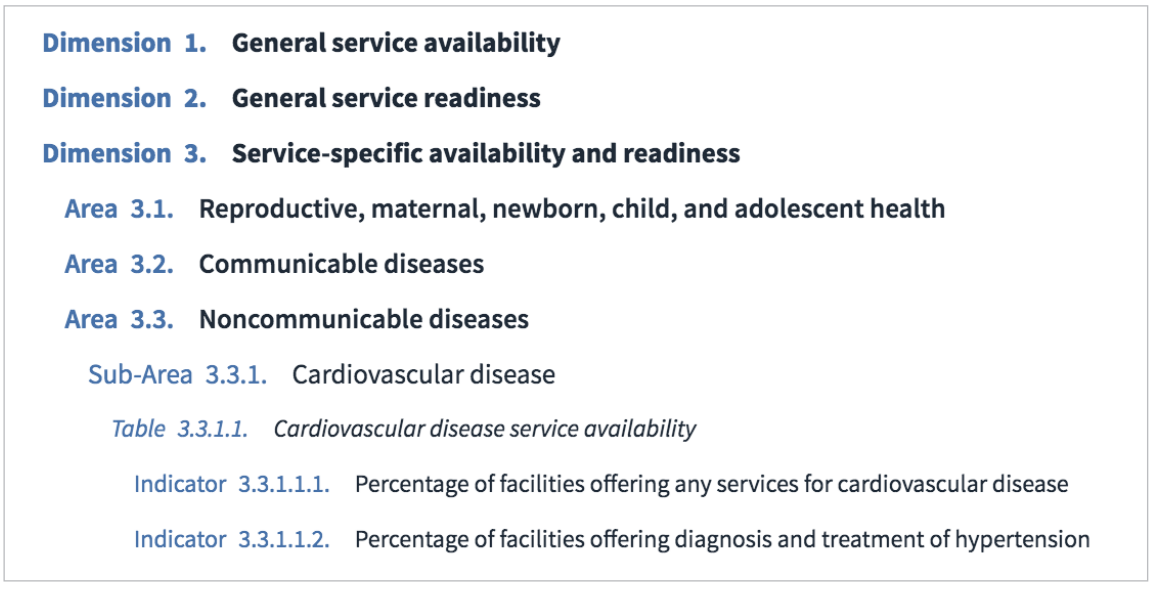
There are five service dimensions that broadly correspond to the four HHFA modules:

- Dimension 1. General service availability
- Dimension 2. General service readiness
- Dimension 3. Service-specific availability and readiness
- Dimension 4. Management and finance support systems
- Dimension 5. Clinical quality of care

This system of indicator organization is also used in the HHFA data analysis platform and provides the structure for the HHFA report outline.

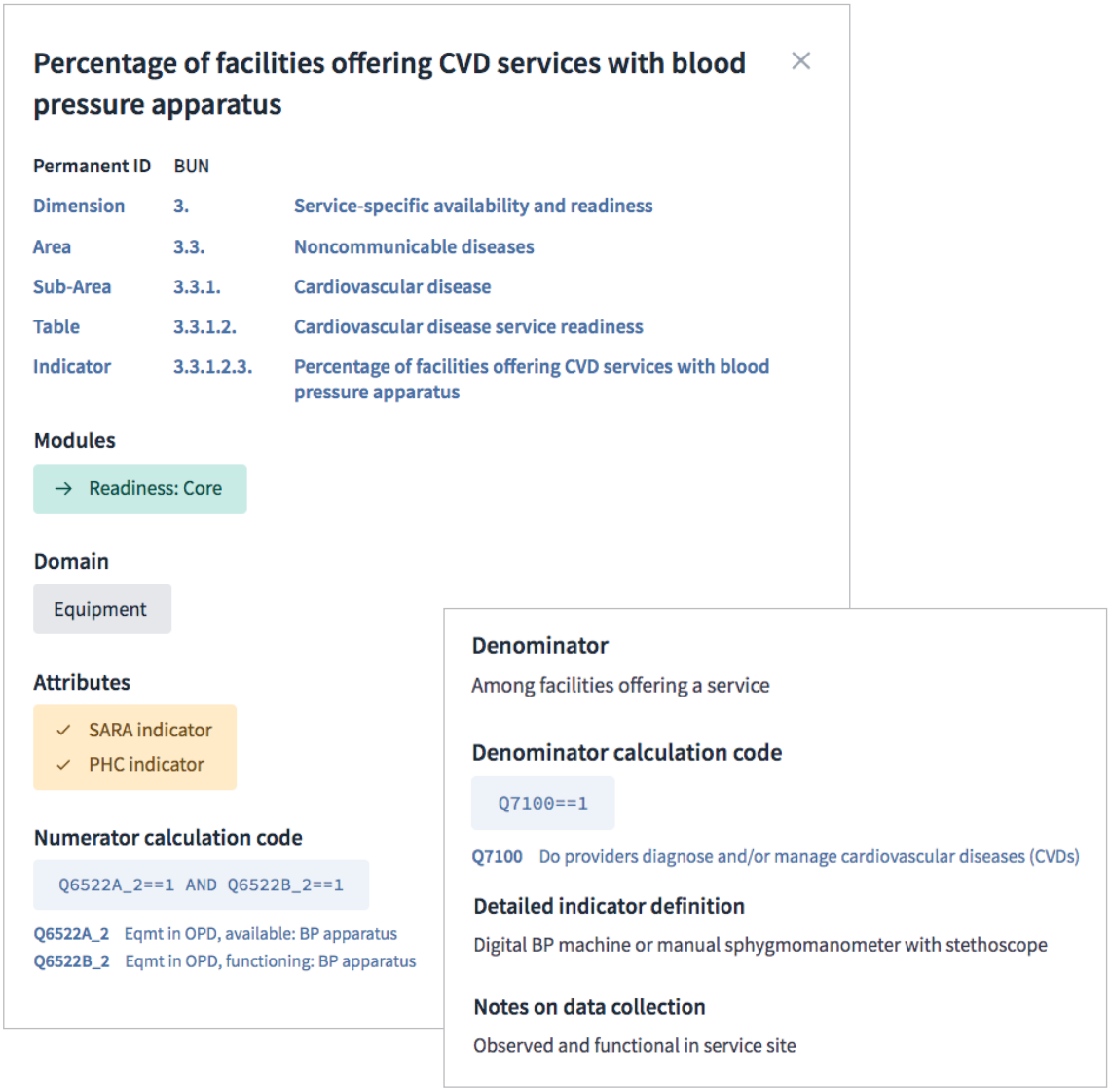
Fig. 7 shows a screenshot of part of the indicator platform, illustrating the indicator organization. Clicking on a dimension will display the areas within the dimension. Clicking on an area will display sub-areas, clicking on a subarea will display tables, and clicking on a table will display all the indicators in the table.

Fig. 7. Indicator organization in HHFA indicator inventory platform



Clicking on an individual indicator will display information that describes the indicator, as shown in Fig. 8. This includes the indicator's permanent, unique three-letter ID code (described in Section 3.3) as well as information on the survey questions and code needed to calculate the indicator.

Fig. 8. Example of indicator metadata as displayed in the HHFA indicator inventory platform



Dimension 1. General service availability

This dimension describes the presence and implementation of facility-level systems and processes that support the continuous availability and quality of facility services. Indices are not calculated for this dimension.	**Service areas** ■ Health infrastructure ■ Health workforce ■ Services available

- **Health infrastructure:** contains indicators of facility and bed density, and supplementary information on building structural conditions and accessibility for mobility-limited persons.
- **Health workforce:** contains health workforce density and median indicators for key occupational categories.
- **Services available:** contains information on general outpatient and inpatient services, specific services, and selected diagnostic and treatment procedures offered. (Most of these indicators are also found in Dimension 3. If a country implements the availability module but not the readiness module, the HHFA report produced by the data analysis platform will contain the services available indicators as part of the general service availability dimension of the report. However, if both availability and readiness modules are implemented, the services available indicators are integrated into the service-specific availability and readiness section of the report.)

Note: The infrastructure and workforce density indicators require population estimates as denominators. Therefore, **these indicators can only be calculated when the survey has involved a census of all facilities (public and private) in the country**. Reliable population estimates are also required, usually obtained by projecting census data.

Dimension 2. General service readiness

General service readiness refers to the overall capacity of facilities to provide basic, general health services, based on the presence and functionality of conditions needed for providing the services. An index is generated for each of the five service areas, based on the number of conditions present. An overall general readiness index is calculated based on the mean of the five service areas.	**Service areas** ■ Basic amenities ■ Basic equipment ■ Standard precautions for infection prevention ■ Basic diagnostic capacity (laboratory) ■ Essential medicines ■ General service readiness index

- **Basic amenities:** assesses power, water source, privacy for consultations, sanitation, communications system, computer with internet and emergency transportation system.
- **Basic equipment:** assesses adult scale, child scale, infant scale, measuring tape, height board, thermometer, stethoscope, blood pressure apparatus, examination light, otoscope, ophthalmoscope and pulse oximeter.
- **Standard precautions for infection prevention:** assesses key items needed for infection prevention and control. Refer to Annex 1 for a detailed list of items. An auxiliary indicator table addresses personal protective equipment.

- **Basic diagnostic capacity (laboratory):** assesses presence of the following onsite testing: haemoglobin; blood glucose; urine dipstick for glucose, protein and ketones; urine test for pregnancy; malaria diagnostic testing; HIV diagnostic testing; and syphilis rapid diagnostic testing.
- **Essential medicines:** assesses a set of general essential medicines that corresponds to the basket of essential medicines for PHC related to SDG 3.8.1. Refer to Annex 1 for a detailed list of medicines.
- **General service readiness index:** each of the five service areas contains a table of indicators representing the presence of basic items. A mean is calculated for each table to provide the readiness index for the service area. This readiness index is the mean percentage of items available among all facilities. A "with all items" index is also provided for each table, indicating the percentage of facilities that have available all the items in the table. In addition, an overall general service readiness index is calculated as the unweighted mean of the five service area readiness indices. Calculation of the readiness indices is described in Annex 1.

Dimension 3. Service-specific availability and readiness

This dimension describes whether health facilities offer a specific service and their capacities to provide that service based on the presence and functionality of tracer items in four domains: trained staff and guidelines, equipment, diagnostic capacity and medicines and commodities.	**Service areas** ■ Reproductive, maternal, newborn, child, and adolescent health ■ Communicable diseases ■ Noncommunicable diseases ■ Surgical services ■ Emergency services ■ Palliative and rehabilitative care

Each service area contains one or more **service subareas** that represent a programme or a programme component. Each service subarea contains one or more of the following types of indicator **tables:**

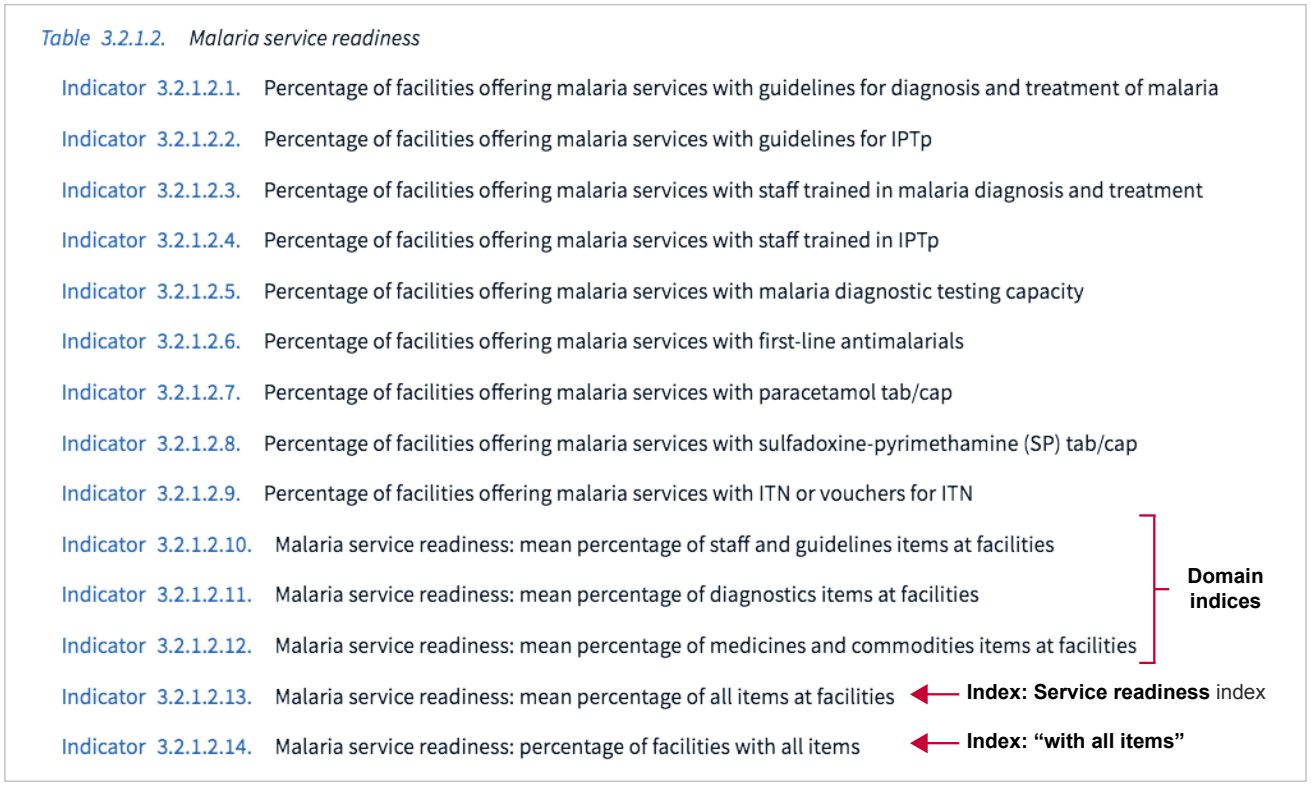
- **Service-specific availability table:** shows the availability of the specific service or service component among all facilities surveyed. The denominator is all facilities.
- **Service-specific readiness table:** contains indicators for a set of items considered essential for the provision of quality care for the specific service. The denominator is the number of facilities offering the specific service. The items are grouped into the four domains: staff and guidelines; equipment; diagnostics; and medicines and commodities. A mean is calculated for each of the four domains and the mean of all the indicators in the readiness table provides the overall readiness index (or score) for the service. A "with all items" index is also produced for readiness tables, indicating the percentage of facilities that have available all the items in the table. Annex 1 describes the calculation of the indices.
- **Auxiliary indicator tables:** there may be one or more auxiliary tables containing further indicators relevant to the programme. The denominator is the number of facilities offering the specific service. Indices are usually not calculated for auxiliary indicator tables.

Fig. 9 shows the availability, readiness and auxiliary indicator tables for family planning services.

Fig. 9. Tables for family planning service availability, readiness and auxiliary indicators

Fig. 10 shows the indicators in the malaria readiness table, including the index for each domain, the overall readiness index for the service and the "with all items" index.

Fig. 10. Malaria service readiness table



Dimension 4. Management and finance support systems

This dimension describes the presence and implementation of facility-level systems and processes that support the continuous availability and quality of facility services. Indices are not calculated for this dimension.

Service areas

- Facility governance and management
- Facility finances and accounting
- Systems to support staff
- Systems for staff and patient safety
- Quality monitoring systems
- Health information systems

- **Facility governance and management:** includes facility management committees, support services for routine facility functioning (e.g. staff transport, laundry) and maintenance systems.
- **Facility finances and accounting:** includes budgets, user fees (if relevant) and accountability systems.
- **Systems to support staff:** includes staff credentials, supervision, training, well-being, and benefits.
- **Systems for staff and patient safety:** includes systems for infection prevention and control, cleaning and emergency preparedness.
- **Quality monitoring systems:** includes quality assurances systems and quality monitoring systems for inpatient care, prescribing, adverse reactions to medicines, immunization services, infection prevention and control, and laboratory services.
- **Health information systems:** includes information management and reporting systems, unique identifiers and patient record systems.

Dimension 5. Clinical quality of care

This dimension describes the provider's adherence to standards in the patient care process, as documented in individual patient records. Clinical quality of care indicators are available for each of six[1] service areas: antenatal care, HIV testing and counselling, antiretroviral therapy, preventing mother-to-child transmission of HIV (PMTCT), malaria and tuberculosis.

Service areas

- Antenatal care
- HIV testing and counselling
- Antiretroviral therapy
- Preventing mother-to-child transmission (PMTCT) of HIV
- Malaria
- Tuberculosis

The clinical quality of care dimension is different from the other dimensions in that the data are not collected through the HHFA facility audit methodology, but through a separate record review methodology, using a sample of individual patient/client records per service area (refer to Section 4.3 for further details).

[1] Record review questionnaires and indicators may be developed for additional service areas in the future.

For each service area, there are indicators representing various steps in the care process that can be used to inform quality improvement strategies. An index is also calculated for each individual patient to represent the complete care process (comprised of a package of assessment, diagnosis and treatment). This enables assessment of the quality of a complete service, as well as the individual components of the care process. Finally, a facility average score or index is calculated for each service, based on the mean of the individual patient indices.

The indicators were developed in collaboration with subject-matter experts and aligned with standards of care for each service area. Countries may need to adapt the indicators to ensure that they align with the service packages used in the country. However, efforts should be made to maintain the standard HHFA indicator definitions and calculations, to facilitate interpretability of results as well as comparisons over time and across countries. If necessary, country-specific additional indicators can be added, rather than substantially modifying the standard HHFA indicators.

The sampling method for the HHFA record review was developed to provide a picture of the routine care process for a specific service at the facility level. Analysis is carried out using the facility as the unit of measure[1] and is based on the facility-level index, and not on the individual patient indices. When using the facility as the unit of measure, facility averages for results are calculated, and subnational or national results represent the averages across facilities. The results are weighted so that the facility results are representative of the facility distribution in the country. (This weighting method is also used for the facility audit part of the HHFA.) This ensures that the practices in facilities with smaller caseloads (often remote and less-supervised facilities) are proportionally represented – an important consideration when working to ensure equitable access to quality services. Using the facility as the unit of measure may also provide information that helps to identify management and supervision issues that are present across the service.

If a record review has been conducted in conjunction with an HHFA facility audit that measures service availability and readiness, efforts should be made to combine the findings from both assessments to provide an in-depth, meaningful understanding of the service. If the record review reveals that certain processes have not been conducted (as evidenced by lack of recording in the patient record), it is important to know whether the items required to conduct the processes were available or not. For example, failure to conduct a malaria blood test may be the result of a provider not following protocols, but also may occur because the test is not available in the facility. Combining service availability, service readiness, and technical care process data provides a unique opportunity to deepen understanding of the challenges and potential intervention points for improving health service quality.

3.6 HHFA data analysis platform

The HHFA data analysis platform is an online web application that automatically calculates the HHFA indicators, and displays them in tables and graphs, as well as providing the outline of the HHFA report. The analysis platform can be used to undertake all the analysis required to produce the HHFA report for each indicator. A country team creates a country-specific version (or “instance”) of the analysis platform and decides who will be given access to it. The team can then work collaboratively in the platform on the analysis of the HHFA dataset and report.

[1] The HHFA uses the facility as the unit of measure (rather than the individual patient/client) for the following reasons: 1) individual-level analysis provides information on the quality of care received by the majority of patients/clients, but may mask issues in low-volume facilities which are often isolated or under-resourced; 2) individual-level analysis requires data on facility caseloads which would require a level of time and effort not feasible in the context of the HHFA; and 3) the HHFA aims to evaluate the care process rather than an individual service provider, as the care process is frequently influenced by factors other than provider expertise, such as availability of specific tools or commodities to provide a service and by service expectations of supervisors at facility or higher level. However, as a future development, WHO may also explore options for using the individual as the unit of analysis for an in-depth assessment.

The platform can be accessed at https://analysis-platform.hhfa.online. Anyone with a WHO e-mail address (@who.int), or who has already been given guest access to the WHO sign-in system, can access the platform by navigating to the website and clicking "Sign in". Country team members who do not have WHO guest access should contact the HHFA coordination team at hhfa@who.int to request access. An invitation will then be sent to the team members which will grant them access to the platform through the WHO single sign-in system (Microsoft SSO). Once an invitation has been accepted and access has been granted, users can navigate to the website and click "Sign in" to start using the analysis platform.

The data analysis platform is designed to make it as easy and quick as possible to analyse an HHFA dataset. Ideally, a team uploads a cleaned dataset (cleaning completed in CSPro), configures a minimal set of options in the platform, runs all the analyses required to produce the default (standard) indicators and tables, and generates a template report in Word (and PDF) within 1–2 hours. A dataset that has been collected using the official standard HHFA CSPro package, and has been vetted and cleaned, should not require significant configuration on the platform. A dataset collected using a customized tool, or with renamed questions or variables, will require some additional configuration and additional time. Also, any country-specific indicators will need to be added to the platform.

The platform runs fully online and does not require any software to be downloaded or installed. The user can simply navigate to the website and start to use the tool.

Also, the platform does not require users to be familiar with any statistical software (R, Stata, SPSS, SAS, etc.). However, users will need a foundational understanding of descriptive statistics to be able to interpret and interrogate the platform's analytical outputs, and to configure any required custom indicators.

How the HHFA data analysis platform works

The platform works by running analysis scripts that are auto-generated from indicator-specific codes in the HHFA indicator inventory, and from information that the platform infers from the uploaded dataset. Using the pre-established code for each indicator's numerator and denominator, and the list of questions and variables in the dataset, the platform internally creates an analysis script to produce individual tables and figures. These scripts are then run on the dataset, and the outputs are presented to the user in the form of tables and graphs. The basic process for using HHFA analysis platform is described in Box 2.

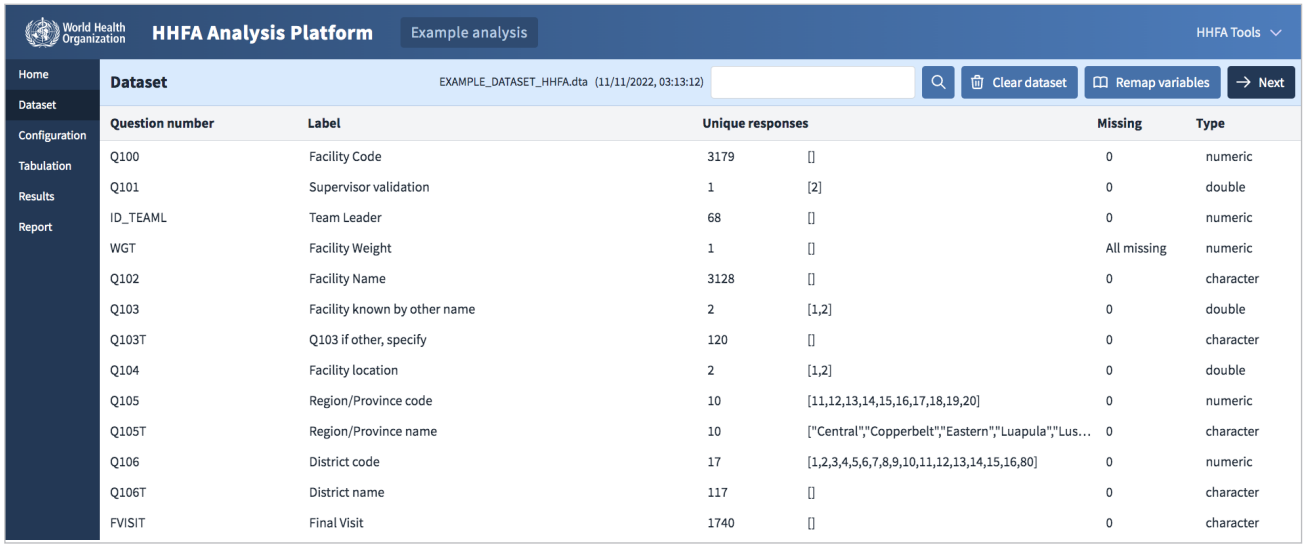
Box 2. How to use the analysis platform

1. Login to the platform and create a "version" of the report that the team aims to prepare.
2. Upload a cleaned dataset.
3. View the dataset (e.g. Fig. 11), checking that all variables exist as expected, and remap any variables that have different names to those used in the HHFA indicator inventory.
4. Choose the variables that you want to use as "stratifiers", e.g. for administrative region, managing authority, urban/rural, facility type (e.g. Fig. 12).
5. If the dataset represents a census, enter information on population numbers, to enable the platform to calculate density indicators.
6. Import the set of indicators (from the HHFA indicator inventory) that you want to calculate for your analysis and report, and that are appropriate for your dataset (depending on the modules and core and additional questions that were used for data collection).
7. Add, remove, or customize indicators as needed (e.g. Fig. 13).
8. "Run all" tables (which may take some time to complete), or individually click "Run" for each table. Check that each table has run successfully and has been produced. You will be able to see the table on the screen and produce related charts and figures (e.g. Figs 14 and 15).
9. Once all tables have been produced, you have the option to generate a Word document, with all tables contained in it, or to print a PDF version of all tables.

For users who may want additional control over the statistical analyses, there are options to view and check the indicator code and the underlying scripts, and to download the intermediate outputs that are created as part of the analyses (for example, indicators calculated at facility-level, before aggregation).

All the codes for indicator calculation can be seen and edited within the platform. Users are encouraged to interrogate these codes and ensure they align with their dataset. Users are also able to download the auto-generated analysis scripts and run them offline, outside of the platform, using R or Stata.

Fig. 11. Viewing the uploaded dataset



Question number	Label	Unique responses		Missing	Type
Q100	Facility Code	3179	[]	0	numeric
Q101	Supervisor validation	1	[2]	0	double
ID_TEAML	Team Leader	68	[]	0	numeric
WGT	Facility Weight	1	[]	All missing	numeric
Q102	Facility Name	3128	[]	0	character
Q103	Facility known by other name	2	[1,2]	0	double
Q103T	Q103 if other, specify	120	[]	0	character
Q104	Facility location	2	[1,2]	0	double
Q105	Region/Province code	10	[11,12,13,14,15,16,17,18,19,20]	0	numeric
Q105T	Region/Province name	10	["Central","Copperbelt","Eastern","Luapula","Lus…	0	character
Q106	District code	17	[1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,80]	0	numeric
Q106T	District name	117	[]	0	character
FVISIT	Final Visit	1740	[]	0	character

Fig. 12. Configuring the stratifier variables

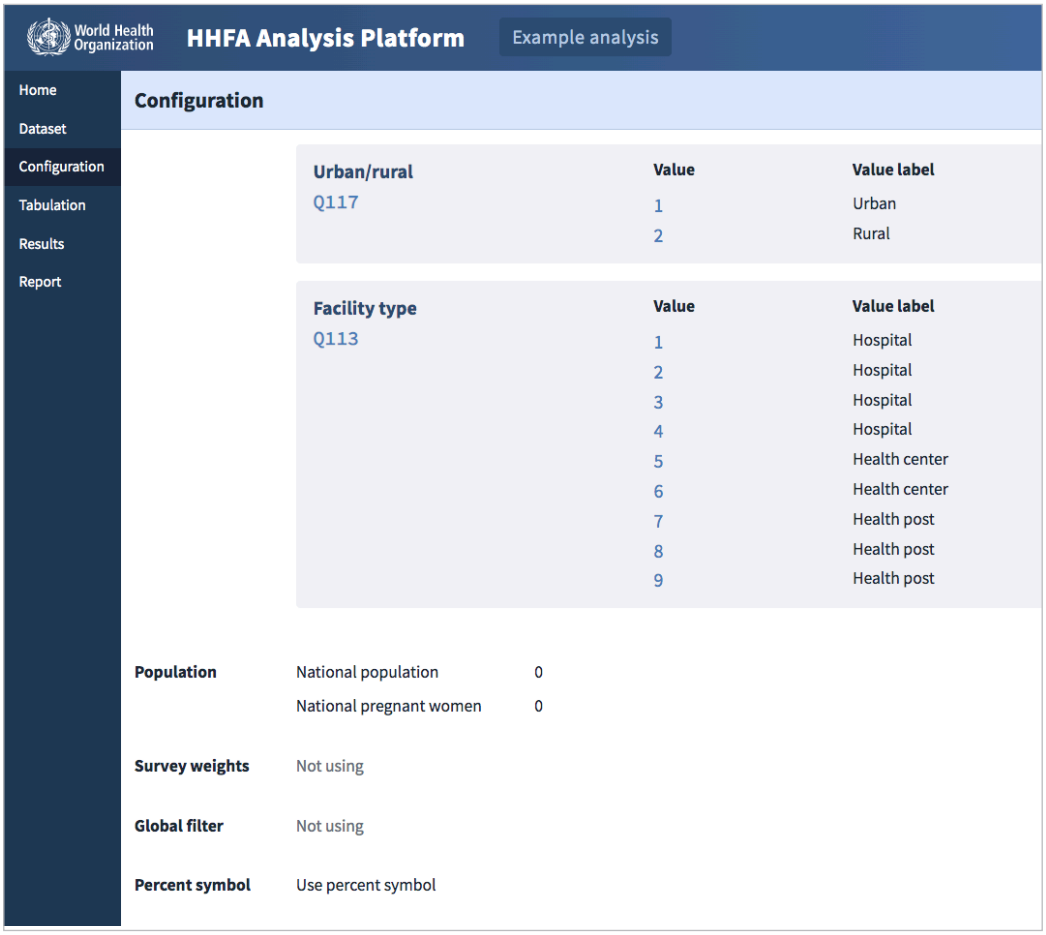


Fig. 13. Customizing indicators

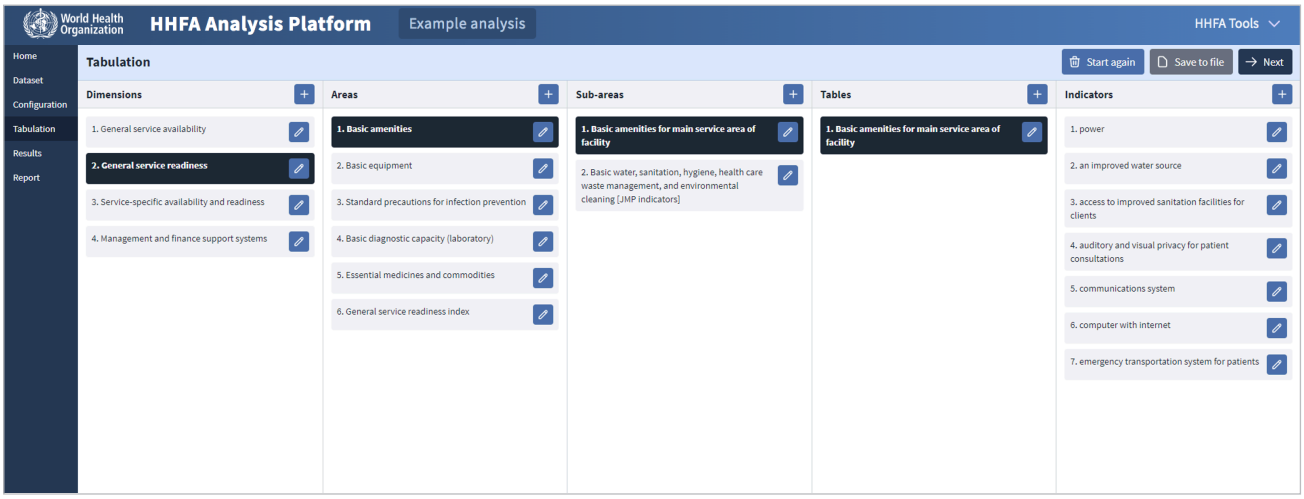


Fig. 14. "Running" individual table analyses to see outputs

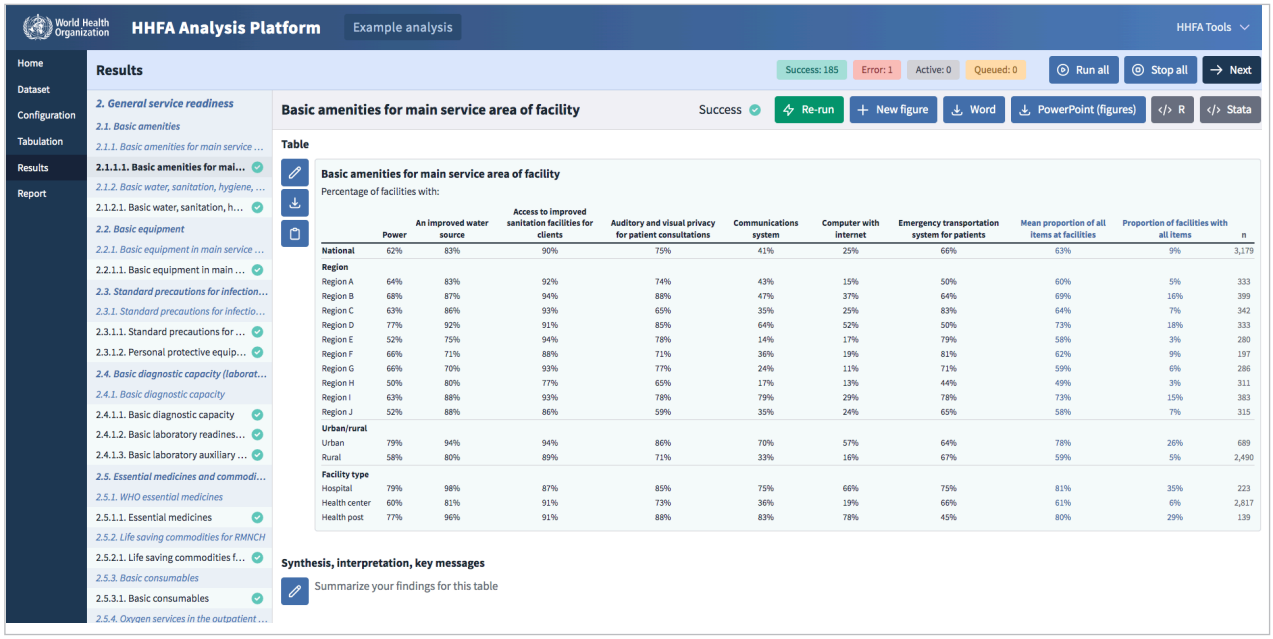
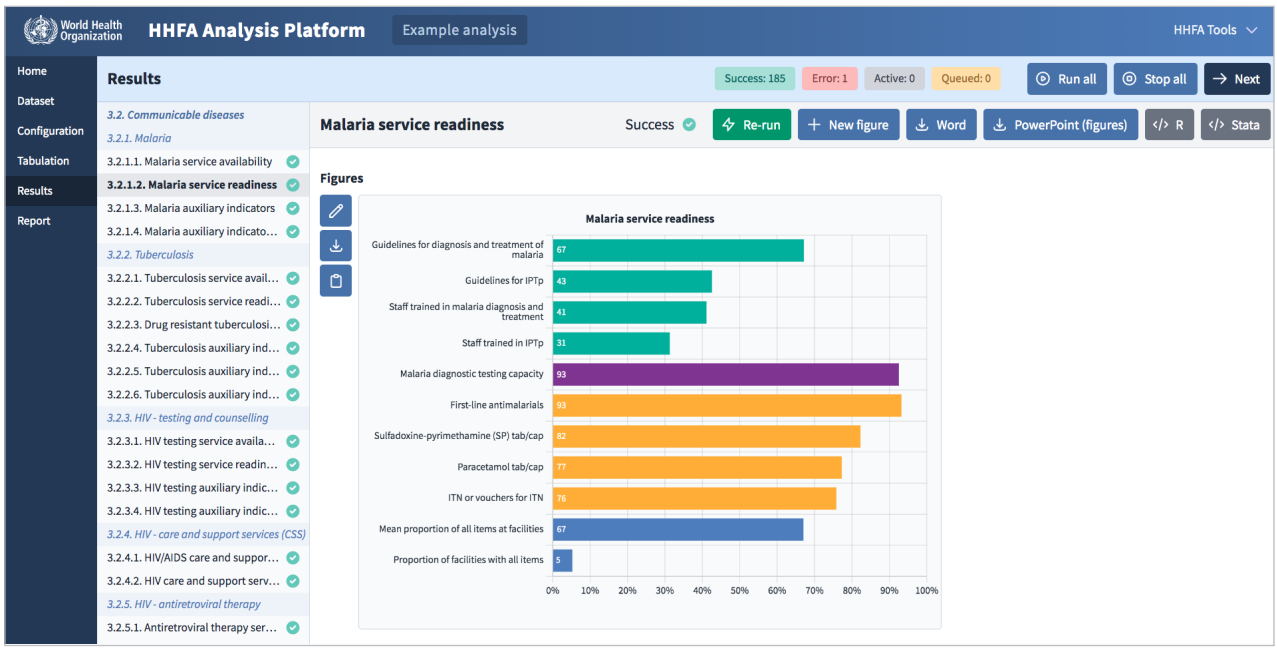


Fig. 15. Options for customizable charts



4. HHFA methodology

4.1 Establishing or updating a master facility list (MFL)

Regardless of the HHFA design methodology selected (census or sample), a complete list of ALL the health facilities in the country is required before the HHFA can be conducted. This is called the master facility list (MFL) or national facility registry. It serves as the basis for constructing the HHFA sampling frame. The MFL is a database that must include all health facilities in all sectors, including public sector, private-for-profit sector, nongovernmental organizations (NGO), faith-based organizations (FBO) and other sectors, e.g. military. It is strongly recommended that countries invest in establishing and maintaining an updated, comprehensive MFL. WHO and partners have developed a guide to support countries in creating and strengthening an MFL *[3]*. A facility census may be needed to establish the MFL. The MFL should be updated annually through facility/district self-reporting, with validation approximately every 5 years through a census.

The MFL assigns a unique identification code to each facility. If a specific ID is not attached to each facility, there is a risk of duplicate data collection in the HHFA. Furthermore, unique facility IDs also enable comparisons across different surveys and over time.

In the MFL, a set of identifying information is provided for each facility, including:

- unique ID code
- facility name
- facility type (from hospital at the highest level to first level primary care facility at the lowest level, according to the country classification system)
- managing authority, e.g. ministry of health, municipal, private-for-profit, FBO, NGO, military
- geographic/administrative area (province, district, town or village)
- urban/rural location
- geographical positioning coordinates (the geographical coordinate collection method should also be recorded, i.e. global positioning system [GPS], digital place names, gazetteers, etc.).

During the HHFA planning phase, the existence and reliability of an official MFL must be assessed. Before the survey can be implemented, it is essential that ALL health facilities in the country are identified. **Failure to ensure an accurate MFL will result in substantial HHFA data quality problems** (e.g. duplicate facility names, omission of important facilities), which may impact the survey results. Time consuming efforts will be needed to correct such errors during the data cleaning phase and will delay the data analysis process.

In some countries, a comprehensive, updated MFL containing all the required information may be available. In many cases, however, this information is not readily available and must be compiled and updated. The ministry of health generally maintains information on public sector facilities and sometimes also on NGO and FBO facilities. Private health facilities are often not consistently included in the MFL, particularly in unregulated contexts. However, they can represent an important share of the health facility network in some countries and efforts should be made to identify as many as possible.

If the MFL is incomplete or out of date, efforts should be made to quickly identify all facilities (public and private) by reviewing the most recent ministry of health lists, databases and other sources, validating them with district/regional health officers, and cross-checking the initial list with other sources (NGOs, private providers, etc.). All available health facility listings must be reconciled to create a single, comprehensive list. If it is impossible to compile an exhaustive list of private facilities, the option of limiting the HHFA to the public sector health facilities may be considered.

4.2 HHFA design methodology

Two potential design methodologies can be used for facility selection in an HHFA:

- A facility census (assessment of all health facilities in the country).
- A sample survey (assessment of a representative sample of health facilities).

A **facility census** aims to assess ALL health facilities (public, private and other providers) in the country. A census can be used to establish or update the national MFL and to establish baseline information on overall service availability in the country. A census may also be used to establish baseline information for other HHFA modules.

In a **sample survey** not all health facilities are assessed, but rather a representative sample, the results of which can then be inferred to represent the target group[1] of health facilities. Sampling typically uses probabilistic selection methods to ensure that the findings are representative of the country or regions/districts in which the survey is conducted. The national MFL is used as the basis for the sampling frame. If the country does not have a comprehensive, up-to-date MFL, the process of drawing a stratified, random sample of health facilities becomes complicated, if not impossible. Therefore, it is necessary to establish or update the MFL before implementing a sample survey. Refer to Annex 2 for details on calculating the sample size and procedures for sample selection.

The **choice between a facility census and a sample survey** depends on a number of factors including the survey objectives, the size of the country, the number of health facilities, the country context, the resources available, the required timeframe of the survey, and the availability of a valid MFL. For example, if the objective is to establish the MFL, and to obtain information on facility density and service availability in all facilities in the country, a census is needed. If the objective is to obtain high-level national estimates, a sample survey that provides accurate estimates of acceptable precision[2] at the national level would be appropriate. However, if the objective is to obtain estimates to assess and compare subnational areas (e.g. provinces/districts), the sampling methodology must be adjusted to use either a stratified sampling design (with increased sample size) or, in some cases, a census of all facilities. Table 5 summarizes the advantages and disadvantages of the two strategies.

The large sample sizes needed for surveys designed to provide estimates for lower subnational levels (or any other possible disaggregation) have significant cost implications. Countries often use a sample survey to assess PHC facilities along with a census of all hospitals. Refer to Table 7 in Section 6.5 for an overview of sampling options and cost implications and to Box 4 in Section 9.1 for a discussion on when it is correct to make comparisons among different strata in sample-based HHFAs.

[1] "Target group" (e.g. the "universe" of health facilities) is used here to designate the statistical term "target population", which might be confusing in the HHFA context, since the objects of the assessment are health facilities and not population members.

[2] Precision is a measure of the uncertainty around the estimate produced by a sample, resulting from the fact that the inclusion of facilities in a sample is partly determined by chance.

Table 5. **Advantages and disadvantages of census versus representative sample**

	Facility census	Sample survey
Advantages	■ Since all facilities are included, there is no statistical variation. ■ It is possible to carry out a geographically disaggregated analysis without bias and uncertainty, e.g. comparing provinces or districts. ■ There is no need for weights; the analysis is straightforward.	■ Data collection from a sample is more efficient than from a census. ■ It is less expensive than a facility census. ■ Concentrated efforts can be devoted to supervision to ensure the collection of quality data, compared with a census.
Disadvantages	■ If the number of facilities is large, the survey will be expensive, and the time required for data collection will be prolonged (but depends on the number of data collectors). ■ The logistics are more complex to reach every facility and hence and implementation problems are more likely to occur. ■ The quality of the data may suffer, due to data collector fatigue, less supervision, etc. ■ Biases can affect the quality of data: undercoverage, nonresponse, interviewer, response and coding biases.	■ Disaggregated analysis of data is limited because of a potentially small number of facilities. ■ The sampling design can be complex and requires special expertise for proper selection and weighting. ■ If disproportionate sampling is used, weights need to be applied. ■ There is statistical uncertainty inherent in probabilistic samples which introduces imprecision of findings.

4.3 Data collection methodologies in facilities

HHFA data collection methods currently include:

- facility audit (key informant interviews and observation).
- record review.

Facility audit

A facility audit assesses facility adherence to minimum criteria for service availability, readiness, and management and finance. (The key prerequisites for service quality are integrated within the readiness, and management and finance modules.) Facility audit methodologies used in the HHFA include observation of key items, systems and service conditions, and interviews with key informants. These methodologies are used in all the HHFA questionnaires currently available for the availability, readiness, and management and finance modules.

Record review

The HHFA quality of care module requires a record review methodology. A technical quality of care record review (also called a “clinical audit”) involves review of a sample of individual patient records to see if the provider followed the appropriate standards throughout the care process, based on the information documented in the patient record.

Record review sampling occurs in two stages. First, facilities are sampled. The procedures for this first stage of sampling (selection of facilities) are the same as those for selecting the sample of facilities for the facility audit part of the HHFA (refer to Annex 2). Second, at each facility visited, client records are sampled.

Once at the facility, a sample of five[1] eligible patient records are identified and reviewed for each targeted service area *[4]*. (Further details on the process for identifying the sample of patient records are available in the HHFA quality of care record review questionnaire.[2] The process should be adapted to the record-keeping systems in use in the country.)

The record review methodology has some recognized limitations, including:

- Documentation does not necessarily prove that a service was provided, and conversely, a lack of documentation does not necessarily mean that a service was not provided.
- Record review relies on service providers knowing which aspects of patient care are expected to be documented.
- A register or other database that provides the information for selecting a sample must exist, and patient information (either in individual patient records or in registers) must be available.

Quality of care can also be assessed through various other methodologies including direct observation, gold standard reassessment, health worker vignettes and client simulations. However, the record review methodology was selected for the HHFA for the following reasons:

- Record review does not depend on patients with a specific condition being present on the day of the facility assessment. Records can be identified for patients that have received services for a specific condition, regardless of when the services were provided.
- It is possible to assess the care process over time for patients requiring follow-up (e.g. tuberculosis, antiretroviral therapy, antenatal care), since the details from all visits should be documented in their individual records.
- Knowledge that quality of care will be assessed using recorded evidence may improve documentation by providers, as well as reinforcement by management staff of the expectation for documentation.
- Record review is less resource intensive than other methods for assessing technical quality of care.

4.4 Selecting HHFA modules and questionnaires for implementation

The HHFA is designed to provide flexibility for countries to implement the survey according to their needs. Countries can select the modules and questionnaires they plan to implement, based on the objectives of the survey, the need to address existing data gaps for policy, planning and monitoring purposes, and implementation feasibility.

Countries may choose to implement a **single HHFA module**. In this case, the relevant **Stand-alone questionnaire** within the module is used for data collection. For example, a country may only need information on service availability and may decide that only the core indicators are needed. The availability Core questionnaire will then be used.

Within a single HHFA module, EITHER the Core OR the Core+Additional questionnaire should be selected, as the Core+Additional questionnaire includes all the questions in the Core questionnaire as well as the additional questions. A Supplementary questionnaire can be used along with a Core or a Core+Additional questionnaire or may be implemented on its own.

[1] A sample of five records per facility is used for feasibility reasons.

[2] https://www.who.int/publications/m/item/harmonized-health-facility-assessment-(hhfa)quality-of-care--record-review

Countries may also choose to implement **multiple HHFA modules**, in various combinations. Whenever more than one module will be implemented, the **Combined Core questionnaire** should be used. The Combined Core questionnaire is available in various versions with various combinations of modules and is designed to ensure that data across the various modules are collected in an integrated way.

Implementation of all the core and additional questions from all the modules during a single survey is not recommended. The resulting combined facility audit questionnaire would be excessively long, with long and complex data collector training requirements, a very high data collection burden (with risks to data quality), increased costs, and an overwhelming volume of data to analyse, interpret and report. Furthermore, the quality of care module requires specific expertise and training, and may be time consuming to implement. Therefore, it is not recommended to implement this module along with multiple facility audit modules.

If a country chooses to implement all three facility audit modules (availability, readiness, and management and finance), it is recommended that only the Core questions are used. If a country needs the detailed information provided by the additional questions, it is recommended to implement only one module at a time, using the relevant Stand-alone Core+Additional questionnaire. However, countries implementing the Combined Core questionnaire may wish to review the additional questions in the stand-alone questionnaires and to incorporate selected additional questions into their country-adapted Core Combined questionnaire as needed.

The decision on the modules and type of question to implement should also be considered along with the decision on sample size. For example, if high-level information is needed across a range of service aspects to inform policy decisions, the core questions of all three facility audit modules may be implemented using a sample that provides information of acceptable precision at national level. If the need is to compare service delivery capacities among subnational areas, the core availability and readiness modules may be implemented using a sample size that allows subnational comparisons. If in-depth information is needed related to a specific module, the Stand-alone Core+Additional questionnaire and/or the Additional/ Supplementary questionnaire may be implemented for the specific module.

4.5 Adapting the questionnaire to the country context

Questionnaire adaptation involves making a limited number of changes to the selected questionnaire, based on policies, practices and terms used in the country, and on country needs. Adaptations, as well as the addition or removal of questions, should be done with care, as changes have cascade implications (e.g. skip patterns, calculation of indicators, etc.).

The HHFA provides standard questionnaires, of which a stand-alone or combined version should be selected as described in Section 4.4. However, some degree of adaptation will be needed in each country context. Questionnaire adaptation is conducted by the HHFA technical committee, in close collaboration with national stakeholders and the key resource persons from the appropriate technical units/programmes.

It is essential that questionnaire adaptation is done very thoroughly:

- to ensure that the questions and terms used in the questionnaire reflect the country context and are understood by local health workers;
- to avoid confusion or controversy during data collector training;
- to minimize the risks of inconsistent data collection and poor-quality data; and
- to avoid problems resulting from data quality issues at the data analysis stage.

The adaptations should initially be made in the Excel version of the paper questionnaire. Once the adaptations are finalized in the Excel version, the final version, with **all changes clearly indicated**, is provided to the data manager/CSPro expert for adaptation of the CSPro tool.

Adaptations to reflect the specificities of the health care system in the country

All countries typically require adaptations for some or all of the following components:

- cover page and facility identifiers
- facility types
- managing authorities of facilities
- health worker occupational categories/qualifications
- national guidelines for services
- national policies for medicines (e.g. for HIV, tuberculosis)
- immunization schedules, etc.

The above components often include the term "COUNTRY ADAPT" in the questionnaire. However, during the questionnaire adaptation process, all questions should be reviewed to see if additional country adaptation may be needed.

If implementing a **quality of care record review**, the questionnaire instructions must be adapted to the country context. The record review questionnaire contains basic instructions for obtaining a sample of patient records for review. However, different countries and different programmes may have different systems of patient lists/registers and patient records. Therefore, the instructions for obtaining the sample and the individual patient information should be adapted according to the country context and should be adequately addressed during the data collector training. Refer to Sections 3.5 (Dimension 5) and 4.3 (Record review) for further details on record reviews.

Adding questions

Countries may choose to include additional country-specific questions. The recommended way to number these specific country questions is to use the country ISO.2 code.[1] For example:

SL_01: where SL corresponds to the ISO.2 code for Sierra Leone. The new questions are numbered sequentially. For a list of country ISO codes please refer to:

http://www.iso.org/iso/country_codes/iso_3166_code_lists/country_names_and_code_elements.htm

Before adding a question, countries should check that it does not already exist in a different part of the questionnaire. If adding a question to a Core questionnaire, it is also important to check if the question already exists as an additional question in the Core+Additional questionnaire. The relevant question can then be copied into the Core questionnaire. Any new question should be considered in terms of the analysis outputs. Before a question is added, the related indicator(s) should be developed, and it should be made clear how the indicators will be included in the tables of the analysis platform.

Deleting questions

It is possible that certain questions might not be relevant to a specific country. In this case, the question can be removed from the questionnaire. Both the question and the question number should then be deleted. A single question may be linked to multiple indicators in the indicator inventory and data analysis platforms. Therefore, before deleting a question, it is important to review the indicators to which it is linked, to assess how the indicator might be impacted.

[1] ISO codes are internationally recognized codes used when referring to countries and their subdivisions.

Deleted question numbers must not be re-used and the subsequent questions must not be re-numbered. Each question number is linked to a specific indicator in the data analysis platform; therefore, if the question numbers are changed, the analysis platform will not work properly.

Important tips for questionnaire adaptation

- **Documentation of changes to questions:** Always use and keep a questionnaire version with both the original version and the changes clearly indicated.
- **Order of questions:** Do not change the order of the questions.
- **Numbering:** Do not change the question numbering; the original numbering structure of the standard questionnaire should be maintained. Changing the numbering will affect links to the existing tools for automated data processing and analysis output production.
- **Skip patterns:** Adding or deleting a question may require changes to skip patterns. Always check the skips and adapt if needed.
- **Modifying questions:** The standard question text should not be replaced by revised question text that will change the meaning of the question. If needed, clarification can be added in parenthesis to help the respondent understand the question. It is very important to keep each question with its original numbering; therefore, do not change the content of existing questions. Rather than substantially modifying an existing question, it should be deleted, and a new question added.
- It is also recommended to **maintain the standard HHFA indicator definitions and calculations**, to facilitate interpretability of results as well as comparisons over time and across countries. If necessary, country-specific additional indicators can be added to the data analysis platform, rather than substantially modifying the standard HHFA indicators.
- New questions will require **new indicators** to be created and added to the **data analysis platform**.
- Changes to the questionnaire will require **modification of the CSPro tool**. Therefore, sufficient time should be allocated to this process between finalizing the paper questionnaire and starting the data collection training.
- It is important to remember that the HHFA focuses on providing information on key items and systems that should be available in health facilities. It is not intended to provide comprehensive data on all aspects of health service functioning. It is important not to stray from this HHFA concept by adding a long list of additional questions.

It is also important to remember that the HHFA questionnaires are extensive and that adding questions will increase the time and costs of the data collection training, the data collection, and the data analysis and interpretation processes. A questionnaire of excessive length may also risk compromising the quality of the data collection.

5. HHFA planning and implementation steps

The HHFA is a complex survey that requires the collection of a large amount of data in a large number of facilities. As in all complex surveys, the HHFA must address both methodological issues and potential operational constraints. Thorough planning and preparation of the survey can minimize such constraints. Table 6 summarizes the steps for planning and implementing an HHFA. The steps may not necessarily follow the sequence presented here and some may occur in parallel. The details of each step are discussed in the following chapters.

Table 6. Summary of HHFA planning and implementation steps

Survey planning and preparation (Chapter 6)
1. Establish a survey coordination group of country stakeholders
2. Define key roles and responsibilities to oversee and facilitate the survey
3. Define the survey objectives
4. Establish or update the national MFL
5. Determine the geographic scope and design methodology of the survey
6. Select modules and questionnaires for implementation (technical scope)
7. Prepare a survey proposal and secure funding
8. Prepare a detailed implementation plan and survey schedule
9. Recruit survey personnel
10. Map facilities and prepare data collection logistics
11. Adapt questionnaire to country context and needs
12. Adapt HHFA CSPro tool and set up server and synchronization method
13. Train collectors and supervisors
Data collection (Chapter 7)
1. Prepare materials and tools for data collectors
2. Plan data collection visits in collaboration with local authorities
3. Arrange for transport and regular communication during fieldwork
4. Confirm appointments with health facilities
5. Assign questionnaire sections to data collectors, visit facilities and collect data
6. Transfer electronic files to team leader, combine and synchronize to server
7. Area supervisors oversee data collection and conduct data validation checks
8. Data manager reviews data on server throughout data collection process
Data processing and analysis (Chapter 8)
1. Edit, validate and clean the dataset within CSPro
2. Export the final dataset from CSPro
3. Configure HHFA data analysis platform and upload final dataset
4. Conduct data analysis using standard HHFA indicators and any country-specific indicators
Data review and description (Chapter 9)
1. Prepare the data analysis team
2. Review the standard HHFA data analysis platform outputs
3. Conduct additional analyses as needed

Data interpretation and communication (Chapter 10)

1. Review guiding principles for data interpretation
2. Organize the interpretation process
3. Review principles for effective communication of HHFA findings
4. Prepare the final HHFA report
5. Prepare additional communication products
6. Disseminate the HHFA findings

HHFA data curation (Chapter 11)

1. Introduction
2. Best practices for survey documentation and archiving
3. Creating metadata for the HHFA
4. HHFA global data archive

6. Survey planning and preparation

6.1 Establish a survey coordination group of country stakeholders

Bringing partners together and mobilizing them around the survey is a key initial step toward successful HHFA implementation. One of the first activities in planning the HHFA is to establish a group of core country-level stakeholders to oversee, coordinate and facilitate the planning, implementation, and follow-up of the survey process. The survey coordination group, led by the ministry of health, may include:

- various units of the ministry of health (e.g. planning, finance, health workforce, health information systems, epidemiology, health services, programme-specific units, etc.);
- academic institutions and public health institutes involved in health systems research;
- NGOs;
- United Nations health-related organizations; and
- international donors active in the country.

The coordination group provides leadership and oversight throughout the survey process. The roles of the group include:

- defining the objectives of the survey, and identifying important policy issues that should inform the objectives and the underlying questions that the survey should address;
- supporting the survey manager in planning and implementing the survey;
- mobilizing adequate financial resources and technical assistance for conducting the survey;
- advising on any matters that arise during survey preparation, data collection and data analysis;
- assisting in interpreting data, identifying priority issues and developing policy recommendations;
- disseminating the findings of the survey and advocating for actions.

It is important that the survey coordination group meets regularly throughout the survey process.

6.2 Define key roles and responsibilities to oversee and facilitate the survey

The following section outlines the roles and responsibilities of the key parties involved in the planning and implementation of the HHFA, in addition to the survey coordination group.

Ministry of health: The ministry of health has overall responsibility for leadership and coordination of the survey. Ministry of health roles include: approving the survey objectives defined by the coordination group; obtaining permission to conduct data collection where relevant; informing subnational authorities and facilities of the survey; helping to coordinate data analysis and interpretation; endorsing the final report; coordinating results dissemination meetings by inviting all appropriate stakeholders; and promoting use of the data for policy-making and planning.

Technical committee: A technical committee, consisting of a small group of experts from the ministry of health and partners, supports the survey manager in technical issues such as survey design, questionnaire adaptation, and data analysis and reporting. The technical group may be a subgroup of the survey coordination group.

Implementation agency: Once the coordination group is established, it is important to define an entity that will be in charge of the survey field implementation. It is recommended to identify a national institute (e.g. research unit within the ministry of health, national statistical office, school of public health, external organization etc.) or other entity that has experience in conducting such surveys. The selection is done in agreement with the ministry of health. The implementation agency works closely with the coordination group and is responsible for planning and conducting data collection and for supporting data cleaning, analysis, report writing and results dissemination.

Agency providing quality assurance and technical support: Involvement of an independent party such as an independent consultant or national institute is recommended. This agency provides technical assistance to the implementing agency and quality assurance for the implementation process to ensure due processes are followed during training, data collection, cleaning and analysis (including validation visits in 10% of the facilities). The agency advises the survey coordination group of any concerns. The agency also provides technical support for development of the HHFA report.

It is critical to ensure that adequate technical expertise is available to support the survey. Specific expertise is needed for some aspects. Section 6.9 describes key survey personnel that need to be recruited or assigned from within the ministry of health or the implementing agency.

6.3 Define the survey objectives

The survey objectives should be determined by the ministry of health's key information needs for policy-making, planning and monitoring of the health facility service delivery system.

Defining the HHFA objectives is a key strategic decision that will guide decisions on the design, scope of content (modules and questions) and geographic scope (geographic/administrative areas included), and will influence the required resources, the analysis and interpretation of data, and the use of the findings. This decision inevitably involves trade-offs among the desired information to be collected, analysed and interpreted; the quality of the data; the time needed for completing the survey; and the capacity and resources to carry out the survey with the desired quality standards.

When defining the survey objectives, it is important that the high-level decision-makers requesting the survey have a solid understanding of the types of information that can and cannot be obtained through an HHFA, and how the survey results can be used.

The HHFA is designed to provide a snapshot of the functioning of the overall health service delivery system, based on the services that health facilities offer, and items and systems that exist within the facilities. Hence, the HHFA indicators reflect the "percentage of facilities offering…" (a specific service) or the "percentage of facilities with…" (a specific item, system or set of items/systems).

The HHFA indicators are not designed to represent individual facilities or to provide comparisons among individual facilities. The HHFA does not collect data on the numbers of people using the services, diagnoses, interventions received or health outcomes. Therefore, the HHFA cannot provide information on service utilization, disease incidence or prevalence, mortality, coverage of interventions or treatment outcomes. Furthermore, the HHFA does not collect information on the quality of routine facility data.[1]

The HHFA objectives should be formulated as precisely as possible, avoiding broad and vague statements. Refer to Section 3.1 for considerations in defining survey objectives.

[1] The WHO Data Quality Assurance (DQA) toolkit can be used for this purpose (https://www.who.int/data/data-collection-tools/health-service-data/data-quality-assurance-dqa).

6.4 Establish or update the national MFL

Before beginning the HHFA process, it is crucial to assess the availability of a comprehensive, updated MFL, covering ALL public and private facilities in the country, and including a unique ID code for each facility. This is an essential prerequisite for the survey, as it will be used to construct the sampling frame for a sample-based assessment or to identify facilities for a census-based assessment. Updating or establishing the MFL may be a time-consuming process and **should be started well in advance of the HHFA**. Failure to ensure an MFL that is as accurate as possible may substantially impact the quality of the HHFA. Refer to Section 4.1 for further details.

6.5 Determine the geographic scope and design methodology of the survey

The **geographic scope** of the survey refers to whether the survey will be implemented country-wide or will only cover a selected geographic/administrative area, such as a selected province or number of provinces.

The **design methodology** refers to whether the survey will involve a census of all facilities or a representative sample, the sample size and the sampling methodology used. The design methodology chosen has significant implications on the complexity and costs of the survey. Refer also to Section 4.2.

Determining the sample size (i.e. the number of health facilities from which data will be collected) and then selecting the sample of facilities is a crucial and complex part of the survey design process. Since the projected sample size is a key component of the total HHFA cost, it is important to determine a sample size and a sampling methodology that minimize costs, while at the same time ensuring the desired precision.

The sample size is inversely related to the variance[1] of the indicator estimations. The smaller the sample size, the higher will be the variance and the lower the precision of the estimations. Because of this imprecision, if too small a sample size is chosen, the findings can become irrelevant. Conversely, too large a survey sample will result in a waste of resources and time to complete the data collection. The higher precision of the larger sample will be at the expense of higher financial costs, substantial opportunity costs for the facility staff, long duration of data collection, and often a lower quality survey. It is difficult to maintain a high quality of data in a very large survey because of data collector and informant fatigue as well as challenges in ensuring adequate supervision.

Standard formulas based on probability theories are available and various statistical packages (SPSS, SAS, Stata, etc.) have functions for calculating sample size. However, there will be considerable variation among different surveys, depending on the desired precision and type of estimates required, as well as the number of facilities in the country, and the objectives of the survey. To ensure appropriate decisions on the sample size and sampling methodology, it is strongly recommended to involve a statistician with expertise in sampling. Annex 2 provides further details on the HHFA sampling methodology.

Table 7 presents a summary of various HHFA sampling options. The most commonly used is Option 1: a nationally representative sample obtained by taking a stratified random sample of facilities within each stratum (facility type, location and managing authority) at the national level, with a census or oversampling of hospitals.

The overall costs of an HHFA will be influenced by multiple factors, including the scope and design methodology of the survey (e.g. length and content of the questionnaire), the number of data collectors, the distance and terrain between selected facilities, and the ease with which information can be accessed within the facility. Table 7 provides a rough average cost estimate, based on US$ 1000 per facility for a national level estimate.

[1] The variance of an estimate measures how far data are spread out, or the differences that occur between measurements, some of which may be explained by known factors, and the remainder attributed to chance.

When planning for a nationally representative sample of 200 to 250 facilities, budgeting for around US$ 1000 per facility provides a reasonable general estimate. However, as the sample size increases in a large survey, the cost per facility is expected to decrease, i.e. there is usually an economy of scale.

Table 7. Sampling options for conducting an HHFA, with estimated costs

Domains of estimation	Sampling method	Sample size[1] (estimate)	Estimated cost
Option 1: **National estimates only** National-level estimates with disaggregation by facility type (three levels) and managing authority (public/private)	***Small country*** Stratification by facility type, location and managing authority, simple/systematic random sampling within each stratum with census or oversampling of hospitals (design effect [deff] = 1)	150–250 facilities	US$ 150 000–250 000
	Medium country Blend of list and area sampling: list sampling for large health facilities, and area sampling for small facilities (census of facilities in the sampled area PSUs[2]) (deff = 1.2)	250–500 facilities	US$ 250 000–500 000
Option 2: **Subnational estimates** Regional and national estimates with disaggregation by facility type (three levels) and managing authority (public/private)	***Small country*** Stratification by region, facility type and managing authority, simple/systematic random sampling within each stratum, with census or oversampling of hospitals (deff = 1)	5 regions: 250–500 facilities 10 regions: 500–800 facilities	US$ 250 000–350 000 US$ 350 000–460 000
	Medium/large country Blend of list and area sampling: list sampling for large health facilities, and area sampling for small facilities (census of facilities in sampled area primary sampling units PSUs) (deff = 1.2)	***Medium country*** 4 regions: 300–500 facilities ***Large country*** 4 regions: 400–800 facilities	US$ 120 000–200 000 US$ 300 000–500 000 US$ 400 000–800 000
Option 3: **Subnational estimates** Regional estimates for a subset of regions, with disaggregation by facility type (three levels) and managing authority (public/private) for selected regions; no national estimates	***Large country*** Purposive sample of regions, simple/systematic random sample with oversampling of hospitals for each region	4 regions (150 facilities per region): 600 facilities	US$ 150 000 per region

[1] Sample size estimates assume a margin of error of 0.1 and 95% level of confidence.

[2] Administrative units that form the PSUs (primary sampling units) for the area sample should contain approximately one to five health facilities each (communes, subcounties, villages).

Option 4: **District sample** District estimates for sampled districts; national estimates if sufficiently many facilities are sampled	***Small, medium and large countries*** List sampling for regional and national hospitals plus sampling of districts (two-level cluster sample: selection of districts as first level, selection of facilities within these districts as the second level) (deff = 2)	***Small country*** 300–500 facilities (10–30 districts1) ***Medium country*** 400–800 facilities (20+ districts) ***Large country*** 600–1000 facilities (30+ districts)	US$ 250 000–500 000 US$ 400 000–800 000 US$ 600 000–1 000 000
Option 5: **Facility census** All possible domains of estimation	***Small, medium and large countries*** Census of all facilities		Very expensive

Small country: 50–100 hospitals, 1000–2000 health facilities total, 10–80 districts (e.g. Burkina Faso, Sierra Leone, Togo).
Medium country: 100–500 hospitals, 2000–5000 health facilities total, 80–500 districts (e.g. Uganda, United Republic of Tanzania)
Large country: 500–1000 hospitals, 5000–10 000 health facilities total, 50–1000 districts (e.g. Democratic Republic of the Congo, Nigeria).

6.6 Select modules and questionnaires for implementation

The selection of modules, indicators and questionnaires defines the **scope of the technical content** to be included in the survey. The selection should be guided by the survey objectives, the intended geographic scope and the design of the survey. These decisions will have an impact on the number of data collectors and the time needed to complete the survey at each facility, as well as the volume of information needing to be analysed and interpreted, and the related cost implications. In general, the larger the volume of information collected, the greater the risks to data quality.

Review of the indicators in the HHFA indicatory inventory platform will inform the decisions on which modules to implement and which indicators are needed. Decisions on the module(s) for implementation are followed by decisions on the need for core indicators only, versus also needing the additional indicators. It is important that the HHFA content selection is driven by information needs (i.e. indicators) rather than questions. Questions are then selected based on the indicator needs. Refer to Section 4.4 for further details on module and questionnaire selection.

6.7 Prepare a survey proposal and secure funding

The survey proposal should include a short description of the rationale, general objectives, proposed geographic scope, design methodology, scope of technical content, timeframe and preliminary budget. It is essential to ensure that the items in Table 8 are included in the budget.

The proposal should be submitted to the ministry of health for approval and, if necessary, to external agencies to agree on a partnership and to secure funding for implementation.

[1] The number of districts in the sample depends on the number of facilities per district.

Table 8. **HHFA resource requirements**

Human resources	
■ Survey manager ■ Technical advisory/quality assurance entity ■ Area supervisors ■ Data collectors	■ Drivers ■ Data managers/CSPro experts ■ Data analyst(s) ■ Report writer(s)
Technical resources	
■ Mobile electronic data collection devices (one for each data collector plus one back-up device per data collection team), e.g. tablets or mobile phones ■ Chargers for devices ■ Memory cards for devices	■ GPS devices (if used): one per data collection team ■ Batteries for GPS devices ■ Computer(s) for data analysis ■ Internet access ■ Server and data synchronization method
Training	
■ Training venue ■ Daily allowance (accommodation, meals, transport) ■ Equipment (projectors, screens, microphones, etc.)	■ Printing/photocopying (e.g. participant guide, exercises, paper questionnaire) ■ Expenses related to field practice day
Data collection and validation	
■ Daily allowance (accommodation, meals) for data collectors, area supervisors and drivers ■ Transport (vehicles, fuel)	■ Materials (e.g. notebooks, pens) ■ Communication (e.g. telephone/internet charges)
Analysis, report writing and dissemination	
■ Data cleaning, processing and analysis ■ Meetings of the survey coordination group and analysts ■ Interpretation workshop ■ Report production and dissemination	■ Validation workshop ■ Additional communication products ■ Advocacy and communications
Other	
■ Contingency fund for foreseen events/expenses	■ Overheads

6.8 Prepare a detailed implementation plan and survey schedule

A well-developed implementation plan (or protocol) is key to ensuring the success of the survey. It serves as the binding reference on all aspects of how the survey will be carried out and overseen to ensure that it will be completed according to appropriate quality standards, on time and within budget.

The implementation plan is a detailed document that expands upon the survey proposal document. The plan defines the rationale for the survey and is developed based on the objectives, the geographic scope and design methodology, and the scope of technical content to be covered. It provides a detailed description of the methodological and operational aspects, with clear responsibilities and troubleshooting procedures, and a budget, including contingency funds for unexpected problems. The design methodology will drive much of the operational planning and budget.

The survey timeframe should be defined and a survey schedule (e.g. Gantt chart) developed, detailing the amount of time allotted for each step in the process. This serves as a timeline for all survey activities.

Completion of the entire survey process generally requires at least 6 months, from preparation to report production. Further time will be needed for dissemination and follow-up activities. As the information produced from the HHFA should be used to inform decision-making, it is important that data collection is conducted rapidly (ideally within a period of 1–2 months), and that the report is generated as soon as possible after data collection is complete. This will ensure that the survey results are relevant and up-to-date for decision-makers.

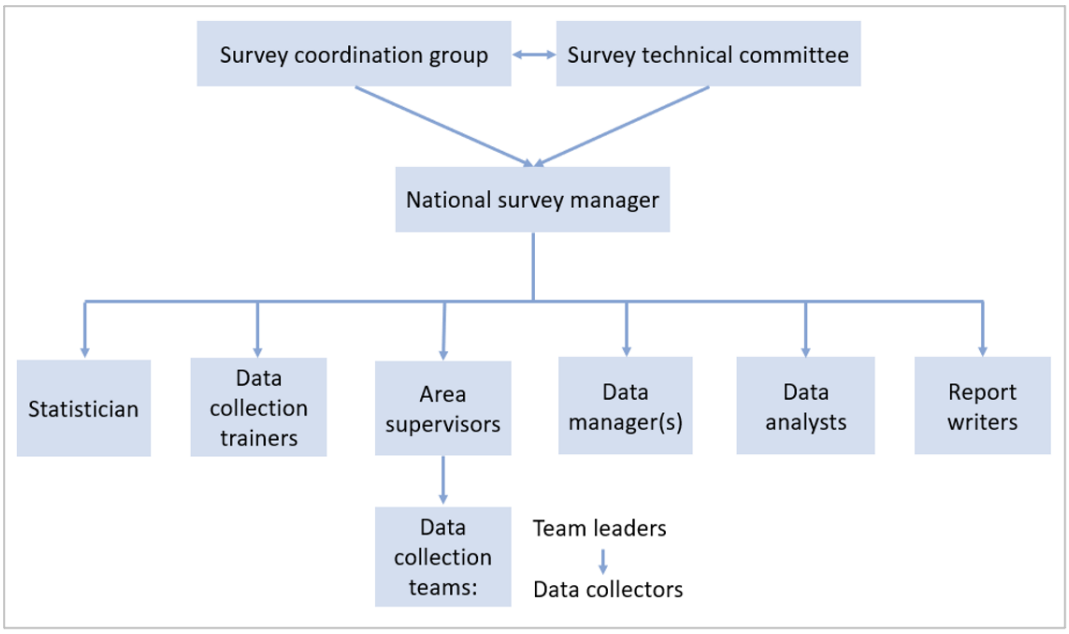
Seasonal issues and other contextual factors that may affect data collection, such as rainy seasons, holiday periods, country elections, etc., should be considered when establishing the timeframe. Also, an HHFA implies high opportunity costs, keeping several ministry of health officials busy with aspects of the survey for significant periods of time. Therefore, if there are likely to be concurrent demands on the time of these officials, (e.g. large grant proposals, reports, planning processes, etc.), these factors should also be considered when establishing the HHFA timeframe and schedule.

It is essential to allocate adequate time for survey preparation activities, including questionnaire adaptation, CSPro adaptation, hiring of staff, preparation of tablets for data collector training, and the data collection training workshop. The duration of the data collection will depend on the availability of resources, the number of teams, the number and size of facilities to be visited, the modules implemented, and the size and terrain of the country (refer also to Section 6.10). The estimated duration of the survey is calculated during the planning phase and is unique to each country. Adequate time should also be allowed for data cleaning prior to analysis. The schedule should be consulted regularly by the survey coordination group to ensure that activities are proceeding according to plan.

6.9 Recruit survey personnel

Key survey personnel include a national survey manager, statistician, area supervisors, data collectors, data managers/CSPro experts, data analysts and report writers. These personnel are supported by the survey coordination group and the survey technical committee. Refer to Annex 3 for job description examples for selected key staff.

Fig. 16. Survey personnel and reporting lines



National survey manager

The survey manager plans and coordinates the survey at national level. This includes planning the technical and logistical aspects, recruiting and training survey personnel, overseeing the data collection process, conducting data quality assurance and data analysis, interpreting results, preparing a survey report, and communicating findings, as well as troubleshooting when unplanned events disrupt the implementation of the survey. The survey manager should have experience in conducting surveys, be very familiar with the health care system, and have basic statistical and data interpretation skills. Successful communication of the survey results also requires an understanding of the policy-making process and different advocacy strategies. Where the survey manager does not possess all of these qualities, technical committee members should be selected with the necessary health, surveying, statistics, policy and advocacy skills.

Statistician

The statistician/sampling expert assists the survey manager in designing the survey. Their role is crucial for ensuring a high-quality survey. Ideally, the statistician should be integrated into the HHFA technical team as early as possible, when strategic decisions on survey methods need to be made. They should be involved in the decisions concerning:

- sampling methodology and sample size, including adjustments for expected non-response, estimation of change, finite target group (population) correction, and design effect (in cluster sampling);
- selection of health facilities from the sampling frame; and
- calculation of sampling weights.

Data collector trainers

The data collection training workshop requires at least two main facilitators: one with a clinical background and another with a data management background. The clinical background facilitator should have experience working in local health facilities. The data management facilitator should have a high level of CSPro expertise and should lead the sessions related to the CSPro tool. Some sessions may require the presence of both facilitators. As there is a large volume of questionnaire content, is it recommended that the main facilitators are supported by a number of co-facilitators. They may also assign specific sessions to co-facilitators. Refer to Section 6.13 and Annex 4 for further details on the data collection training workshop.

Area supervisors

An area supervisor oversees several teams of data collectors and is responsible for all aspects of data collection in a specified geographic area; they report to the survey manager. An area supervisor should be designated for each geographic area that will be surveyed. The number of area supervisors depends on the sample size, the geography of the country and the survey timeframe. Area supervisors play a crucial role in ensuring data quality. They should be experienced in data collection, familiar with health terminology and, if possible, familiar with the health system in the area for which they are responsible. They are also instrumental in gaining access to facilities; if any area supervisor is unfamiliar with their designated area, a local contact may be needed to assist. It is essential that all area supervisors attend the full duration of the data collection training workshop.

Data collector team leaders

In each data collection team, one of the data collectors is assigned the role of team leader. The team leader also collects data (and undergoes the same training as the data collectors) but has additional responsibilities including assigning sections of the questionnaire to team members, receiving the completed sections from team members, ensuring that all required sections have been completed, combining the different questionnaire sections for each facility, and synchronizing the completed questionnaire to the server.

Data collectors

Data collectors visit facilities to collect the HHFA data. The number of data collectors depends on the sample size, the survey timeframe, the modules/questionnaires to be implemented, the locations of the facilities and the travel conditions. Data collectors should preferably have a health qualification (nurse, midwife, doctor or medical student) and familiarity with the organization and functioning of health facilities, as well as the local area and language. Ideally, data collectors should also have some previous experience in conducting surveys. It is preferable to have a smaller number of better qualified data collectors than a larger number where some data collectors may lack the necessary skills. Data collectors must be available to work full time for the duration of the fieldwork. They should be willing to work long hours if necessary and be able to stay away from their homes for extended periods of time.

Data collection requires an aptitude for concentration and attention to detail. The best data collectors combine the discipline of collecting data in a standardized way with the ability to identify unusual situations that require advice from the area supervisor or survey manager. Further details on interviewer skills are available in the HHFA data collection training materials.

Data manager/CSPro expert

The data manager configures and adapts the CSPro tool to country needs, sets up the server and data synchronization method, ensures the set-up of the CSPro tool on the data collection devices, supports the training of the data collectors, supports the data collector teams during data collection, assesses data for completeness and quality, addresses quality issues where needed, and conducts data cleaning to prepare the dataset for the analysis phase. Once all data collection is complete and the data have been cleaned, the data manager exports the dataset from CSPro and transfers it to the data analysis team for upload to the HHFA data analysis platform. The data manager may also be given the responsibility of configuring and adapting the analysis platform. Refer to the *HHFA Data manager guide* for a detailed description of the data manager's tasks.

Data analysts

The data analysis team consists of a small group of individuals which conducts the initial review of the standard HHFA analysis platform outputs. They prepare selected outputs for review and interpretation by programme-specific experts, engage in discussion with these experts, and prepare any additional analyses needed. The data analysis team presents a summary of the initial findings in a data interpretation workshop and guides further interpretation processes during the workshop (refer to Chapters 9 and 10). The team must include an analyst with advanced knowledge of the HHFA data analysis platform (or other analysis software that the country chooses to use). The analyst adapts the analysis platform to country needs, inspects the completed dataset, uploads the data to the data analysis platform and runs the analysis. (These tasks may also be assigned to the data manager.) The team must also include senior-level individuals with strong analytical skills, technical knowledge of health service delivery resources and processes, and knowledge of the country's health system and overall context.

Survey report writers

It is recommended to assign or recruit a lead report writer to ensure clear responsibility for drafting and finalizing the main HHFA report. The lead writer may be assisted by only a limited number of analysis team members with strong writing skills, to avoid inconsistencies in language and writing style. The writer(s) may use the report outline template provided by the HHFA data analysis platform to structure the report (refer to Chapters 8–10 for further details). The report writer compiles the final key findings, interpretation of content and recommendations, based on the feedback from the analysis team, programme experts, and the interpretation workshop, to produce the survey report.

6.10 Map facilities and prepare data collection logistics

Map the facilities to be surveyed

It is recommended to map all the facilities in the survey sample (or all the facilities in the census) to facilitate the planning of logistics for data collection. This map can be made on paper or electronically. The map should include information useful for gaining familiarity with the survey areas, such as roads, topography, basic geographical features, elevation and location of health facilities. Teams should be assigned to facilities based on the geographical distribution of the selected facilities. A draft schedule of visits to the health facilities is then planned, for informing the district managers and receiving their approval.

Time required per facility

If all three core facility audit modules (service availability, service readiness, and management and finance) are implemented, approximately 3–4 hours may be needed on average for a team of two data collectors to complete data collection in one medium-sized PHC facility. A team of four data collectors requires approximately 1–3 working days to complete data collection in a hospital, depending on the size of the hospital and the services offered. Additional time is needed for travel, briefing of facility staff, checking of questionnaire sections, etc.

Survey team requirements

For all surveys, logistics planning must consider the following:

- vehicles and fuel for the duration of the survey;
- daily living allowances for drivers, data collectors and area supervisors; and
- lodging for drivers, data collectors and area supervisors, when necessary.

It is also important to include contingencies in the logistics plans, such as extra time, staff and resources.

Equipment requirements

Equipment requirements are determined according to country-specific needs, as well as the availability of resources and budget. Equipment is needed centrally as well as for data collection in the field, and for operations as well as for training. A guiding principle when preparing survey equipment is to have backup components and a contingency plan in case equipment fails or is lost. If feasible, paper questionnaires and printing capabilities can provide a contingency plan for the worst-case scenario of mobile device failure.

An electronic data collection device (tablet or mobile phone) is required for each data collector, with at least one back-up device available for each team. Data collection teams and area supervisors require internet access for regular upload to a server of the completed data for each facility. They also require mobile phone connectivity. The data manager and survey manager require a reliable, continuous internet connection and mobile phone connectivity.

Box 3 provides examples of some of the planning required for an HHFA.

Box 3. Examples of HHFA planning

Example 1: A small country wants to implement the three HHFA facility audit modules (service availability, service readiness, and management and finance), using the Core questionnaire. The country has 20 districts with an average of 30 PHC facilities (health centres, health posts and clinics) per district. The country also has 12 hospitals. The survey is intended to provide national-level estimates for PHC facilities, using a sample of 156 PHC facilities. A census of all 12 hospitals will be conducted.

One team of two data collectors can complete two PHC facilities per day. Using six teams for PHC facilities, they can complete 12 facilities per day. Therefore, 156 / 12 = 13 working days are needed to complete data collection in the PHC facilities (excluding travel time between districts/facilities and other potential time constraints).

One team of four data collectors can complete one hospital in 2 days. Three four-person teams can complete data collection in all hospitals in 8 working days (excluding travel time between districts/facilities and other potential time constraints). Alternatively, two four-person teams can complete all hospitals in 12 working days.

Example 2: An HHFA will be implemented in one rural district with five hospitals and 96 PHC facilities. The core questions for two modules (service availability and service readiness) will be implemented but resources are limited so only three teams of four data collectors are available.

Using two teams of four data collectors, eight PHC facilities can be completed per day. However, as they are implementing only two modules, the teams could potentially complete more than two facilities per day, which reduces the time needed for survey completion. However, travel time between facilities must be considered. The PHC facilities could possibly be completed in 12–14 working days. The remaining team of four can survey the hospitals, which could take as little as 5 days, but will probably take longer due to distances between hospitals in rural areas. The entire survey could potentially be completed in 2–3 weeks.

However, if fewer teams are used, the implementation will not only take longer due to the reduced number of facilities that can be surveyed per day, but also due to the travel time between facilities. These trade-offs in number of teams, team size and geographic spread, must be considered carefully according to country-specific factors such as data needs, data collector availability and budget limitations.

6.11 Adapt questionnaire to country context and needs

The questionnaire is adapted based on practices and terms used in the country and on country needs (refer to Section 4.5 for further details). Changes are finalized and clearly indicated in the paper questionnaire. The **final version (with clearly indicated changes) is provided to the data manager** for adaptation of the CSPro tool. If implementing a quality of care record review, the instructions for selecting the sample of individual patient records must also be adapted to the country record-keeping systems.

6.12 Adapt HHFA CSPro tool and set up server and synchronization method

The HHFA CSPro application must be configured for the country before it can be used. This includes: defining the administrative areas that will be used when defining the locations of facilities; defining and assigning roles to the individuals that will have access the HHFA data collection application; and defining the facilities for which data will be collected. All the HHFA questionnaires have been programmed into

the CSPro tool. The tool allows selection of the modules and questionnaires (Core, Core+Additional, Supplementary) to be implemented by the country. Individual questions can also be turned on or off as required, and country-specific questions can be added. After country adaptation of the paper questionnaire has been finalized, the CSPro tool is adapted accordingly.

Survey preparation also involves the setting up of a server and a method for synchronizing to the server the data collected on all the data collection devices. Refer to the *HHFA Data manager guide* for further details.

6.13 Train data collectors and area supervisors

Training of data collectors, team leaders and area supervisors is a critical element of survey preparation because it helps to ensure the accuracy and consistency of the data collected. Any trainees that do not exhibit the necessary competency in using the questionnaire by the end of their training should be excluded from fieldwork.

It is essential to allow sufficient time for thorough training on the questionnaire, interview skills and use of the CSPro application. Depending on the modules and questionnaires selected for implementation, 8–10 days of training (including a field practice day) are recommended. It is also recommended to conduct a training of trainers (TOT) workshop prior to the data collector training workshop, to ensure that all trainers have a thorough and consistent understanding of the questionnaire.

As the data collectors will be trained on use of the CSPro application, it is **essential that the tool is configured on the data collection devices in advance of the training** and that sufficient devices are available for use during the workshop.

Data collectors and area supervisors are trained on:

- a brief overview of the HHFA;
- the consequences of poor data quality;
- an overview of data collection processes;
- general guidance on interviewing practices and techniques;
- questionnaire structure and content;
- the purpose and meaning of the questions and how to ask the questions and record responses for different types of questions;
- identification of medical devices, medicines, commodities and other items for which direct observation is required;
- ethical issues related to conducting a health facility survey;
- how to collect geographical coordinates using GPS;
- use of the CSPro tool.

Data collector team leaders and area supervisors receive additional training based on their responsibilities. Refer to Annex 4 for further details on data collection training.

The final part of data collection training involves a field practice day. Data collector teams visit local health facilities (easily accessible from the training venue) and collect data in the same way as they would do during the actual survey. The data collected during this exercise are not included in the final survey dataset, as the data may not be of appropriate quality and there may be some final revisions to the questionnaire

after the field practice day. If the HHFA is a sample survey, the field practice day should target facilities not included in the survey sample. If the HHFA involves a facility census, the data should be recollected in the field practice facilities.

The data collectors note any questions or problems that come up during the field practice and these are clarified during the training workshop the following day. This exercise serves to identify any problems in wording of questions, misunderstanding of instructions, and weaknesses in interview skills that need to be addressed, and may also highlight any aspects of the questionnaire that require a final revision. The exercise should also identify data collectors who lack the required competence and who need to be excluded from field work. The field practice day also tests field logistics, supervisory capacity and CSPro tool functionality. Based on the field practice experience, if necessary, the survey manager revises the paper questionnaire and provides the updated version to the data manager/CSPro expert for updating of the CSPro application. The survey implementation plan should include sufficient time between the field practice day and the start of data collection in the field, to enable completion of any final revisions and uploading of the final CSPro tool to all the mobile devices that will be assigned to the data collectors and supervisors.

7. Data collection

Data collection should start as soon as possible after the data collection training has been completed to maximize the retention of knowledge and skills gained by the data collectors and area supervisors. Data collection involves the deployment of multiple teams to multiple (including remote) locations and requires careful planning, preparation and monitoring.

7.1 Prepare materials and tools for data collectors

The survey manager and area supervisors ensure that all data collectors receive the following:

- contact details of area supervisor;
- official identification document with photograph;
- electronic data collection device (tablet or mobile phone) with finalized CSPro tool;
- back-up device for each team;
- mobile phone, internet access and airtime for each team;
- electronic copy of the final paper questionnaire for reference;
- list of facilities to be visited, with facility opening hours and map showing the facility location;
- name and contact details of person in charge at each facility;
- blank consent forms for each team; and
- notebook and pen to record any significant events, findings, or challenges.

Note that each data collector must use the same electronic data collection device throughout the survey and that **devices must not be shared** among data collectors. If a device fails during fieldwork, the area supervisor is informed and will advise on next steps.

7.2 Plan data collection visits in collaboration with local authorities

Each area supervisor is responsible for ensuring the planning of data collection visits in their assigned survey area. The area supervisor may oversee multiple geographic areas within their assigned area. A list of the sampled health facilities in each geographic area is prepared, including the location of each health facility, and the name and contact details of the person in charge at the facility.

A data collection team is assigned to each geographical area. The number of days required to collect the data is estimated based on the number and types of facilities to be visited in each geographic area, the distances between them, the mode of transport available, and the number of data collectors in each team.

The data collector team leader receives from the area supervisor the list of facilities to be surveyed in their assigned geographic area. The team leader contacts the person in charge at each facility in advance to establish an appointment date and time for the data collection visit. Before data collection starts, the team leader prepares a written schedule of facility visits for their data collection team and shares this with their team members and area supervisor.

A letter of introduction is circulated in advance by the ministry of health through the various administrative levels (regional offices, district offices, facilities) so that facilities are aware that an assessment will take place. The letter provides reassurance that the anonymity of the respondent will be maintained, and contains:

- name of the organization conducting the survey and the survey manager;
- contact details;
- purpose of the survey;
- estimated time required for data collection per facility.

The area supervisor should also provide the data collector team leaders with sufficient copies of the letter of introduction for use when scheduling facility visits and during the data collection visits.

7.3 Arrange for transport and regular communication during fieldwork

After receiving the schedule of visits from the team leaders, the area supervisor arranges transport according to the facilities to be visited, the number of teams involved, and the number of people per team. The area supervisor supports the data collector teams through regular communication and should be available to help resolve issues that may arise in the field. Team members should meet at the end of each day to discuss the data collection process, resolve any problems and ensure the transfer of completed data files to the team leader.

7.4 Confirm appointments with health facilities

The team leader contacts each health facility the day before the scheduled data collection visit to confirm the appointment.

7.5 Assign questionnaire sections to data collectors, visit facilities and collect data

Before visiting the facility, the team leader assigns the required sections of the questionnaire for that facility to the data collectors, ensuring that the sections are appropriately distributed among the data collectors. This is done in the CSPro tool through an internet or Bluetooth connection. (The CSPro tool prevents assignment of the same section to more than one data collector in a facility.) The data collectors then confirm that they have receive their section assignments. Refer to the *HHFA Data manager guide* for further details on questionnaire section assignment.

On arrival at the facility, the data collector team leader introduces the team to the facility in-charge and explains the purpose of the visit. After the team leader has obtained signed consent to conduct the survey, the data collectors complete the survey based on the questionnaire sections assigned to each data collector. (If the facility in-charges advises that, due to unforeseen events, data collection is no longer possible on this day, the team leader obtains a new appointment date for the visit.) Further details of the facility visit are described in Annex 5.

Before leaving the facility, each data collector uses the “View reports” feature in the CSPro tool to check that all the sections of the questionnaire assigned to them have been completed and resolve any missing or incomplete sections.

7.6 Transfer electronic files to team leader, combine and synchronize to the server

After completion of each facility visit, the completed questionnaire sections are transferred from each data collector to the team leader. The team leader then creates a complete facility record by combining the files from all the data collectors. (This process is called data concatenation and is automated within the CSPro tool.) The team leader checks that the data for the facility are complete and ensures the synchronization of the data to the server.

If the data are incomplete, the team should return to the facility the following day to complete the questionnaire. It is also particularly important to check that the facility ID items such as facility number, name, location, type and managing authority have been entered correctly, and that there are no inconsistent or missing data. While data collection remains ongoing, all edits to the data must be made on the device originally used to collect the data. Wherever possible, any data quality problems should be addressed while the data collectors are still in the field. Refer to the *HHFA Data manager guide* for further details.

7.7 Area supervisors oversee data collection and conduct validation checks

The CSPro "View reports" function can be used by the area supervisors to track the progress and completeness of data collection across their teams. The supervisors should visit the surveyed facilities regularly with the data collection teams, on a rotational basis, to ensure that procedures are followed as required.

Area supervisors also validate data collection by repeating a subsection of the questionnaire in about 10% of the surveyed facilities. (This validation can also be conducted by the entity in charge of survey quality assurance.) Facilities visited for validation should be selected at random. Ideally, the validation should be done soon after the data collectors' visit (within the same week) to avoid changes in the availability of items at the facility. The supervisor's results are checked against those of the data collectors using the CSPro Compare Data tool. If the validation reveals possible data quality problems, the area supervisor consults the survey manager for next steps. Refer to the *HHFA Data manager guide* for further details.

7.8 Data manager reviews data on server throughout data collection process

Ensuring high-quality data requires review and editing of the data both in real time during data collection, as well as after the data collection is complete. Data are reviewed by data collectors, team leaders, area supervisors and the data manager throughout the data collection process. If quality issues are discovered, they should be corrected as soon as possible on the original tablet. Edits can be made by data collectors during the data collection process, with ongoing synchronization of completed facility data to the server.

8. Data processing and analysis

8.1 Edit, validate and clean the dataset within CSPro

The HHFA dataset must be checked and cleaned within CSPro before exporting for analysis. Once all data collection is complete, further data edits can be made to the combined dataset (containing all the data from all the facilities) by the data manager, using a batch edit application. It is good practice always to preserve an unedited copy of the dataset and to document in detail the data editing process.

Data managers use a tracking sheet to track the progress across all teams toward completing data collection for all facilities. They are also responsible, throughout the data collection process, for recording information on facilities that are missing, inaccessible, closed, replaced, etc. After all the data collection is complete, the data manager should synchronize the data from each tablet one final time when the tablets are returned from the field.

The CSPro tool's data manager menu includes a reporting function to generate information on the completeness of facility records in the final combined dataset. Using the tracking sheet, the data manager also identifies and resolves any potential duplicate records.

The CSPro tool contains built-in functionality to minimize the risk of data quality problems. However, the last step in the data management process is to conduct a review of key variables in the final combined dataset. This involves downloading the dataset, reviewing key fields, and creating a list of any errors or inconsistencies that should be corrected in the final combined dataset that will be used for analysis. Key variables for review include: consent; final result code; facility type; managing authority; urban/rural; administrative areas; GPS coordinates; and "Other" response options.

After the data have been reviewed and all necessary edits have been identified, the final combined dataset can be edited.

At this stage, if a sample survey was conducted, sampling weights should be calculated, and a weighted variable generated using the CSPro batch application for editing the HHFA dataset. Refer to Annex 2 for further details on sample weights and to the *HHFA Data manager guide* for details on the use of the batch application.

8.2 Export final dataset from CSPro

CSPro has a built-in "Export data" tool that allows quick and easy export of data in a variety of formats. The exported data can then be imported into different software programs as needed, including the HHFA data analysis platform.

8.3 Configure HHFA data analysis platform and upload final dataset

The HHFA data analysis platform automatically calculates all the HHFA indicators and produces a standard set of indicator tables and graphs as well as a standard HHFA report outline. These products are called the analysis platform outputs. The modules and questionnaires implemented by the country are selected within the analysis platform and indicators can also be turned on or off as required. The platform also allows countries to adapt the tables and reports according to their needs, and to create additional country-specific indicators and tables.

After configuring the analysis platform based on the country requirements, the final HHFA dataset is uploaded to the platform and the analyses can be run (refer to Section 3.6 and the instructions tab in the online HHFA data analysis platform for further details).

8.4 Conduct data analysis using standard HHFA indicators and any country-specific indicators

Many different types of analyses can be obtained from surveys. The types of analyses that can be produced depend to a large extent on the survey design (e.g. census or sample; sample size allowing acceptable precision at national level only or also at subnational levels). The HHFA data analysis platform produces descriptive analyses, based on the distribution of indicator frequencies and on comparisons using selected stratifier variables.[1] The analysis platform provides **a standard set of tables, graphs and maps, as well as a report outline**. These products are the analysis platform outputs.

It is important to start the analysis and review process by running **a complete analysis of all the survey data**, generating the full range of analysis platform outputs. The analysis team should then review ALL the standard tables and graphs produced by the analysis platform. It is important that the scope of this initial phase is comprehensive, in order to:

- generate a preliminary overview of the major findings across the survey;
- avoid overlooking any important findings;
- identify any data quality issues that were not found during data cleaning in CSPro; and
- identify issues that should be highlighted to specific programme or management experts in the next phase of the review.

The tables, graphs and maps generated by the analysis platform display the frequencies of each indicator in standard formats. These outputs can be adapted and formatted within the platform according to the needs of the analysts.

Tables, graphs and maps represent complementary ways of displaying data.

Tables present details of the data and allow the viewer to scrutinize the numbers directly. They are also useful for displaying many indicators at once. “Heatmaps” can be useful for identifying patterns within tables by using colours to highlight, for example, the highest and lowest values. Table 9 shows a screenshot of a diabetes service availability table in the HHFA data analysis platform, using conditional formatting to create a “heatmap”. The screenshot also show how users can select stratifier variables they want to include in the table.

[1] Stratified random sampling allows estimates to be obtained separately for desired, mutually exclusive subgroups of facilities called strata. A specific type of subgrouping is called a stratifier variable. Stratifier variables commonly used in the HHFA are: facility type, managing authority, urban/rural location and geographic area (e.g. region/province).

Table 9. Diabetes service availability

Table editor

Decimal places: 0

Table splitting: Split table at specific points

Stratifiers to include: Province, Location, Ownership, Facility Type

Conditional formatting: None, Two: ~ 50% ~, Two: ~ 90% ~, Three: ~ 90% ~ 80% ~

Diabetes service availability

Percentage of facilities offering:

	Any services for diabetes	Diagnosis of diabetes	Treatment for diabetes	Patient follow-up for diabetes	Counselling for diabetes self-management	n
National	40%	29%	34%	34%	38%	3,179
Location						
Urban	59%	55%	55%	48%	57%	689
Rural	35%	22%	28%	30%	33%	2,490
Ownership						
Public	38%	26%	32%	32%	37%	2,949
Private	60%	59%	57%	51%	58%	230
Facility Type						
Third Level Hospital & Above	80%	80%	80%	75%	80%	20
Second Level Hospital	94%	94%	94%	77%	94%	31
First Level Hospital	84%	80%	82%	72%	84%	152
Health Centre	48%	37%	42%	41%	46%	1,544
Health Post	25%	12%	17%	20%	24%	1,432

Graphs summarize data and attract the interest of viewers. Graphs are particularly useful for displaying, in a summarized way, large, complex sets of data and for enabling comparison across categories. Most HHFA data are of the categorical type and can be presented in bar charts. Whenever possible, the bars should be shown in ascending or descending order for easy visualization. Figs 17 and 18 are both derived from Table 9 above and illustrate presentation options available within the analysis platform.

Fig. 17. Diabetes service availability: national-level estimates for multiple indicators

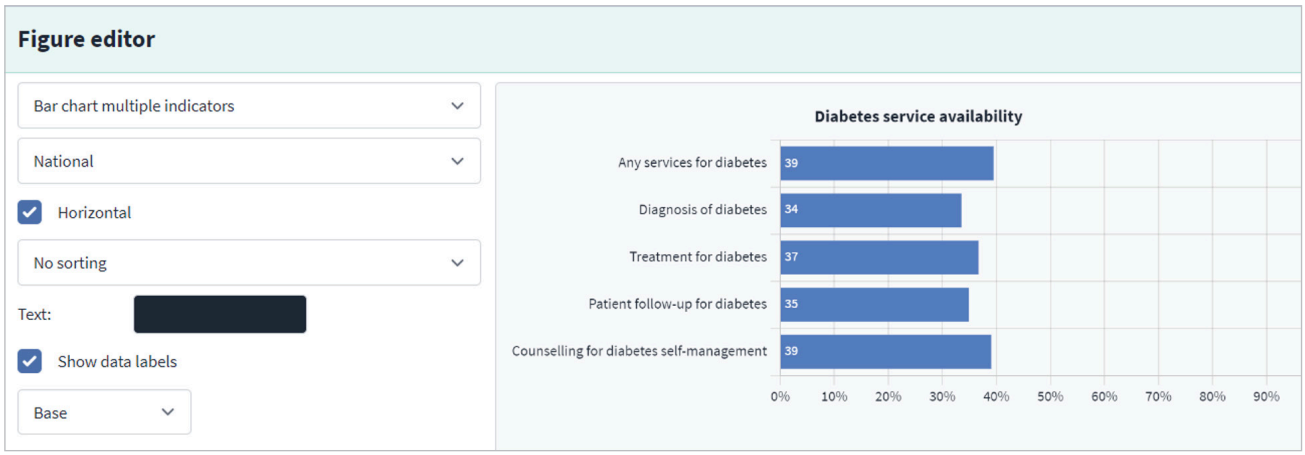
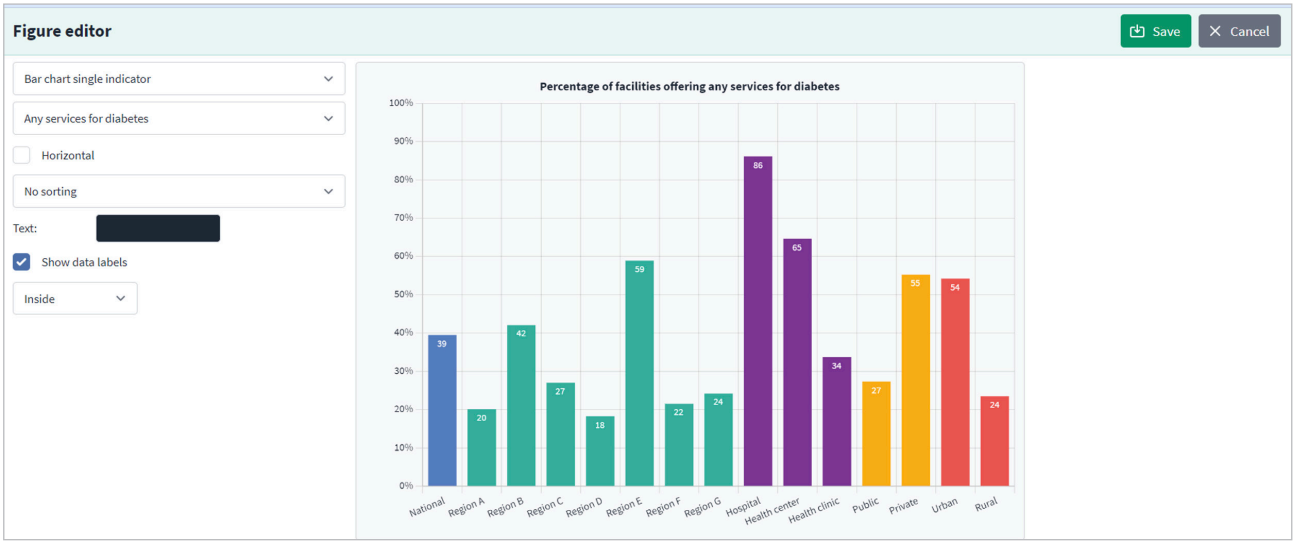
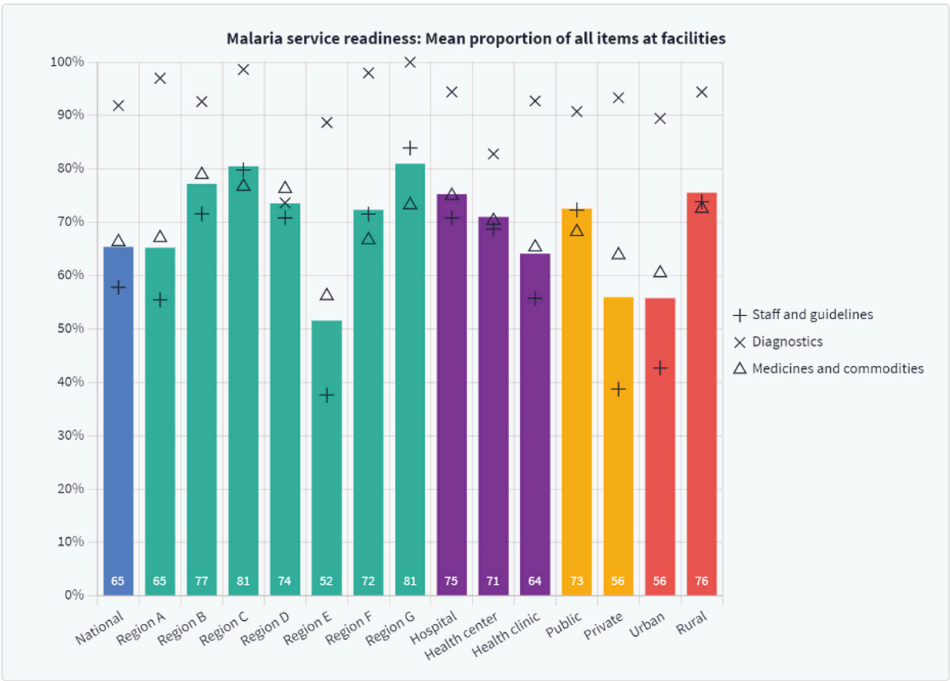


Fig. 18. **Diabetes service availability: national-level estimate and estimates by stratum for two stratifier variables for a single indicator**



In the data analysis platform, for graphs where the bars represent an index, the components of the index can be displayed with symbols superimposed on the bar. This provides additional information on how the individual components are affecting the overall index score. Fig. 19 provides an example from the analysis platform of a readiness index where bars are grouped according to stratifier variables and the strata are shown in descending order within each stratifier variable. Symbols are used to show the domain scores for each stratum.

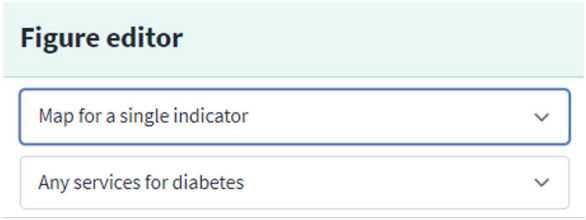
Fig. 19. **Malaria readiness index: national level and by stratifier variable**



When data from previous SARAs/HHFAs are available, it is possible to use line graphs to display trends over time. For the final HHFA report, only the most relevant and informative graphs should be included.

Maps provide an alternative way of displaying data that are geographically related. They can display administrative borders and locations of health facilities and can also show patterns by displaying regions in different colours according to the values of the indicator. Maps are visually attractive, as they allow quick comparisons among regions. However, they usually hide the data details that tables offer. Maps can therefore complement, but not substitute for, other forms of data visualization, such as tables and graphs. The “Figure editor” in the HHFA data analysis platform provides an option for creating maps, as shown in Fig. 20. If the sample size and sample methodology of the country HHFA allow meaningful subnational estimates, maps showing comparisons by subnational area can be created.

Fig. 20. Figure editor in HHFA data analysis platform showing option for creating a map



9. Data review and description

9.1 Prepare the data analysis team

Before starting the process of reviewing and describing the analysis platform outputs, the data analysis team should review **key aspects of the HHFA design and implementation** that may have impacted the availability and quality of data and are therefore relevant to their analysis and interpretation. Analysts require an overview of the HHFA objectives and their underlying questions, the modules and questionnaires that were implemented, and the changes that were made to the standard HHFA questionnaire. Furthermore, the analysts need to consider the limitations and potential errors of the HHFA and their consequences on data accuracy and representativeness. Special attention should be paid to issues related to the MFL, sampling, and undercoverage, and to challenges in the data collection (e.g. non-response) and data management.

Box 4 warns against making comparisons among strata when the sample size is not sufficiently large to allow disaggregated analysis.

Box 4. Comparisons in the HHFA – a need for caution

Comparing indicators across strata: some methodological notes

HHFA data analysis and interpretation are essentially descriptive. In short, analysts look at the frequency distributions of the indicators and make comparisons with the national average, a target (when existing), previous comparable assessments (if available), and across the different strata, e.g. facility types, subnational areas such as regions or districts.

Is it correct to make comparisons across strata? It depends … There are three possible scenarios:

a. The HHFA is conducted on a census of all health facilities. Provided that the MFL is complete, accurate, and updated, and no significant biases are introduced in data collection and management, these **comparisons are correct, since there is no sampling variation**.

b. The HHFA is based on a sample size that allows sufficiently **precise estimates at national level**, but is not large enough for sufficiently precise estimates at a disaggregated level (i.e. according to strata). **Comparisons across strata could be incorrect** (see below).

c. The HHFA is based on a sample for which the size was calculated in order to perform **disaggregated analyses with sufficient precision**, taking into account the number of strata. The sample size enables analysts to make comparisons across strata. The larger the number of strata within a stratifier variable, e.g. provinces, the larger is the sample required.

The objectives of the HHFA, the desired precision of the estimates, the number of strata, and the resource constraints are the main principles guiding the statistician in deciding on the best sampling design (usually stratified sampling) and the appropriate sample size. Because of the complexity of these decisions, it is recommended to involve the statistician from the start of the survey design process.

In a sample (as in scenarios b. and c.) there is always uncertainty about estimates, because of sampling variation, even when the sample is drawn according to correct probabilistic methods (refer to sampling in Annex 2). The smaller the sample, the larger is this uncertainty, or lack of precision. When the estimates are imprecise, it can be impossible to judge if the difference in the estimates in the sample (e.g. among strata) represent real differences in the target group or is due to sampling variation (i.e. chance).

In statistics, the uncertainty of estimates is measured by the confidence intervals (see the definitions in Annex 2). Manual calculation of the confidence intervals for all the indicator values would be time-consuming; in the future, the HHFA will include this calculation as part of the data analysis platform outputs. (It is also possible to export the HHFA data to a statistical program, e.g. STATA, to calculate

the confidence intervals.) In the absence of the calculated confidence intervals, caution is needed in comparisons across strata in sample-based HHFAs where the size was not specifically calculated for disaggregated analysis (scenario c. above).

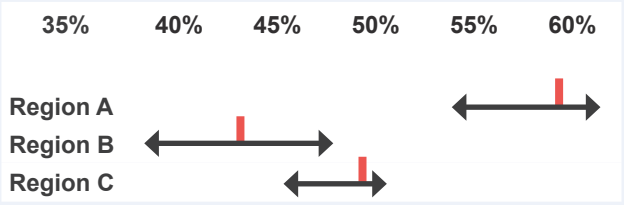
In scenario b. comparisons can be made in the description stage only, but it could be incorrect to stretch the observations to hypotheses, e.g. to claim that the availability of some items is higher in some provinces than in others. The observation can only be due to chance. When observations of many indicators, e.g. availability of several items, go in the same direction and a clear pattern emerges, we can be more confident that there are real differences in availability in certain provinces. It is recommended to limit the analysis and interpretation to the national level, in order to stay on the safe side.

The **differences in sample sizes** among strata is important, not only their sizes. In scenario c. it is possible that some strata (e.g. hospitals, private facilities, some provinces) have much smaller sample sizes than others. In this case, comparisons could be misleading, due to the uneven precision of the estimates across the strata.[1] A difference in the estimates between two or more strata might only be due to chance, and not reflect a real difference in the target groups to which the results are extrapolated. The same problem affects the comparison of the indicators over time, when HHFA findings are contrasted with those of previous similar assessments. In disproportionate stratified sampling, therefore, the number of facilities to be allocated to the strata is not proportional to the representation in the target group, in order to achieve a comparable precision across the strata. Equal or similar precision among strata is not always the solution, if the size in the strata is small, as there is a risk of comparing similarly imprecise findings.

The availability of the confidence intervals through the data analysis platform in the future will facilitate the analysis and interpretation of HHFA findings.

The following table and figure show a fictitious example of the estimates of the indicator “Percentage of facilities offering family planning services” (FP) with their respective confidence interval (CI) for three regions.

Percentage of facilities offering FP services			
	Region A	Region B	Region C
%	58%	48%	50%
95% CI-UL	60%	44%	51%
95% CI-LL	55%	40%	46%



It can be seen that region A has a higher percentage of facilities offering FP than regions B and C and that region C has a higher percentage of facilities offering FP than region B. Are these differences due to the uncertainty of the estimates (i.e. due to chance)? Are there other sources of uncertainty, like differential biases? Or, excluding substantial biases, is it possible to be confident at the 95% level that these differences represent real differences within the target group of health facilities? According to statistical theory, if 95% confidence intervals do not overlap (e.g. region A vs region B and region A vs region C), it is possible to be confident (at the 95% level) that there is a true difference. If the CI overlap (e.g. regions B and C), however, **it is not possible** to be confident that there is no difference in FP availability between the two regions *[5, 6]*. The reasons for this counterintuitive fact reside in statistical theory, beyond the scope of this guidance.

[1] See example in Annex 2.

Before proceeding, it is also important for the data analysis team to **gain an overview of the analysis platform outputs:** the indicator tables, the indices and the way they are organized. This can be achieved quickly by reviewing the five service dimensions within the indicator inventory platform, as the indicators are organized in the same way in both the indicator inventory and the data analysis platform.

9.2 Review the standard HHFA data analysis platform outputs

Review of the analysis platform outputs involves three stages:

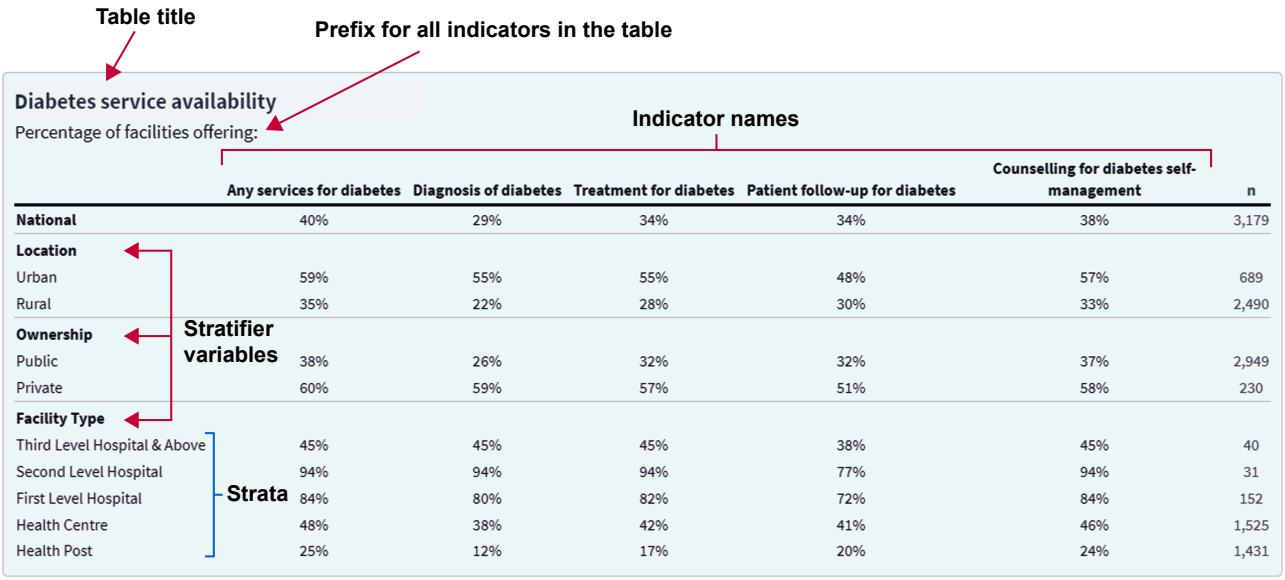
- Stage 1: Understanding the structure of the outputs.
- Stage 2: Understanding what the data reveal.
- Stage 3: Describing and summarizing the findings.

Following these stages, the next phase is interpretation (or “making sense of the data”) and looking for explanations for the findings (refer to Chapter 10). Table 10 suggests steps that may be followed through the stages of reviewing an HHFA indicator table. Similar steps may be followed when reviewing an HHFA graph. The proposed sequence of steps is not rigid and may involve several rounds of iteration.

Table 10. Steps for reviewing HHFA indicator tables

Stage 1: Understanding the structure of the indicator table (refer to example in Fig. 21)
1.1 Title The table title provides information on the service aspect/programme (e.g. malaria) and the HHFA indicator dimension (e.g. service availability).
1.2 Indicator prefix The indicator prefix Is located immediately below the title and provides the prefix for all the indicators in the table, e.g. “Percentage of facilities offering”.
1.3 Column headings The column headings show the indicator names. It is useful to quickly scan the indicator names to gain a sense of the number and types of indicators in the table.
1.4 Row headings The row headings refer to the stratifier variables (and their strata) selected for the analysis. Stratifier variables often used in the HHFA include: administrative area (if the survey design allows subnational comparison), location (urban/rural), managing authority/ownership and facility type/level.
1.5 Table denominator All the indicators in a single HHFA table always have the same denominator. For most HHFA indicators, there are two possible denominators: all facilities assessed, or only those facilities offering the specific service represented by the table. This information is often evident from the table title or may be found in the indicator inventory platform by clicking on any indictor in the table. In the example shown in Fig. 21, the denominator is "all facilities".
1.6 Number of facilities in denominator This is shown by (n) in the last column of the table. The total national level “n” will be the same value in all tables with the “all facilities” denominator. Also note the different “n” values according to the stratifier variable.

Fig. 21. Example: diabetes service availability table



Diabetes service availability

Percentage of facilities offering:

	Any services for diabetes	Diagnosis of diabetes	Treatment for diabetes	Patient follow-up for diabetes	Counselling for diabetes self-management	n
National	40%	29%	34%	34%	38%	3,179
Location						
Urban	59%	55%	55%	48%	57%	689
Rural	35%	22%	28%	30%	33%	2,490
Ownership						
Public	38%	26%	32%	32%	37%	2,949
Private	60%	59%	57%	51%	58%	230
Facility Type						
Third Level Hospital & Above	45%	45%	45%	38%	45%	40
Second Level Hospital	94%	94%	94%	77%	94%	31
First Level Hospital	84%	80%	82%	72%	84%	152
Health Centre	48%	38%	42%	41%	46%	1,525
Health Post	25%	12%	17%	20%	24%	1,431

Stage 2: Understanding what the indicators reveal

2.1 Indicator names

The indicator names provide a quick idea of the main service aspects or items presented in the table.

2.2 Indices

Observe if the table contains any indices. (All readiness tables contain indices.) For example, review any domain indices (e.g. staff and guidelines), the overall readiness index of the table (mean proportion/average percentage of all items available in facilities) and the "with all items" index. Then review how the individual indicator values within the index contribute to the index.

2.3 Individual indicators

The following set of considerations may be helpful in describing what the indicator has revealed:

- Note the national level value.
- Within each of the stratifier variables, note the minimum and maximum value and the range for each indicator (the difference between the maximum and minimum values).
- Assess the data for plausibility:
  - Do the values "make sense" and seem reasonable?
  - Are there any exceptionally high or low values ("outliers")? Outliers may reflect data errors, but may also point to unusual, but real and interesting findings that require further investigation, and should therefore not be discarded automatically.
- Compare the values within each stratifier variable (e.g. facility type) with the national average:
  - Are there large differences?
  - Identify/facility types/managing authorities/locations/regions that perform significantly below or above the average.

- Compare the indicator values against any existing targets, e.g. national or subnational targets:
  - Is the value close to the target or not?
  - Is the target realistic for the country/subnational area?
- Compare findings with those of previous facility surveys:
  - If a SARA or an HHFA was conducted in the past, with comparable methods and indicators, it is important to assess if there have been improvements or not. Monitoring the trends and then investigating the underlying causes of changes is one of the most important uses of HHFA data.
- Consider absolute numbers as well percentages:
  - While most HHFA indicators are expressed as percentages, absolute numbers also require consideration. For example: assume that 60% of facilities offer PMTCT services. The implications of scaling up services to achieve 100% are different in a context where the total number of facilities is 500 versus 5000.

2.4 Patterns

- Look for patterns emerging across indicators or strata. For example:
  - Do some indicators consistently show poor performances across multiple regions, or across both urban and rural locations?
  - Is there a difference in the performance of public facilities versus private facilities for particular indicators?
  - Does a particular region show poor performance across multiple indicators?
  - Do some high-performing regions emerge, with very satisfactory findings across several indicators?
- Note that when there are several indicators and several strata, the review becomes time-consuming and complex, and it may be helpful to review and describe one stratifier variable at a time. (The data analysis platform enables the selection of an individual stratifier variable to create a table that only displays the estimates for this stratifier variable.)

2.5 Linkages between related tables

- Consider together the various tables within a single service subarea or programme (e.g. malaria): availability, readiness, auxiliary indicators.
- Compare the results of the different tables. For example:
  - Is there good performance in service availability but poor performance in readiness?

2.6 Linkages across the service delivery system

- Health systems consist of multiple interlinked subsystems, programmes, resources and functions (governance, management, quality assurance, etc.).
- Therefore, in addition to analysing specific programmes or services, it is important also to maintain an overall view across all service components and indicator dimensions to identify potential common patterns or differences and potential influences of different components upon each other. For example:
  - How does the availability of noncommunicable disease services component compare with the availability of HIV services?
  - Are there similarities in the availability and readiness of HIV, tuberculosis and malaria?
  - Are certain items (e.g. guidelines) unavailable across all programmes?
  - Are there associations between performance related to management and finance components, and service availability and/or readiness?

Stage 3: Describing and summarizing the findings

3.1 Document systematically

- The report outline format produced by the data analysis platform contains boxes where analysts can insert key findings.
- Selected findings for each table should be documented systematically, below each table or groups of tables, guided by the points suggested in Stage 2.
- Note that this documentation process describes what the data show but does not attempt to explain the findings.

3.2 Summarize key findings

- A single table may yield a large amount of information and a lengthy description. Therefore, it is useful to select for each table, or group of related tables, a few key findings that will help to identify the conclusions and recommendations for the HHFA report.
- These findings will be further summarized and prioritized when identifying the key findings and recommendations for specific programmes or sections of the HHFA report, and once again when identifying the key messages of the overall report.
- The analysis and interpretation of the vast amount of information collected in an HHFA requires a progressive selection of the main findings, for which realistic recommendations are available.
- Selection of key findings and priorities to be addressed should be guided by some criteria, such as relevance, importance, evidence, plausibility of explanations, and existence of policies and interventions to address the identified issues within the current capacity.
- It is important to be parsimonious with the priority findings and recommendations: *too many priorities = no priorities.*

9.3 Conduct additional analyses as needed

As discussed, data analysis, review and interpretation represent an iterative cycle. Based on the initial set of results from the standard analyses, there may be a need for further analysis in areas of interest. For example, unexpected patterns of data and relationships between variables can emerge and may provide new insights and generate new questions that require further exploration of the data. These additional analyses can be generated within the data analysis platform, or the HHFA dataset can be exported to different formats that allow the use of statistical software and further customized analysis (e.g. Stata, SPSS, R, etc.).

Box 5 illustrates a simple additional analysis of two HHFA indicators looking at the availability of malaria service subcomponents among facilities that offer malaria services (fictitious data), highlighting key findings and providing comments and questions that require further exploration and interpretation.

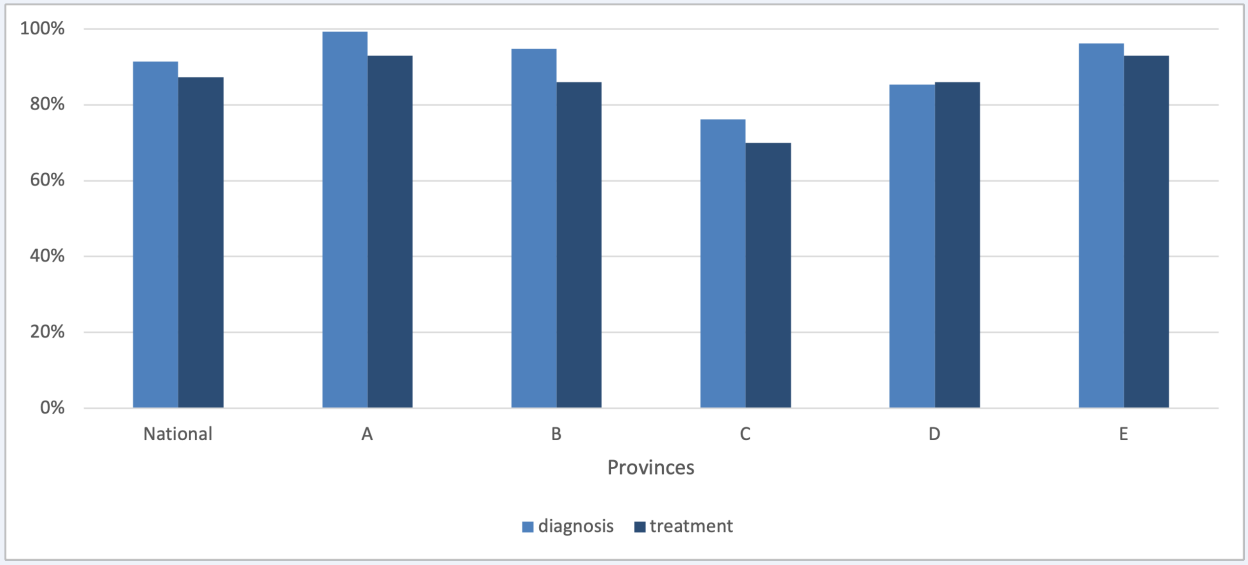
.

Box 5. **Examples of additional HHFA analyses and questions that may emerge: malaria service subcomponent availability**

Percentage of facilities offering diagnosis and treatment of malaria - table

	Malaria diagnosis	Malaria treatment	n
National	91%	87%	1377
Province:			
A	99%	93%	405
B	95%	86%	326
C	76%	70%	189
D	85%	86%	328
E	96%	93%	129
Location:			
Urban	82%	72%	326
Local	94%	93%	1051
Managing authority:			
Public	95%	92%	1034
Private	80%	65%	343

Percentage of facilities offering diagnosis and treatment of malaria - graph



Observations and questions

- There is a difference in the availability of malaria diagnosis services compared with treatment services. What are the reasons for this difference?
- The table shown in Box 5 uses a simple heatmap to enable easy identification of the regions with lower values of availability of malaria diagnosis and treatment: regions C and D. What are the reasons for this?
- The value of malaria treatment in region D is higher than the corresponding value for diagnosis in the same region. Is this an error?
- Health facilities in urban areas have lower availability of malaria diagnosis and treatment services than in rural areas. Why?
- Private health facilities have lower availability of malaria diagnosis and treatment services than public facilities. Why?

10. Data interpretation and communication

10.1 Review guiding principles for data interpretation

Introduction to interpretation

Data interpretation is a crucial but often neglected part of the HHFA process. It is often reduced to a simple description of the data through summary statistics (frequencies, means, etc.) that are presented in tables and graphs.

The purpose of the initial review of the analysis platform outputs (Section 9.2) is to provide a description of what the data show. Interpretation goes a step further and **aims to understand the reasons behind the findings**. Interpretation is therefore "an explanation of what something means" *[7]*.

In the context of the HHFA, interpretation is the process by which reviewers "learn from the data" in order to identify the most critical issues in health facility services. It involves knowing what the HHFA data are saying about the status of the services, understanding what the data mean within the country context, exploring the underlying causes, and making relevant and realistic recommendations for addressing the identified issues.

Ideally, data interpretation produces convincing and coherent "stories" that go beyond a purely descriptive narrative of numbers and facts. For example, the findings of an HHFA may show that several essential items for service delivery (equipment, diagnostic tests, essential medicines) are lacking in a high proportion of facilities. Mere reporting of these findings is not enough, given their importance for service delivery: it is crucial to understand the causes. Contextual information might point to inadequate procurement and weak logistic systems. Other explanations might include a devaluation of the local currency that has impaired international procurement, or the withdrawal of support from key funding agencies.

The overall process of interpretation usually involves an iterative cycle of analysis, review and interpretation, with several rounds of data exploration, interspersed with reviews by knowledgeable people of the insights gained at each stage. It may involve going back to the raw data, developing new analyses, supplementing the data with contextual information and data from other sources, and using prior knowledge and experience to understand findings and to ask new questions.

Guiding principles for interpretation

The volume of data produced by an HHFA is extremely rich but can also be overwhelming. The need for an initial broad scope of the analysis (as discussed in Chapter 8) and interpretation should, therefore, be balanced with the challenge of dealing with the large amount of data produced. Furthermore, there are no formal methods on how to interpret data, only guiding principles based on best practices. It is therefore important to approach the interpretation process systematically. Attention to the following aspects can help analysts to focus the analysis:

- data that are relevant to the HHFA objectives, and their underlying questions;
- critical aspects studied in previous surveys that need to be monitored (e.g. programmes that performed poorly, weaknesses in support systems, etc.);
- service delivery aspects that may have been affected by contextual changes (e.g. new policies, new programme components, changes in resource availability, outbreaks, etc.);

- new questions or hints that emerge from the initial analysis (e.g. exploration of reasons for significant improvements in indicators for some programme areas);
- important and potentially controversial issues (e.g. differences in the performance of key indicators among regions);
- unexpected or potentially overlooked findings that require further exploration; and
- gaps in understanding.

Stages of interpretation

Following the initial three stages of data review as described in Chapter 9, the interpretation process can also be outlined in stages (sometimes overlapping) which involve:

1. Understanding the HHFA data within the overall context of the country and its health system.
2. Triangulating HHFA data with data from other sources.
3. Finding explanations for HHFA data:
   - formulating initial hypotheses/explanations;
   - examining alternative explanations;
   - validating or discarding the initial hypotheses through other information sources and consultation with experts.
4. Reaching conclusions.
5. Defining priorities and recommendations.

Understanding the HHFA data within in the overall context of the country and its health system

Data interpretation does not start with a blank slate. It is always context dependent. The analysts' knowledge of the health system, its strengths and weaknesses, and the "external" factors that affect, either positively or negatively, health service delivery, are essential to understanding HHFA findings.

Therefore, data are best interpreted by local analysts and programme experts familiar with the health system and local conditions and that can use their knowledge, experience and judgment to make meaning of the findings. In addition, the relevance of the interpretation context used by the analysts contributes to the acceptance or rejection of conclusions by users.

Useful information sources for providing context include country profiles, health sector reviews, country statistics bulletins, WHO and World Bank health indicators, household survey reports, qualitative studies, etc. It is also important to review the HHFA findings in relation to country policies and plans, to assess their implementation progress.

Triangulating HHFA data with data from other sources

The interpretation of data and the formulation of hypotheses are strengthened by "triangulation": the process by which HHFA indicators are compared with data from different sources and/or methods, e.g. household surveys, RHIS, evaluation studies, etc. When data from different sources and methods point to the same interpretation and conclusions, confidence in the interpretation of the HHFA data is reinforced. For example, the HHFA data may highlight gaps in availability of trained staff and medicines in some programmes or regions; at the same time, RHIS data may reveal poor utilization of these services because users may seek services elsewhere.

Conversely, triangulation can expose inconsistencies that may require revision of the initial hypotheses and further analysis. For example, two regions show satisfactory values for most availability and readiness indicators for the most prevalent communicable diseases. The results of a household survey show, however, poor health outcome indicators for the same diseases. The contradictory findings raise questions. Is the population not able to access the services or choosing not to use the services? Are there significant issues related to service quality that were not captured by the HHFA? Do the findings of the household survey refer to the past? Have the past household survey findings perhaps triggered improvements in service delivery in the two regions, that may result in improved health outcomes shown in future household surveys?

Finding explanations for HHFA data

A core principle of interpretation is the transparent and honest use of data. Interpretations are neither facts nor truths. Data are filtered through the lenses of those who interpret them. Every interpretation is underpinned by assumptions about the aspects of service delivery considered important, based on previous experience, professional interests and knowledge, personal judgments, cultural perspectives, etc.

Interpretation can produce sound judgments based on strong evidence but can also produce errors resulting from “cognitive biases”. Such errors occur when analysts interpret the data based on their own wrong assumptions rather than on the facts. A common error, called “confirmation bias”, involves selecting only the evidence that supports one’s own convictions and theories, while ignoring the remaining data. The consequences of such errors can be serious, as they can result in wrong decisions.

The elements of subjectivity in data interpretation cannot, therefore, be completely eliminated. Consequently, interpretations are only suggested explanations, or **hypotheses**, about what the data mean. The hypotheses then need to be confirmed by contextual insights, information from other sources and expert opinions.

Interpretation perspectives: comprehensive overview as well as in-depth programme-specific views

The need for an initial broad analysis of all the HHFA data has been discussed. In addition to this comprehensive overview perspective provided by the initial analysis, it is important that the interpretation process also includes detailed programme-specific perspectives. The insights of programme experts are essential for achieving a thorough understanding and contextualization of programme-specific data, and for defining detailed programme-specific recommendations.

However, focusing only on specific indicators and programmes, in isolation from the other components of health services, can result in important patterns in the data being missed, and relationships between indicators of different service areas being overlooked. Further, common problems that require a unified strategy may be missed and the resulting picture may end up being narrow or even distorted. For example, analysts who focus only on malaria will not understand the issues that malaria services may have in common with other service areas (e.g. a weak supply system, lack of in-service training, insufficient supervision, poor measures of infection prevention and control, etc.).

Health services are the result of dynamic interactions among interconnected elements: the various programmes and services, as well as the workforce, infrastructure, equipment, medicines, financial resources, regulations, guidelines and procedures, management systems, etc. Therefore, for example, management and finance indicators may provide insights into performance in the service readiness dimensions. Service readiness performance may help to explain results within the quality of care dimension. There may be significant differences in performance among different programmes, or problems common to multiple programmes. Maintaining an overview of all the programmes and dimensions of health services is therefore crucial to achieving a solid understanding of what the data can reveal and can help to highlight and prioritize issues for the attention of managers and policy-makers.

10.2 Organize the interpretation process

The interpretation of HHFA data and the communication of findings are led by the data analysis team, supported by the survey coordination group, with inputs from programme-specific experts and other stakeholders. The analysts have often been involved in all phases of the HHFA: its design, planning and implementation. If not, they should be thoroughly briefed on the HHFA objectives and their underlying questions, the modules and questionnaires implemented, any changes to the standard questionnaire, the survey design, and any issues that may have impacted on the quality of data and, therefore, on their analysis and interpretation. These issues include the limitations and potential errors of the HHFA design (e.g. problems related to the MFL, sampling, undercoverage, etc.) and challenges and errors in the survey implementation (e.g. non-response).

Broadly, the process involves a comprehensive initial review of the standard analysis platform outputs by the data analysis team, followed by in-depth programme-specific reviews, with further rounds of exploration as needed, and an interpretation workshop that brings together multiple stakeholders for final review of the findings, conclusions and recommendations.

- **Conduct a comprehensive initial review of the standard analysis platform outputs**

  Refer to Section 9.2.

- **Engage programme experts for in-depth review of specific analysis outputs**

  Ideally, this phase involves an in-depth exploration of specific sections of the analysis outputs by groups of programme/technical experts. The analysis team provides an overview of the findings of the comprehensive initial review, and also prepares the specific sections of the analysis outputs for each expert group, highlighting any key initial programme-specific findings. The expert group reviews and interprets the findings.

  Ideally, the analysis team then engages in discussions with individual expert groups to refine the analysis and interpretation, and to develop recommendations. This process may require several rounds of discussion and further analysis, including the review of information from various sources to help explain the findings. If it is not feasible to engage with individual expert groups in this way, this phase could also take place in the context of an extended interpretation workshop.

- **Conduct an interpretation workshop**

  The HHFA interpretation workshop brings together the data analysis team, representatives of the various technical programmes, and other key technical stakeholders. The data analysis team provides an overview of the HHFA process and an introduction to the HHFA data analysis and interpretation concepts, guiding the participants in the process of reaching conclusions and developing key recommendations that will support decision-making. The participants then work in programme groups to review the HHFA findings, provide further interpretation and recommendations, and develop key messages. The outputs of these discussions will form the basis of the survey recommendations and report.

10.3 Review principles for effective communication of HHFA findings

The way in which the HHFA findings and recommendations are presented is important for their acceptance and use by decision-makers. Communication can attract the attention of users by presenting interesting findings packaged in an attractive and user-friendly way. Good communication aims to improve understanding among stakeholders about key service delivery issues and different policy options, and to inform decisions in a transparent, non-manipulative way.

The following points provide some guiding principles for communication of HHFA findings:

- Communication should be limited to informing, rather than persuading. Facts and data should be presented separately from interpretation. The readers of an HHFA report should be able to judge the conclusions and recommendations in relation to the evidence provided and the plausibility of the interpretations.
- The limitations and quality issues of the assessment should be acknowledged upfront.
- If there is no consensus on the interpretation of some indicators or on some recommendations, the writers of the report should be honest in recognizing the lack of evidence or the gaps in understanding.
- A balance must be found between the need to be concise and the complexity of the issues discussed. This may be challenging. On the one hand, complexity can make it difficult to tell a clear and convincing story, particularly to non-technical audiences. On the other hand, over-simplification can compromise the integrity of the message. *"Everything should be made as simple as possible, but no simpler*" (Albert Einstein).
- The findings of an HHFA are relevant to different audiences – policy-makers, national managers, subnational managers, programme managers and partners – each with their own needs, capacities and interests. Communication should therefore be adapted to the audience to which is addressed. Further, some of the users of an HHFA report may be "data illiterate" (e.g. politicians), lacking familiarity with technical jargon and methodological issues. Others may be too busy to delve into details (e.g. policy-makers, senior managers). A balance should be found between providing sufficient key information to support informed decision-making, and providing excessive technical detail. It is therefore important to:
  - understand the decisions that each audience needs to make and the specific information that would support those decisions;
  - summarize key findings, but also provide access to detailed data for reference if needed; and
  - use language appropriate to the audience, avoid jargon, and provide definitions of key technical terms.
- The conclusions and recommendations are, arguably, the most difficult and influential aspects in the communication of HHFA findings. Many readers will focus only on these sections of the report. The conclusions are the result of a complex process of selection of the most important findings. Recommendations are often not straightforward: different policy options may need to be provided, along with their respective advantages and disadvantages. Useful recommendations require:
  - clear relationships between the main findings (highlighted in the conclusions) and the recommendations;
  - in-depth knowledge of the health system and of the technical and financial capacity of the ministry of health and its partners for implementation of the recommendations; and
  - clear prioritization and sequencing of interventions, within a stated timeframe and with defined responsibilities.

10.4 Prepare the final HHFA report

Based on the results of the data interpretation workshop, the report writer prepares the final draft of the survey report. The final draft report is submitted to the survey coordination group and to the ministry of health for approval. Sometimes, a "validation" workshop is held to engage a wide range of stakeholders, achieve their buy-in and receive further inputs.

The report should present all the HHFA findings but focus on communicating the most important and relevant findings, along with recommendations for action. In addition to addressing new findings, the report should also address the objectives of the survey, i.e. the questions to which the data should provide answers.

The data analysis platform produces a very large number of tables and graphs. While all of these analysis outputs should be reviewed as discussed, it is not necessary to include the full set of outputs in the main body of the final survey report. Therefore, the data analysis team, in collaboration with the survey coordinating group, should select a subset of key tables and graphs to include in the report. The complete set of analysis outputs should be made available as an annex.

A suggested HHFA report structure is presented in Table 11. The structure can be adapted to country needs.

Table 11. Suggested HHFA report structure

Report section	Contents/use
Front page	Title, country, year and organization(s), usually the ministry of health and partners.
Foreword	Optional; often signed by a high-level ministry of health official.
Contents	Provides an overview and helps to navigate the document.
Abbreviations	Should be included if abbreviations/acronyms are used.
Executive summary	An important section, and often the only one that is read by busy managers and politicians; should focus on key findings, conclusions and recommendations; should encourage reading of other key sections of the report.
Introduction	Provides the rationale for the survey and the relevant country and health system context that inform the interpretation of the findings.
Objectives	Objectives should have been defined in the survey planning phase, based on the questions the survey aimed to address; the objectives help to direct the focus of the reader.
Methods	A detailed description of the survey methods should be provided, including: ■ the HHFA modules and questionnaires implemented, and any adaptations ■ the survey design: census or sample (with sampling design), sample size, sampling methodology ■ numbers of data collectors and area supervisors ■ duration of data collection training ■ duration of data collection ■ data management, quality review and data cleaning processes ■ practical challenges encountered and methodological limitations of the survey, and their influences on data collection, including the level of completeness of the MFL, the response rate (if low, with reasons), etc.
Findings	Selected tables and graphs generated by the data analysis platform are presented in sections according to the structure of analysis platform's report outline. Tables and graphs are briefly described, highlighting indicator patterns, unexpected findings and possible associations with other indicators. Selected key findings and conclusions are presented for each section and subsection.
Conclusions	This section is critical and should present a narrative of the overall key findings, with an honest balance among: ■ the evidence provided by the data ■ acknowledgment of the data limitations ■ contextual knowledge ■ interpretive insights. It should provide a synthesis of key findings with their interpretations, with reference to key issues of service delivery that need to be addressed. The conclusions should be linked to the recommendations of the next section.

Recommendations	Recommendations should: ■ be realistic in number and scope, considering existing and anticipated system capacities; ■ provide a sense of priority, with a timeframe for implementation; ■ define clear responsibilities for action; ■ possibly include an implementation roadmap and a monitoring framework; and ■ highlight important evidence gaps and controversial issues, that may require further analyses, surveys and/or other research.
References	Lists articles, reports and other documents referenced in the text, with web links where available.
Data annex	The data annex should provide the complete set of data analysis platform outputs. Given the potential length, it is not necessary to print the data annex, but it should be available electronically.
Other annexes	Additional materials used in the survey that may be useful for understanding the survey or for future surveys (questionnaires, training materials, etc.); they can also be stored in an easily accessible online workspace.
Data curation	An optional link to a password-protected online repository may be provided, where the dataset, the metadata and other materials are available, upon request to the ministry of health, for research or other purposes.

10.5 Prepare additional communication products

Table 12 provides a summary of various products and methods that may be used to communicate HHFA findings, conclusions and recommendations to different target audiences.

Table 12. HHFA communication products and methods

Product/method	Audience	Contents
Full report	Ministry of health national and subnational staff, programme managers (ministry of health and partners), planners	Executive summary; detailed description of objectives, methods, limitations, findings, conclusions and recommendations
Short report	Policy-makers and high-level managers (ministry of health and partners)	Expanded executive summary: short description of objectives, methods and limitations; focus on main findings, conclusions and recommendations
Subnational report, e.g. provincial (if adequate sample size)	Ministry of health provincial and district staff, programme managers (ministry of health and partners)	Focus on provincial dataset, with key findings, conclusions and recommendations; comparisons with other regions and national averages for selected indicators; district comparisons if adequate sample size
Programme-specific report	Programme staff at health system various levels	Focus on detailed programme-specific findings, conclusions and recommendations
Policy brief	Policy-makers	Focus on a few key policy issues; concise summary of findings and conclusions; focus on policy recommendations
Oral presentation	Participants in planning meetings, health sector reviews and other events	Summary of objectives, methods and limitations; main findings; summary of conclusions and recommendations

Dashboard	Policy-makers, high-level managers, planners	Summary of key information
Press release	Journalists, news media	Main findings related to the conclusions and recommendations
Social media	General public; followers	Brief information that HHFA findings are available; short video; link for downloading the report

Preparing a policy brief

The following points summarize the main features of a policy brief:

- A policy brief is a short document, addressed to politicians and decision-makers, that focuses on a limited number of important health service issues and provides recommendations on decisions and policies that would address those issues.
- It usually conveys a sense of urgency about the decision(s) to be made.
- It usually consists of headings, followed by short sentences. It should be easy to read and useful to busy policy-makers and should avoid details of survey methods.
- It succinctly provides the context and scope of the issue(s), in order to highlight the importance of the decision to be made.
- It can have either an advocacy tone, arguing in favour of a particular policy, or it can provide balanced information and different policy options, that aim to assist the policy-makers in making decisions. In the latter case, it should discuss the pros and cons of the different courses of action, with recommendations based on the evidence from the HHFA findings, and information from relevant contexts, past experience and the scientific literature.

Preparing dashboards

Key data are increasingly presented in dashboards. **"**A dashboard is a visual display of the most important information needed to achieve one or more objectives, consolidated and arranged on a single screen so the information can be monitored at a glance" *[8]*.

Dashboards are multivisual displays, useful for viewing several indicators at a glance. They help to highlight important issues, and to detect patterns and relationships that may otherwise remain unnoticed. The choice of indicators and the design are crucial: they should guide the attention of the user to the most important information. A mix of graphs and simple tables can often offer the best visual display of data. An optimal balance must be found between the quantity of information provided and the capacity to absorb and understand it: a dashboard that is excessively complex may obscure key information.

10.6 Disseminate the HHFA findings

Dissemination of HHFA findings is key to the success of the survey and should start as soon as possible. The survey findings will be useful only if data are received in a timely manner by the intended recipients and if the strengths, potential uses and limitations of the findings are well understood by the target audience.

The purpose of dissemination is to ensure that the right people receive survey results in a format that is targeted specifically to their needs. Target audiences for the HHFA are usually decision-makers at national, subnational and facility levels, ministry of health as well as partners.

The survey manager plans meetings to present the HHFA results to key stakeholders within the ministry of health and among key donors and implementing partners. After the HHFA report is finally approved, it should be circulated among key stakeholders and made available on the ministry of health website.

Any of the following activities may be undertaken to disseminate the HHFA results:

- national and/or subnational dissemination workshops;
- dedicated meetings with presentations to key donors and other stakeholders;
- presentations of policy briefs to high-level ministry of health decision-makers;
- presentations in annual health sector reviews and planning meetings;
- web dissemination, e.g. ministry of health website, country observatory, etc.;
- publication of reports, presentations, brochures and scientific articles; and
- press releases and use of newspapers and social media to inform the public of the release of the HHFA report.

11. HHFA data curation

11.1 Introduction

Much of the content of this chapter is derived from the website of the International Household Survey Network (IHSN) *Guidelines: Data archiving and dissemination*. Please refer to this website for detailed guidance on survey data curation.[1]

It is important that a country's final HHFA dataset and related documentation are stored in a secure location where they can be accessed by authorized users for future reference and further analysis. This is called **data curation** and involves the documentation, anonymization, cataloguing, dissemination and preservation of microdata and metadata, as summarized in Section 11.2.

Archiving refers to the storage of documents or records so that they can be accessed again in the future. A data archive is a collection of data that is moved to a specific repository for future reference or compliance reasons.

Microdata are "…unit-level data obtained from sample surveys, censuses, and administrative systems. They provide information about characteristics of individual people or entities such as households, business enterprises, facilities, farms or even geographical areas such as villages or towns…".[2] For the HHFA, microdata refers to individual facility level data or, in the case of the quality of care record reviews, to individual patient-level data.

Metadata are data that provide information about other data. Metadata help researchers and other users to find the data, understand what the data are measuring and assess the quality of the data.

HHFA documentation and archiving should include the storage and related processes of:

- the final HHFA report;
- the final HHFA dataset (microdata); and
- other information (metadata) useful for understanding the survey, using the data, and/or for conducting future surveys (e.g. questionnaires, training materials, documentation of survey methodology and processes followed, explanations of survey challenges and limitations, etc.).

11.2 Best practices for survey documentation and archiving

Survey microdata and documentation are irreplaceable assets and valuable resources for government departments and academic researchers. They should be managed in a way that encourages their widest possible use and re-use, while at the same time protecting confidentiality. The absence of effective procedures for preserving such data has led to important data being lost. Furthermore, lack of transparent procedures for accessing the data limits their use for present and future users. It is therefore important to establish and maintain a national HHFA archive based on international recommendations and best practices. Today, data archives are almost always digital and are ideally web-based or made publicly available through the internet.

[1] Adapted from: Data archiving and dissemination | IHSN (accessed 1 August 2022).

[2] https://datahelpdesk.worldbank.org/knowledgebase/articles/228873-what-do-we-mean-by-microdata (accessed 1 August 2022).

The IHSN[1] provides detailed guidance and best practices related to data archives, including the documentation, anonymization, cataloguing, dissemination, and preservation of microdata, as well as the institutional arrangements and technical requirements for operating a data archive.

Documentation (metadata)

Appropriate information about the survey data helps users to:

- **Find the data they are interested in:** Names, abstracts, keywords, year and other important metadata elements enable users to locate specific datasets and variables. Any cataloguing and resource location system is based on metadata.
- **Understand what the data are measuring and how the data have been created:** Descriptions of the survey design and the methods used when collecting and processing the data enable users to understand the data and their context.
- **Assess the quality of the data:** Information about the data collection standards, as well as any deviations from the planned standards, help users to know whether data are useful for specific purposes.

"…Metadata can be stored in a variety of places. Where the metadata relates to databases, the data is often stored in tables and fields within the database. Sometimes the metadata exists in a specialist document or database designed to store such data, called a data dictionary or metadata repository…The most common types of metadata…useful for data analysis are: value labels and missing value codes (the domain), variable labels, variable types, relationship to other data, variable sets, change logs, weights, strata…"[2]

A set of international metadata standards has been developed to facilitate data communication between organizations and software systems and to improve the quality of statistical documentation provided to data users. These include the Data Documentation Initiative (DDI) and the Dublin Core Metadata Initiative (DCMI). The DDI and DCMI are based on Extensible Markup Language (XML), a type of regular text file that tags for meaning – rather than appearance – and can be viewed and edited using any standard text editor. XML files can be searched and queried like a regular database. The IHSN recommends the adoption of the DDI metadata standard for the documentation, cataloguing and dissemination of survey microdata.

Anonymization

Anonymizing a microdataset refers to removing or modifying its identifying variables. For the HHFA, this involves removing identifiers of individual health facilities such as name, unique ID and geographic information system (GIS) data; for HHFA quality of care record reviews, individual patient identifiers such as name, address, patient registration number and national ID number as also removed.

Cataloguing

The objective of a catalogue is to provide easy access to data and documentation in a format most convenient for users, preferably through a searchable online catalogue. The IHSN and the World Bank have developed a free, open-source, DDI-compliant application for the cataloguing and dissemination of microdata, the National Data Archive (NADA).[3] The NADA is a web-based survey cataloguing system that serves as a portal for researchers to browse, search, apply for access and download relevant survey data and metadata. It is available as a pre-packaged but fully customizable website that can be used by countries for archiving and disseminating their HHFA data. The NADA does not provide tools for data tabulation or analysis, but aims to provide users with detailed and searchable documentation of microdatasets, along with information on policies and procedures for their access and use.

---

1 http://ihsn.org/ (accessed 1 August 2022).

2 https://www.displayr.com/what-is-metadata/

3 NADA | Microdata Cataloging Tool (ihsn.org)

Dissemination

Providing researchers and other users with access to microdata can promote the diversity and quality of analyses, broaden the use of existing data, and increase the return on data collection investments. Disseminating microdata, however, also involves costs and risks, including the risk of disclosure of confidential information. It is therefore important to establish formal policies and procedures defining the conditions of access to microdata within the framework of the national legislation.

Preservation

Survey data represent a significant investment by producers, have considerable value for present and future users, and must be preserved. A sustainable preservation programme addresses organizational infrastructure, technological infrastructure and funding.

11.3 Creating metadata for the HHFA

Metadata can be created through various media such as simple word processing programs and software application programs. This section provides guidance on creating metadata for the HHFA by identifying key elements that need to be included, based on the five sections of the DDI framework. Much of this information will have been generated as part of the data processing steps and the HHFA report production.

Survey description

Section 1. Document description

The document description serves as an introduction to the survey metadata as a whole. It provides background information such as the study title, document producer(s), date of production and version number, as well as the documentation process. It provides “metadata about the metadata”.

Section 2. Study description

The study description serves to identify the study (survey) itself and to provide overview information. This includes the survey scope, coverage and sampling, who collected, compiled and disseminates the data, the data collection methods and processing, editing, review and access, as well as a summary (abstract) of the content of the data. This section also names producers and sponsors, and describes points of contact, disclaimers, copyrights and how the study should be cited.

Dataset description

Section 3. Data file description

The file description of a dataset provides the dataset contents, its producer, and the version. It should also include record and variable counts, and an explanation of how missing data are coded or accounted for, as well as any other relevant notes. When applicable, a section on processing checks should be included. This element serves to provide information about the types of checks and operations that have been performed on the data file to make sure that the data are as correct as possible, e.g. consistency checking.

Section 4. Variable description

This section consists of descriptions of the actual data: detailed information on each variable, including the literal question text; universe, variable and value labels; and any derivation and imputation methods.

The **variables list** is typically a table listing every variable in the dataset and providing for each the variable number, name and label. This list also provides the literal question associated with the variable, the variable format (character or numeric, number of units), and the number of valid and invalid cases (see Table 13 for an example).

Table 13. Variables list

#	Name	Label	Type	Format	Valid	Invalid	Question
1	V_001	Facility Name	Discrete	Character-12	97	0	Record the name of the facility

The **variables description** is more detailed than the variable list. It includes variable information (type, format, missing value coding), statistics (valid and invalid), literal question, derivation and imputation methods, and any notes (see Table 14 for an example).

Table 14. Variables description

#1 V_001: Facility name	
Information	[Type= discrete] [Format=character] [Missing=*]
Statistics	[Valid=97 /-] [Invalid=0 /-]
Literal question	Record the name of the facility
Notes	

Section 5. Other material

This section allows for the description of other materials related to the survey. These can include resources such as documents (e.g. questionnaires, coding information, technical and analytical reports, interviewer's manuals), data processing and analysis programs, photos and maps. The DCMI[1] provides a standard for documenting digital resources such as questionnaires and reports.

11.4 HHFA global archive

WHO has developed a global SARA/HHFA data archive where countries may choose to store their survey reports and metadata. The archive applies the DDI/DCMI standards for describing the data produced by the surveys. The NADA software is used as the underlying platform for the data archive.

Information is made available in the global archive based on authorization provided to WHO from individual countries. Members of the public can create an account to request access to data, based on country authorization.

[1] https://www.dublincore.org/

References

1. SCORE for health data technical package: global report on health data systems and capacity, 2020. Geneva: World Health Organization; 2021 (https://www.who.int/publications/i/item/9789240018709, accessed 16 November 2022).

2. Data quality review: a toolkit for facility data quality assessment. Module 3: Data verification and system assessment. Geneva: World Health Organization; 2017 (https://www.who.int/data/data-collection-tools/health-service-data/data-quality-assurance-dqa?ua=1, accessed 16 November 2022).

3. Master facility list resource package: guidance for countries wanting to strengthen their master facility list. Facilitator guide for the MFL training. Geneva: World Health Organization; 2019 (https://apps.who.int/iris/handle/10665/329492, accessed 16 November 2022).

4. Turner AG, Angeles G, Tsui AO, Wilkinson M, Magnani R. Sampling Manual for Facility Surveys for Population, Maternal Health, Child Health and STD Programs in Developing Countries. MEASURE Evaluation Manual Series, No. 3. MEASURE Evaluation. Carolina Population Center, University of North Carolina at Chapel Hill. July 2001. Pages 59 and 60. (https://www.measureevaluation.org/resources/publications/ms-01-3/at_download/document, accessed 16 November 2022).

5. Austin PC, Hux JE. A brief note on overlapping confidence intervals. J Vasc Surg. 2002 Jul;36(1): 194-5. doi: 10.1067/mva.2002.125015. (https://www.jvascsurg.org/article/S0741-5214(02)00030-7/pdf, accessed 16 November 2022).

6. Schenker, N., & Gentleman, JF. On Judging the Significance of Differences by Examining the Overlap Between Confidence Intervals. Am Stat., 2001; 55(3), 182–186. (http://www.jstor.org/stable/2685796, accessed 16 November 2022).

7. Cambridge Dictionary. Cambridge University Press. (https://dictionary.cambridge.org, accessed 16 November 2022).

8. Few, S. Information Dashboard Design. El Dorado Hills (CA): Analytics Press, second edition; 2013.

Annex 1: Calculating HHFA indices

A1.1 General service readiness domain scores and index

General service readiness is described by the following five domains of tracer indicators:

- Basic amenities
- Basic equipment
- Standard precautions for infection prevention
- Diagnostic capacity (laboratory)
- Essential medicines.

Each domain consists of a set of tracer items. Table A1.1 lists the tracer items for each domain and shows the calculation for the domain score. In the HHFA indicator inventory platform, each item is presented as a percentage. The domain score is presented at the end of the list of indicators in each readiness table, as the mean percentage of items in that domain available across all facilities. Note that the lists of tracer items in the HHFA are updated versions of the SARA lists. The essential medicines list has been updated for consistency with the basket of essential medicines for PHC related to SDG 3.8.1.[1]

Table A1.1. General service readiness tracer items, domain scores and general readiness index

General service domains	Tracer items	Domain score (mean percentage of items available)
(1) **Basic amenities**	**1.** Power **2.** Improved water source facility premises **3.** Access to improved sanitation facilities for clients **4.** Room with auditory and visual privacy for patient consultations **5.** Communication system (phone or short-wave radio) **6.** Access to computer with internet **7.** Emergency transportation system for patients	*n* / 7 × 100, where n is the total number of items available in the domain
(2) **Basic equipment**	**1.** Thermometer **2.** Stethoscope **3.** Blood pressure apparatus **4.** Pulse oximeter **5.** Examination light **6.** Otoscope **7.** Ophthalmoscope **8.** Adult weighing scale **9.** Child scale **10.** Infant scale **11.** Measuring tape **12.** Height board/stadiometer	*n* / 12 × 100, where n is the total number of items available in the domain

[1] https://unstats.un.org/sdgs/metadata/?Text=&Goal=3&Target=3.8

General service domains	Tracer items	Domain score (mean percentage of items available)
(3) **Standard precautions for infection prevention**	**1.** Guidelines for standard precautions **2.** Guidelines for health care waste management **3.** Staff trained in health care waste management) **4.** Hand hygiene items (soap and running water or alcohol-based hand rub) **5.** Latex gloves **6.** Single-use, standard disposable or auto-disable syringes **7.** Sterilization equipment in facility or system for sending items outside for sterilization **8.** Environmental disinfectant **9.** Appropriate storage of sharps waste (sharp container **10.** Appropriate storage of non-sharp infectious waste (waste receptacle with lid and plastic bin liner) **11.** Safe final disposal of sharps **12.** Safe final disposal of non-sharp infectious wastes	*n* / 12 × 100, where n is the total number of items available in the domain
(4) **Diagnostic capacity**	**1.** Haemoglobin **2.** Blood glucose **3.** Urine dipstick – glucose **4.** Urine dipstick – protein **5.** Urine dipstick – ketones **6.** Urine pregnancy test **7.** Malaria diagnostic capacity **8.** HIV diagnostic capacity **9.** Syphilis rapid diagnostic test (RDT)	*n* / 9 × 100, where n is the total number of items available in the domain
(5) **Essential medicines**	**1.** Salbutamol inhaler **2.** Beclomethasone or other corticosteroid inhaler **3.** Gliclazide or other sulphonyl urea oral **4.** Metformin tab/cap **5.** Insulin – regular injection **6.** Any two of the following oral anti-hypertensives: beta blocker, calcium channel blocker, thiazide-like diuretic, ACE inhibitor **7.** Simvastatin or other statin tab/cap **8.** Furosemide oral or injectable **9.** Aspirin tab/cap **10.** Ibuprofen tab/cap **11.** Paracetamol tab/cap **12.** Morphine (oral or injection) **13.** Fluoxetine or other SSRI tab/cap **14.** Phenytoin or carbamazepine tab/cap **15.** Gentamicin injection **16.** Ceftriaxone injection **17.** Amoxicillin tab/cap **18.** Procaine penicillin or benzathine penicillin injection	*n* / 30 × 100, where n is the total number of items available in the domain

	19. Artemisinin-based combination therapy (ACT) tab/cap **20.** Artesunate (injection or suppository) **21.** Antiretrovirals (ARVs) for first line combination treatment regimen **22.** Combination therapy for tuberculosis **23.** Oral rehydration salts (ORS) **24.** Zinc sulphate tab/cap **25.** Ready-to-use therapeutic food (RUTF) **26.** Hormonal contraceptives (oral, injection and/or implants) **27.** Folic acid tab/cap **28.** Magnesium sulphate injection **29.** Oxytocin injection **30.** Chlorhexidine	*n* / 30 × 100, where n is the total number of items available in the domain
General service readiness index		**Mean score of the five domains:** ***(a + b + c + d + e) / 5***

Fig. A1.1. **Example of general readiness index calculations**

Service area	Indicator table	Readiness index per service area	General service readiness index
Basic amenities	Basic amenities for main service area of the facility	90%	70%
Basic equipment	Basic equipment in the main service area of the facility	70%	
Standard precautions for infection prevention	Standard precautions for infection prevention	80%	
Basic diagnostic capacity (laboratory)	Basic diagnostic capacity	50%	
Essential medicines	WHO essential medicines	60%	

A1.2 Service-specific readiness domain scores and indices

For each specific service that is assessed in the HHFA, there is a readiness table containing indicators that show the availability of tracer items within each of four domains:

- Trained staff and guidelines
- Equipment
- Diagnostics
- Medicines and commodities.

Table A1.2 provides an example of a service-specific readiness table for diabetes services. In the HHFA indicator inventory platform, each item is presented as a percentage. In service-specific readiness tables, the denominator is the number of facilities offering the specific service. The indicators are therefore expressed as "Percentage of facilities offering diabetes services with…".

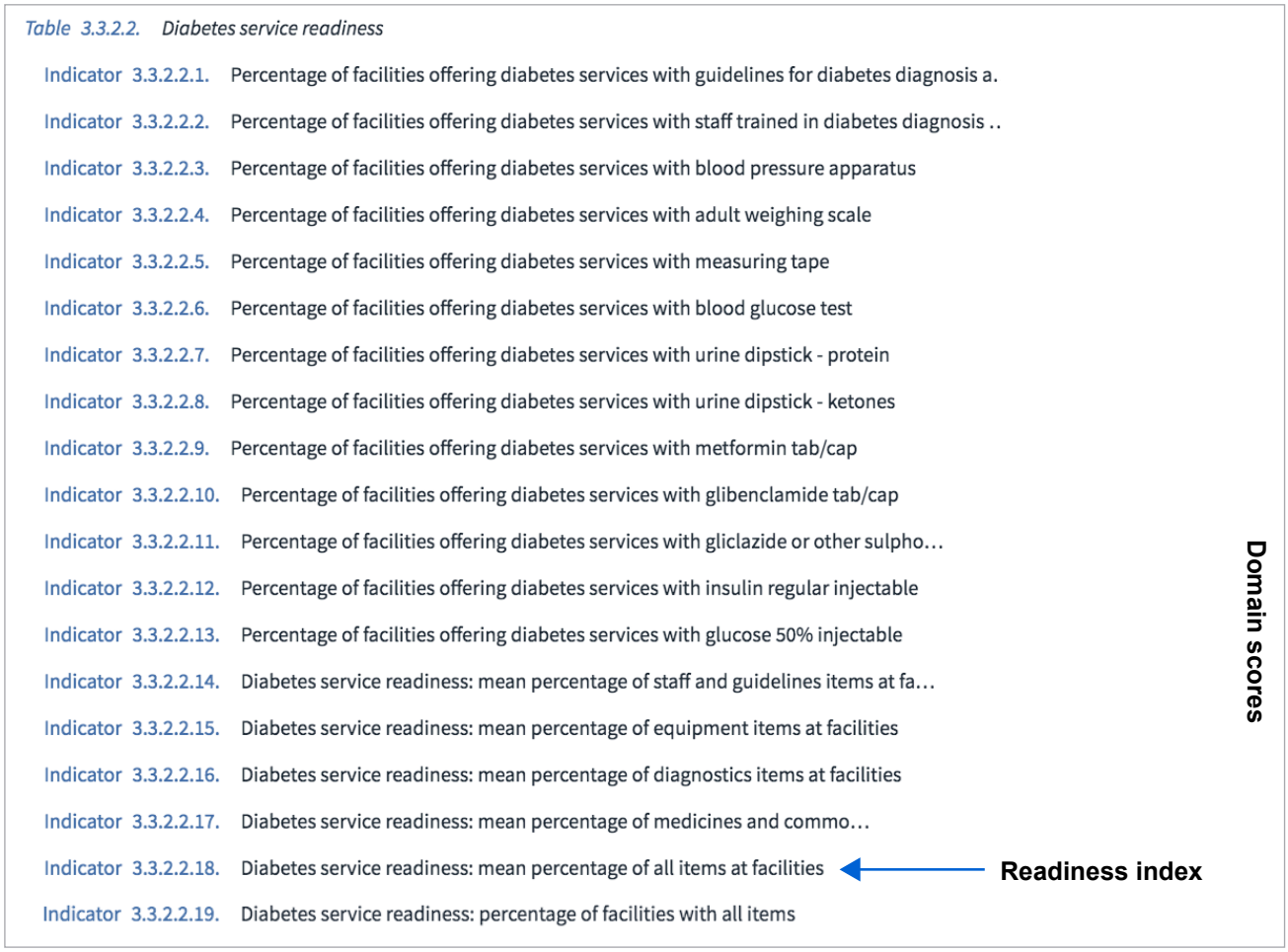
The score for each domain as well as the overall readiness index are presented as "mean" indicators at the end of the list of indicators in each readiness table. Note that the number of items per domain varies among different services. For some services, some of the domains may not be represented.

Table A1.2. Diabetes service readiness indicators, domain scores and index

Service specific domains	Tracer items	Domain score (mean percentage of items available)
(a) **Trained staff and guidelines**	a1) Guidelines for diabetes diagnosis and treatment	*n* / 2 × 100, where *n* is the total number of items available in the domain
	a2) Staff trained in diabetes diagnosis and treatment	
(b) **Equipment**	b1) Blood pressure apparatus	*n* / 3 × 100 where *n* is the total number of items available in the domain
	b2) Adult weighing scale	
	b3) Measuring tape	
(c) **Diagnostics**	c1) Blood glucose test	*n* / 3 × 100 where *n* is the total number of items available in the domain
	c2) Urine dipstick – protein	
	c3) Urine dipstick – ketones	
(d) **Medicines and commodities**	d1) Metformin oral	*n* / 3 × 100 where *n* is the total number of items available in the domain
	d2) Glibenclamide, gliclazide or other oral sulphonyl urea	
	d3) Insulin regular injectable	
Service-specific readiness index		**Mean score of all the items:** (***a1 + a2 + b1 + b2 + b3 + c1 + c2 + c3 + d1 + d2 + d3***) / 11

Fig. A1.2 shows the diabetes service readiness table in the HHFA indicator inventory platform.

Fig. A1.2. **Diabetes service readiness table**



Annex 2: Sampling methodology

A2.1 Introduction

Much of the content of this annex is adapted from *Sampling manual for facility surveys for population, maternal health, child health and STD programs in developing countries, July 2001*. Please refer to this manual for further details *[1]*.

The HHFA design methodology includes the following key components:

- Deciding whether the survey will involve a census of all facilities or a sample. If the decision is to use a sample of facilities:
  - determining the sampling methodology (the method used for selecting the sample of facilities);
  - determining the sample size (the number of health facilities that are selected and from which data will be collected).

The first component is addressed in Section 4.2. This annex addresses the remaining components. The annex does not aim to provide detailed instructions on sampling but rather to highlight key considerations of which the HHFA survey coordination group, the technical committee, and the survey manager should be aware. Table A2.1 provides definitions of key terms relevant to this annex.

Decisions concerning sampling are complex and have substantial resource implications. A statistician with expertise in survey sampling design should therefore always be involved from the early phases of HHFA planning.

The statistician leads the technical aspects of sampling. However, the HHFA technical committee must provide the statistician with detailed information on the survey objectives, information needs and context (e.g. availability of an updated MFL, number of health facilities in the country, desired precision of estimates, subnational comparisons, stratifier variables[1] to be used, comparison with previous health facility assessments, time constraints, and what variables are required at what level of disaggregation etc.), as well as the available budget and other resources. A dialogue is then needed between the technical committee and the statistician to ensure an HHFA design methodology that meets information needs and quality standards, in the most efficient way and within the available survey budget.[2]

[1] Stratified random sampling allows estimates to be obtained separately for desired, mutually exclusive, subgroups of facilities, called strata. A particular type of subgrouping into strata is called a “stratifier variable”. Refer to Section A2.3.

[2] When financial resources are scarce, it may be necessary to try to maximize precision within a given budget. For this purpose, the Neyman “optimum allocation” method can be used. This method determines the sample size of the different strata according to their costs and variance. A statistician’s advice is recommended, due to the complex sampling design required.

Table A2.1. **Key terms relevant to sampling**[1]

Term	Definition
Accuracy	"The degree to which a measurement or an estimate based on measurements represents the true value of the attribute that is being measured" *[2]*. Sometimes the term validity may be used, with the same meaning.
Bias	"Deviation of results or inferences from the truth, or processes leading to such deviation. An error in the conception and design of a study—or in the collection, analysis, interpretation, reporting, publication, or review of data—leading to results or conclusions that are systematically (as opposed to randomly) different from truth" *[2]*. Biases are systematic errors that can result in inaccurate findings and interpretations.
Census	A census is a study of every unit in a population. For the HHFA, this means inclusion of all the health facilities in the country in the survey.
Confidence interval	"A range of values, calculated from the sample observations, that is believed, with a particular probability, to contain the true parameter value. A 95% confidence interval, for example, implies that were the estimation process repeated again and again, then 95% of the calculated intervals would be expected to contain the true parameter value" *[3]*. The confidence interval measures the degree of uncertainty, or certainty, around a sample estimate, with respect to the true value in the target group.[2] It may also be described as a range of values around the point estimate that is likely to include (usually with 95% probability) the unknown target group value. The confidence interval is related to the precision of the estimate, and not to systematic errors (biases).
Coverage bias	Coverage bias occurs when some facilities of the target group are erroneously missing from the sampling frame because, for example, the sampling frame was built on an outdated or incomplete list of facilities. The amount of bias depends on the proportion of the facilities not covered and whether the characteristics of facilities not covered differ from those of the facilities in the target group.
Design effect	The design effect for a cluster survey is the ratio of the variance for that design to the variance calculated from a simple random sample of the same size. It is an adjustment that should be used to determine the survey sample size in cluster sampling. It varies from survey to survey.
Errors (in surveys)	"Deviations from the true values applicable to the target group studied. Deviations of what is desired in the survey process from what is obtained." *[4]*. Errors in surveys are differences between the sample estimates and the true values of the target groups. There are various types of error, as described in Annex 2.
Non-response bias	Non-response bias occurs when facilities included in the sample cannot be surveyed (e.g. due to refusal of the facilities' managers to participate in the HHFA). If the non-surveyed facilities differ systematically from the other facilities in the sample, especially with respect to some of the factors under study, this will result in non-response bias.

[1] Quotes, from reported sources, are included within inverted commas, sometimes followed by further explanation. A more extensive glossary for HHFA is included as Annex 6.

[2] "Target group" (e.g. the "universe" of health facilities) is used here to designate the statistical term "target population", which might be confusing in the HHFA context, since the objects of the assessment are health facilities and not population members.

Precision	"The quality of being sharply defined through exact detail. Relative lack of random error. In statistics, the measure of precision is the inverse of the variance of a measurement or estimate" *[2]*. Precision is a measure of the uncertainty around the estimate produced in a sample, due to the fact that the inclusion of facilities in a sample is partly determined by chance: the composition of two or more samples (obtained using the same sampling procedure) will differ in part by chance. (In contrast, accuracy refers to how close a measurement in a sample is to the true value in the target group.) The larger the sample, the higher will be is the precision of the estimates.
Probability sampling	Probability sampling is a sampling method by which every element in the target group (every health facility in HHFA) has a known, non-zero probability of being selected. If correctly applied, it enables selection biases to be avoided.
Representative sample	A representative sample is an unbiased sample, selected with probability methods, which to a large extent reflects the characteristics of the target group of interest. In the HHFA, the target group is all the eligible health facilities that constituted the sampling frame.
Sample	"A selected subset of a population. A sample may be random or non-random and may be representative or nonrepresentative" *[2]*. In the HHFA, samples are selected randomly, using probability methods.
(Sample) weights	Sample weights are adjustments used to correct for the unequal sampling probabilities between units. For example, in disproportionate sampling, when the strata allocations in the sample are not proportional to the strata population sizes, weights must be used.
Sampling error	The error that results from the random (chance) variation in samples, even when they are selected using correct procedures. It can be controlled by increasing the sample size: the larger the sample size, the smaller the sampling error.
Sampling frame	The sampling frame is a listing of the target group from which the sample is selected. In the HHFA, it is based on the MFL; after ineligible facilities are excluded from the MFL, the remaining facilities constitute the sampling frame.
Selection bias	"The introduction of bias into the results of a study because those selected differ from those not selected in some systematic way" *[5]*. Selection bias occurs when some subgroups of the sampling frame have a higher or lower chance of selection than others.
Target group	The target group is the group (or "population") of interest that the sample aims to represent. In the HHFA, the target group consists of all eligible health facilities from which a sample is selected.
Variance	"A measure of the variation shown by a set of observations" *[2]*. The variance of an estimate measures how far data are spread out, or the differences that occur between measurements, some of which may be explained by known factors, and the remainder attributed to chance.

A2.2 Selecting a sampling methodology

Any survey, including the HHFA, aims to produce results that are relevant to the needs of the users, accurate, and credible (giving users confidence that results are accurate). The survey sample should provide an accurate representation of the target group,[1] with an acceptable level of precision, to enable the users to generalize the results from the sample to the target.

Selecting the sample for a facility survey and determining its size is a complex subject, which will vary considerably from case to case depending on the desired precision and type of estimates, the number of facilities in the country, the available resources, the timeframe for producing the required information, and the specific objectives of the assessment. For example, an HHFA conducted to produce district-level estimates will require a much larger sample size (to achieve estimates with acceptable precision) than if only national estimates are required.

A representative sample is a sample that to a large extent resembles the target group of health facilities. The findings based on a representative sample can then be applied ("generalized") to the target group. For the HHFA, a representative sample of facilities has the following characteristics:

- It is selected from a sampling frame that is derived from an MFL that is as complete and up to date as possible. Ineligible facilities are excluded from the MFL to produce the target group of facilities that form the sampling frame.
- Probability sampling is used. This is a sampling methodology that uses some form of probabilistic selection from the target group, where each unit (facility) of the target group has a known, non-zero chance of selection. Probability sampling reduces the risk of errors (selection bias) in the estimation of indicators.
- Mechanisms to ensure that key subgroups of the target group, (e.g. hospitals, private health facilities, regions, etc.), which may differ in important ways, are adequately included in the sample. These subgroups are referred to as "strata" (refer to Section A2.3 Step 4. Determination of strata).
- The sample is large enough to minimize the uncertainty (imprecision) around the estimates, which is due to random (chance) variation in the sample composition. The larger the sample, the smaller the chances of producing results that are imprecise.

Most facility surveys use one of the following sampling methodologies: "stratified random sampling" or "cluster sampling". Table A2.2 summarizes the main advantages and disadvantages of each methodology. **Stratified random sampling is the recommended and most commonly used methodology for the HHFA.**

Costs are a critical consideration when determining the HHFA sampling methodology. Table 7 in Section 6.5 provides a summary of different HHFA sampling options with estimated costs.

[1] "Target group" (e.g. the "universe" of health facilities) is used here to designate the statistical term "target population", which might be confusing in the HHFA context, since the objects of the assessment are health facilities and not population members.

Table A2.2. Advantages and disadvantages of facility survey sampling methodologies

Sampling design	Stratified random sampling	Cluster sampling
Advantages	■ Sample selection is relatively simple if a complete and updated MFL is available. ■ Sampling error is lower than in simple random sampling, enabling a smaller sample size. ■ Allows grouping according to key subcategories (strata) within the target group and obtains separate estimates for each stratum. ■ Ensures representation of the target group's key subgroups (e.g. hospitals, private facilities, etc.). ■ Sample size per stratum can be controlled to maximize precision within the available budget, for example in disproportionate stratification. ■ Enables comparisons between strata when there are sufficient numbers in each stratum to achieve acceptable precision. ■ There are no "cluster effects" (refer to Section A2.4).	■ Reduces costs and duration of the survey by reducing travel time in contexts where facilities are scattered across the country with large distances between them.
Disadvantages	■ If the allocation of samples within strata is disproportionate (i.e. strata sample sizes are not proportional to the strata sizes of the sampling frame), the analysis is more complex, as sample weights need to be used. ■ Information on stratification variables is required for each health facility.	■ A larger sample size is required than in stratified random sampling, due to the "cluster effects", due to the intra-cluster correlation (the fact that units in the same cluster tend to be homogeneous); a correction factor, the "design effect" is used to adjust the sample size.

A2.3 Stratified random sampling

Stratified random sampling allows estimates to be obtained separately for desired, mutually exclusive, subgroups of facilities, called **strata.** A particular type of subgrouping of strata is called a **"stratifier" (or stratifier variable). Stratifiers commonly used in the HHFA** are**: facility type, managing authority, urban/rural location and geographic area (e.g. region).** Stratified random sampling consists of independent selection, through simple random or systematic random sampling procedures[1] of facilities within each pre-defined stratum within a stratifier. Independent estimates are obtained for each stratum. These findings are then combined in order to obtain the aggregate, national estimates.

Selection of a nationally representative sample stratified by facility type and managing authority will generally involve the following six steps:

Stage 1: Verification, updating or establishment of the national MFL

Stage 2: Determination of eligible health facilities

[1] Refer to Step 6 following for a description of random and systematic sampling.

Stage 3: Construction of the sampling frame (from which the sample will be selected)

Stage 4: Determination of strata

Stage 5: Determination of sample size

Stage 6: Selection of the sample.

Step 1. Verification, updating or establishment of the national MFL

This is an essential early step in the HHFA planning process, regardless of whether a census of all facilities will be conducted, or a representative sample will be selected. The updated MFL serves as the basis for determination of eligible facilities and for constructing the HHFA sampling frame.

Step 2. Determination of eligible health facilities

The next step is to determine the characteristics of the facilities that will form the target group from which the sampling frame will be constructed. The definition of the target group is a key decision in the HHFA. The sampling frame will consist of all facilities that meet defined eligibility criteria for inclusion in the HHFA and from which the sample will be selected. Examples of eligibility criteria include:

- Managing authority (public, private-for-profit, NGO, FBO, military, etc.):
  - e.g. will only public sector facilities be included, or also private sector facilities?
- Facility type (from PHC facilities at the lowest level to tertiary care hospitals at the highest level):
  - e.g. will health posts staffed only by community health workers be included?
- Geographical area:
  - e.g. will all regions/provinces of the country be included, or only a selected region or regions?
- Urban versus rural location:
  - e.g. will both urban and rural facilities be included?

Often, a combination of several such criteria is used to construct the sampling frame. For the HHFA, it is generally recommended to include in the target group all facility types and all managing authorities. However, specialized facilities such as eye hospitals, dental clinics, etc. may be excluded. Countries sometimes choose to include only public (ministry of health) facilities and private not-for-profit facilities if it is not possible to achieve a comprehensive and accurate list of private-for-profit facilities in the MFL.

Step 3. Construction of the sampling frame

After the ineligible facilities have been removed from the MFL, based on the defined criteria, the remaining facilities constitute the sampling frame. The sampling frame consists of all the facilities from which the sample will be selected. Whenever a sampling frame is constructed for any survey, some principles must be kept in mind: the sampling frame must be, in so far as is practicable, complete, accurate and up to date.

Step 4. Determination of strata

Once the sampling frame has been established, probability sampling principles are used to select the facilities that will be assessed. Usually, a stratified sampling plan is followed to ensure representation across various strata (subgroups) of the eligible facilities.

In stratified random sampling, the sampling frame is partitioned into strata based on the stratifiers and their pre-defined strata. Each stratum is then independently sampled, usually through simple or systematic random sampling. For example, within the facility type stratifier, three strata may have been defined: hospital, health centre and health post. A sample is then selected independently within each of these three strata. Usually, a combination of stratifiers is used (see below).

When the HHFA data are analysed, the results from the different strata are combined to form estimates that are representative for the entire target group.[1]

There are a number of advantages to using a stratified sample for the HHFA rather than a simple random sample of all facilities:

- A stratified sample guarantees that a defined number of facilities from each stratum will be assessed, whereas a simple random sample of all facilities might result in under-representation of certain types of facilities. For example, the number of hospitals in a country is generally small compared with the number of PHC facilities, and thus a simple random sample of all facilities in a country is likely to include only a very small number of hospitals or might miss them completely. By stratifying the sample by facility type, the number of hospitals and PHC facilities can be controlled to ensure that a sufficient number of hospitals are included in the sample.
- In a stratified sample, more precise estimates can be obtained in cases where facilities within each stratum are relatively homogeneous and the variation between strata is relatively large.

The recommended sampling methodology for the HHFA is to select all the hospitals in a country plus a simple random sample of the lower level facilities stratified by a combination of facility type, managing authority, urban-rural distribution and region.

There is a limit, however, to the number of strata that can feasibly be used. The more numerous the strata, the larger the sample will need to be, to the extent that it can approximate the total number of health facilities in the target group. Therefore, for example, if a country needs to obtain data disaggregated to lower subnational levels (e.g. district), it should consider conducting a census-based HHFA.

Stratified sampling can be of two main types:

- **Proportional stratification**, where the size of the sample from each stratum is proportional to the relative size of that stratum in the target group.
- **Disproportionate stratification**, where the strata sample sizes are not proportional to the sizes of target group strata. This is convenient when some subgroups (e.g. higher level health centres) are fewer in number compared with other types of health facilities, and it is important that they are adequately represented in the survey's sample, with adequate precision of the estimates. Smaller countries also typically require proportionally larger samples. Oversampling of the small strata will yield a sufficient number of health facilities to achieve acceptable precision.

[1] If it was not possible to obtain a complete list of facilities to include in the sampling frame for a stratum, it is recommended to not produce estimates for the entire target group, but rather to maintain separate estimates per stratum.

Disproportionate stratification is also useful when one of the survey's aims is to make comparisons between the stratum estimates (e.g. among regions), provided that the overall sample size is sufficiently large to achieve adequate precision in the strata. This comparative analysis is impossible when the number of facilities sampled varies largely across strata. Comparing a large sample in stratum X with a small sample in stratum Y would be incorrect, because the sampling error in Y would be larger than the sampling error in X[1] (see Box 4, Section 9.1 for a more detailed discussion).

In proportionate stratified sampling, the number of facilities allocated to the strata is proportional to their representation in the target group, as illustrated in Table A2.3.

Table A2.3. Proportionate stratified sampling: an example

Region	Target group		Proportionate stratified sample	
	Frequency	Percentage	Frequency	Percentage
A	612	51%	77	51%
B	60	5%	7	5%
C	528	44%	66	44%
Total	1200	100%	150	100%

It is clear in the example above that for stratum B, with a sample size of 7 facilities, the sampling error, and therefore the imprecision, will be much larger than the corresponding sampling errors for strata A and B. A comparison of the indicators could therefore, be misleading, due to the uneven precision of the estimates across the strata.

In order to enable a comparison of equally precise estimates would require sampling the regions disproportionally, as shown in Table A2.4. In disproportionate stratified sampling, the number of facilities allocated to the strata is not proportional to the representation in the target group.

In some cases, e.g. in small regions, or in strata with a small number of health facilities, a census of those particular strata may be the best option

Table A2.4. Disproportionate stratified sampling: an example

Region	Target group		Proportionate stratified sample	
	Frequency	Percentage	Frequency	Percentage
A	612	51%	50	33%
B	60	5%	50	33%
C	528	44%	50	33%
Total	1200	100%	150	100%

[1] There is an inverse relationship between sample size and sampling error: the larger the sample, the lower the sampling error and the higher the precision of the estimate.

In the example above, an equal number of health facilities was selected in each region, to maximize the sample size of each stratum.

In disproportionate stratification, sample weights need to be applied when analysing the data, to calibrate for national representation and obtain an overall estimate.

Sampling weights

In disproportionate stratification sampling, adjustments are used to correct for the unequal sampling probabilities among sample units (facilities). These adjustments are called **sampling weights.**[1] Failure to apply the sample weights in the analysis can lead to biased estimates of the indicators. Sampling weights are calculated and applied to the indicators in the analysis stage, after the data have been processed and cleaned. The weight is the inverse of the probability of selection of the sample units (facilities) by stratum; it is obtained by dividing the number of facilities in the sampling frame by the number of facilities in the sample for each stratum (refer to the example in Table A2.5).

The following information is needed to calculate sample weights:

- the stratifiers (i.e. region, facility type, managing authority, etc.)
- the strata within each stratifier
- the number of facilities in the sampling frame by stratum
- the number of facilities in the selected sample by stratum.

Table A2.5 provides an example of stratifiers, their strata ("stratification variables") and the calculation of weights

Table A2.5. Stratifiers, strata and weight calculation

A	B	C	D	E
Stratifier 1 (region): Stratification variable	Stratifier 2 (facility type): stratification variable	Number of facilities in the sampling frame	Number of facilities in the sample	Weights (Column D/ Column E)
Northern region	Hospital	9	9	1.00
	Health centre	129	25	5.16
	Health post	285	42	6.78
Southern region	Hospital	5	5	1.00
	Health centre	89	23	3.86
	Health post	124	28	4.42

[1] Other adjustments are needed for non-response, finite population or estimating changes from a previous assessment. Some of these adjustments also represent ways of weighting the findings, in order to correctly extrapolate them to the target group of facilities.

The weights that have been calculated need to be added to the final dataset. The HHFA data analysis platform automatically includes weights and adjusts the dataset, based on the information provided on the sampling frame, the stratification variables and the sample size.

The sampling weight adjustment allows the sample units (facilities) to contribute to estimates for the total target group of facilities in a correct proportion: when estimates are combined across strata, the weights compensate adequately for the unequal sampling probabilities. The weighting of variables must be documented in the HHFA report.

Step 5. Determination of sample size

Determining the sample size is a complex subject for any survey and **requires the technical assistance of a statistician.** The overall sample size for a facility survey will vary from country to country, depending upon the survey objectives, context, precision requirements, sampling methodology and available resources. Since the sample size is a key component of the total HHFA cost, it is important to determine a sample size and a sample design that minimize costs, while at the same time ensuring the desired precision.

The sample size is inversely related to the variance of the estimated indicators. The larger the sample size, the greater the precision of the estimates. The smaller the sample size, the higher the variance and the lower the precision of the estimates. Because of this imprecision, if too small a sample size is chosen, the findings can be insufficiently robust to use, e.g. to inform decision-making. Comparisons among strata (e.g. regions or provinces) can also be incorrect when the sample size was not calculated in order to allow disaggregated analysis (refer to Box 4, Section 9.1). Conversely, too large a survey sample will result in a waste of resources and time to complete the data collection. The higher precision of the larger sample will be at the expense of high financial costs, substantial opportunity costs for the facility staff, long duration of the data collection, and the risk of a lower quality survey, if insufficient supervision is available.

Given a desired level of precision and confidence interval, it is possible to determine the necessary sample size using standard mathematical formulas based on probability theories, assuming that some reasonable assumptions about the unknown parameters (the estimates of key indicators) can be made.

Various statistical packages (SPSS,[1] SAS,[2] Stata[3] etc.) have functions for calculating sample size.

The HHFA produces hundreds of estimates, each of which would require a different sample size according to the sample size formulas. In such cases, it is customary to choose a small number of the most important indicators and their assumed estimates, calculate the sample size requirements for each of these, and then to choose, conservatively, the largest one. For example, five indicators are selected as the most important; their assumed estimates are provided, along with their estimated proportions in the target group, desired level of confidence (usually 95%), and precision; the indicator requiring the largest sample size among them will determine the survey sample size.

In addition, the sampling design, the design effect (when cluster sampling is used), as well as the available resources need to be factored in, to adjust the sample size. Finally, adjustments for anticipated non-response, for estimating change, and for “finite population correction” need also to be applied, to obtain the final sample size.

Adjusting sample size for non-response

Non-response refers to facilities that exist and are functioning, but for which no data could be collected. The sample size calculation assumes that data will be collected for all facilities in the sample. However,

[1] Statistical Package for Social Sciences (https://www.ibm.com/products/spss-statistics).

[2] Statistical Analysis System (https://www.sas.com/en_us/home.html).

[3] https://www.stata.com

in practice, a 100% response is rarely attainable (e.g. some facilities may refuse to participate, or cannot be accessed due to weather or security constraints). Non-response, particularly when high, can affect the quality of the survey, by increasing imprecision and introducing a bias (systematic error). If the cause of non-response is related to key indicators, the bias can distort the results. For example, if there is high non-response in hard-to-reach areas that also have difficulties in receiving supplies of medicine and regular supervision, these areas may erroneously appear better off, since a high number of low-performing facilities could not be assessed and, therefore, could not contribute to the estimation of the indicators.

The sample size should therefore be increased by a proportion reflecting the anticipated non-response rate. For example, if the non-response rate is estimated at 10%, the sample size should be increased by 10% (oversampling).

This adjustment, however, does not rule out the risk of non-response bias. A high non-response should be considered carefully in the analysis, since it can negatively affect the accuracy of findings.

Sampling to estimate change

Facility surveys often aim to monitor changes in indicators over time to assess if they have improved, declined, or remained stable since previous, comparable surveys. The need to estimate change has implications for survey operations and sampling methodology and requires an increased sample size.

When making decisions for selecting the sample for a repeat HHFA in a country, three methodologies may be considered:

- Use of the same sample of facilities on each occasion, provided that there are no large changes in the number and distribution of health facilities (e.g. in post-conflict situations, when many health facilities are destroyed or new health facilities are built).
- Use of rotating or replacement groups of facilities.[1]
- Use of new, different samples each time.

Refer to Chapter 6 of the Measure Evaluation manual for additional technical details on this subject.

Adjusting the sample size for “finite population correction”

When the sampling frame contains relatively few facilities (e.g. in a small country, or when the HHFA covers only one region of the country), the sample size can be a significant proportion (e.g. 5% or more) of the total number of facilities in the sampling frame. In this case, the calculated sample size (n) should be reduced by a correction factor, 1-(n/N), where N is the total number of facilities in the country or the region. For example, if N = 2145 and n = 175 (with the sample size representing 8.25% of the target group), the correction factor to reduce the sample size will be: 1-(175/2145), i.e. 0.92; the sample size will then be 175 x 0.92, i.e. 161.

Adjusting the sample size in stratified sampling

As discussed in Step 5, adjustment of the sample size is needed when the facilities in some subgroups (e.g. higher level health centres or health facilities in urban areas) are fewer in number compared with those in other subgroups.

[1] This refers to using the same facilities as previously used as part of the sample, but replacing some of the previous facilities with new facilities.

Step 6. Selection of the sample

Stratified sampling

Once the stratification and the sample size have been determined, the final step is to select the sample of facilities to be assessed from within the sampling frame.

Selecting the sample of facilities from the sampling frame using random sampling

Two procedures are commonly used:

- **Simple random sampling**, in which all health facilities in the sampling frame have the same probability of selection. Each selection is made from the sampling frame, excluding those facilities already selected. Each facility is assigned a number.

Using a random number generator (e.g. an online tool or the RAND function in Microsoft Excel), a sample is selected by extracting random numbers until achieving the required sample size.

- **Systematic sampling**, in which a sample is selected by taking every kth facility in the sampling frame target group (where kth is ratio of target group size to sample size, a constant interval). Starting with a randomly selected facility between 1 and k, every kth facility is selected. For example, if the target group consists of 120 facilities (randomly ordered and numbered 1 to 120) and the sample size is 10, the interval is 12. A number is randomly selected between 1 and 12 (e.g. 2). The facilities selected will be those numbered 2, 2+12=14, 14+12=26, etc. and the sample will consist of the facilities numbered 2, 14, 26, 38, 50, 62, 74, 86, 98 and 110.

However, if the sampling interval is related to a periodic/cyclical pattern within the sampling frame (i.e. the facilities are not randomly ordered), a bias may result. For example, if the list of facilities in the sampling frame is ordered by type (hospitals, large health centres, etc.), the facilities that are adjacent to each other have a low probability of being selected. In the previous example, if the hospitals in the list are numbered 1 to 13, only one of them will be selected.

Therefore, in this case, systematic sampling does not generate a truly random sample. In order to avoid this bias, the sampling frame list should be sorted randomly before the sample is selected.

Replacement facilities

In any facility survey there will be sampled facilities that cannot be assessed because they are not functional or do not exist (e.g. have closed or relocated). This information was not known at the time of constructing the sampling frame. These facilities therefore erroneously remained in the sampling frame and were included in the sample. The facilities therefore need to be replaced in the sample.

The replacement facilities should be selected in the same way as the sampled facilities were selected. There are various options of identifying replacement facilities:

- During the original sampling process, within each stratum, select an additional number of facilities (e.g. 5%) that can be used as replacements if needed.
- After the facilities that cannot be assessed have been identified during the data collection process, identify the next 10 facilities listed after these facilities in each stratum of the sampling frame. Then select randomly the required number of replacement facilities.
- Oversample during the original sampling process, i.e. increase the sample size by an estimated percentage.

A2.4 Other sampling designs

Cluster/area sampling

Clusters are groupings of facilities that are geographically or otherwise related. In the former case, the terms cluster and area sampling are used interchangeably. Cluster sampling is used when facilities are scattered across large areas of the country, in order to reduce the distances between the sampled facilities, and hence reduce travel time and data collection costs. The approach can be used in very large countries or countries where travelling is time-consuming for other reasons.

Cluster sampling for the HHFA usually requires two stages. In the first stage, geographical areas (e.g. provinces, districts) are randomly sampled. These areas are called primary sampling units (PSU). Then, a random sampling procedure is used to sample the geographical areas, except that the names of the areas are selected, instead of the facility names. In the next stage, facilities are sampled from within each PSU, using the same random sampling procedure. Information on the PSU to which a facility belongs, will generally be available from the MFL. If the approximate number of facilities in a PSU (e.g. province or district) is known, sampling proportional to the number of facilities in each PSU is recommended. This means that facilities located within PSUs containing many facilities will have a higher probability of being sampled.

Cluster random sampling reduces the costs and time required for the survey, but the disadvantage is increased sampling error compared with a simple random sample of the same size, because of intra-cluster correlation: in each cluster, facilities located close to each other tend to have similar characteristics. This is called the “design effect” and can be adjusted for by increasing the sample size by the design effect factor. It is strongly recommended to seek the advice of a statistician for the determination of the sample size, which is a complex subject that includes the number of PSUs, the number of facilities in each PSU, and the selection procedures for each PSU, as well as the estimate of the “design effect”.

Blend of list and cluster sampling

Details on the blend of the list and area sampling methodology can be found in the *Sampling manual for facility surveys for population, maternal health, child health and STD programs in developing countries* [1].

A2.5 HHFA quality: potential design errors

The HHFA is based on a robust methodology, is supported by detailed guidance and has been tested in several countries. However, as with any survey, errors can occur and can impact the quality of the HHFA. Errors can be minimized by sound design of the survey and thorough planning and preparation, within the existing cost and time constraints. However, they cannot be completely eliminated: no survey can achieve perfect results. It is important, therefore, to understand the limitations of each HHFA and how these limitations can affect the results and their use, and to document this information in the survey report.

The following section describes common survey errors and ways of minimizing their occurrence. The errors are considered in two groups: errors related to sampling, and non-sampling errors.

Errors related to sampling

There are two types of errors related to sampling:

- **Sampling error**, which results from the fact that a sample does not always reflect the target group’s true characteristics because of random (chance) variation in the sample composition, even if the sample is selected using correct procedures. Sampling error results from “luck of the draw” and can be minimized by increasing the sample size: the larger the sample, the smaller the chances of producing results that are imprecise. Conversely, a small sample size produces imprecise results, with wide confidence intervals.

- **Sampling bias**, which occurs when the facilities of a sample differ from the facilities of the target group in a **systematic way**. There are three types of sampling bias:
  - **Coverage bias** may involve undercoverage or overcoverage. Undercoverage occurs if some facilities of the target group are excluded because they are erroneously missing from the sampling frame. This may occur, for example, if the sampling frame was built on an outdated or inadequate MFL. Overcoverage occurs when some ineligible facilities, excluded by the sampling frame (e.g. pharmaceutical depots, small health posts staffed by community health workers, etc.), are included erroneously. The amount of bias depends on the proportion of the facilities erroneously missing or erroneously included, and whether their characteristics differ from those of the facilities in the target group. Ways to minimize coverage bias include ensuring that the sampling frame is based on a comprehensive, updated MFL.
  - **Selection bias** occurs when some subgroups of the sampling frame have a higher or lower chance of selection than others (e.g. if the facilities of some regions have a lower probability of being selected than those in the other regions and have different characteristics, such as weaker supply systems due to their inaccessibility). The sample's estimates will then not be representative of the target group (or cannot be generalized to the target group).
  - **Non-response bias** occurs when facilities that are included in the sample and are functional, cannot be assessed (e.g. refusal of the facility manager to participate in the HHFA, inability to access the facilities due to flooding). If the non-assessed facilities differ systematically from those that could be assessed, especially with respect to some of the factors under study, non-response bias will occur. The sample from which the information is collected will therefore not be the same as the sample originally selected; these differences can be important. Unlike sampling error, sampling bias cannot be controlled by increasing the sample size. Ways to minimise non-response bias include:
    - Obtaining the collaboration of district health managers, by informing them in advance of the survey objectives and processes, and the importance of collecting the information.
    - Contacting the health facilities in advance, informing them of the survey and its importance, the expected dates of the facility visits, and the type and duration of the data collection.
    - Ensuring that adequate logistics are available for the data collectors and supervisors to travel to hard-to-reach areas.
    - Ensuring that the data collection timeframe considers local weather patterns and other issues that could affect the data collectors' access to the facilities.

Non-sampling errors

Non-sampling errors refer to all sources of error that are not related to sampling, such as:

- **Interviewer bias**, related to the improper administration of the questionnaire.
- **Response bias**, related to the systematic inaccuracy of responses.
- **Coding bias**, related to inaccuracy in recording responses.
- **Interpretation bias**, related to drawing conclusions about the data based on wrong assumptions and theories, rather than on facts, e.g. stretching hypotheses into explanations, in the absence of robust evidence from the HHFA data and other sources of information (also called "cognitive biases", see Section 10.1).

References Annex 2

1. Turner AG, Angeles G, Tsui AO, Wilkinson M, Magnani R. Sampling Manual for Facility Surveys for Population, Maternal Health, Child Health and STD Programs in Developing Countries. MEASURE Evaluation Manual Series, No. 3. MEASURE Evaluation. Carolina Population Center, University of North Carolina at Chapel Hill. July 2001. Pages 59 and 60. (https://www.measureevaluation.org/resources/publications/ms-01-3/at_download/document, accessed 16 November 2022).

2. Porta M (ed.). A dictionary of epidemiology. 6th ed. New York: Oxford University Press; 2014.

3. Everitt BS, Skrondal A. The Cambridge Dictionary of Statistics, 4th edition. Cambridge: Cambridge University Press; 2010.

4. Adapted from: Groves RM, Fowler FJ, Couper MP et al. Survey methodology. Wiley-Interscience; 2004.

5. Webb P, Bain C. Essential epidemiology. Cambridge; Cambridge University Press; 2011.

Annex 3: HHFA job description examples

A3.1 HHFA Survey manager – example job description

Scope of work

Plan and coordinate the HHFA survey at the central (national) level

Responsibilities/tasks

- Contribute to the design of the survey
- Prepare the HHFA implementation plan and schedule
- Facilitate the overall coordination of the survey
- Recruit, brief and contribute to the training of survey personnel
- Supervise the procurement of equipment and materials
- Manage the survey budget
- Support the data manager(s) in the adaptation of the CSPro tool
- Oversee, plan and monitor the data collection process
- Lead and contribute to:
  - data quality assurance and data analysis processes
  - interpretation of results
  - preparation of a survey report
  - communication of findings
- Provide troubleshooting for unplanned events
- Ensure the engagement of relevant stakeholders from the inception to the dissemination phases of the HHFA process
- Contribute to the preparation of the HHFA report and other communication products and to their dissemination

Qualifications and skills

- Health professional (medical doctor, nurse, health technician), preferably with a postgraduate degree in public health, epidemiology or health systems
- Minimum of 5 years' working experience in a ministry of health department at central or provincial level (preferably PHC, planning, epidemiology, etc.), with deep knowledge of the country health system
- Professional experience in the assessment and monitoring of health services and systems
- Management experience in a ministry of health programme at central or subnational level
- Proven experience in conducting surveys
- Familiarity with basic statistics and data interpretation
- Familiarity with policy-making processes

Supervised by/reporting to:	Assisted by:	Supervises/coordinates with:
HHFA Technical Committee	Administrative assistant Logistics officer	Statistician, Data collector trainers, Area supervisors, Data managers, Data analysts, Report writers

A3.2 HHFA Statistician/Sampling expert – example job description

Scope of work

Provide technical guidance on the survey design

Responsibilities/tasks

- Provide technical guidance to decisions that must balance the HHFA objectives, the desired analysis (national or subnational level; groups of facilities constituting the strata) and the desired level of precision, against the availability of resources and the time constraints
- Provide technical guidance in designing a sample based HHFA, including:
  - construction of the sampling frame
  - the sampling methodology
  - determination of the sample size
  - adjustments to the sample size for expected non-response, estimation of change, finite target group size (if applicable), and design effect (in cluster sampling)
  - the selection of health facilities from the sampling frame
  - the calculation of sampling weights
- Contribute to the analysis and interpretation of the HHFA data

Qualifications and skills

- Statistician/sampling expert with at least 5 years' working experience
- Experience in the design of large, complex surveys
- Familiarity with the country context and health system
- Knowledge of the HHFA objectives and methods

Supervised by/reporting to:	**Supervises/coordinates with:**
HHFA Survey manager	HHFA Technical Committee

A3.3 HHFA Area supervisor – example job description

Scope of work

Responsible for all aspects of data collection in an assigned geographical area and for ensuring the accuracy and consistency of data collection

Responsibilities/tasks

- Oversee and coordinate the teams of data collectors in the assigned area
- In close consultation with the survey manager:
  - plan the work of data collectors, assigning them to the health facilities, and preparing a schedule of their deployment
  - facilitate access to health facilities to be assessed
  - ensure that the equipment for data collection is available and functioning
  - ensure that logistic means are effectively deployed to transport data collectors to assigned health facilities
  - monitor the process of data collection using the required CSPro functions
  - accompany data collector teams to facilities on a rotational basis to ensure that required processes are followed
  - troubleshoot when problems arise during data collection
  - validate data collection by repeating a subsection of the questionnaire in about 10% of the surveyed facilities

Qualifications and skills

- Health professional
- Previous experience in surveys/data collection
- Familiarity with medical terminology
- Deep knowledge of the HHFA questionnaire used in the survey
- Working knowledge of the HHFA CSPro tool
- Familiar with the geographical area of assignment and its health system

Supervised by/reporting to:	**Supervises:**	**Coordinates with:**
Survey manager	Data collectors	Data manager(s)

A3.4 HHFA Data manager – example job description

Scope of work

Ensure configuration of the CSPro tool to country needs, and that all HHFA data management procedures, from data collection to data processing, are in place and that data collectors know how to use the mobile devices for data collection

Responsibilities/tasks

- Install the HHFA CSPro application and configure and adapt it according to the country context and needs
- Set up the data synchronization method
- Configure the mobile devices for data collection and install the HHFA CSPro application
- Contribute to the adaption of the data collection training materials
- Facilitate training of data collectors and area supervisors on use of the HHFA CSPro application
- Assess the data for completeness and quality during and after data collection
- Solve problems related to data entry and synchronization during data collection
- Ensure data cleaning and create the final dataset
- Export the final dataset for analysis and archiving
- Configure and adapt the HHFA data analysis platform as needed
- Run the analyses in the platform and generate HHFA outputs
- Generate additional analyses as requested by the data analysis team

Qualifications and skills

- University degree in information technology
- Familiarity with CSPro
- Working experience in the health sector
- Good knowledge of the country health system
- Good facilitation/training skills

Supervised by/reporting to:	Supervises/coordinates with:
Survey manager	Data analysis team

A3.5 HHFA Data collector team leader – example job description

Scope of work

Lead a team of HHFA data collectors, ensure that data are collected according to quality standards, and ensure synchronization of data to the server

Responsibilities/tasks

In addition to the tasks of a Data collector:

- Prepare a schedule of facility visits for the data collection team and share it with the team members and area supervisor
- Fix an appointment with the assigned health facilities for the data collection visit
- For each facility, assign all the required sections of the questionnaire to the data collection team
- Introduce the team to the facility in-charge and to explain the purpose of the visit
- Create a complete facility record by combining the data files from all the data collectors on completion of the data collection for each facility
- Check that the data for the facility are complete and ensure the synchronization of the data to the server
- Contact the area supervisor concerning any problems encountered during data collection

Qualifications and skills

- Health professional (nurse, midwife, doctor, or medical student)
- Familiarity with the organization and functioning of health facilities and with medical terminology
- Familiarity with the local area and language
- Ideally, some previous experience in conducting surveys
- Availability to work full time for the duration of the data collection
- Understanding of HHFA questionnaire and competence in use of the HHFA CSPro application demonstrated by end of HHFA data collection workshop
- Capacity to lead a small team of data collectors in field conditions

Supervised by /reporting to:	Supervises/coordinates with:
Area supervisor	Team of data collectors

A3.6 HHFA Data collector – example job description

Scope of work

To collect high-quality HHFA data based on training received during the data collection training workshop

Responsibilities/tasks

- Collect the HHFA data in the designated health facilities, according to the questionnaire sections assigned to him/her, applying the interviewing skills learned during the data collection training, and using the mobile data collection device and HHFA CSPro application
- Record the GPS coordinates of each visited health facility
- Check that all the questionnaire sections assigned to him/her are completed (using a special feature of the CSPro application) and fill any gaps before leaving the health facility
- Transfer the completed questionnaire sections to the team leader according to the established data management procedures

Qualifications and skills

- Preferably health professional (nurse, midwife, doctor or medical student)
- Familiarity with the organization and functioning of health facilities and with medical terminology
- Familiarity with the local area and language
- Ideally, some previous experience in conducting surveys
- Availability to work full time for the duration of the data collection
- Understanding of HHFA questionnaire and competence in use of the HHFA CSPro application demonstrated by end of HHFA data collection workshop

Supervised by/reporting to:	Supervises/coordinates with:
Data collection team leader	Data collection team members

A3.7 HHFA Report writer – example job description

Scope of work

Based on a thorough understanding of the HHFA objectives and reporting requirements, ensure that relevant findings, conclusions and recommendations are effectively communicated in the HHFA report

Responsibilities/tasks

- As member of the data analysis team:
  - Contribute to the analysis and interpretation of HHFA data, with interactions with relevant experts and programme managers
  - Contribute to the selection of the most relevant and important findings
  - Contribute to the development of conclusions and recommendations
  - Draft the main HHFA report and summary report, aligned with the ministry of health standards and formats, as requested by the HHFA coordination group
  - Ensure that data and information are communicated in language and formats that are easily understood and appropriate to the users of the reports
  - Submit draft reports, revise them based on feedback received, and prepare final reports
  - Deliver oral presentations and other related tasks as assigned by the survey manager

Qualifications and skills

- University degree in a health or social science discipline
- Basic knowledge of statistical methods used in surveys
- Familiarity with the HHFA content and methodology
- Strong analytical, writing and reporting skills
- Excellent command of the country language used for official documents
- Working experience in the health sector, with good knowledge of the country health system
- Ability to work in a team
- Proven experience in delivering documents of high quality within the required timeframe
- Preferably, experience in developing survey reports

Supervised by/reporting to:	**Supervises/coordinates with:**
Survey manager	HHFA Data analysis team, Technical Committee

Annex 4: Training of data collectors, team leaders and area supervisors

Participants should include all personnel involved in HHFA data collection and supervision. All require training, regardless of whether they have attended any similar training in the past. A data collection training workshop is a critical part of the survey preparation process. This annex provides practical guidance on conducting a training workshop for data collectors, team leaders and area supervisors.

Importance of data quality

The quality of HHFA data must be ensured for several reasons:

- solid data support reliable conclusions and recommendations;
- future policy decisions and resource allocation may rely on the evidence generated by the survey;
- critics and opponents will tend to look for weaknesses in the survey methods and results;
- results will be publicly accessible and may be used by multiple stakeholders and researchers, e.g. in conducting international comparisons.

There are several potential reasons for HHFA data quality problems:

- data collectors, team leaders and area supervisors receive insufficient or poor-quality training;
- the field practice of the data collection training workshop is not conducted properly;
- data collection is of poor quality (misunderstanding of instructions, time pressure, insufficient supervision, no quality control for submission of completed forms, etc.);
- interviewers ask questions incorrectly and/or influence responses;
- data are entered incorrectly;
- data are not checked at every stage of the survey process;
- there are problems with uniquely identifying facilities;
- some facilities refuse to participate in the survey (non-response).

Data quality problems can be minimized by:

- carefully studying the survey tools, guidance documents and training materials at every step, and following instructions;
- selecting capable and reliable personnel and ensuring that they are well trained in the survey methodology;
- encouraging personnel to communicate openly about uncertainties in survey procedures and questionable data;
- double-checking all sections of the questionnaire for accuracy and completeness after each data collection visit, at the end of each day of fieldwork;
- ensuring adequate supervision of data collectors by area supervisors.

Thorough training of survey personnel is one of the most important ways of ensuring accurate data collection and good-quality data. Experience from previous surveys has shown that poor survey preparation, including inadequate training of survey personnel, results in onerous and time-consuming data checking

and cleaning that can significantly delay the survey's completion. It is therefore much more effective and efficient to conduct thorough training and to apply rigorous data collection methods than to attempt to correct data once they have already been collected.

Overview of the data collection training

The overall objective of the data collection training workshop is to provide data collectors, team leaders, and area supervisors with the knowledge, skills and attitudes required to carry out an HHFA survey in an accurate, reliable and ethical manner. The members of the survey coordination group should be invited to the introductory session of the training workshop to meet survey personnel and discuss the survey methodology.

Upon completion of the training, participants should:

- be familiar with the key aspects of the HHFA and how it is conducted;
- be able to define their roles and responsibilities in the survey, including specific tasks, timelines and reporting requirements;
- be able to explain the critical content required to do their job effectively and possess the skills required to undertake each of their activities;
- be aware of common issues that may arise during survey activities, and understand troubleshooting/ problem-solving strategies to address these issues;
- recognize the intrinsic value of good-quality data and be motivated to ensure data quality as part of their activities.

A comprehensive **HHFA data collection training package** is available, including a facilitator's guide, presentations, exercises and supporting materials.[1] The training materials were created as templates that should be adapted to the country context and the HHFA modules selected by the country for implementation. Adaptation may also be needed based on the level of experience of the target audience and the resources and time available for the workshop. The facilitator's guide is intended to provide the workshop facilitators with an overview of each session and key information that will assist during facilitation.

Preparing for the training workshop

Preparation of the training workshop can require substantial planning, time and effort. Workshop preparations should begin early in the survey development process and should run in parallel with other survey planning and preparation activities. In preparing the training, it is essential to ensure that there is an adequate budget to cover costs for the training venue, materials, transport, and a daily allowance and accommodation for participants.

Selecting a training venue

A training venue should be selected based on the following criteria:

- availability of a room of appropriate size for the number of participants;
- sufficient space in the main room or additional rooms to accommodate small-group work;

[1] The training package is available upon request to hhfa@who.int.

- adequate ventilation and cooling/heating as needed;
- means of darkening the room to ensure visibility of projected presentations;
- availability of essential technical resources (printer, photocopier, projectors for presentations, large screens, microphones, electricity, extension cords, etc. In a large room with a large number of participants, multiple projectors and screens may be needed to ensure that all participants can see the screen clearly);
- proximity to health facilities that can be surveyed during the field practice day;
- accessibility by routine modes of transport;
- on-site or nearby refreshments and accommodation for out-of-town participants;
- reasonable cost.

Scheduling the dates of the training workshop

The training workshop should be held as close as possible to the initiation of data collection, to ensure that data collectors and supervisors retain maximum recall of the questionnaire administration process.

The workshop should not be scheduled at a time when weather or other conditions may delay the initiation of the subsequent data collection. However, at least 1 week should be planned between the end of the workshop and the start of data collection, to allow time for any final updates to the questionnaire and CSPro tool. All survey personnel must attend the workshop and should be advised of the dates as early as possible. Invitations to attend the introductory session of the workshop should also be sent to survey coordination group and technical committee members.

Identifying the training facilitators

The data collection training workshop requires at least two main facilitators: one with a clinical background and another with a data management background. The data management facilitator should have a high level of CSPro expertise and should lead the sessions related to the CSPro tool. Some sessions may require the presence of both facilitators. As there is a large volume of questionnaire content, is it recommended that the main facilitators are supported by a number of co-facilitators. They may also assign specific sessions to co-facilitators. An OpenWHO HHFA Data collection TOT e-learning course is available to prepare the facilitators and to ensure that a common understanding of the questionnaire and data collection processes is communicated during the workshop.

Preparing materials for the workshop

The HHFA data collection training package includes a sample agenda and a list of required equipment and materials. The workshop agenda, sample PowerPoint presentations, exercises and data collectors test must be adapted to the country HHFA context. All facilitators must familiarize themselves thoroughly with the sections of the questionnaire related to their assigned sessions and with the related training materials, adapting the materials as needed.

Planning the field practice day

During the field practice day, each data collection team should visit at least one public health facility and one private health facility (if private facilities are included in the survey) and collect data according to the survey procedures. The participation of facilities in the field practice day should be secured well ahead of the training workshop. The appointments should be made in advance, avoiding peak periods when health facilities may be busy with patients, and reconfirmed before the field practice day.

Annex 5: Procedures during the health facility data collection visit

Several general procedures should be followed by the data collector team during a facility visit, as outlined below.

Locating and verifying the survey facility

The area supervisor has provided each data collection team with a list of the facilities in which they are responsible for conducting the survey. Every effort should be made to conduct the survey at each facility on this list. If, after contacting local authorities, the team cannot locate a facility on their list, or are not sure about whether the facility that they have found is the one identified on the list, the team leader contacts the area supervisor for advice. If a facility included in the list has closed, the facility need not be visited, but the area supervisor is informed, and a replacement facility will be assigned to the team. Finally, the team should not visit and survey any facility that is not included on the team's list, unless specifically approved in advance by the area supervisor and survey manager.

Validating the cover page of the questionnaire

Before starting data collection, the team leader should check that the identifying information in questionnaire section 1 (that is pre-filled in CSPro) is complete and correct. If there are any errors, the team leader informs the area supervisor at the end of the day.

GPS data collection

Upon arrival at the facility, the geographic coordinates part of the questionnaire is completed according to the instructions received during the data collection training. The GPS coordinates, which use a space-based satellite system, are used to precisely locate the geographic position of the facility.

Introductions and obtaining consent to survey the facility

Data collection teams will visit facilities operated by the government and, depending on the survey design, potentially also facilities operated by NGOs and other private health care providers. All facilities must give consent for the survey to be conducted on their premises.

The first contact at the facility should be made by the team leader asking to speak with the person in charge. If the official "in-charge" is not present on the day of the survey, the team leader should ask to speak with the person acting as "in-charge" for the day. The team leader will then:

- introduce the team
- explain the purpose of the visit, the activities that are part of the survey, and the estimated duration of the visit
- give the "in-charge" the letter of introduction provided by the ministry of health
- reassure facility staff that individual respondents will not be identified
- respond to any questions and concerns from the "in-charge" or other facility staff
- request consent to conduct the survey
- request the "in-charge" to sign the informed consent form provided by the ministry of health.

If the “in-charge” refuses consent for the team to conduct the survey, the team leader contacts the area supervisor and provides the name of the facility, its managing authority, location and the reason for refusal. The area supervisor will then attempt to contact appropriate persons that may be able to assist in obtaining consent to proceed with the survey in this facility.

Identifying facility staff to interview

An important objective of the survey is to obtain correct and consistent answers to the questions. As the questions relate to the facility and not to a specific person, the information can be obtained from a variety of respondents as long as they are knowledgeable about the topic. Data collectors may need to speak with a number of respondents in order to obtain complete and correct information.

The data collectors are responsible for working out a plan for completing all the questionnaire sections assigned to them at each facility. They should discuss this plan with the “in-charge”. It may be helpful to meet with relevant supervisors (at large facilities) and other staff who may be requested to facilitate interviews and observations during the team’s visit. For a small facility this may be relatively easy, since most services are in the same general area. For large facilities, this may involve visiting several departments.

Annex 6: Glossary

Term	Definition
Assessment	"A formal process of evaluation of a process or system, preferably quantitative, but sometimes necessarily qualitative" *[1]*.
Accuracy	"The degree to which a measurement or an estimate based on measurements represents the true value of the attribute that is being measured" *[2]*. Sometimes the term validity may be used, with the same meaning.
Audit	"An examination or review that establishes the extent to which a condition, process, or performance conforms to predetermined standards or criteria" *[2]*.
Bias	"Deviation of results or inferences from the truth, or processes leading to such deviation. An error in the conception and design of a study—or in the collection, analysis, interpretation, reporting, publication, or review of data—leading to results or conclusions that are systematically (as opposed to randomly) different from truth" *[2]*. Biases are systematic errors that can result in wrong findings and interpretations.
Census	A census is a study of every unit in a population. For the HHFA, this means inclusion of all the health facilities in the country in the survey.
Confidence interval	"A range of values, calculated from the sample observations, that is believed, with a particular probability, to contain the true parameter value. A 95% confidence interval, for example, implies that, were the estimation process repeated again and again, then 95% of the calculated intervals would be expected to contain the true parameter value" *[3]*. The confidence interval measures the degree of uncertainty, or certainty, around a sample estimate, with respect to the true value in the target group. It may also be described as a range of values around the point estimate that is likely to include (usually with 95% probability) the unknown target group value. The confidence interval is related to the precision of the estimate, and not to systematic errors (biases).
Coverage bias	Coverage bias occurs when some facilities of the target group are erroneously missing from the sampling frame because, for example, the sampling frame was built on an outdated or incomplete list of facilities. The amount of bias depends on the proportion of the facilities not covered and whether the characteristics of facilities not covered differ from those of the facilities in the target group.
Data	Raw material: facts and figures, not analysed. "A collection of items of information" *[2]*. Primary data: data collected ad hoc, first-hand for a specific purpose, from an original data source, through a survey, interview, etc. Secondary data: data already available, collected by someone else, often for other purposes than that of the user.
Design effect	The design effect for a cluster survey is the ratio of the variance for that design to the variance calculated from a simple random sample of the same size. It is an adjustment that should be used to determine the survey sample size in cluster sampling. It varies from survey to survey.
Errors (in surveys)	"Deviations from the true values applicable to the target group studied. Deviations of what is desired in the survey process from what is obtained" *[4]*. Errors in surveys are differences between the sample estimates and the true values of the target groups. There are various types of error, as described in Annex 2.
Evidence	"Any form of knowledge, including, but not confined to research, of sufficient quality to inform decisions" *[5]*.

Index or composite indicator	"A rating scale, e.g. a set of numbers derived from a series of observations of specified variables" *[2]*. It is used to summarize and communicate information about a selected group of indicators.
Indicator	"An attribute that can be used to measure and/or record an event, process, or phenomenon. Indicators are tools for quantifying, through direct or indirect measures, a significant aspect of a health issue" *[2]*.
Information	"Facts that have been arranged and/or transformed to provide the basis for interpretation and conversion into knowledge" *[2]*. Information includes any data that may inform understanding, presented in a context that gives it meaning.
Metadata	"A set of data that describes and gives information about other data" *[6]*. Metadata document information about the survey and serve for future reference and analysis.
Module	In the HHFA: "A set of questions that provide information about a main topical area, e.g. availability, readiness, etc."[1]
Non-response bias	Non-response bias occurs when facilities included in the sample cannot be surveyed (e.g. due to refusal of the facilities' managers to participate in the HHFA). If the non-surveyed facilities differ systematically from the other facilities in the sample, especially with respect to some of the factors under study, this will result in non-response bias.
Outlier	"Observations with values differing widely from the rest of the data. This may suggest that an error was committed in their measurement or recording, or that the values come from a population different from that giving rise to the bulk of the observations. Yet, the values may be valid and precise" *[2]*.
Precision	"The quality of being sharply defined through exact detail. Relative lack of random error. In statistics, the measure of precision is the inverse of the variance of a measurement or estimate" *[2]*. Precision is a measure of the uncertainty around the estimate produced in a sample, resulting from the fact that the inclusion of facilities in a sample is partly determined by chance: the composition of two or more samples (obtained using the same sampling procedure) will differ in part by chance. (In contrast, accuracy refers to how close a measurement in a sample is to the true value in the target group.) The larger the sample, the higher will be the precision of the estimates.
Probability sampling	Probability sampling is a sampling method by which every element in the target group (every health facility in HHFA) has a known, non-zero probability of being selected. If correctly applied, it enables selection biases to be avoided.
Reliability	"The extent to which repeated measurements on units yield similar results" *[3]*.
Representative sample	A representative sample is an unbiased sample, selected with probability methods, which to a large extent reflects the characteristics of the target group of interest. In the HHFA, the target group is all the eligible health facilities that constituted the sampling frame.
Sample	"A selected subset of a population. A sample may be random or non-random and may be representative or nonrepresentative" *[2]*. In HHFA, samples are selected randomly, using probability methods.

[1] Everitt BS, Skrondal A. The Cambridge Dictionary of Statistics, 4th edition. Cambridge: Cambridge University Press; 2010.

(Sample) weights	Sample weights are adjustments used to correct for the unequal sampling probabilities between units. For example, in disproportionate sampling, when the strata allocations in the sample are not proportional to the strata population sizes, weights must be used.
Sampling error	The error that results from the random (chance) variation in samples, even when they are selected using correct procedures. It can be controlled by increasing the sample size: the larger the sample size, the smaller the sampling error.
Sampling frame	The sampling frame is a listing of the target group from which the sample is selected. In the HHFA, it is based on the MFL; after ineligible facilities are excluded from the MFL, the remaining facilities constitute the sampling frame.
Selection bias	"The introduction of bias into the results of a study because those selected differ from those not selected in some systematic way" *[7]*. Selection bias occurs when some subgroups of the sampling frame have a higher or lower chance of selection than others.
Service readiness	In the HHFA, readiness measures the extent to which the resources and conditions are in place to provide services according to defined minimum standards, including the presence and functionality of basic amenities, trained staff, guidelines, equipment, diagnostic capacity, and medicines and commodities.
Stratifier and stratum	Stratified random sampling allows estimates to be obtained separately for desired, mutually exclusive subgroups of facilities called strata. A specific type of subgrouping is called a stratifier (or stratifier variable). Stratifiers commonly used in the HHFA are: facility type, managing authority, urban/rural location and geographic area (e.g. region). The facility type stratifier may, for example, include the following strata: referral hospital, district hospital, health centre and health post.
Survey	"An observational investigation, usually descriptive, in which information is systematically collected" *[7]*.
Target group	The target group is the group (or "population") of interest that the sample aims to represent. In the HHFA, the target group consists of all eligible health facilities (sampling frame) from which a sample is selected.
Triangulation	Triangulation is the practice of improving the confidence in findings and interpretation by contrasting and integrating information from diverse data sources and methods, possibly independent of each other.
Variance	"A measure of the variation shown by a set of observations" *[7]*. The variance of an estimate measures how far data are spread out, or the differences that occur between measurements, some of which may be explained by known factors, and the remainder attributed to chance.

References Annex 6

1. Last JM. A dictionary of public health. Oxford: Oxford University Press; 2007.
2. Porta M (ed.). A dictionary of epidemiology. 6th ed. New York: Oxford University Press; 2014.
3. Everitt BS, Skrondal A. The Cambridge Dictionary of Statistics, 4th edition. Cambridge: Cambridge University Press; 2010.
4. Adapted from: Groves RM, Fowler FJ, Couper MP et al. Survey methodology. Wiley-Interscience; 2004.
5. Buse K, Mays B, Walt G. Making health policy. Open University Press; 2005.
6. Oxford English Dictionary. Oxford University Press; 2022. (https://www.oed.com, accessed 16 November 2022).
7. Webb P, Bain C. Essential epidemiology. Cambridge University Press; 2011.

Department of Data and Analytics
20 Avenue Appia
CH-1211 Geneva 27
Switzerland

hhfa@who.int

9789240060302