



Revisiting Health Financing Technical Initiative Efficiency Collaborative

MEASURING HEALTH SYSTEM EFFICIENCY IN LOW- AND MIDDLE-INCOME COUNTRIES: A RESOURCE GUIDE





**JOINT
LEARNING
NETWORK**
For Universal Health Coverage

**Revisiting Health Financing
Technical Initiative
Efficiency Collaborative**

MEASURING HEALTH SYSTEM EFFICIENCY IN LOW- AND MIDDLE-INCOME COUNTRIES: A RESOURCE GUIDE

Resource Guide Disclaimer

This Resource Guide was produced by the Joint Learning Network for Universal Health Coverage (JLN), an innovative learning platform where practitioners and policy makers from around the globe co-develop global knowledge that focuses on the practical “how-to” of achieving universal health coverage. For questions or inquiries about this manual or other JLN activities, please contact JLN at jln@msh.org or Naina Ahluwalia at nahluwalia@worldbank.org.

The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of The World Bank, its Board of Executive Directors, or the governments they represent. The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of The World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries. Nothing herein shall constitute or be considered to be a limitation upon or waiver of the privileges and immunities of The World Bank, all of which are specifically reserved.

The findings, interpretations, and conclusions made by country contributors do not necessarily reflect the views of the organization, institutions or governments they represent.

Rights and Permissions

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0). To view a copy of this license, visit <https://creativecommons.org/licenses/by-sa/4.0/legalcode>.

The content in this document may be freely used and adapted in accordance with this license, provided it is accompanied by the following attribution:

Measuring Health System Efficiency in Low- and Middle-Income Countries: A Resource Guide, Copyright © 2020, Joint Learning Network for Universal Health Coverage, International Decision Support Initiative (IDSI), The World Bank Group.

Product and company names mentioned herein may be the trademarks of their respective owners.

Suggested citation

Hafez, R., ed. *Measuring Health System Efficiency in Low- and Middle-Income Countries: A Resource Guide*. Joint Learning Network for Universal Health Coverage, 2020

Acknowledgments

The Resource Guide was co-produced with the following country members of the JLN Efficiency Collaborative, and facilitated by the Joint Learning Network's Efficiency Collaborative technical team.

TECHNICAL EDITOR

Reem Hafez

COUNTRY CONTRIBUTORS:

Bangladesh

Ayesha Afroz Chowdhury, Ministry of Health & Family Welfare.

Subrata Paul, Ministry of Health & Family Welfare.

Ethiopia

Mideksa Adugna, Oromiya Regional Health Bureau.

Ghana

Kwabena Boakyee Boateng, Ghana Health Services.

Kwakye Kontor, Ministry of Health.

Titus Sorey, National Health Insurance Authority.

Vivian Addo-Cobbah, National Health Insurance Authority.

India

Kavita Singh, Ministry of Health and Family Welfare.

Indonesia

Ahmad Ansyori, Council of National Social Security (DJSN).

Eka Yoshida, Ministry of Health.

Pandu Harimurti, The World Bank.

Mongolia

B. Munkhtsetseg, Ministry of Health.

Nigeria

Bolaji Aduagba, Federal Ministry of Health.

Kayode A. Obasa, Ministry of Budget and National Planning.

Nneka Orji, Federal Ministry of Health.

Philippines

Andrea Margreth S. Ora-Corachea, Department of Health.

Lavinia A. Oliveros, Department of Health.

Martha de la Paz, Department of Health.

Paulus Magnus Bacud, PhilHealth.

Vietnam

Nguyen Khanh Phuong, Health Strategy and Policy Institute.

Special thanks are owed to the contributors of the two country pilots:

Kenya team: Ahmed Omar, Ministry of Health, Esther Wabuge, JLN CCG Coordinator, Isabel Maina, Ministry of Health, Kenneth Munge, KEMRI Wellcome Trust, Mercy Mwangangi, Ministry of Health

Malaysia team: Abdul Hakim Bin Abdul Rashid, Ministry of Health, Lee Kun Yun, Ministry of Health, Mohammad Najib Bin Baharuddin, Ministry of Health, Nor Izzah binti Hj. Ahmad Shauki, Ministry of Health, Yussni Binti Aris A. Haris, Ministry of Health

Other country contributors to the Guide include:

Md. Abul Bashar Sarker, Md. Anwar Sadat and Md. Sarwar Bari from Bangladesh; Genet Mulugeta Hirpesa, Eyerusalem Animut and Tseganeh Guracha, from Ethiopia; Daniel Adin Darko, Emmanuel Ankrah Odame, and Ernest Asiedu Konadu from Ghana; Alok Saxena and B K Datta from India; Alice Wangui, Anne Musuva, Andrew Mulwa, Cyrus Matheka and Joseph Githinji from Kenya; Noraryana Bt Hassan from Malaysia; A. Khishigbayar, Bayanjargal Ganbol, L. Munkhzul and Ts. Tsolmongerel from Mongolia; Okechukwu Okwudili, Oritseweyimi Ogbe, Shamsudeen Saad and Uchenna Eugene Ewelike from Nigeria; Arturo Alcantara and Israel Pargas from the Philippines; Nguyen Lan Huong from Vietnam

The Revisiting Health Financing Technical Initiative team would like to thank the Country Core Groups across JLN countries, as well as the JLN Steering Group for continued support for leveraging existing resources for health, as 'efficiency' is a very high technical priority for the JLN.

Valuable feedback was received by the Health Financing Global Solutions Area under the leadership of Christoph Kurowski and other members of the World Bank HNP Leadership team. The technical facilitation team wishes to express special thanks to David Wilson for chairing the peer review process and Owen Smith, David Evans and Laura Di Giorgio for their time and effort in reviewing the guide.

The Resource Guide could not have been co-produced without the financial cooperation of the Government of Japan, the Rockefeller Foundation, the Bill and Melinda Gates Foundation and Australian Aid. The International Decision Support Initiative (iDSI), represented through the Center for Global Development (CGD), were technical partners to the World Bank for facilitation of the Efficiency Collaborative. Management Sciences for Health (MSH) provided continued support and external perspective as the JLN Network Manager.

A special thanks also to the World Bank JLN focal points from Bangladesh, Ethiopia, Ghana, India, Indonesia, Kenya, Malaysia, Mongolia, Nigeria, Philippines and Vietnam in facilitating and encouraging this work.

The Guide technical facilitation team was guided by Reem Hafez from the World Bank and included Somil Nagpal, Danielle Bloom, Naina Ahluwalia and Lydia Ndebele. Melissa Ouellet provided research and data support. Dorothee Chen, Marvin Ploetz, Alessia Thiebaud, Moritz Piazzi, Federica Secci, Manuela Villar Uribe provided data and input to several fact sheets. Ruth Young and QUO Global provided editing, data visualization, layout and production support.

Genesis of the efficiency collaborative

The Joint Learning Network (JLN) is a global network connecting practitioners and policy makers from 31 countries around the globe to co-develop knowledge that focuses on the practical ‘how-to’ of universal health coverage reform in low- and middle-income country contexts.

Over a series of consultations, JLN members reported that one of the most common issues facing their health care systems was how to assess the efficient use of resources. As a result, the Efficiency Collaborative (EC) was launched in April 2017. Participants from 11 JLN member countries (Bangladesh, Ethiopia, Ghana, Indonesia, Kenya, Malaysia, Mongolia, Nigeria, Philippines, Sudan, Vietnam) decided on two strategic work streams:

- i) the measurement and information stream (MIS) which aims to provide a framework for identifying and measuring efficiency, and
- ii) the systematic priority setting stream (SPS) which aims to support countries in maximizing their stated health sector priorities within a given resource envelope.

There are three complementary products tied to these workstreams – [Health Priority Setting: A Practitioner's Handbook*](#) and a [Health Priority Setting and Resource Allocation Tool and Database**](#) under the SPS stream – and this guide, which introduces readers to the concept of efficiency and gives guidance on how to assess efficiency in a practical setting. It includes a list of indicators most often used for tracking health system performance and several ‘fact sheets’ in Annex 2 meant to inform practitioners on how best to use, visualize, and interpret each indicator.

The development of this guide and the final list of proposed indicators was the result of an iterative process between technical facilitators and country participants over four in-person meetings, two webinars, and individual discussions with four countries testing out the validity and feasibility of the proposed approach.

Health Priority Setting:
A Practitioner's Handbook



Health Priority Setting and Resources
Allocation Benchmarking Database



* <https://www.jointlearningnetwork.org/resources/health-priority-setting-a-practitioners-handbook/>.

** <https://www.jointlearningnetwork.org/resources/health-priority-setting-and-resource-allocation-tool/>

CONTENTS

Acknowledgments	iii
Genesis of the efficiency collaborative	iv
Introduction	I
Concepts and principles of efficiency	7
A framework for measuring efficiency: a benchmarking plus approach	9
Why this approach?	13
From indicators to decision making	15
Conclusion	23
References	24
Annex 1: Indicator lists reviewed	25
Annex 2: Indicator fact sheets	26

List of boxes, figures and tables

Boxes

Box A	What are the main sources of inefficiency in health?	1
Box B	What makes health spending particularly prone to inefficiency?	2
Box C	A hypothetical application	11
Box D	What potential sources are available to collect information for efficiency analysis?	16
Box E	Kenya pilot	18
	Exhibit E.1 Ratio of nurses per 100,000 population by county (latest year available)	19
	Exhibit E.2 Maternal health coverage indicators, %, (2014–2018)	19
	Exhibit E.3 Maternal service readiness index by county (2013)	19
	Exhibit E.4 Maternal service readiness index by ownership and tier of care (2013)	19
Box F	Malaysia pilot	20
	Exhibit F.1 Diabetes prevalence (%), 2017	21
	Exhibit F.2 Diabetes care cascade in Malaysia (absolute number and %), 2017	21
	Exhibit F.3 Adherence to clinical practice guidelines for diabetes mellitus in Malaysia (%), 2017	21

Figures

Figure 1	Ways to benchmark	8
Figure 2	Life Expectancy vs. Government Health Expenditure per capita (2014)	9
Figure 3	The results chain or theory of change	10
Figure 4	Using data to inform decision making	15

Tables

Table 1	Common indicators used to assess efficiency in hospital and pharmaceutical sub-systems	13
Table 2	Common indicators that can be used in an efficiency analysis	14

Introduction

Universal health coverage (UHC) will inevitably require governments to find additional resources for health; and efficiency is critical to achieving that objective. Health expenditure growth¹ has exceeded gross domestic product (GDP) growth in nearly every country in the world over the past two decades. While there are several ways in which additional fiscal space for health can be met,² the pressure to extract greater value for money from health spending is mounting as health systems increasingly absorb a larger share of government, employer, and household incomes. But perhaps more important than how much countries are spending on health is whether they are using their limited resources efficiently (World Bank, 2019). The 2010 World Health Report highlighted that between 20–40% of all resources spent on health are wasted. It identified human resources, hospitals, and pharmaceuticals as the biggest sources of inefficiency (Box A) which result in potential savings of ~USD 1.2 billion³ globally from these sources alone (Chisholm & Evans, 2010; Couffinhal & Socha-Dietrich, 2017). A more comprehensive description of sources of inefficiency can be found [here](#).⁴

Box A: What are the main sources of inefficiency in health?

The main sources of inefficiency can be grouped into two types:

1. Inefficiencies arising from system-level resource allocation decisions – from poor planning or slow response to the changing health needs of the population:
 - Inappropriate or costly input/staff mix;
 - Inappropriate hospital size;
 - Sub-optimal deployment of health workers and facilities.
2. Inefficiencies that result from facility- or physician-level decisions linked to poor incentives or lack of accountability measures:
 - Inappropriate hospital admissions or length of stay;
 - Over-use of health care technology;
 - Sub-optimal quality of care and medical error;
 - Under-use of generic drugs;
 - Irrational use of drugs.

¹ Health expenditure growth is driven by rapidly ageing populations, growing burdens of noncommunicable and chronic diseases, technological progress, and rising population expectations.

² A parallel JLN Collaborative on Domestic Resource Mobilization for health covers four of these methods in greater detail. In short, they are: i) conducive macroeconomic conditions and increases in overall government revenue; ii) earmarking and specific commitment devices for health; iii) re-prioritization of health in the overall budget; and iv) access to health sector grants and donor aid.

³ Inflated to current 2019 USD from USD 965 billion in 2009.

⁴ <http://pubdocs.worldbank.org/en/616561494511379987/UHC-Efficiency-Background-Paper-V10-170419-1100.pdf>.

Some inefficiencies will be easier to address than others. Several characteristics of the health sector make health spending particularly prone to inefficiency. Uncertainty in the demand for health, informational asymmetries between patients and providers, difficulties in linking inputs to outcomes, and fragmented sources of financing often lead to lower actual spending on health than the allocated budget (Box B). Improving allocative efficiency is generally more difficult, as many stakeholders outside the Ministry of Health are involved and decisions are often constrained by civil servant regulations and budget rigidities.⁵ The SPS produced a Handbook and a Resource Allocation Benchmarking Database to help tackle these issues. In comparison, addressing inefficiencies that result from facility- or physician-level decisions are more straightforward as accountable entities are easier to identify and fall directly under the supervision of the Ministry of Health (for example, regional hospital managers, hospital managers, hospital departments, or individual practitioners). It is the latter type of inefficiency that this guide focuses on.

Box B: What makes health spending particularly prone to inefficiency?

There is a common perception amongst Ministries of Finance and Budget and Planning that health is more inefficient than other sectors. However, there are many factors that are beyond the control of the Ministry of Health that impact health system efficiency:

- i. **Uncertainty:** The demand for health services is unknown. It is difficult to predict when and where people get sick. This makes the planning and management of resources challenging. Public financial management (PFM) systems often restrict the ability to make mid-year adjustments to respond to the changing health needs of the population.
- ii. **Information asymmetry:** The average patient generally has less information than providers about the need for and quality of health care. This leads to complicated principal-agent relationships in which a patient (the principal) delegates decision making authority to a better-informed doctor (the agent). The assumption is that the doctor will act in the best interests of the patient, but this can sometimes be at odds with the doctor's own interest of wanting to maximize income. The end result is an increase in unnecessary (and total) health expenditures.
- iii. **Difficulty in linking inputs to outcomes:** Health outcomes are more difficult to observe than outcomes in other sectors (e.g., a road being built or the number of students completing secondary education). In health, budgeting is often linked to inputs such as human resources or infrastructure (e.g., hospitals) and not to what is being purchased or delivered (e.g., health care services or treatment) and even less to health outcomes (e.g., health status or financial protection). Budget format and PFM rigidities also mean there is little flexibility to move funds around between line items.
- iv. **Fragmented financing:** The health sector in low- and lower-middle income countries is marked by the presence of several specific programs (such as HIV/AIDS, malaria, tuberculosis, immunization, and family planning). Many of these programs are funded through external assistance. Often, this means that there are multiple donors present in the sector; health funds do not always flow through government budgets; and there is a lack of clarity on the full resource envelope. This not only poses a challenge for budgeting and planning but also leads to the duplication and waste of resources, the misalignment of resources with country priorities, and the use of parallel procurement, financial management, and monitoring and evaluation systems – placing a significant burden (financial and time) on the Ministry of Health.

This compromises Ministry of Health officials' ability to argue for additional resources for the sector. The JLN's Domestic Resource Mobilization Collaborative is producing a forthcoming messaging guide on how to make the case for health. It aims to facilitate discussion between Ministry of Health and Ministry of Finance officials highlighting the particularities of the health sector.

⁵ Stakeholders involved in the priority setting or benefit package selection process may include Ministries of Finance, Ministries of Health, health insurance agencies, market authorization and quality control agencies for drugs, national procurement agencies, professional medical associations, health technology agencies or academia, pharmaceutical lobbies, industry, the media, and the public.

While defining efficiency is relatively simple, measuring efficiency is challenging. The most basic definition of efficiency is maximizing outcomes relative to inputs. Therefore, the first key characteristic of efficiency analysis is that it requires information on both inputs and outputs. Efficiency analysis attempts to explain the unexplained variation across accountable entities – that is, why some individual providers, facilities, or health systems perform better than others. Therefore, a second key characteristic of efficiency analysis is that it involves making a comparison. The standard methods for economists to assess the efficiency of health care systems – stochastic frontier analysis (SFA) or data envelopment analysis (DEA) – empirically measure the productive efficiency of country health systems at the macro-level, or facilities at the micro-level. They take a production function approach, modeling the maximum attainable output achieved given a set of inputs. Yet, they struggle to produce actionable policy recommendations on how to improve efficiency. There are notable challenges to this approach. First, efficiency scores and rankings relative to the best performers are highly sensitive to model specification and the data used. Second, the results rarely help policy makers identify how they can improve efficiency, as efficiency scores – on their own – do little to tell a policy makers why the health system is inefficient.⁶ Finally, they are data intensive, time consuming, and require particular expertise. In low- and middle-income countries SFA and DEA are almost always carried out by external consultants as a one-off exercise, which does not help institutionalize routine monitoring of the efficiency of health care resources.

This guide proposes an alternative ‘benchmarking plus’ approach that can be readily used by practitioners and policy makers – the primary audience of the Efficiency Collaborative. The focus is less on empirical measurement and more on enabling the routine assessment of health system performance from an efficiency perspective. Benchmarking is simple to perform and easy to interpret; but on its own, it too will not tell policy makers how to improve efficiency. To be able to pinpoint the cause of variation, it will almost always be necessary to look at several indicators along the process of how inputs get transformed into particular outcomes and complement that with additional contextual knowledge in order to identify the appropriate policy action.

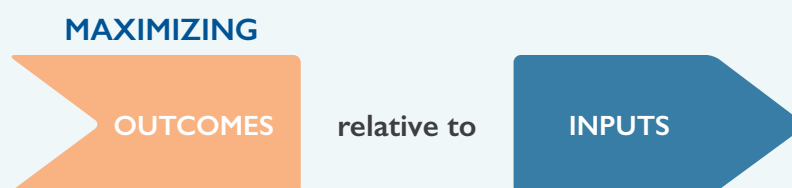
This guide gives a brief overview of concepts and principles of efficiency, and provides a framework for identifying and measuring efficiency in a practical way. It provides a list of indicators most often used for tracking health system performance and gives guidance on how they can be used to measure efficiency. It does not try to reiterate the work of the more comprehensive references from which it draws (World Bank, 2017) (Smith, 2012) (Cylus, et al., 2016) (Papanicolas & Smith, 2017) but rather attempts to operationalize the approach first suggested by (Smith, 2012)⁷ for low- and middle-income country settings.

⁶ Health systems have many goals – be it maximizing health status, financial protection, and/or patient satisfaction. There are also many sub-systems or sectors that contribute to the efficiency of the overall health system – such as preventive and promotive health, primary health care, the hospital sector, and the procurement, distribution, and use of pharmaceuticals and medical equipment. One part of the system may operate better than others.

⁷ Smith describes a “partial indicator” approach, which suggests benchmarking indicators across the entire results chain from inputs, to outputs, and outcomes.

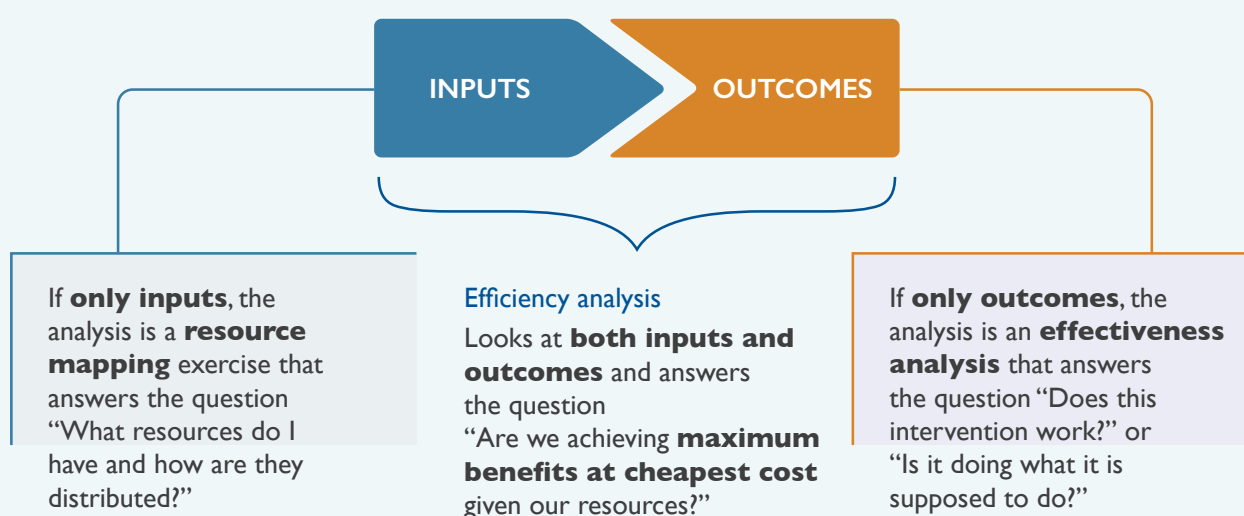
A BENCHMARKING PLUS APPROACH TO EFFICIENCY ANALYSIS

The most basic **definition of efficiency** is



There are **two key characteristics** of efficiency analysis:

1. Efficiency requires information on both inputs and outcomes



2. Efficiency involves making a comparison

Efficiency analysis attempts to explain the unexplained variation across accountable entities – that is why some individual providers, facilities, or health systems perform better than others.



Benchmarking makes a comparison based on average performance, relative to the best performer, relative to a clinical norm or target, or relative to past performance.

A BENCHMARKING PLUS APPROACH TO EFFICIENCY ANALYSIS (CONTINUED)

Inefficiencies can occur at any point in the results chain which transforms inputs into outcomes



The 'plus' in benchmarking plus...

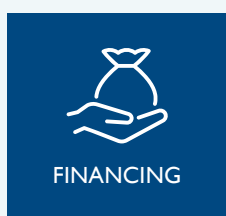
Benchmarking is helpful at detecting emerging issues such as stagnating outcomes, outliers, or significant changes from one period to the next. However, on its own, it is not enough.



As specific issues emerge, additional indicators and key informant interviews may be necessary to better understand whether an issue is truly an inefficiency or not.

...allows more actionable policy recommendations

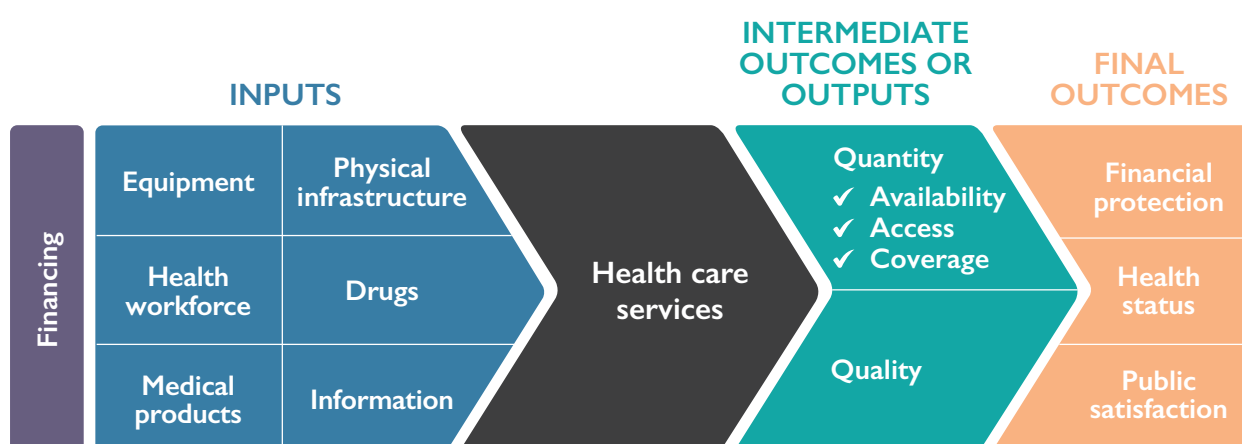
Policy makers have a range of policy levers to influence the behavior of accountable entities. Together, they affect how resources are allocated to different goods and services and how inputs are used to produce a given set of goods and services.



Concepts and principles of efficiency

The most basic definition of efficiency is maximizing outcomes relative to inputs.

- **Inputs:** Most often, people use inputs to refer to the costs, resources, or investments used to buy or produce health care inputs; and while financing for the health care system is one of the most fundamental inputs, others include health workforce (e.g., doctors, nurses, midwives, community extension workers); physical infrastructure (e.g., health care facilities, medical supply stores); drugs and medical products, equipment (e.g., MRI machines), and information – often the most overlooked input (e.g., data on civil registration and vital statistics, disease-specific registries, patient reported health outcomes, drug stocks).
- **Outcomes** can refer to the consequences, effectiveness, or benefits of service delivery interventions. In most health care systems around the world, the main outcomes of interest concern health status, financial risk protection, and public satisfaction. However, in practice, most efficiency metrics use intermediate outcomes or outputs. Outputs include information on the quantity (e.g., availability, access, coverage) and quality (e.g., diagnostic accuracy, treatment success rates) of the goods and services provided.



A key characteristic of efficiency analysis is that it requires information on both inputs and outcomes.

If only inputs are examined, the assessment is essentially a resource mapping exercise. The question being asked is ‘what resources do I have and how are they distributed?’ If only outcomes are examined, the assessment is called an efficacy or effectiveness analysis.⁸ This asks the question ‘does this intervention work?’ or ‘is this intervention doing what it is supposed to do?’ Efficiency looks at both resources used (inputs) and effectiveness of interventions (outcomes), and assesses:

- i) whether the maximum benefit to society is reached given the current use of inputs (‘doing the right things’), and
- ii) whether the same benefits to society can be reached at cheaper cost (‘doing things in the right way’).

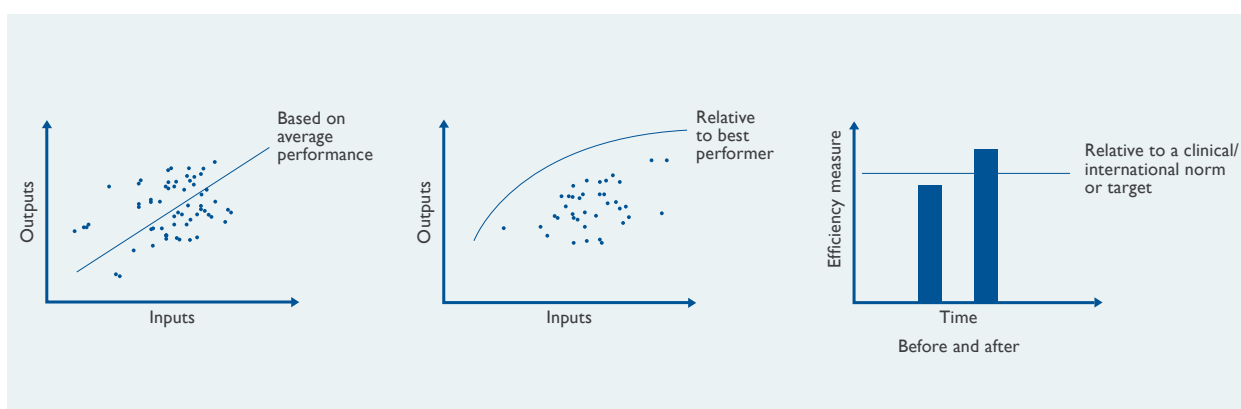
⁸ Efficacy refers to whether an intervention produces the expected result under ideal circumstances among a particular test group (such as a randomized control trial). Effectiveness refers to whether an intervention is beneficial to the wider public under ‘real world’ clinical settings.

Ideally, policy makers want to do the right things and do them in the right way. To improve efficiency, decision makers can either:

- Reorganize the current level of inputs (direct resources to different priorities), or adjust how services are delivered to increase benefits, which is sometimes referred to as allocative efficiency; or
- Find a way to produce the same amount of benefit but at a cheaper cost, which is sometimes referred to as technical efficiency.

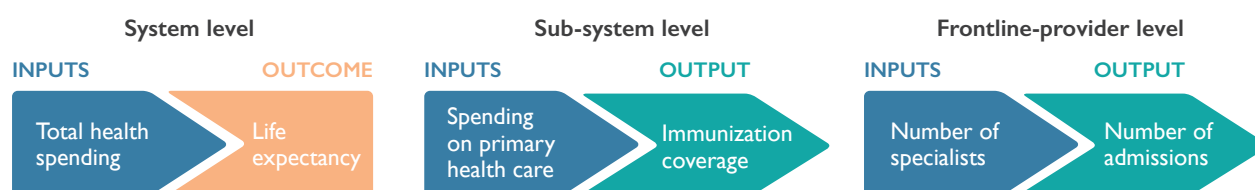
A second key characteristic of efficiency analysis is that it involves making a comparison. To identify whether resources are being used efficiently, there needs to be some standard by which a comparison can be made. Benchmarking makes comparisons based on average performance, relative to the best performer, relative to a clinical norm or target, or relative to past performance.

Figure 1: Ways to benchmark



A framework for measuring efficiency: a benchmarking plus approach

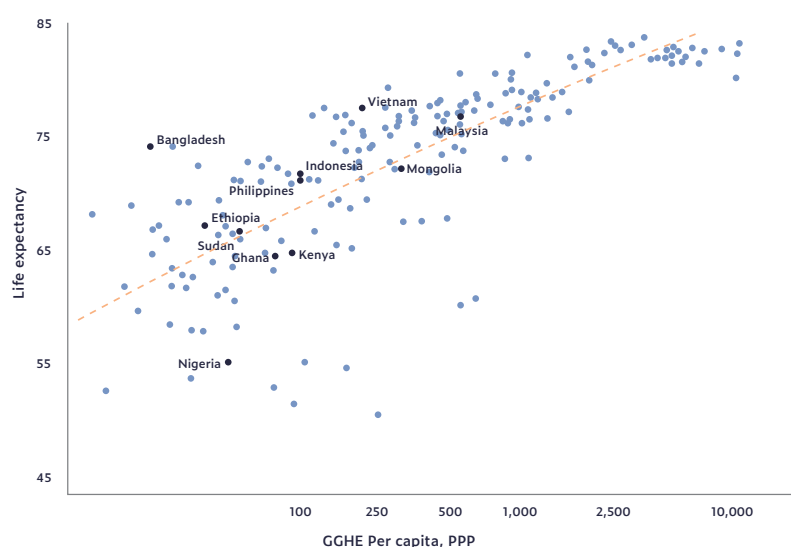
Typically, the first step to benchmarking is to define the scope of the analysis. At the system level, the focus is on how health sector investments are contributing to overall health system objectives. At the sub-system or frontline-provider level, the focus might be on whether spending on broad health categories improves intermediate health outcomes, or on how facility resources are increasing service utilization.



Defining the scope of the analysis will help inform the choice of comparator. This could be countries of similar socioeconomic status or level of health system development; geographic regions or population groups within a country; or individual facilities or providers. Internal benchmarking is likely the most relevant for countries as understanding distribution across regions, hospitals, clinics, and population groups can help policy makers and program managers target resources more effectively.

While benchmarking can help identify outliers or raise flags, looking at discrete indicators does not tell policy makers how to improve efficiency. Figure 2 assesses efficiency at the health system level by plotting government health spending versus the average life expectancy of several countries. Countries can assess their efficiency relative to the average (red line). However, there are a number of drawbacks to this approach. First, it focuses on a single input (government health spending) and a single outcome (life expectancy) but many other inputs contribute to life expectancy, and the health system may have many other important outcomes of interest. Second, it does not tell us why for the same level of expenditure, one country is performing better than another given the same level of government spending (for example, Nigeria vs. Sudan). Information on the processes required to translate inputs into outcomes is missing; this is sometimes referred to as the ‘black box’ of service delivery.

Figure 2: Life Expectancy vs Government Health Expenditure per capita (2014)



Source: (World Bank, 2019)

Inefficiencies can occur at any point in the results chain that transforms financing into outcomes. Instead of focusing on financing and final outcomes – the two ends of the results chain – it is more informative to benchmark indicators along an entire results chain to help pinpoint areas of inefficiency. Indicators that provide information on the management of resources, and on access to and quality of services are especially informative. Pinpointing areas of inefficiency will help narrow down the appropriate accountable entity for follow-up action – for example, regional hospital managers, hospital departments, or individual practitioners.

It is recommended to have discussions with key informants who can provide additional contextual knowledge and help get a sense of the relevance and extent of a particular source of inefficiency. They can help fill in the gaps when data is unavailable, incomplete, or unreliable, and be the starting conversation for thinking through appropriate policy actions and helping to determine which areas to tackle first. Addressing inefficiency is politically complicated, as there will be winners and losers in the process. This highlights the importance of having a transparent process around prioritization. For more information, see [here](#),⁹ and [here](#).¹⁰

Figure 3: The results chain or theory of change



Having identified the areas of inefficiency, policy makers have a range of policy levers to influence the behavior of accountable entities. We use the systems framework presented in listing financing, provider incentives, organization, and regulation as the key policy instruments to improve efficiency (Yip & Hafez, 2015).

- i. **Financing** is one of the most powerful levers for influencing the efficient delivery of care, especially through the design of the benefits package. For example, a common practice for health insurance schemes in South and East Asian countries has been to focus on low probability, high-cost inpatient care over high probability, low-cost outpatient care, reasoning that catastrophic and impoverishing health expenditures are more likely to occur from expensive hospital services. But by covering only hospital care, the incentive is to admit patients even for simple conditions. Instead, offering free health promotion activities and primary health care can prevent the development of more serious and costly conditions. Another example of a financing lever is tiered cost-sharing arrangements that incentivize the use of lower-level facilities over tertiary hospitals (such as higher co-payments at hospitals versus primary health care facilities).
- ii. **Provider incentives** – financial or otherwise – are another powerful policy lever that influences the efficiency (and quality) of care. For example, in prospective and bundled payments (such as capitation and case-based payments) greater risk is borne by the provider, incentivizing them to manage resources more efficiently. However, they may also incentivize other undesirable behaviors such as skimping on costs and avoiding chronic or sicker patients. For these reasons, most countries use a combination of payment methods to balance out the strengths and weaknesses of various methods. Non-financial provider incentives include relative social ranking, frequent interim feedback, and information availability. For more on provider payment methods please see [here](#).¹¹
- iii. **Organization** covers a broad set of policies for managing and coordinating the delivery system. For example, for clinical interventions and diagnostic services, purchasing from both the private and public sector reinforces competition, and generally leads to lower prices and improved quality where effective regulatory and monitoring capacity are in place. But for goods and services with low-profit margins (for example, public health interventions), the public sector is recommended to ensure the socially efficient level of provision as these types of services are labor intensive, have low-profit margins, and are generally underprovided by the private sector. Also, policies that encourage disease prevention, gatekeeping, coordinated care, and a shift from hospital-based to ambulatory care help promote the cost-effective use of care.
- iv. **Regulation** involves setting the rules, standards and operating guidelines within which the system is meant to operate. These may include medical and clinical protocols, national essential drug lists, procurement

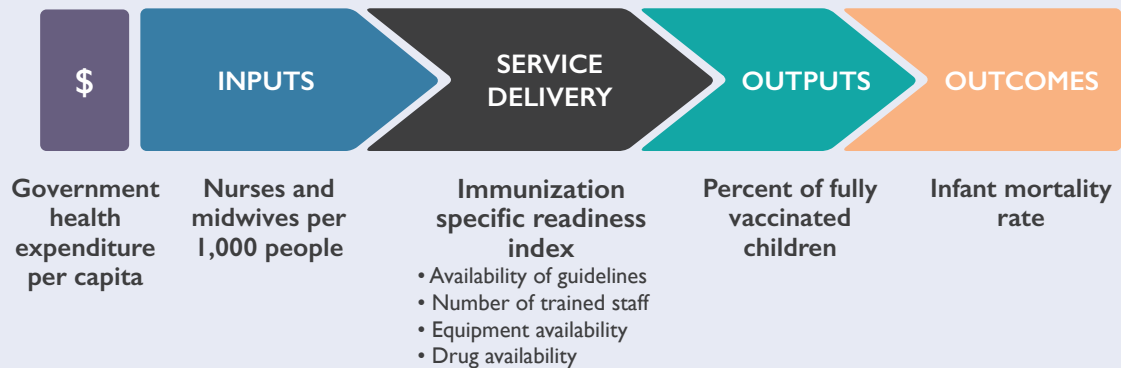
⁹ <https://www.jointlearningnetwork.org/resources/health-priority-setting-a-practitioners-handbook>

¹⁰ <https://www.jointlearningnetwork.org/resources/health-priority-setting-and-resource-allocation-tool>

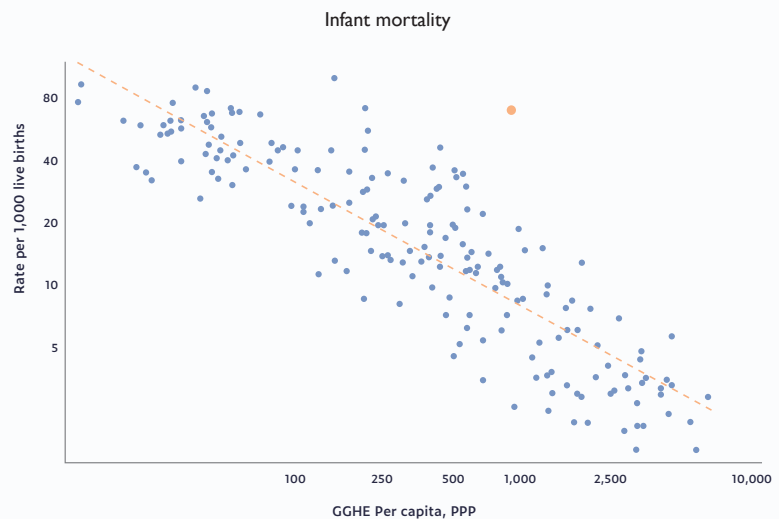
¹¹ <https://www.jointlearningnetwork.org/resources/assessing-health-provider-payment-systems-a-practical-guide-for-countries-w/>

Box C: A hypothetical application

By benchmarking several indicators across the entire results chain, JLNNotopia hopes to better understand why it underperforms on infant mortality compared to countries with similar government health spending.



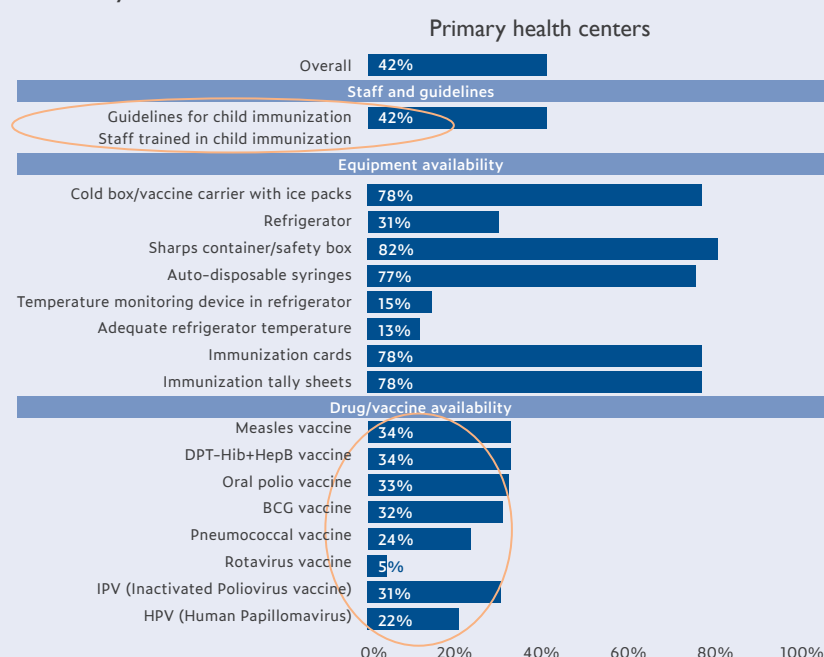
To narrow down the cause of high infant mortality, JLNNotopia needs to look at other indicators such as availability of providers and capacity to deliver immunization services.



The problem does not appear to be staff numbers.



However, facilities have on average only 42% of all service-specific readiness requirements for immunization. Low availability of drug and vaccines and a lack of staff trained in child immunization are likely impacting service delivery.



As a result, the number of fully vaccinated children is low, and the infant mortality is high.

	Children	
	% fully vaccinated*	IMR per 1,000 live births
Residence		
Urban	40.5	57
Rural	14.8	88
Wealth quintile		
Lowest	4.4	90
Highest	56.1	51
Total	25.7	70.3

Note: *BCG, measles, and three doses each of DPT and polio vaccine

By benchmarking several indicators across the entire results chain the analysis was able to pinpoint issues in the supply chain and in provider training. However, there is not enough information to know what specifically in the supply chain or training of providers is the issue. Taking supply chain as an example, the issue could be poor inventory management and capacity to forecast needed drugs, in which case the implementation of an inventory management or tracking system may be a solution. Or, the issue could be delays in the distribution of medicines, in which case outsourcing the distribution function to third parties might yield better outcomes. The 'plus' in the 'benchmarking plus' approach refers to the fact that it will almost always be likely to dig deeper. It might be necessary to look at additional indicators used to measure supply chain efficiency (for example, unit price of vaccines compared to international reference, time to process orders, number of days items were unavailable, etc.) and/or interview key informants at the national procurement agency, medical storage centers, and facilities to narrow down the right accountable entity for follow-up action. Only then can a conversation be had on how to reprioritize resources towards the right intervention based on a country's prioritization process.

guidelines for hospital equipment and drugs, quality control of drugs, manuals for claims management and audit (including fraud detection), certification and accreditation of facilities, licensing of the health workforce, price controls, and private sector service delivery.

For a more in-depth discussion of various policy instruments to address inefficiencies, see [here](#).¹²

¹² <http://pubdocs.worldbank.org/en/616561494511379987/UHC-Efficiency-Background-Paper-V10-170419-1100.pdf>

Why this approach?

The approach reflects the realities of many low- and middle-income countries attempting to measure health system efficiency. Every JLN deliverable is co-produced and driven by country participants over a series of face-to-face meetings. Technical experts help facilitate meetings to produce an end product that is based on topical expertise and country demand. The initial request was to produce a list of indicators to routinely assess the efficiency of health sector spending, especially in areas that are known to be major sources of inefficiency and consume the most resources, such as hospitals and pharmaceuticals (Table 1). However, many ‘true’ efficiency measures that combine inputs and outputs in a ratio, such as average consultations per doctor or unit cost per episode of care, are not readily available. Even if data is available, collection is difficult as datasets are often paper-based and housed in different units or administrative levels requiring a significant investment to standardize, aggregate, and clean data for analysis. In the absence of clear-cut efficiency metrics, benchmarking health system performance indicators along a results chain could help look at service delivery bottlenecks through an efficiency lens.

Table 1. Common indicators used to assess efficiency in pharmaceutical and hospital sub-systems

Pharmaceuticals	Hospitals
<ul style="list-style-type: none"> • Pharmaceutical spending as % of total health expenditure (THE) • Antibiotics spending as % of total pharmaceutical spending • Unit price of drugs/medical consumables • Unit price compared to international reference prices (especially for high-cost/use items) • Cost of freight/distribution to facilities • Order/use of high-cost items • High-use items • Number or % of expired items • Value of expired items • Stock-outs • Antibiotic prescription rates • Percent of encounters that end up in antibiotics being prescribed • Time to process orders • Time to pay suppliers • Drug availability • Rate of anti-microbial resistance 	<ul style="list-style-type: none"> • Spending by function (e.g., outpatient, inpatient, pharmaceutical, primary health care, public health or prevention, curative care) as % of government health expenditure (GGHE) • Hospitals per 100,000 population, hospital bed density, bed occupancy rate • General service readiness • Number of visits/admissions per day/month/year/per capita • Share of outpatient/inpatient • Diagnostic accuracy for tracer condition • Adherence to clinical guidelines • Number of incidents per 1,000 patient days (e.g., center line-associated bloodstream infections, standardized infection ratio) • Avoidable admissions for chronic obstructive pulmonary disease (COPD), asthma, hypertension, diabetes • Referral rate • Average length of stay • Readmission rate • Caesarean section (C-section) rates

While many of the indicators considered are more commonly used to assess health system performance, any health system performance indicator can be viewed from an efficiency angle. For example, quality indicators such adherence to clinical guidelines, diagnostic accuracy, and treatment success rates would become the outputs, and a doctor’s visit or consultation the input. These could be benchmarked against a target (such as 100% diagnostic accuracy) relative to other providers, or compared to own past performance. In this way, indicators help fulfill the two key characteristics of efficiency analysis – first, it considers both inputs and outcomes and second, it involves making a comparison.

The final list of indicators presented in this guide reflects a balance of current data availability in participating countries as well as some aspirational indicators. There are many health indicator lists that have been developed by international organizations, academics, advocacy groups, and others (including the JLN) grouped for different purposes – for tracking progress towards universal health coverage (UHC), assessing primary health care (PHC), ensuring hospital quality, and/or evaluating health provider payment systems (Annex I). These existing lists were used as a starting point to compile a consolidated list of over 200 indicators. Indicators were organized according to the levels of the results chain (financing, inputs, outputs, and outcomes) and by domains of interest, such as access to care, quality, management, health status, risk factors, and financial protection. During the collaborative process, countries were asked to narrow down indicators based on:

- i) countries' health system priorities;
- ii) the availability of data in-country; and
- iii) the frequency of collection.

Indicators that were not frequently collected included those that covered the quality dimension, as these are more difficult to assess and time consuming to collect via facility surveys, vignettes, or patient record review. Hospital-centric indicators (e.g. average length of stay, bed occupancy rates, avoidable admissions, referrals) were also absent. As a result, the indicators that were initially chosen by participants focused only on financing, outputs related to the access dimension, and outcomes, as definitions were standardized and availability widespread. However, it was important to include indicators that may not be currently collected by health information systems in all participant countries. These are the indicators that will help countries move from assessing 'coverage' of key interventions towards assessing 'effective coverage'. It may also help generate demand for collecting these indicators. Table 2 summarizes the final list of chosen indicators.

Table 2. Common indicators that can be used in an efficiency analysis

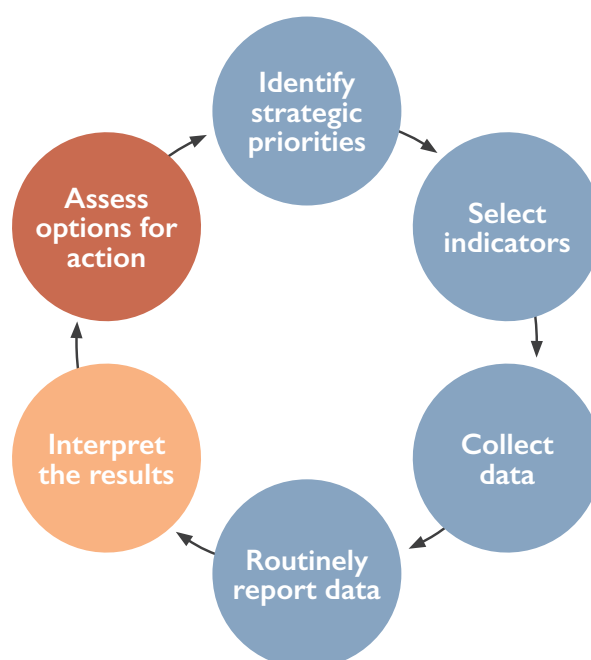
Financing (management)	Inputs (management)	Outputs (access)	Outputs (quality)	Outputs (risk factors)	Outputs (management)	Outcomes (health status)	Outcomes (financial protection)
<ul style="list-style-type: none"> • Total health expenditure (THE) as % of GDP or in per capita terms • Government health expenditure (GGHE) as % of GDP, as % of budget, as % of THE, or in per capita terms • Share of pre-paid/pooled spending as % of THE; out-of-pocket expenditure (OOP) as % of THE; external funds as % of THE • THE, GGHE, OOP spending by function (e.g. outpatient, inpatient, pharmaceutical, primary health care, public health or prevention, curative care) as % of THE, as % of GGHE, as % of OOP expenditure, or in per capita terms 	<ul style="list-style-type: none"> • Density of doctors, nurses, midwives • Hospital density or hospital bed density • Availability of basic facility infrastructure, essential medicines, equipment • General services readiness or service-specific readiness for tracer condition(s) 	<ul style="list-style-type: none"> • Service utilization • Antenatal coverage • Births attended by skilled health personnel • Immunization coverage • TB case detection 	<ul style="list-style-type: none"> • TB treatment success rate • Diagnostic accuracy for tracer condition(s) • Adherence to clinical guidelines for tracer condition(s) • Diabetes control • % of incidents per 1,000 patient days • Avoidable admissions 	<ul style="list-style-type: none"> • Raised blood glucose/diabetes among adults • Children under 5 who are stunted • Tobacco use among persons aged 18+ 	<ul style="list-style-type: none"> • Caseload • Absenteeism • TB notification rate • Average length of stay in hospital • Bed occupancy rate • Referral rate • Hospital readmission rates • C-section rates • Expired drugs • Stock-outs • Claims ratio • Budget execution rates 	<ul style="list-style-type: none"> • Life expectancy at birth • Under-five mortality rate • Maternal mortality rate • TB burden • Mortality between 30–70 years age from cardiovascular diseases, cancer, diabetes or chronic respiratory diseases 	<ul style="list-style-type: none"> • Catastrophic health expenditure

Note: Based on stated country priorities, indicators for immunization and maternal health were used as tracer conditions for assessing efficiency of primary health care, tuberculosis as a tracer condition for assessing efficiency of communicable diseases and a functioning surveillance system, and diabetes as a tracer condition for assessing efficiency of non-communicable diseases. Maternal health and diabetes can also be used as proxy conditions for assessing the efficiency of a functioning referral or integrated system.

From indicators to decision making

The intention is not to have countries track all indicators listed above. The choice of indicators should be driven by country priorities. Given the mix of country issues and varying level of data availability, countries should view Table 2 as a menu of indicators from which to choose. Figure 4 describes a process for using data to inform decision making. This section provides additional guidance on how to get started. It also includes case studies from two participant countries – Kenya and Malaysia – which have piloted the benchmarking plus approach.

Figure 4: Using data to inform decision making



Step 1 – Identifying strategic priorities: As mentioned in the introduction, the main sources of inefficiency have been widely documented in the literature; having discussions with key informants around the top ten can help get a sense of the relevance and extant of a particular source of inefficiency. This can be a starting point for countries thinking through which are the most important in their settings, and which they should tackle first. Addressing inefficiencies can be politically sensitive. Therefore, it is helpful to be transparent around the process. Health sector priorities listed in a National Health Strategic Plan, and routine reviews of health sector performance and budget implementation are natural entry points.

Step 2 – Selecting indicators: Having identified a health priority area, for example, maternal health, focus on choosing the right mix of inputs, outputs, and outcomes that clearly highlight the theory of change across the entire results chain – that is, indicators that are relevant for improving maternal health. For example, spending on maternal health, the number of nurses and midwives, the availability of minimum level of equipment and medicines for basic obstetric care, antenatal coverage, adherence to clinical guidelines for antenatal care activities, diagnostic accuracy for post-partum hemorrhage or hypertensive disorders such as eclampsia and preeclampsia, and maternal mortality.

Step 3 – Collecting data: Policy makers tend to underestimate the availability of data and evidence in their countries. Routine administrative and health financing data are amongst the most valuable sources of information for conducting efficiency analysis. However, the lack of standardized reporting and accounting formats, the low prevalence of electronic health records and patient reported health outcomes, unreliable internet connectivity, and poor reporting compliance mean that significant effort and investment are needed to collect and process data. Until transactional data, administrative data, and/or health management and information systems are able to routinely generate dashboards or alerts for timely analysis and action by decision makers, countries can rely on several other less-frequent sources such as population and health facility surveys. In the meantime, it is helpful to document any data quality issues (e.g., completeness, validity, timeliness) so that an action plan to improve the quality and use of data for decision making can be implemented. Box D summarizes the main sources of data used in efficiency analysis.

Box D: What potential sources are available to collect information for efficiency analysis?

National health accounts, budget documents, and public expenditure reviews/public expenditure tracking surveys are used to assess how public resources – financial, human, in-kind – are allocated and managed and how they flow across administrative levels (e.g., from central government to frontline providers).

Routine or administrative data is generated as part of the operation of the health care system for purposes other than research – typically for administrative reasons or to support care. They include infectious disease surveillance data, medical records, hospital episode data, claims data, and supply chain data on medicines and commodities.

Health facility surveys are used to measure different dimensions of health facility performance supply-side readiness, provider training/knowledge, and provider effort.

Household or population surveys can be used to better understand health-seeking behavior and health status based on household characteristics (such as income, education, location, risk preferences, etc.).

Qualitative surveys or key informant interviews are also valuable instruments to better understand the political, organizational, and institutional environment in which decisions are made and interventions implemented.

The strengths and weaknesses of various data sources are summarized in Annex I of the SPS's **Health Priority Setting: A Practitioner's Handbook**.¹³

Step 4 – Reporting data: Routine reporting provides the first indication that something may be wrong – reports are pushed to users who are then expected to ask the right follow-up questions to extract meaning from the report and take appropriate action. Ideally, the data reporting process is automated; however, standard reporting formats and regular review meetings to discuss areas of concern and action are also effective.

Step 5 – Interpreting the results: It is important to review and question the data to gain better insight. Fact sheets were developed to help guide practitioners in the use of indicators. For each indicator, the fact sheet provides a simple definition, identifies potential sources of data, suggests potential comparators, and gives guidance on benchmarking and interpretation. This includes suggestions on other indicators that might help provide additional insight. The fact sheets are presented in Annex 2.

¹³ <https://www.jointlearningnetwork.org/resources/health-priority-setting-a-practitioners-handbook/>

- i. **Indicators may not always have a recommended benchmark.** Where there are clear recommendations, targets or clinical guidelines, benchmarking provides a nice signpost. For example, stock-out rates should ideally be zero, and C-section rates, of higher than 10% at the population level are not associated with reductions in maternal and newborn mortality. But where there are no set norms, benchmarks should be used with caution. For example, numerous health-spending targets have been recommended – 15% of government budgets, or 5% of GDP. While these targets can serve as global benchmarks, they are usually not helpful for determining appropriate levels of spending at the country level. Many countries spend more than these targets and have yet to provide a basic package of services to their population, while others spend less and achieve near universal levels of population coverage. This again highlights the importance of considering both inputs and outputs when interpreting indicators through an efficiency lens.
- ii. **It is important to make the right comparisons.** Cross-country comparisons and national averages may mask regional variations or socioeconomic inequalities – especially in decentralized contexts. For most output indicators related to service delivery it is best to compare across facility types or providers within a country or region (internal benchmarking) as clinical protocols may vary from country to country. Looking at trends and past performance is also a helpful comparison – especially if there is no obvious comparator – as it informs policy makers on whether there is improvement, no change, or worsening performance.
- iii. **Inefficiency is difficult to attribute.** As already mentioned, benchmarking is helpful for detecting emerging issues such as stagnating outcomes, outliers, or significant changes from one period to the next; however, on its own, it is not enough. Many other factors need to be considered such as time lags, data availability and quality, issues in costing, disease burden, and economies of scale. For example, the cost of service delivery can be much higher to achieve the same outcomes in remote areas that require significant outreach or in districts where the disease burden is more pronounced. As specific issues emerge, additional indicators and key informant interviews may be necessary to delve deeper and understand whether an issue is truly an inefficiency or not. The fact sheets in Annex 2 give additional guidance on benchmarking, choice of comparator, and interpretation for each indicator.

Step 6 – Assess options for action: Once inefficiencies have been identified, decision makers need an action plan to take recommendations forward. This requires having clear accountability processes in place and knowing the responsible unit or budget holder. Most countries reported that they typically collect hundreds of indicators as part of their regular monitoring and evaluation of the health sector – indicators that feed into annual budget reviews, annual program reports, key performance indicator dashboards, and statistical yearbooks. However, few were able to articulate whether and how this information was being used. These ‘reporting’ tools fall shy of the ‘analysis’ that might generate insight to guide policy and decision-making. While participant countries are already undertaking steps 1 through 4, only a few are assessing results, and even fewer are coming up with an action plan to address issues identified.

Ghana, Kenya, Malaysia, and the Philippines volunteered as pilot countries to test the validity and feasibility of the benchmarking plus approach. Boxes E and F summarize the Kenya and Malaysia case studies, respectively, applying the steps from identifying strategic priorities and choice of indicators to assessing options for action.

Box E: Kenya pilot

What was the main issue? Maternal mortality remains a challenge in Kenya, where the maternal mortality rate (MMR) has been stagnant since 1993, prompting the government to set reduction of MMR as a flagship program in the Kenya Health Sector Strategic Plan 2014–2018 (KHSSP). The aim was to reduce maternal mortality by half or decrease from 400 at baseline to 150 maternal deaths per 100,000 live births between 2014 and 2018. In the current KHSSP 2018–2022, the MMR baseline was 250 with a target reduction of 30% (Kenya Ministry of Health, 2018). Participants from Kenya's JLN team applied the 'benchmarking plus' approach to identify inefficiencies in the service delivery chain.

What indicators were used and what were the findings? The team looked at the number of doctors and nurses, access to antenatal care (ANC) services and skilled birth attendance (SBA), and readiness to deliver basic obstetric care services as a proxy for quality of care. According to the 2018 Kenya Economic Report, Kenya has a significant shortage of health workers, exacerbated by poor distribution in the 47 counties in the country. There are 0.25 doctors, 1.82 registered nurses and 3.08 enrolled nurses per 1,000 population compared with the WHO recommendations of 3.0 doctors, 2.6 registered nurses and 5.4 enrolled nurses (Exhibit E.1) (KIPPR, 2018). The percentage of women attending at least one ANC is high (82%), however, those attending four or more ANC visits was much lower at 49%. Data on SBA stood at 65% (Exhibit E.2). On average facilities only had 32% of items necessary to provide maternal health services with readiness highest at hospitals and private facilities; however, national averages mask wide variations (Exhibits E.3 and E.4).

What were the limitations? Was additional contextual information provided? Regular and up-to-date data for strategic planning, resource prioritization, and routine performance monitoring is not easily available. For example, administrative data does not routinely capture disaggregated data on the availability of personnel, medicines, and equipment, or utilization of services. However, population and facility surveys such as the Demographic and Health Surveys, the Service Availability and Readiness Assessment, and the Service Delivery Indicators were able to provide snapshots at a point in time. In addition, it was shared that only 36% of health facilities conducted audits for maternal deaths between July 2015–June 2016; only 43.7% of providers adhered to clinical guidelines; and only 44.6% of providers were able to manage maternal and neonatal complications based on clinical vignettes (World Bank, 2013). The team also shared that nurse strikes in December 2016 and June 2017 significantly affected the percentage of women receiving ANC and SBA services.

What recommendations are suggested? The main bottleneck happens early in the service delivery chain with workforce shortages. Therefore, next steps should focus on mapping out the distribution of providers, considering the labor market dynamics, and assessing the cost-effectiveness of various strategies, such as curriculum reform in medical schools and training, to increase the supply and retention of primary health care workers; introducing new and alternative cadres of health workers including options for telemedicine; monetary and non-monetary incentives for providers working in rural and primary health care practices; and regulatory measures that require workers to practice in remote areas for a certain number of years. In addition, increased supervision and oversight may enhance adherence to clinical protocols, but non-compliance may also be due to overworked personnel given shortages.

Source: Dr. Isabel Maina, Kenneth Munge, Dr. Mercy Mwangangi, Dr. Ahmed Omar, Esther Wabuge

Exhibit E.1: Ratio of nurses per 100,000 population by county (latest year available)

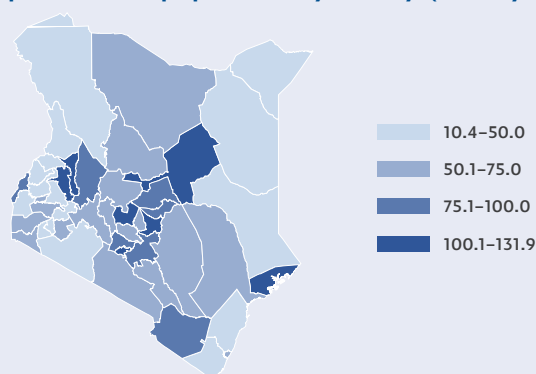
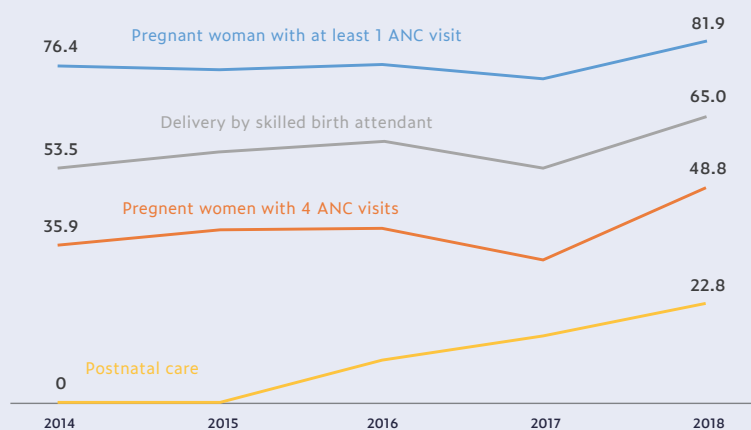
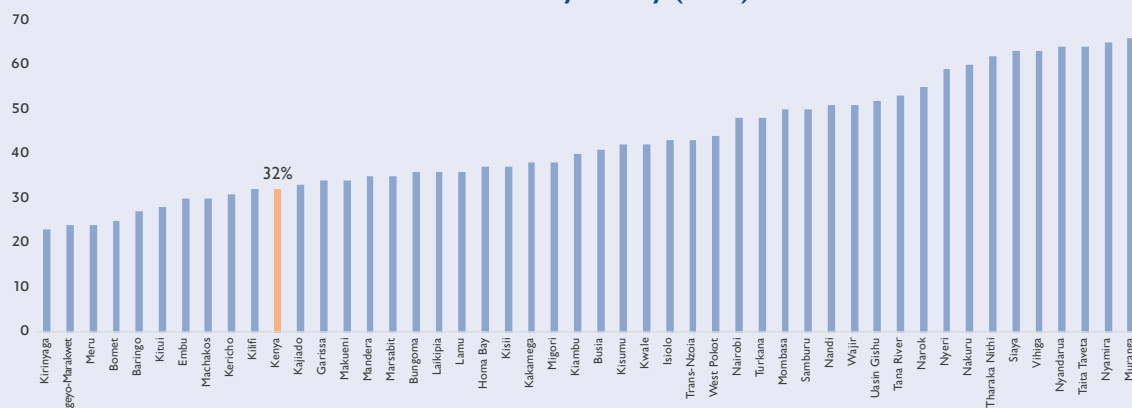


Exhibit E.2: Maternal health coverage indicators, %, (2014-2018)



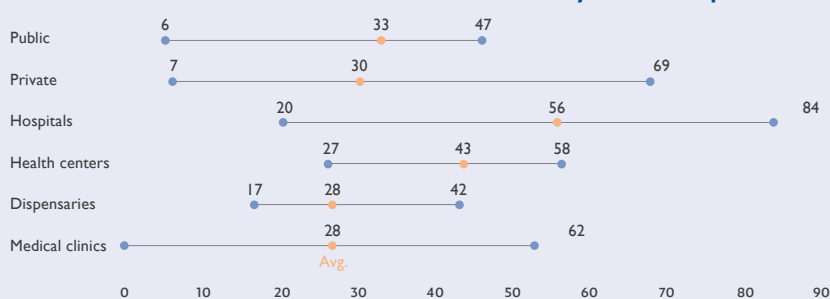
Source: (Kenya Ministry of Health, 2018)

Exhibit E.3: Maternal service readiness index by county (2013)



Source: (Government of Kenya, 2014)

Exhibit E.4: Maternal service readiness index by ownership and tier of care (2013)



Source: (Government of Kenya, 2014)

Box F: Malaysia pilot

What was the main issue? In 2017, Malaysia had the highest diabetes prevalence (16.9%) in the East Asia Pacific region (Exhibit F.1) (International Diabetes Federation, 2019). Compared to other countries, Malaysia has high rates of hospital admission for long-term complications of diabetes (such as premature heart disease and stroke, blindness, amputations, and kidney failure), which are 10–12 times costlier to treat than regular follow-up diabetes care (Mustapha, et al., 2017). In Malaysia, diabetes is also the fifth highest cause of death and disability combined (IHME, 2019). For this reason, the JLN Malaysia team sought to explore potential efficiency gains along the results chain that could lead to improved outcomes for diabetes.

What indicators were used and what were the findings? The team looked at indicators along the diabetes care cascade – diabetes prevalence, percent of diabetes patients who are diagnosed, and among diagnosed the percent of patients who are on treatment and who have their diabetes under control (Exhibit F.2). Findings from an annual clinical audit of type 2 diabetes patients provided additional insight into service delivery by looking at adherence to clinical guidelines at public health care facilities and the rate of complications (Exhibit F.3). The breakdown happens early in the continuum of care, as more than half (59%) of diabetes patients remain undiagnosed. While 86.5% of diagnosed patients receive annual A1C screening, fundoscopy and renal and liver function tests are less prevalent. Among diabetes patients 58.2% had good glycaemic control A1C <8% while 1 in 5 patients had an HbA1c >10%. In terms of complications, in 2017, 11.5% of diabetes patients developed nephropathy, 9.6% developed retinopathy, 5.1% developed ischaemic heart disease, 1.5% developed cerebrovascular disease, 1.3% developed a diabetic foot ulcer, and 0.7% had a limb amputation.

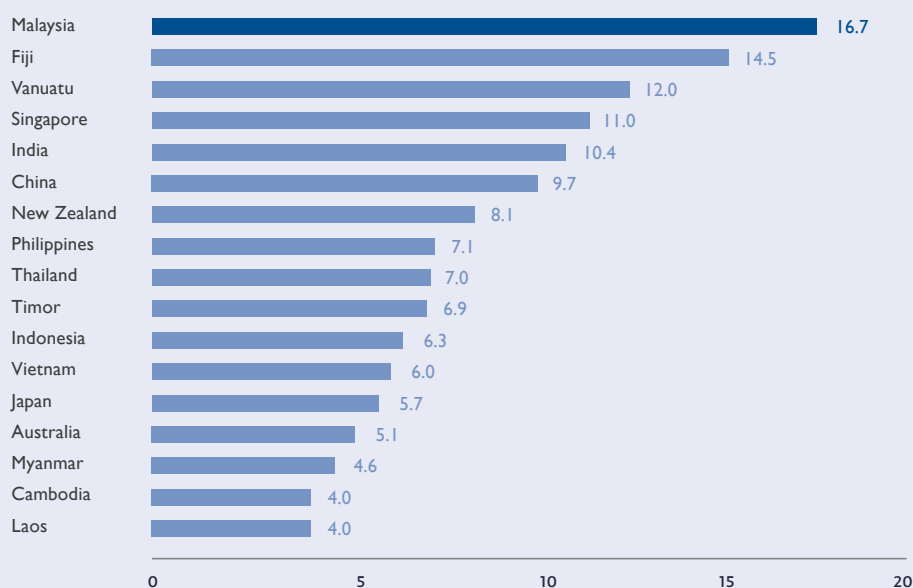
What were the limitations? Was additional contextual information provided? Population-level data on the percent of diagnosed diabetes patients that are on treatment, the percentage whose treatment adheres to clinical guidelines, and the rate of complications was not available. In the absence of such data, the percentages from the National Diabetes Registry (NDR) Clinical Audit sample-based study were used. While the data from the NDR does not include patients seen in the private sector, the 2015 National Health and Morbidity Survey reported that only 15.1% and 3.6% of diabetes patients sought care in private clinics and private hospitals respectively, making these reasonable estimates for the broader population. The Malaysia team complimented the analysis with additional insight on how an increasing number of patients were referred to tertiary hospitals due to certain limitations at primary health care facilities in diagnosing and managing diabetes care. However, there is now a push to have more fundus cameras available and more Assistant Medical Officers certified and trained to carry out fundus exams at district-level primary health care clinics. Multiple efforts have been undertaken by various sectors to overcome the inadequate adherence to clinical guidelines for diabetes in health care facilities as well as the mismatch of knowledge on diabetes, poor compliance of medications, and aversion to insulin delivery methods among patients.

What recommendations are suggested? Based on these findings, recommendations are divided into four perspectives – healthcare providers, patients, policy makers, and researchers. For healthcare providers, additional training is required to improve the knowledge and skills to detect diabetes early on. This will also help promote best practices in managing diabetes and its related complications for improved quality of care. For patients, continued effort by the health promotion team to improve awareness and increase health literacy among diabetic patients. Policy makers should consider alternative approaches to service delivery such as financial incentives to the case managers who oversee long-term care especially among high-risk patients, community empowerment programs that conduct screening, or a complete patient-centered integrated care model that incorporates community resources, patient self-management, integrated information systems, and innovative delivery systems such as clinical decision support applications, mobile health devices, and bundled payment options between primary health care and hospital settings. Researchers should focus on the effectiveness, efficiency, and cost-effectiveness of diabetes interventions to provide evidence and assist policy makers in the prioritization and decision-making process.

The authors would like to thank the Director General of Health Malaysia for his permission to publish this article.

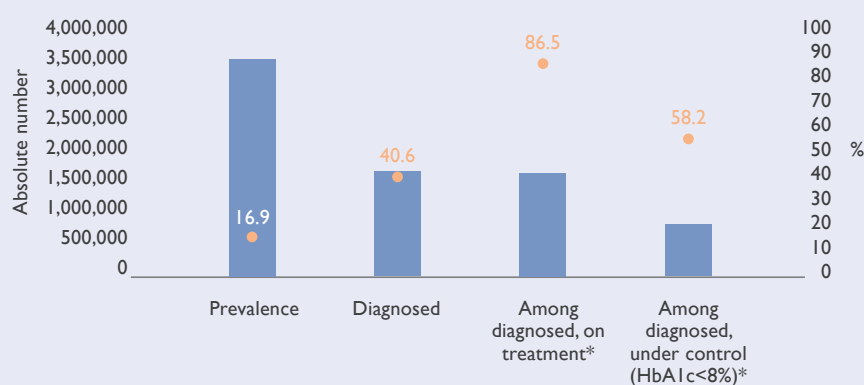
Source: Dr Nor Izzah Hj Ahmad Shauki, Dr Lee Kun Yun, Dr Abdul Hakim b Abdul Rashid, Dr Mohd Najib b Baharuddin, Dr Yussni bt Aris

Exhibit F.1: Diabetes prevalence (%), 2017



Source: International Diabetes Federation (<https://idf.org/our-network/regions-members/western-pacific/members/108-malaysia.html>)

Exhibit F.2: Diabetes care cascade in Malaysia (absolute number and %), 2017



Note: *Based on sample-based clinical audit of diagnosed diabetes patients. <http://www.iku.gov.my/images/IKU/Document/REPORT/nhmsreport2015vol2.pdf>

Sources: International Diabetes Federation (<https://idf.org/our-network/regions-members/western-pacific/members/108-malaysia.html>) and *National Diabetes Registry of Malaysia

Exhibit F.3: Adherence to clinical practice guidelines for diabetes mellitus in Malaysia (%), 2017

Had exam in past year	%
Patients who underwent HbA1c	86.5
Patients who underwent fundoscopy	51.8
Patients who underwent foot exam	76.3
Patients who underwent proteinuria	67.0
Patients who underwent urine microalbumin	52.9

Note: Sample size 160,075

Source: National Diabetes Registry of Malaysia

The findings from the pilot exercise highlighted challenges to applying the proposed framework as well as lessons learned.

Challenges:

- **Availability of data:** A key limiting factor was the availability of data at the right level of disaggregation to make meaningful comparisons (e.g., by geographical region, across population groups) – especially for decentralized countries. It was also difficult to get data across different parts of the health system as information is held in different departments or programs within the Ministry of Health or was spread out across several Ministries/Agencies or levels of government. However, when pressed, all pilot countries were able to bring in anecdotal evidence or refer to country studies to help explain their data.
- **Defining the scope of the efficiency analysis:** All four countries attempted to present country data for the entire list of priority indicators. This made it harder to focus on any one area and to delve deeper into identifying where along the results chain inefficiencies might lie.
- **Interpreting the data:** There was little interpretation of what the data might mean and/or packaging the information into a policy-relevant story line.

Lessons learned:

- **Focus the scope of the analysis:** That is, looking at hospital efficiency, pharmaceuticals or a specific tracer condition within primary health care, such as maternal health or tuberculosis, where relevant indicators were more likely to be available under responsible units or budget holders.
- **Additional indicators should be analyzed when specific issues emerge in the course of routine monitoring:** This includes going beyond indicators and drawing on practitioner's valuable knowledge of the sector to help fill in the gaps in the storyline when data is not available. This will help with the interpretation of findings and the assessment of options for further action.
- **Efficiency analysis needs to be embedded into the formal decision-making process, not a one-off exercise.** Many health system performance indicators are already regularly collected and reported as part of the routine monitoring and evaluation of the health sector. Building on existing processes and applying an efficiency lens to reviewing indicators that question why some regions, hospitals, or providers perform better than others is the first step to institutionalizing efficiency analysis.

Conclusion

There are inefficiencies in all health care systems, the scale of which can be significant. But while the main sources of inefficiency have been widely discussed in the literature, what is less available is a practical approach that policy makers can use to inform change. Stochastic frontier analysis (SFA) or data envelopment analysis (DEA) are the standard methods for economists to assess efficiency. However, these methods are often time consuming, data intensive, and require particular expertise. In addition, they do not necessarily identify the truly efficient behavior, nor do they tell policy makers how to improve efficiency.

This guide presents an alternative ‘benchmarking plus’ approach that can be readily used by practitioners and policy makers. In setting up routine monitoring systems, countries should ideally benchmark a few common efficiency metrics in areas that consume the most resources or are prone to inefficiency such as caseload, average length of stay, avoidable readmissions, unit cost per episode of care, stock-out rates, etc. However, if not available, benchmarking indicators along the process of how inputs are used to deliver services, which produce outputs that contribute to outcomes, can help pinpoint service delivery bottlenecks for further inspection. Discussion with key informants provides additional contextual information on the relevance and extent of identified problems. Only then can appropriate recommendations be considered based on countries’ prioritization criteria and resource constraints.

However, even this simpler approach presented challenges for participating countries given data limitations. Almost all participating countries lacked indicators to measure those aspects of service delivery that are related to quality. Instead, available outputs mostly focused on availability, access, and coverage. But quality metrics are needed to help monitor countries’ progress from assessing ‘coverage’ of key interventions towards assessing ‘effective coverage’. In 2018, poor quality of health services contributed to more amenable deaths (5 million) than non-utilization of services (3.6 million) across lower-middle income countries. Hospital quality metrics are especially important given recent findings that 10% of hospitalized patients acquire infections during their stay; clinical guidelines are followed in less than 45% of cases; and harmful medical errors and preventable complications account for 15% of hospital costs (Kruk, et al., 2018). All of these represent sources of inefficiency because resources are spent on services that produce reduced or no added health benefit.

Countries cannot improve what they cannot measure; however, other resources are likely available to support efficiency analysis. In the absence of regular reporting on key efficiency metrics collected by health management and information systems, almost all countries will be able to fill in the gaps using past studies, ad hoc population and facility surveys, and interviews with key informants, as was demonstrated by the Kenya and Malaysia case studies. For this reason, a benchmarking plus approach is recommended while investments for foundational healthcare analytic systems are made.

Fortunately, countries are increasingly digitizing data which will help fill in these gaps and increase the potential for efficiency analysis. The demand for data-driven decision-making and the rising pressure to reduce healthcare spending while improving patient outcomes is driving the growth of healthcare analytics – a USD 14 billion industry globally that is projected to reach USD 50.5 billion by 2024. Currently, low- and lower-middle income countries still struggle to produce basic descriptive analytics – a precursor to more advanced predictive and prescriptive analytics. Technological advances in text and image storage, cloud computing, and big data are paving the way to extract more value from data by integrating electronic medical records, patient monitoring systems, claims data, and other core administrative data. However, the biggest limiting factor is the broad range of skills necessary to develop a healthcare analytics team that includes information technology specialists to deal with hardware, software, and connectivity issues; clinical coders and data management specialists to enable standardization and interoperability; and policy makers, health care providers, and data scientists to develop applications that help answer policy-relevant questions. All must work together.

References

- Chisholm, D. & Evans, D. B., 2010. *Improving health system efficiency as a means of moving towards universal coverage*, Geneva: The World Health Organization.
- Couffinhal, A. & Socha-Dietrich, K., 2017. Ineffective spending and waste in health care systems: frameworks and findings. In: *Tackling wasteful spending on health*. Paris: OECD Publishing, pp. 17–59.
- Cylus, J., Papanicolas, I. & Smith, P. C., 2016. *Health system efficiency: how to make measurement matter for policy and management*. London: European Observatory on Health Systems and Policies.
- Drummond, M. F. et al., 2015. *Methods for economic evaluation of health care programmes, 4th Edition*. Oxford: Oxford University Press.
- Government of Kenya, 2014. *Kenya Service Availability and Readiness Assessment Mapping (SARAM)*, Nairobi: Ministry of Health.
- Hafez, R., Somanathan, A. & Wilson, D., 2016. *Raising funds for health - Background paper for 1st universal health coverage financing forum*. Washington DC: The World Bank.
- IHME, 2019. *Global Burden of Disease*, s.l.: <http://www.healthdata.org/malaysia>.
- International Diabetes Federation, 2019. *Diabetes Atlas*, s.l.: <https://diabetesatlas.org/across-the-globe.html>.
- Kenya Institute for Public Policy Research and Analysis (KIPPRA), 2018. *Kenya Economic Report 2018: Boosting Investments for Delivery of the Kenya Vision 2030*, Nairobi.
- Kenya Ministry of Health, 2018. *Kenya Health Sector Strategic Plan (KHSSP), July 2018 - June 2022*, Nairobi: Ministry of Health.
- Kruk, M. E. et al., 2018. High quality health systems in the sustainable development goals era: time for a revolution. *Lancet Glob Health*, Volume 6, pp. e1196–252.
- Morgan, D., 2010. Chapter 1: how much is too much? value for money in health spending. In: *Value for money in health spending*. Geneva: OECD, pp. 21–42.
- Mustapha, F. I. et al., 2017. What are the direct medical costs of managing Type 2 Diabetes Mellitus in Malaysia. *Med J Malaysia*, 72(5), pp. 271–277.
- Papanicolas, I. & Smith, P. C., 2017. Theory of system level efficiency in health care. In: *Encyclopedia of Health Economics*. San Diego: Elsevier, pp. vol. 3: 386–394.
- Smith, P. C., 2012. What is the scope for health system efficiency gains and how can they be achieved?. *Eurohealth Observer*, pp. 18(3)3–6.
- Tandon, A. & Cashin, C., 2010. *Assessing Public Expenditure on Health From a Fiscal Space Perspective*, Washington, DC: The World Bank.
- UNFPA, 2014. *The state of the world's midwifery*, <https://www.unfpa.org/data/sowmy/KE>: UNFPA.
- WHO, 2019. *Global Health Observatory Database*, Geneva: WHO.
- World Bank, 2013. *Kenya Service Delivery Indicators*, Washington, DC: World Bank.
- World Bank, 2017. *Greater efficiency for better health and financial protection: Background paper for 2nd Annual UHC Financing Forum, April 20-21, 2017*, Washington DC: The World Bank.
- World Bank, 2019a. *Health Equity and Financial Protection Database*, Washington, DC: The World Bank.
- World Bank, 2019b. *World Development Indicators*, Washington, DC: The World Bank.
- Yip, W. & Hafez, R., 2015. *Reforms for improving the efficiency of health systems: lessons from 10 country cases*, Geneva: The World Health Organization.
- Yip, W., Hafez, R. & Chi, Y. L., 2014. *Strategic purchasing for universal health coverage: policy brief for the Eastern-Mediterranean region*, Geneva: The World Health Organization.

Annex I – Indicator lists reviewed

The World Health Organization-World Bank UHC Measurement Framework

The WHO-World Bank UHC Measurement Framework tracks eight core health service coverage indicators chosen because they involve interventions for which every individual should benefit – mainly maternal and child health, and infectious disease control. Not currently included are tracers for non-communicable diseases. In addition, the framework also tracks two measures of financial risk protection, namely out-of-pocket spending as a share of total consumption and health expenditures that drive households beneath the poverty line. These indicators are generally collected from frequent household surveys such as the demographic and health survey, the multiple indicator cluster survey, and/or household consumption surveys. The limited number of indicators and their relative straightforward measurability allow for comparability across countries making them a useful common framework for tracking progress towards UHC globally.

For more information: https://www.who.int/healthinfo/universal_health_coverage/report/2015/en

The World Health Organization's Global Reference List of 100 Core Health Indicators

This list is a more comprehensive (though not exhaustive) set of indicators organized by four domains i) health status, ii) risk factors, iii) service coverage, and iv) health systems. In addition to a broader range of maternal and child health, and infectious disease control indicators, the global reference list also includes a range of risk factors associated to non-communicable diseases, and outcomes that act as proxies for access and quality of care. Unlike the UHC monitoring framework that looks at a limited section of the results chain (i.e. outputs and outcomes), this expanded list allows policy makers to identify weaknesses along the entire results chain. In fact, they also provide the same list of indicators re-organized by sections of the results chain. However, the focus remains on areas associated with the Millennium Development Goals agenda.

For more information: <https://www.who.int/healthinfo/indicators/2018/en>

Primary Health Care Performance Initiative

The PHCPI focuses exclusively on primary health care. The 36 Vital Signs indicators align and complement the Global Reference List of 100 Core Health Indicators by focusing on important service delivery processes that are crucial for achieving UHC and other global priorities. It looks at indicators across the entire results chain and goes one step further in collecting information related to provider knowledge (e.g., diagnostic accuracy), effort (e.g., caseload), and quality of service delivery (e.g., continuity of care) – indicators that are more complex to produce requiring vignettes, and health facility surveys that are less commonly produced and standardized limit comparability across countries and over time. These might be more useful for a within country comparison (e.g., regional variation).

For more information: <https://improvingphc.org/phcpi-core-indicators>

Organization for Economic Cooperation and Development Health at a Glance Dashboard

The OECD dashboard of indicators is meant to provide an overview of health system performance – organized around similar domains of health care resources, health status, and risk factors. However, unlike the WHO and PHCPI lists, the focus of the access to and quality of care indicators is on non-communicable diseases – a reflection of these countries' epidemiology and level of health system development. The dashboard is used for cross-country comparisons.

For more information: <https://www.oecd.org/health/health-systems/health-at-a-glance-19991312.htm>

Commonwealth Fund Scorecard for State Health System Performance

The 44 indicators on the Commonwealth Fund Scorecard – used to track performance over time and rank US states relative to one another – fall into four domains: access and affordability, prevention and treatment, avoidable hospital use and cost, and healthy lives. Similar to the OECD dashboard, indicators are more focused on hospital use and care, non-communicable diseases and associated risk factors – again a reflection of the US's disease burden and health care system.

For more information: <https://www.commonwealthfund.org/publications/fund-reports/2018/may/2018-scorecard-state-health-system-performance>

Joint Learning Network's Indicators for Monitoring Provider Payment Methods

Finally, the JLN developed a list of indicators specifically geared towards monitoring provider payment methods. Given the diversity of health care systems and provider payment methods across countries, these would ideally be used for within-country comparison at the facility/provider level.

For more information: <https://www.jointlearningnetwork.org/resources/data-analytics-for-monitoring-provider-payment-toolkit>

Annex 2 – Fact sheets

Health Financing Technical Initiative: The Efficiency Collaborative

Indicator Fact Sheets

Financing (management)

- Total health expenditure as % of GDP
- Government health expenditure (GGHE) as % of GDP, as % of budget, as % of THE, or in per capita terms
- Health expenditure by source of financing
- Current health expenditure (CHE) by function

Inputs (management)

- Health worker density
- Hospital bed density

Outputs (access)

- Antenatal care
- Skilled birth attendance

Outputs (quality)

- TB case detection, TB notification rate, and TB treatment success rate
- General service readiness; Service specific readiness
- Diagnostic accuracy for a tracer condition
- Adherence to clinical guidelines
- Diabetes control
- Avoidable hospital admissions for primary care sensitive conditions

Outputs (risk factors)

- Children under 5 who are stunted
- Tobacco use among persons aged 18+

Outputs (management)

- Caseload
- Absenteeism
- Average length of stay (ALOS) vs. Bed occupancy rate (BOR)
- Hospital re-admission rate
- Caesarean section (C-section rates)
- Expired drugs
- Drug stockouts
- Claims ratio
- Budget execution rates (BERs)

Outcomes (health status)

- Life expectancy at birth
- Under-five mortality
- Maternal mortality rate
- TB burden

Outcomes (financial protection)

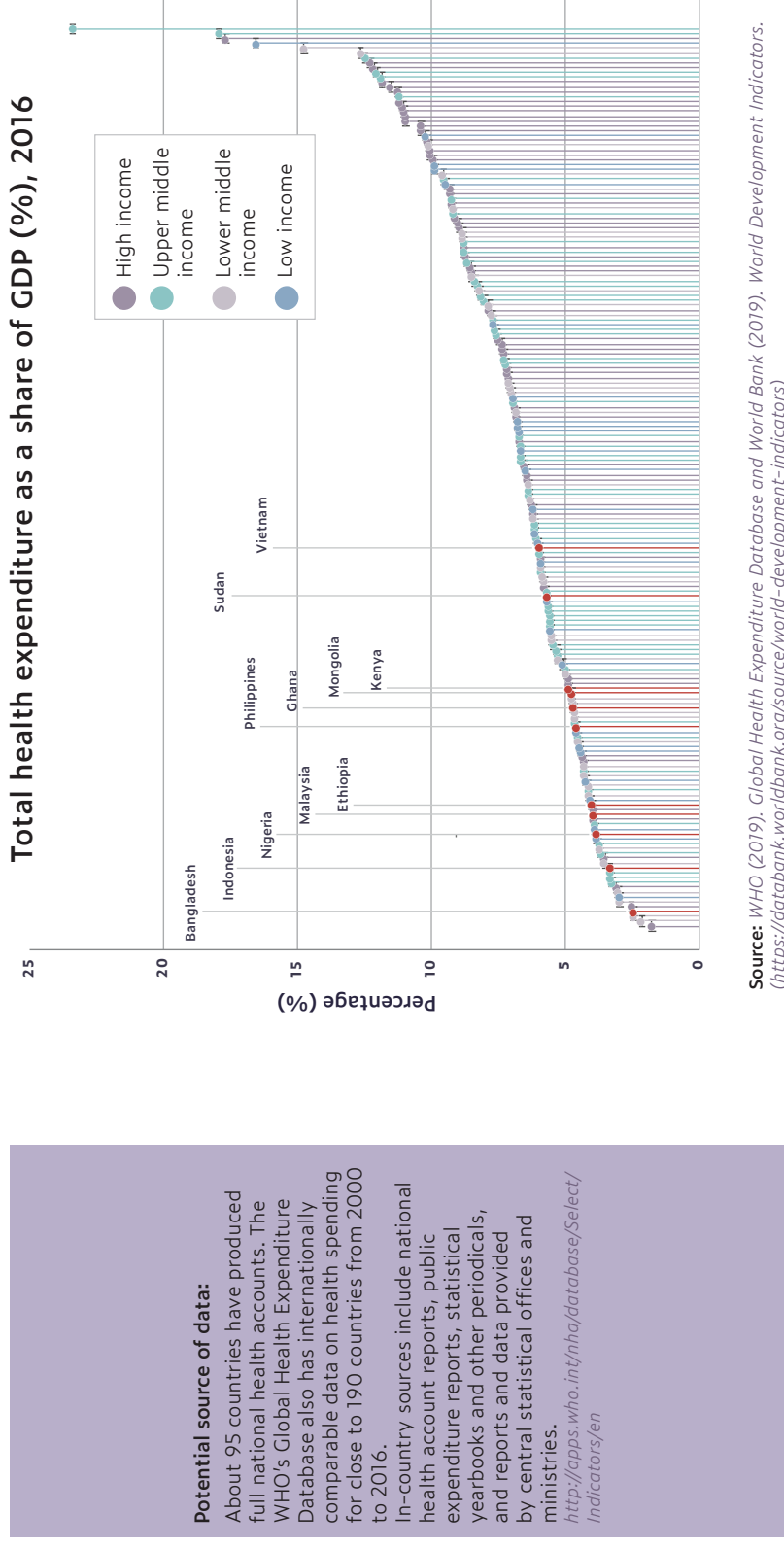
- Catastrophic health expenditure

Total health expenditure as % of GDP

Simple definition: All spending on health care goods and services and capital goods. In the National Health Account terminology this is referred to as the sum of Current Health Expenditures (CHE) and Capital Health Expenditure (HK) where:
CHE is the sum of all government schemes including compulsory contributory health care financing schemes (HF.1), all voluntary health care payment schemes (i.e. voluntary or private insurance) (HF.2), household out-of-pocket (OOP) (HF.3), rest of the world financing schemes (i.e. external donor funding) (HF.4), and other not elsewhere classified (HF.nec) and HK is defined at the acquisition of produced assets; that is assets intended for use in the production of other goods and services.

Frequency of data collection: Annual.

Suggested comparator: Across countries; if country is decentralized subnational comparisons are highly informative; past performance/trend data.



Most recently, the third edition of the Disease Control Priorities initiative (DCP3) estimated the total cost per person for sustaining a high priority package (HPP) and an essential universal health coverage package (EUHC) at 80% coverage would be US \$42 and \$72 respectively in low income (LIC) countries and \$58 and \$110 respectively in lower-middle income (LMIC) countries.

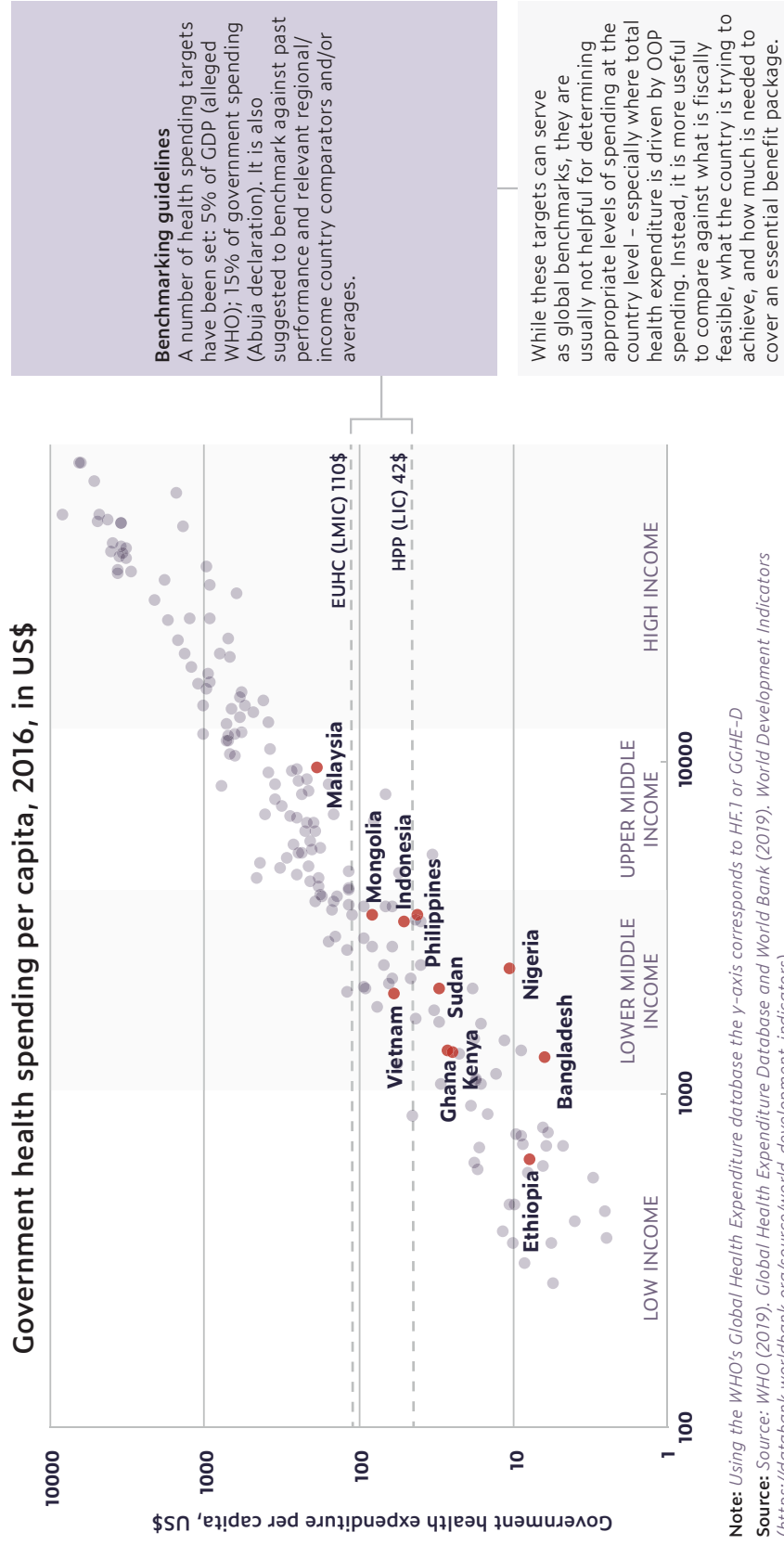
Government health expenditure (GGHE) as % of GDP, as % of budget, as % of THE, or in per capita terms

Simple definition: All spending on health by government schemes and compulsory contributory health care financing schemes.

Frequency of data collection: Annual.

Suggested comparator: Across countries; if country is decentralized subnational comparisons are highly informative; past performance/trend data.

Potential source of data:
About 95 countries have produced full national health accounts. The WHO's Global Health Expenditure Database also has internationally comparable data on health spending for close to 190 countries from 2000 to 2016. In-country sources include national health account reports, public expenditure reports, statistical yearbooks and other periodicals, and reports and data provided by central statistical offices and ministries.
<http://apps.who.int/nha/database/Select/Indicators/en>



Most recently, the third edition of the Disease Control Priorities initiative (DCP3) estimated the total cost per person for sustaining a high priority package (HPP) and an essential universal health coverage package (EUHC) at 80% coverage would be \$US42 and \$72 respectively in low income (LIC) countries and \$58 and \$110 respectively in lower-middle income (LMIC) countries. These estimates suggest that current GGHE will need to double or triple to finance HPP or EUHC packages in LIC/LMIC.

Health expenditure by source of financing

Simple definition: Share of government spending, voluntary and/or private HI, OOP, rest of the world as % of THE.

The overall current spending on health include expenditures of government scheme (HF.1), voluntary schemes (HF.2), out of pocket spending (HF.3), rest of the world (HF.4) and other not elsewhere classified expenditures (HF.nec).

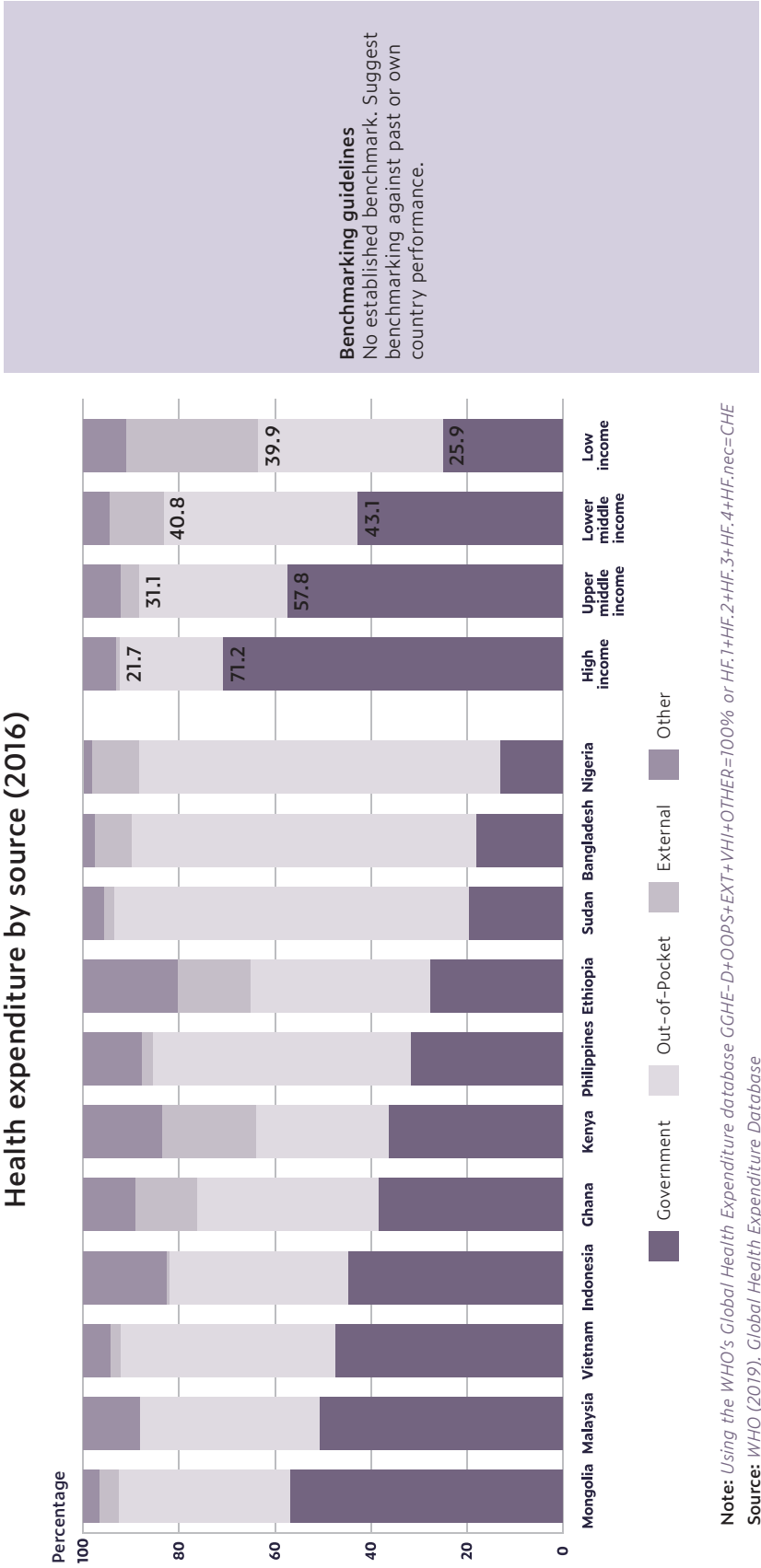
Frequency of data collection: Annual.

Suggested comparator: Across countries; if country is decentralized subnational comparisons are highly informative; past performance/trend data.

Potential source of data:

About 95 countries have produced full national health accounts. The WHO's Global Health Expenditure Database also has internationally comparable data on health spending for close to 190 countries from 2000 to 2016. In-country sources include national health account reports, public expenditure reports, statistical yearbooks and other periodicals, and reports and data provided by central statistical offices and ministries.

<http://apps.who.int/nha/database/Select/Indicators/en>



By construction, the higher the share of government and/or prepaid spending the lower the share of OOP; and the lower the share of OOP spending, the better. Globally, countries that are closest to attaining universal health coverage (UHC) have OOP spending levels that are less than 15–20% of total health spending.

Current health expenditure (CHE) by function

Simple definition: The share of current health expenditure by function (e.g. curative care, pharmaceuticals, preventive care, administrative). Using the WHO's Global Health Expenditure database definitions:

- CHE= Curative care (HC.1) + Rehabilitative care (HC.2) + Long-term care (HC.3) + Ancillary services (HC.4) + Medical goods (HC.5) + Preventive care (HC.6) + Governance and health system and financing administration (HC.7) + Other health care services not elsewhere classified (HC.9)
- Curative care (HC.1)=Inpatient curative care (HC.1.1) + Day curative care (HC.1.2) + Outpatient curative care (HC.1.3) + Home-based curative care (HC.1.4) + Unspecified curative care (HC.1.nec)
- Primary health care expenditure=Outpatient Curative Care (HC.1.3) + Home-based curative care (HC.1.4) + Outpatient long-term care (HC.3.3) + Home-based long-term care (HC.3.4) + 80% of medical goods (HC.5) + Preventive care (HC.6) + 80% of Governance, and health system and financing administration (HC.7). However, it is recommended to adjust this definition to better reflect country context.

Frequency of data collection: Annual.

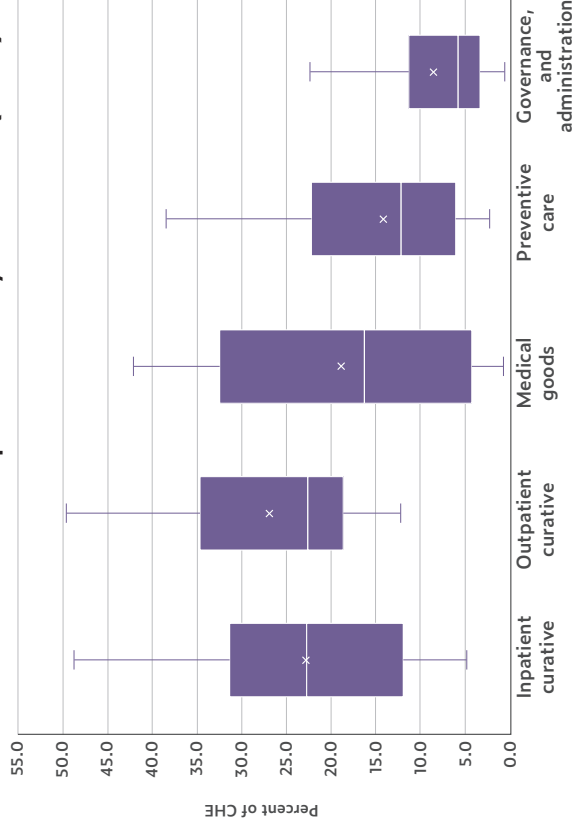
Suggested comparator: Across countries; Past performance/trend data.

Potential source of data:
Beyond OECD countries, only a select number of countries have started reporting expenditures by function to the WHO's Global Health Expenditure Database (<http://apps.who.int/nha/database/Select/Indicators/en>).

The WHO has published the first ever comparable measures of primary health care spending using data from 2016 for 46 low- and middle-income countries (LINK: <https://apps.who.int/iris/bitstream/handle/10665/276728/WHO-HIS-HGF-HF-WorkingPaper-18.3-eng.pdf?ua=1>).

In-country sources may include national health account reports and public expenditure reviews (<http://search.worldbank.org/per>).

Current health expenditure by function (2016)



Notes: Boxplots show the interquartile range of values with the median marked by a line inside the bar and the mean marked by an X. The lines from the bar extend to the maximum and minimum values with outliers excluded. Inpatient curative =HC1.1+HC1.2; Outpatient curative=HC1.3+HC1.4; Medical goods=HC5; Preventive care=HC6; Governance and administration=HC7. Data is for 34 low- and lower-middle income countries for which data is available: Afghanistan; Armenia; Burundi; Burkina Faso; Bhutan; Cote d'Ivoire; Congo, Dem. Rep.; Congo, Rep; Cabo Verde; Djibouti; Ethiopia; Georgia; Ghana; Guinea; Guatemala; Haiti; India; Kenya; Kyrgyz Republic; Cambodia; Liberia; Sri Lanka; Moldova; Mali; Mauritania; Nepal; Philippines; Togo; Tajikistan; Timor-Leste; Tunisia; Tanzania; Uganda; Zambia

Source: WHO (2019). *Global Health Expenditure Database* (<http://apps.who.int/nha/database/Select/Indicators/en>)

Benchmarking guidelines
No established benchmark. Suggest benchmarking against countries of similar income/health system development.

This indicator is difficult to compare across countries. Spending by function is generally driven by factors such as country and health system development, demographics and epidemiology, treatment protocols, and patient preferences, etc. For example, in countries where the quality of primary care is poor or long term care options are limited, spending on curative care/hospitals will be higher. It is also important to note that more than 20% of current health expenditure remains unclassified in some countries suggesting a lack of availability of more granular data for producing health accounts.

Special note on pharmaceutical spending: Four possible factors influence a country's pharmaceutical spending: i) country population and volume of drugs consumed, ii) drug utilization per person, iii) type and mix of drugs consumed (e.g., generics versus brand-name drugs), and iv) prices at which drugs are sold. Comparing drug prices across countries is a complicated and imperfect process, primarily because of the proprietary nature of the rebates drug manufacturers offer different payers.

Health worker density

Simple definition: Number of physicians, nurses, and midwives available (in a region) per 1,000 inhabitants.

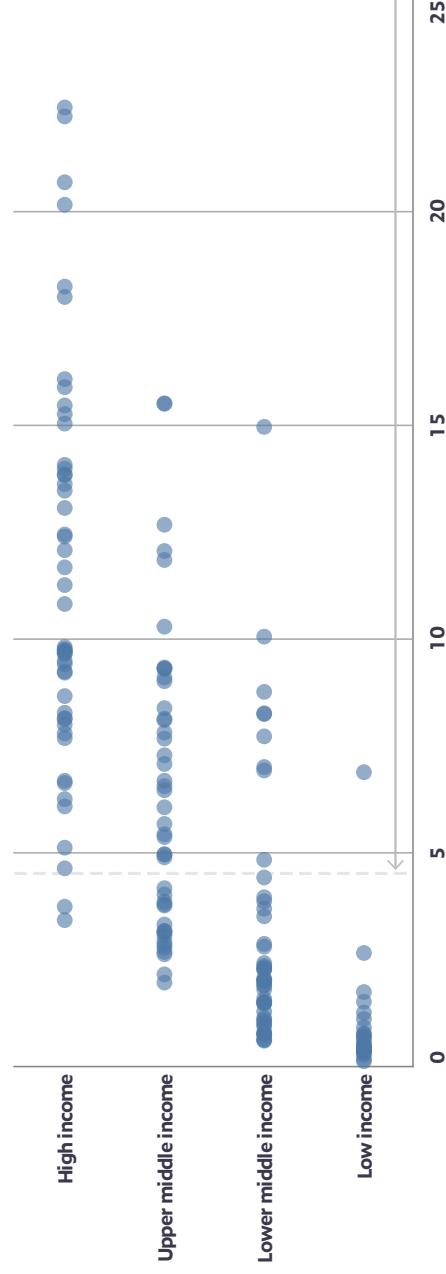
Frequency of data collection: Annual.

Suggested comparator:

- In theory, the relative straightforward and standardized measurability of these indicators should allow for comparability across countries
- It may also be of interest to policy makers to compare by region (e.g. province, urban/rural)

Potential source of data:
The WHO compiles data from household and labor force surveys, censuses, and administrative records (<http://apps.who.int/gho/data/node.main?showonly=HWP>). Data can also come directly from national sources including professional registries and HR systems.

Physicians, nurses, and midwives, by country (per 1,000 people)
(Most recent value in 2010–2016)



Benchmarking guidelines

The WHO estimates that at least 2.5 medical staff (physicians, nurses and midwives) per 1,000 people are needed to provide adequate coverage with primary care interventions. Alternatively, the SDG index threshold is 4.45 physicians, midwives and nurses per 1,000 for UHC and SDG fulfilment.

Source: World Bank (2019), *World Development Indicators*. (<https://databank.worldbank.org/source/world-development-indicators>)

Global data is less frequent than national sources and most recent value in a 5 or 10 year period may be needed for cross-country comparisons.

It is important to check how country definitions differ from global practice if comparing to other countries. In practice, data comparability is limited by differences in definitions and training of medical personnel. In particular, community health workers (CHWs) are generally not included which may be problematic in countries that are increasingly shifting tasks to CHWs and relying on them to provide basic health services.

Here too it may be more useful to consider country context – what is needed to reach health system goals and provide equitable access to health care services. This is especially relevant in sparsely populated and hard to reach areas within a country. It would be good to examine this indicator at the subnational level alongside the caseload indicator to assess allocative and technical (worker productivity) efficiency.

Hospital bed density

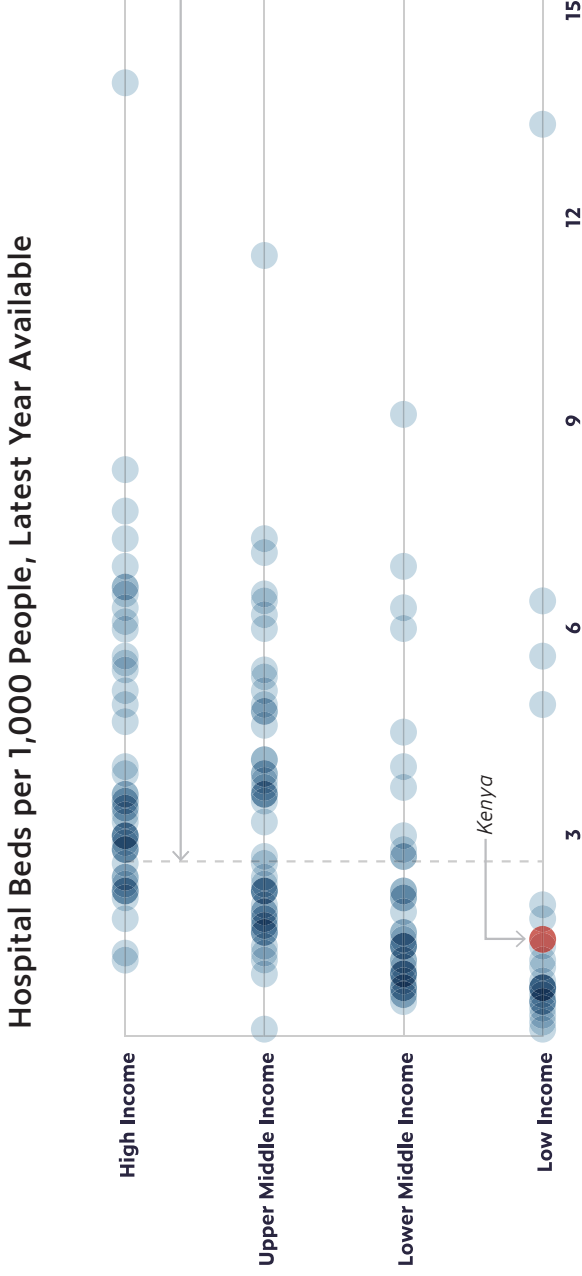
Simple definition: Number of hospitals or hospital beds per 1,000 population.

Frequency of data collection: Annual.

Suggested comparator:

- In theory, the relative straightforward and standardized measurability of these indicators should allow for comparability across countries.
- It may also be of interest to policy makers to compare by region (e.g. province, urban/rural).

Hospital Beds per 1,000 People, Latest Year Available



Source: WHO, various years

Global data is less frequent than national sources and most recent value in a 5 or 10 year period may be needed for cross-country comparisons.

A major limitation of hospital density is that it does not take into account the size of facilities and it may be more informative to look at inpatient bed density. However, differences in definitions and whether maternal beds and/or faith-based beds are included may under/overestimate numbers. Therefore, it is important to check how country definitions differ from global practice if comparing to other countries.

Similar to human resources, it may be more useful to consider country context – what is needed to reach health system goals and provide equitable access to health care services. Here, it would be good to examine bed density at the subnational level alongside the bed occupancy rate indicator to assess allocative and technical efficiency.

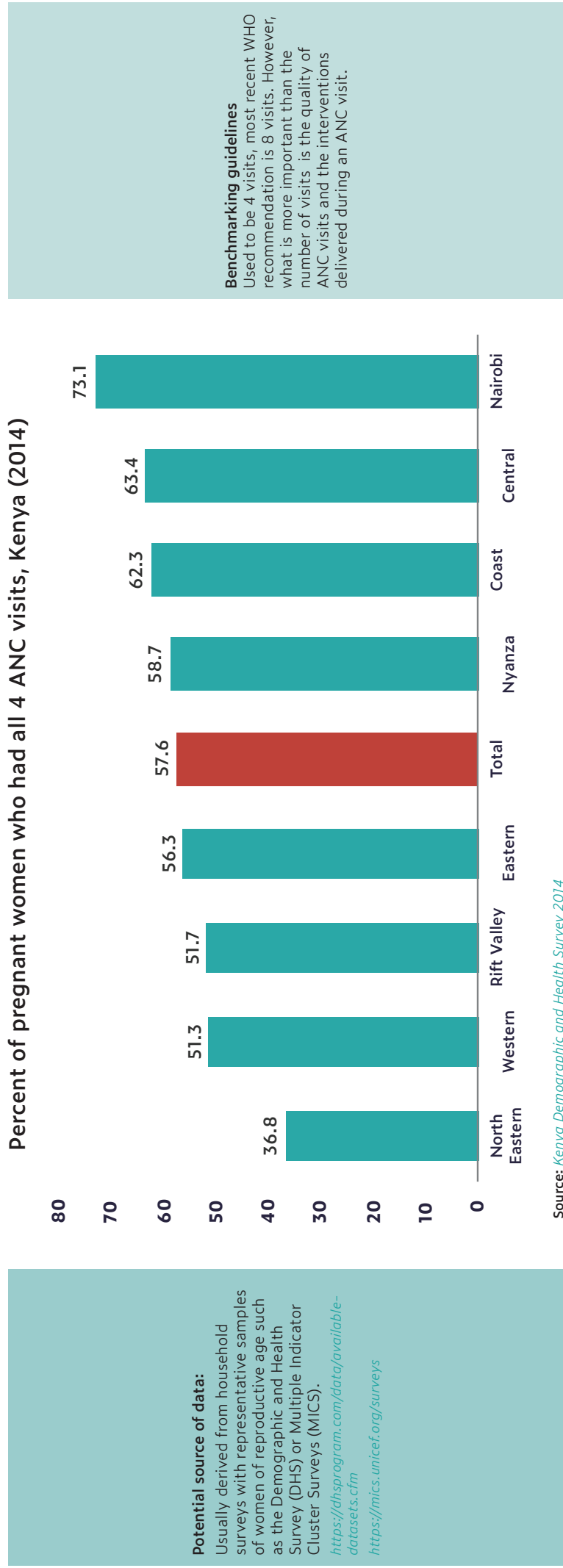
Antenatal care

Outputs (access)

Simple definition: Percentage of women aged 15-49 with a live birth within a given time period that received antenatal care provided by a skilled health professional (e.g. physicians, midwives, nurses) during their pregnancy. Antenatal care includes screening or providing treatment to prevent or reduce the risk of poor pregnancy outcomes.

Frequency of data collection: Usually every 5 years.

Suggested comparator: Across countries. It may also be of interest to policy makers to compare by region (e.g. province, urban/rural) or socioeconomic status (e.g. wealth quintile, level of education). Past performance/trend data.



Source: Kenya Demographic and Health Survey 2014

If available, it would be important to look for information on ANC service-specific availability and readiness and adherence to quality assurance protocols/guidelines from facility surveys such as WHO's Service Availability and Readiness Assessment (SARA) (https://www.who.int/healthinfo/systems/sara_introduction/en/) or the World Bank's Service Delivery Indicators (SDI) (<https://www.sdindicators.org>). DHS surveys also usually have a detailed breakdown of the components of antenatal care.

Skilled birth attendance

Outputs (access)

Simple definition: Percentage of live births attended by skilled health personnel during a specified time period.

Frequency of data collection: Usually every 5 years.

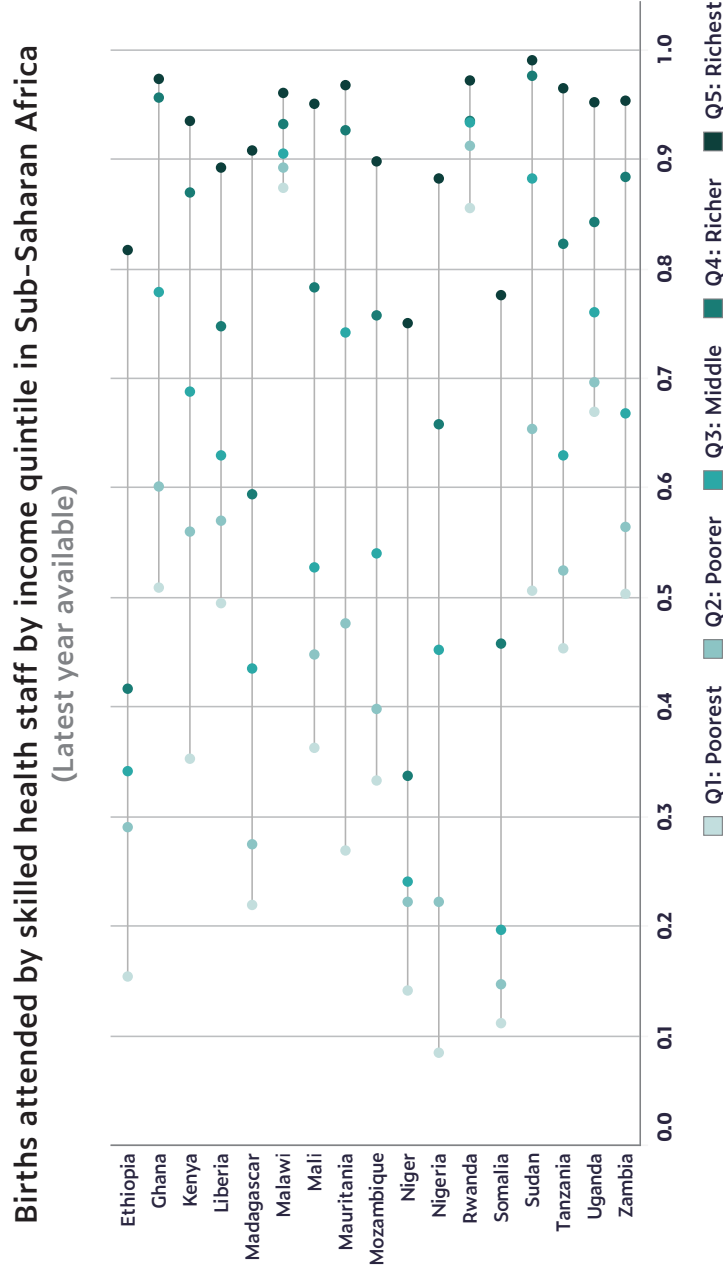
Suggested comparator: Across countries. It may also be of interest to policy makers to compare by region (e.g. province, urban/rural) or socioeconomic status (e.g. wealth quintile, level of education). Past performance/trend data.

Potential source of data:

Usually derived from household surveys with representative samples of women of reproductive age such as the Demographic and Health Survey (DHS) or Multiple Indicator Cluster Surveys (MICS).

<https://dhsprogram.com/data/available-datasets.cfm>

<https://mics.unicef.org/surveys>



Source: *WB (2019). HEPPI database (<http://datatopics.worldbank.org/health-equity-and-financial-protection>)*

Benchmarking guidelines
No established benchmark. Here too, what is more important than the percentage of births attended by a skilled health personnel is the quality of maternal and newborn care.

There are some issues in defining who is included as a skilled health professional. In particular, community health workers (CHWs) are generally not included which may be problematic in countries where CHWs are used as main service provider to deliver babies. When considering whether to include CHWs or not, it is encouraged to look at what training each cadre has and what services they are meant to provide. If available, it would be important to look for information on service-specific availability and readiness for basic obstetric and newborn care, adherence to quality assurance protocols/guidelines, and provider knowledge (e.g. ability to diagnose, treat, and manage life threatening complications such as post-partum hemorrhage and birth asphyxia) from facility surveys such as WHO's Service Availability and Readiness Assessment (SARA) (https://www.who.int/healthinfo/systems/sara_introduction/en/) or the World Bank's Service Delivery Indicators (SDI) (<https://www.sdiindicators.org>)

TB case detection, TB notification rate, and TB treatment success rate

Simple definition:

TB case detection: Percentage of estimated new and relapse TB cases detected in a given year.

TB notification rate: Number of new and relapsed TB cases notified in a given year. A notification means that the case was reported on the national surveillance system.

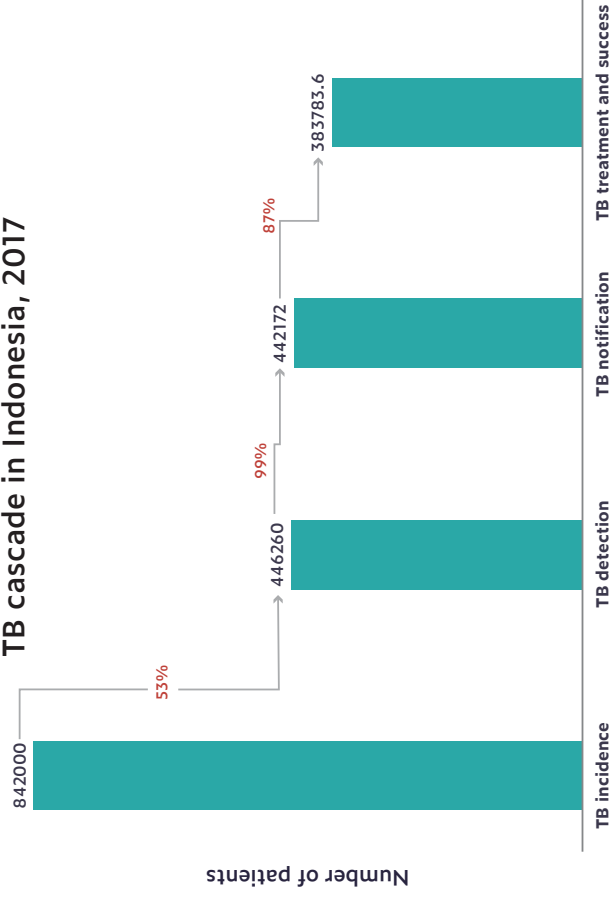
TB treatment success rate: Percent of TB cases that were successfully treated out of all notified TB cases during a specified period of time.

Frequency of data collection: Administrative data could be collected monthly, yearly. Facility surveys are usually ad hoc when funding is available. Inventory or logistics management systems can produce daily/real-time reports.

Suggested comparator: Across countries or compared to past performance/trend data. It may also be of interest to policy makers to compare by region (e.g. province, urban/rural), across facility type (primary health clinic vs hospital) or provider type (public/private).

Potential source of data:
Administrative data from facilities, surveillance systems, TB registers. Annual case notifications are reported by countries to the WHO and then reported in TB annual reports and in the country profiles. <https://www.who.int/tb/country/data/profiles/en/>

TB cascade in Indonesia, 2017



Source: *Global Tuberculosis Report 2018* (https://www.who.int/tb/publications/global_report/en/)

Benchmarking guidelines
For TB case detection and notification rate there is no established benchmark, but the higher the better. Suggest benchmarking against past performance. For TB treatment success rate – at least 90%.

It would be helpful to look at indicators along the entire TB cascade to ensure continuity of care. TB treatment success rates will only be a good measure of a performance if detection of TB is also high. Treatment success rates also rely on the outcome of the treatment being recorded in routine information systems and so treatment success rates will also rely on having robust systems. Routine case notifications may be biased if the private sector does not/under-reports new cases. Surveillance system may also only register bacteriologically confirmed cases – especially if testing and diagnostic capacity is weak.

If available, it would be important to look for information on TB service-specific availability and readiness and provider knowledge such as diagnostic accuracy of TB, and adherence to quality assurance protocols/guidelines from facility surveys such as WHO's Service Availability and Readiness Assessment (SARA) (https://www.who.int/healthinfo/systems/sara_introduction/en/) or the World Bank's Service Delivery Indicators (SDI) (<https://www.sdiindicators.org>)

A high TB notification rate may also act as good proxy for the effectiveness of health surveillance systems overall.

General service readiness; Service specific readiness

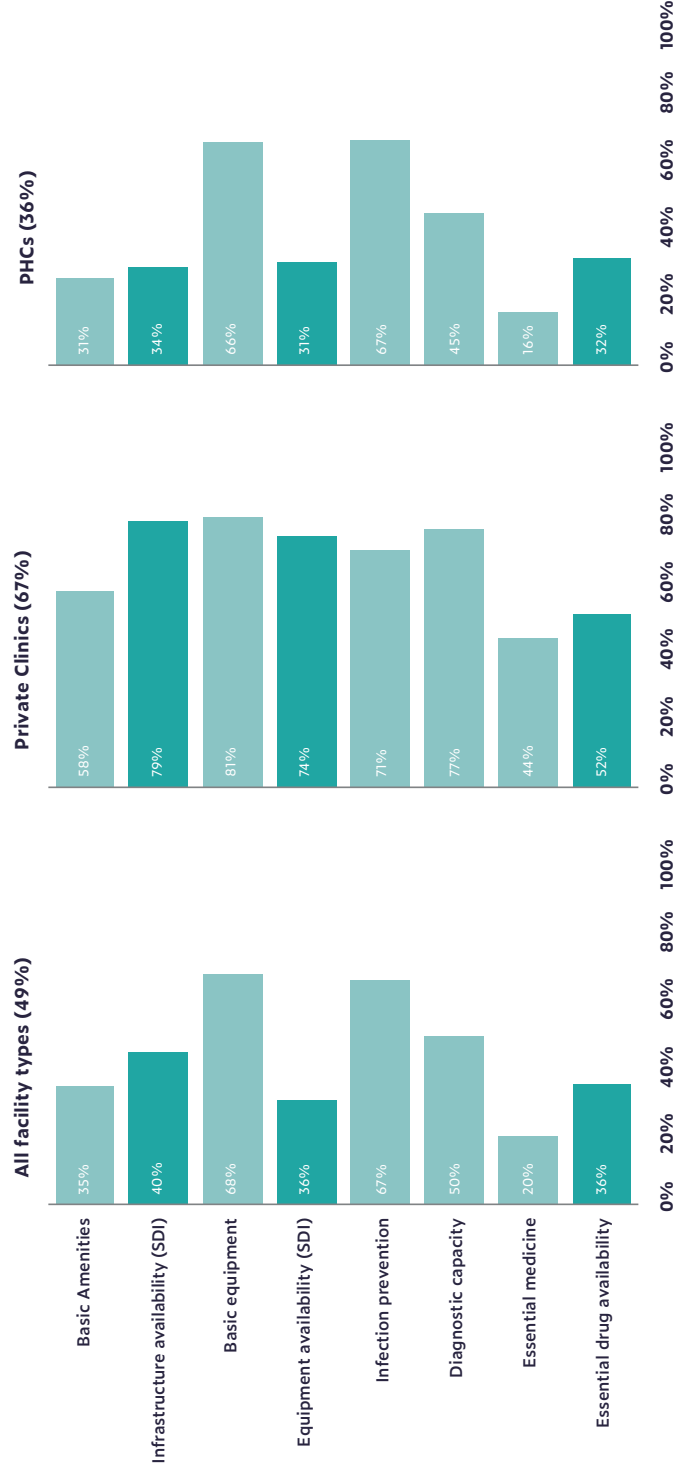
Simple definition: Measures the capacity of health facilities to provide basic services based on the availability of selected tracer items for basic amenities, equipment, essential medicines, diagnostic capacity, and standard precautions for infection prevention.

Frequency of data collection: Survey data is usually done every 5 years. Facility surveys are usually ad hoc when funding is available.

Suggested comparator: Across facilities and/or providers. Across different types of services.

Potential source of data:
Facility surveys such as WHO's Service Availability and Readiness Assessment (SARA) or World Bank's Service Delivery Indicators (SDIs); administrative data.
<https://dhsprogram.com/data/available-datasets.cfm>
<https://mics.unicef.org/surveys>

General service readiness by facility type, Nigeria.



Benchmarking guidelines
No established benchmark, the higher the better.
Suggest benchmarking against past performance, average facility/provider, or best performer.

Note: SARA and SDI surveys have slightly different definitions/checklist items for basic amenities/infrastructure availability, equipment, and essential medicines
Source: Nigeria Health Facility Survey 2015–2016

General service readiness and service specific readiness describes the availability of the components needed to deliver services such as basic amenities, drugs, and equipment. SARA, SDI and SPA protocols are increasingly being standardized to allow greater comparability between countries. Nevertheless, standard definitions may not make sense for the service delivery context and assessment against a country's own standards and protocols is also advised. While often used as a proxy for quality – service readiness only measures availability of tracer items. It is suggested to complement this measure with others such as diagnostic accuracy, adherence to clinical guidelines, etc.

Diagnostic accuracy for a tracer condition

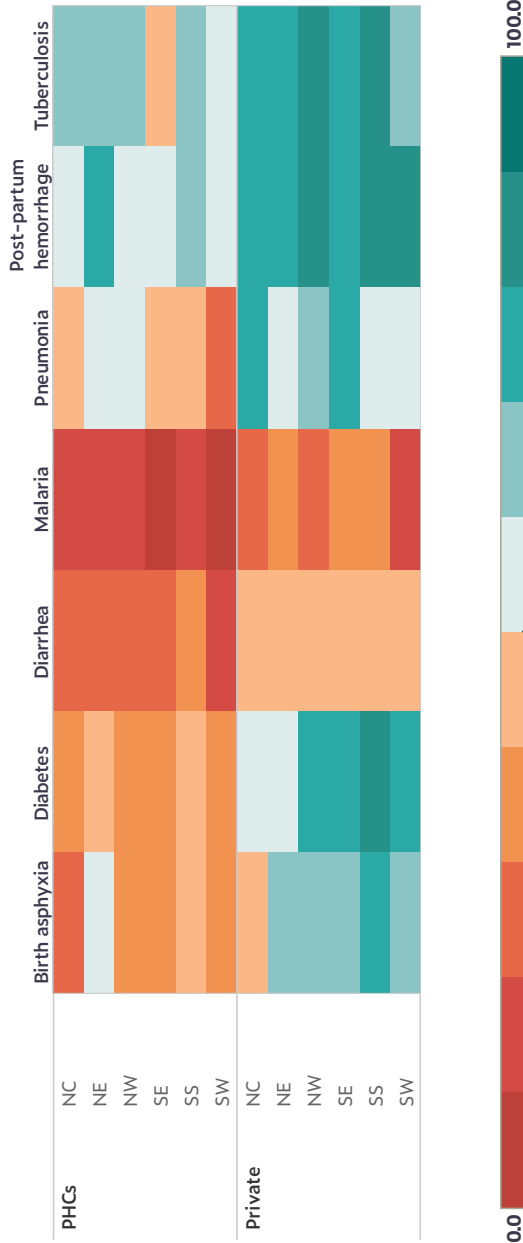
Simple definition: Providers correctly diagnose condition X based on vignette survey data.

Frequency of data collection: Facility surveys are usually ad hoc when funding is available.

Suggested comparator: Across facilities, providers, conditions.

Potential source of data:
Facility surveys such as WHO's Service Availability and Readiness Assessment (SARA) https://www.who.int/healthinfo/systems/sara_reports/en/ or World Bank's Service Delivery Indicators (SDIs) <https://www.sdindicators.org/indicators>

Diagnostic accuracy by condition, provider type, and region (%), Nigeria, 2016



Benchmarking guidelines
No established benchmark, the higher the better. Suggest benchmarking against past performance average provider, or best performer.

Note: Nigeria health facility survey, 2015–2016
Source: Inventory management system, Country X, 2017

Protocols are increasingly being standardized to allow greater comparability between countries. However, given provider training and treatment protocols may vary, countries have also convened focus groups to discuss vignette interpretation and make findings more relevant to local service delivery context. For this reason it may be better to compare diagnostic accuracy within country across regions, facilities, providers.

To better assess provider knowledge, it is also recommended to look at additional indicators such as adherence to clinical guidelines, treatment accuracy, and ability to manage certain tracer conditions which are generally also collected as part of the facility survey vignettes.

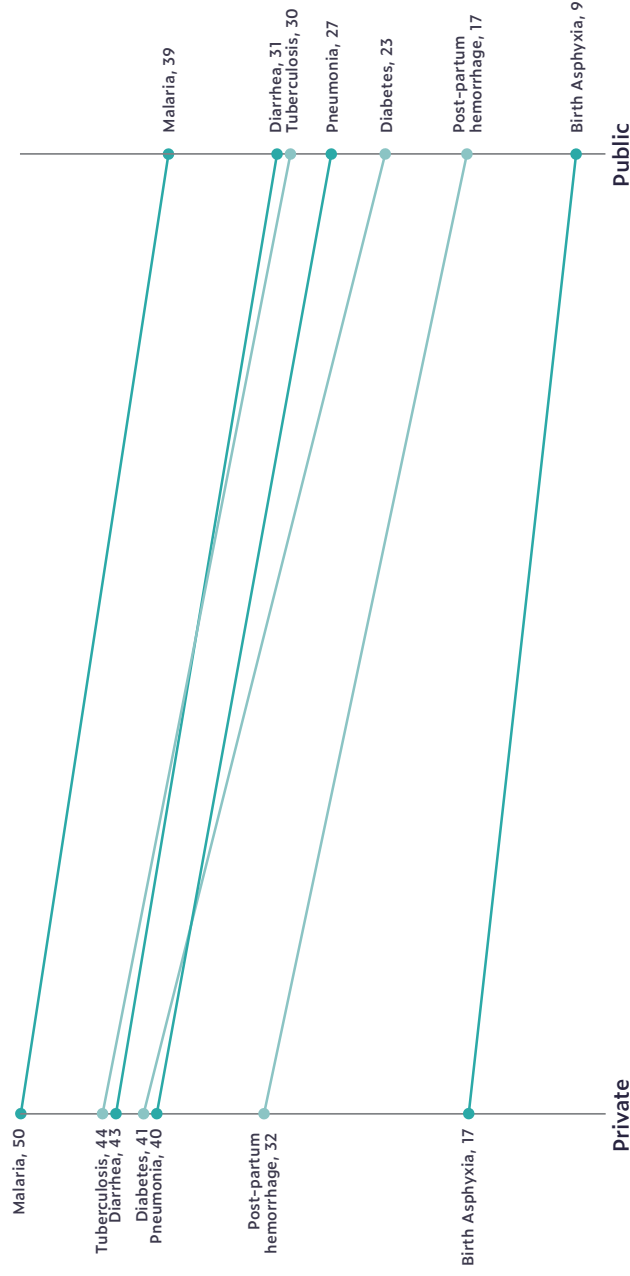
Adherence to clinical guidelines

Simple definition: Whether providers ask or adhere to the questions and exams they should be asking or doing when checking a patient for a particular condition.

Frequency of data collection: Survey data is usually done every 5 years. Facility surveys are usually ad hoc when funding is available.

Suggested comparator: Across facilities and/or providers. Across different types of services.

Adherence to clinical guidelines by condition %, Nigeria.



Source: *Nigeria Health Facility Survey 2015-2016*

Potential source of data:

- Facility surveys such as WHO's Service Availability and Readiness Assessment (SARA) https://www.who.int/healthinfo/systems/sara_reports/ or World Bank's Service Delivery Indicators (SDIs) <https://www.sdiindicators.org/indicators>
- Household surveys such as the Demographic and Health Survey (DHS) <https://dhsprogram.com/data/available-datasets.cfm>. The UNICEF Multiple Indicator Cluster Survey (MICS) <https://mics.unicef.org/surveys> can also be used if guidelines are explicit as to what services should be provided for certain types of services e.g. antenatal care visits.

Benchmarking guidelines
No established benchmark, the higher the better. Suggest benchmarking against past performance, or best performer average facility/provider.

Important precursors to being able to measure this indicator include clearly defined treatment protocols, trained service providers, and good documentation or information systems to be able to monitor adherence. Adherence to clinical guidelines may also be influenced by multiple system factors. It is suggested to complement this measure with service readiness measures such as availability of drugs; diagnostic capacity and provider knowledge measures such as diagnostic accuracy.

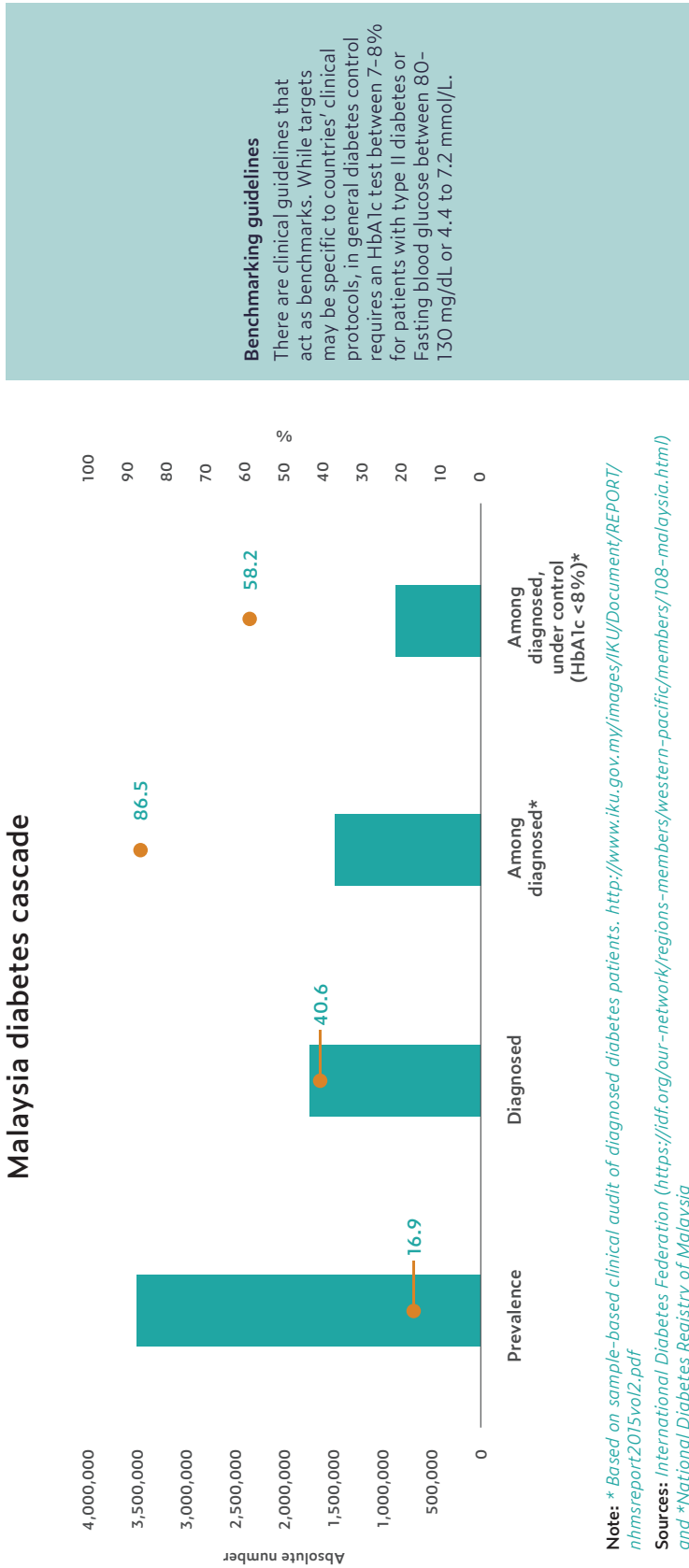
Diabetes control

Simple definition: Percentage of patients who have their diabetes under control. People with diabetes generally have high levels of glucose in the blood measured by a hemoglobin A1c>6.5% and/or blood glucose >126mg/dL.

Frequency of data collection: Analyzed monthly, quarterly, annually. Survey data usually every 5 years.

Suggested comparator: Across facilities/hospitals; against past performance; across regions within country.

Potential source of data:
Medical records, administrative records or surveillance systems like national diabetes patient registries. Household surveys such as the Demographic and Health Survey (DHS), the UNICEF Multiple Indicator Cluster Survey (MICS).



If available, it would be important to look for information on i) diabetes service-specific availability and readiness (i.e. basic items needed to test and treat diabetes), ii) adherence to quality assurance protocols/guidelines, and iii) provider knowledge measures such as diagnostic accuracy from facility surveys such as WHO's Service Availability and Readiness Assessment (SARA) or the World Bank's Service Delivery Indicators (SDI). It would also be helpful to look at a range of indicators along the entire diabetes cascade (e.g. diabetes prevalence; percent on treatment; percent under control; diabetes-related hospital admissions; and mortality due to diabetes)

Avoidable hospital admissions for primary care sensitive conditions

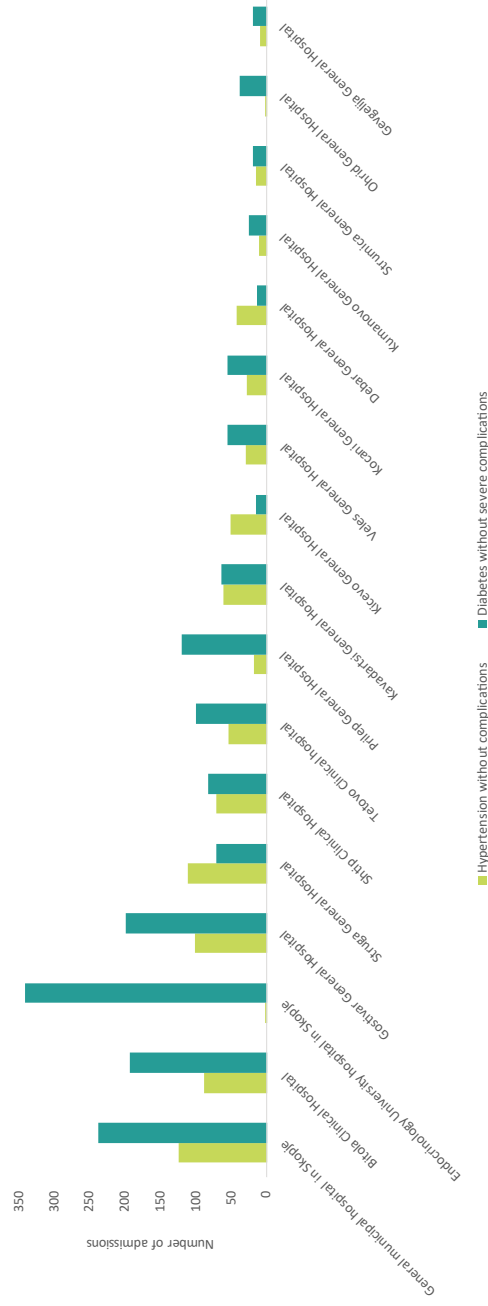
Simple definition: Potentially avoidable admissions for chronic illnesses that could be averted if the patients had quality outpatient care. Conditions for which hospital admissions could be prevented by interventions in community settings.

Frequency of data collection: Usually analyzed monthly, quarterly or annually. Households surveys are less frequently available and not suggested for routine monitoring.

Suggested comparator: Across facilities/hospitals. It may be useful for policy makers to have comparisons aggregated at subnational level (e.g. provinces, districts).

Potential source of data:
Ideally claims data and/or medical records. Some household surveys with a detailed health module.

Hypertension and diabetes admission by facility, North Macedonia, 2016



Source: *Health Insurance Fund of North Macedonia*

Benchmarking guidelines
No benchmark, the lower is better.
Suggest benchmarking against past performance or best performer average facility/provider.

OECD countries use a standardized measure to allow for cross-country comparison. Outside of the OECD, within country hospital/inpatient comparisons are recommended as definitions (i.e. what is considered an avoidable condition) may vary. Conditions generally considered are diabetes, hypertension, and acute respiratory conditions such as bronchitis, asthma, and chronic obstructive pulmonary disease as most are not associated with complications and should not require hospital admissions. Within country comparison is also recommended as admission and treatment protocols and the readiness and quality of primary health care to appropriately manage these conditions may vary. As the quality of claims data improves, case-mix severity should also be taken into account.

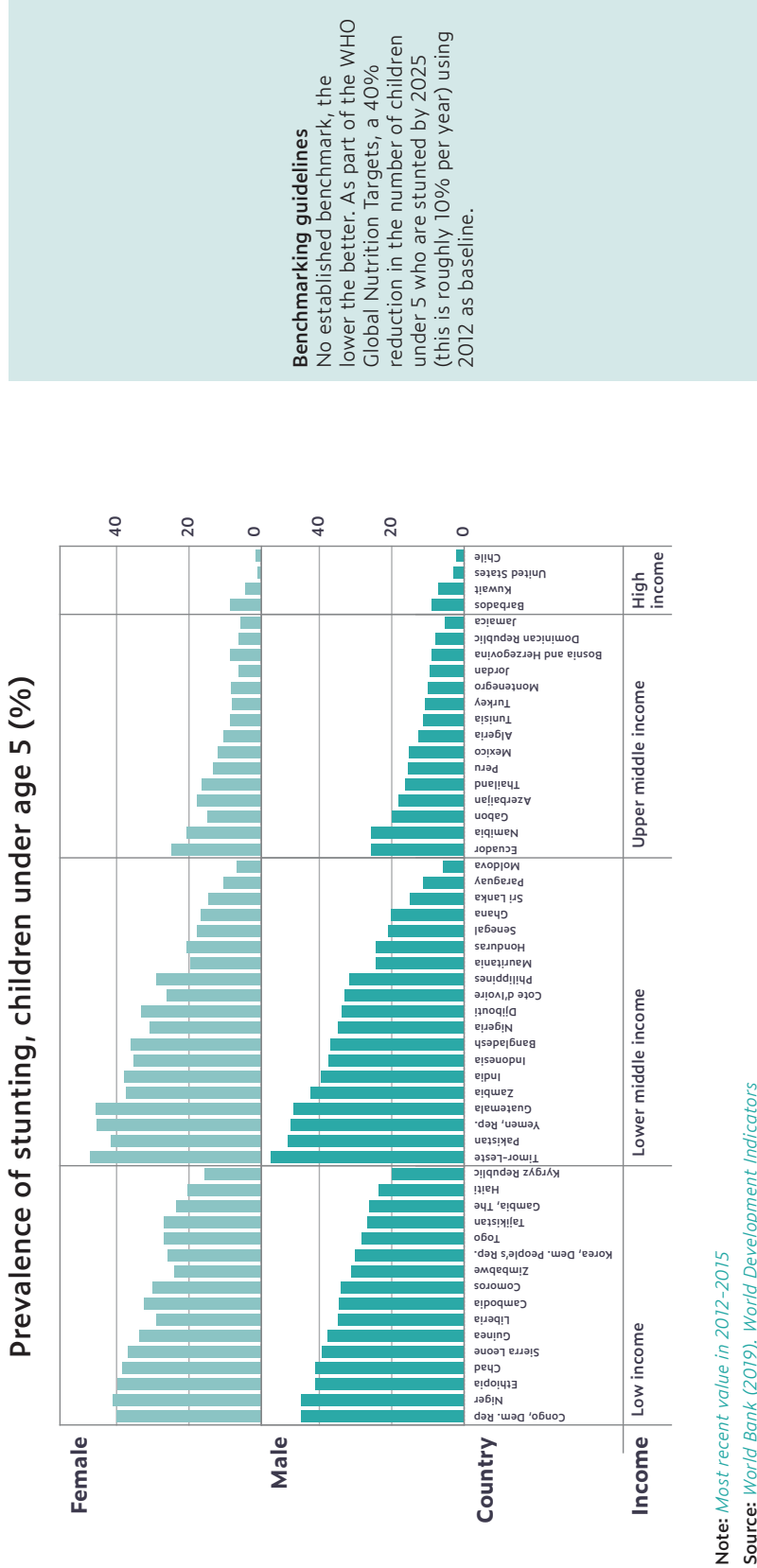
Children under 5 who are stunted

Simple definition: Percentage of stunted (moderate and severe) children aged 0–59 months (moderate = height-for-age below -2 standard deviations from the WHO Child Growth Standards median; severe = height-for-age below -3 standard deviations from the WHO Child Growth Standards median).

Frequency of data collection: Survey data usually every 5 years. UNICEF–WHO–The World Bank: Joint Malnutrition Estimates updated annually.

Suggested comparator: Across countries. It may also be of interest to policy makers to compare by region (e.g. province, urban/rural) or socioeconomic status (e.g. wealth quintile, level of education). Past performance/trend data.

Potential source of data: Potential Source Household surveys such as the Demographic and Health Survey (DHS) (<https://dhsprogram.com>), the UNICEF Multiple Indicator Cluster Survey (MICS) (<https://mics.unicef.org>). National census surveys. Global databases for country comparisons: the Human Capital Project uses data on stunting for construction of its HCI from the UNICEF–WHO–The World Bank: Joint Child Malnutrition Estimates. (<https://www.worldbank.org/en/publication/human-capital>).



Note: Most recent value in 2012–2015

Source: World Bank (2019). *World Development Indicators*

Stunting reflects poor nutrition and infection exposure during pregnancy and after birth. Therefore, if available, it would be important to look for information on ANC and child health service-specific availability and readiness, diagnostic accuracy of common childhood conditions like diarrhea, and adherence to quality assurance protocols/guidelines from facility surveys such as WHO's Service Availability and Readiness Assessment (SARA) (https://www.who.int/healthinfo/systems/sara_introduction/en/) or the World Bank's Service Delivery Indicators (SDI) (<https://www.sdindicators.org>).

Tobacco use among persons aged 18+

Outputs (risk factors)

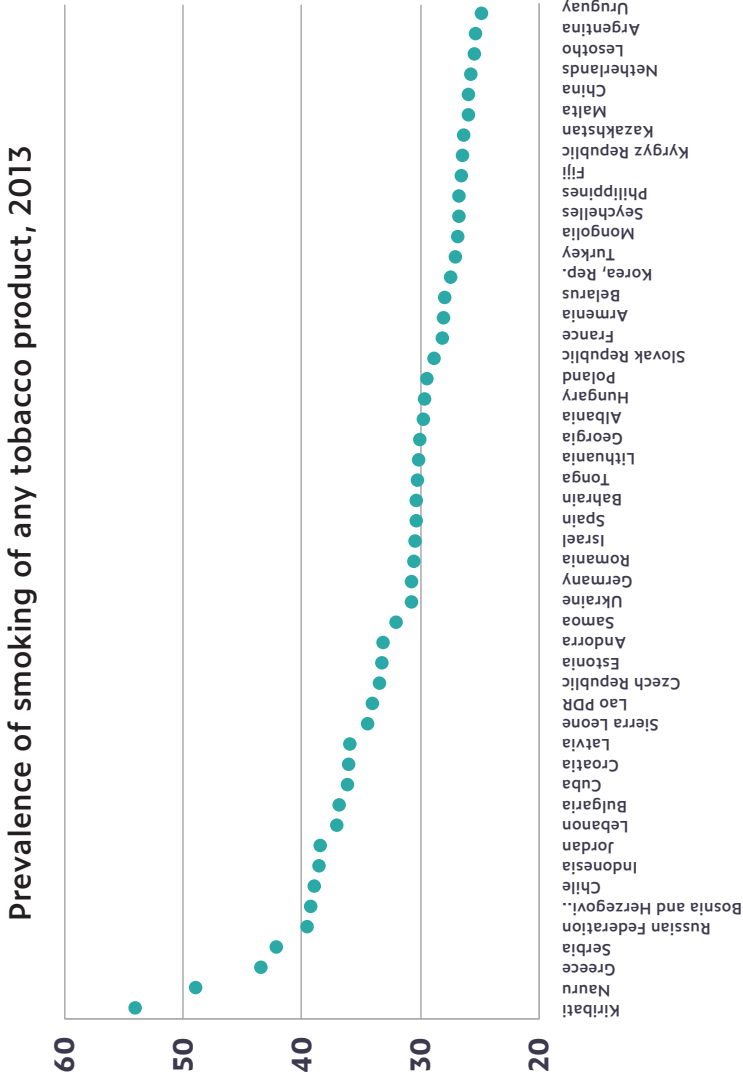
Simple definition: Prevalence of current tobacco use among persons aged 18+ years.

Frequency of data collection: Survey data usually every 5 years.

Suggested comparator: Across countries. It may also be of interest to policy makers to compare by socioeconomic status (e.g. wealth quintile, level of education, gender). Past performance/trend data.

Potential source of data:
Household surveys such as the Demographic and Health Survey (DHS), the UNICEF Multiple Indicator Cluster Survey (MICS), World Health Organization STEPwise (STEPS) survey and national surveys.

<https://dhsprogram.com>
<https://mics.unicef.org>
<https://www.who.int/ncds/surveillance/steps/reports/en/>



Benchmarking guidelines
No established benchmark, the lower the better. Global voluntary target of 30% relative reduction in the use of tobacco in persons over 15 years by 2020 (reference year: 2013).

Note: Top 50 highest countries
Source: World Bank (2019). World Development Indicators (<http://wdi.worldbank.org>)

Surveys may have different inclusion criteria based on age. For instance, STEPS survey will generally ask those 18 years or older whereas DHS asks those who are 15 years or older. Current use definitions may also vary depending on the survey. Tobacco use can be both smoke and non-smoked based used.

Caseload

Simple definition: Number of outpatient visits per clinician per day.

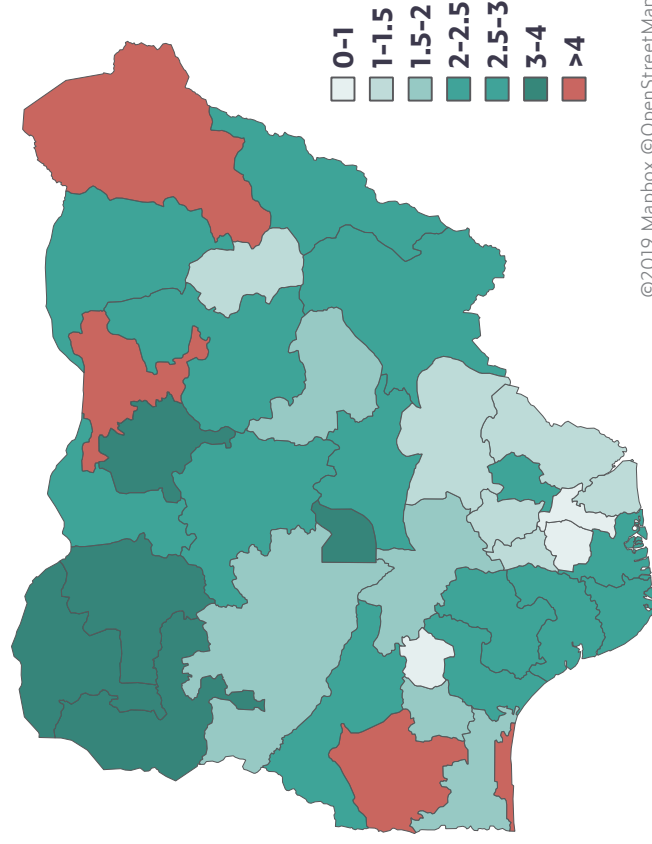
Frequency of data collection: Facility surveys are usually ad hoc when funding is available. Administrative/facility records are generally analyzed monthly, quarterly, or annually.

Suggested comparator: Across facilities, types of providers.

Potential source of data:

Facility surveys such as WHO's Service Availability and Readiness Assessment (SARA: https://www.who.int/healthinfo/systems/sara_reports/en/) or World Bank's Service Delivery Indicators (SDIs: <https://www.sdiindicators.org/indicators>); administrative/facility records.

Caseload



Source: *World Bank (2016). Nigeria National Health Facility Survey*

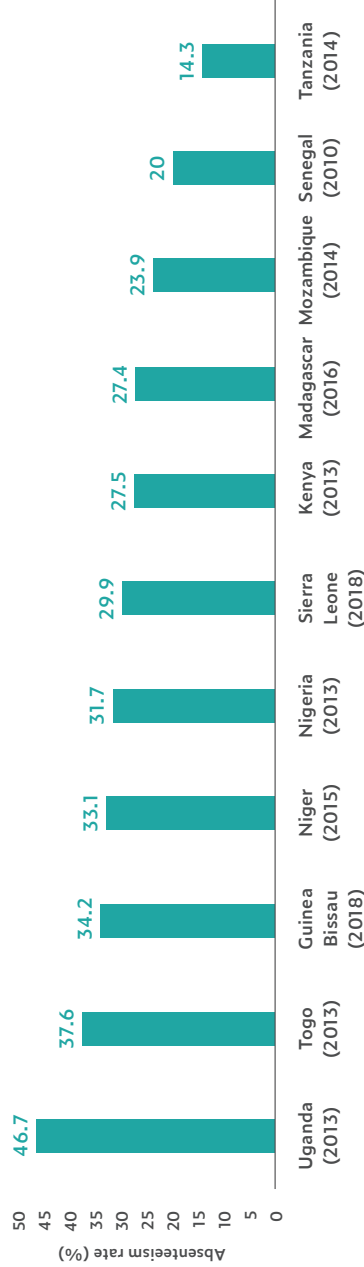
Caseload is a complex indicator to interpret. It could be affected by many supply side issues such as the over- or under-supply of human resources the salary and timeliness of payment to providers, absenteeism rates, the number of patients registered to a primary health care facility, the availability of drugs and equipment, the functionality of facilities, the service provision responsibilities of different cadres or by demand side issues such as a preference for different provider types or settings, the disease burden, the affordability of care, the distance to facilities, or the perceived quality of care. This makes it difficult to compare across countries as the health care context may be different. For example, community health workers (CHWs) are not included in the WB's SDI definition which may be problematic in countries that are increasingly shifting tasks to CHWs and relying on them to provide basic health services. Even within countries, it is recommended to pair this indicator with others (e.g. absenteeism, delay of payment, number of patients registered to a facility or provider, average number of patient visits within a year etc.) to identify the appropriate policy action. While caseload is generally an outpatient/primary health care measure, other measures of productivity such as time per patient per visit (though less common) have also been used.

Absenteeism

Simple definition: Proportion of health professionals who are absent and not off-duty from the facility on an unannounced visit.
Frequency of data collection: Administrative data is usually analyzed monthly, quarterly, and/or annually. Facility surveys are usually ad hoc when funding is available.
Suggested comparator: Across facilities, providers.

Potential source of data:
Administrative data. Facility surveys such as WHO's Service Availability and Readiness Assessment (SARA) (https://www.who.int/healthinfo/systems/sara_introduction/en/).
Or World Bank's Service Delivery Indicators (SDIs) (<https://www.sdiindicators.org/>).

Absenteeism among primary health care providers, country average (%)



Source: Administrative hospital data, Country X, 2017

Benchmarking guidelines
No established benchmark, the lower the better. Suggest benchmarking against past performance or average facility/provider.

Only un-excused absences should be accounted for – for example if provider is on a training, sick, or on leave they do not count as absent. It is also important to check which providers are counted when doing cross country comparators.

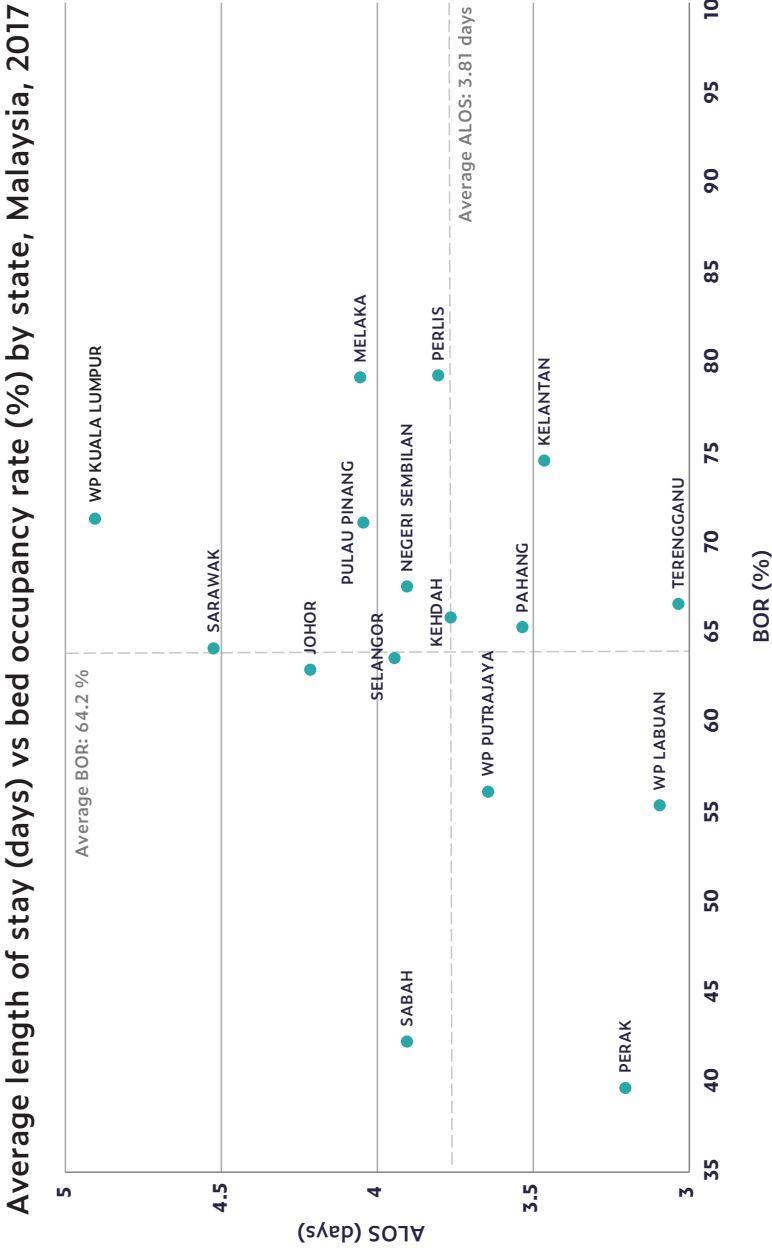
Average length of stay (ALOS) vs. Bed occupancy rate (BOR)

Simple definition: ALOS: Total average number of bed-days for a hospital admission or discharge. Bed-days refer to the number of beds in the facility multiplied by the number of days observed.

BOR: The percentage of bed-days that are occupied over a specified period of time. Occupied bed-days refers to the number of days the beds were occupied.

Frequency of data collection: Administrative data is usually collected daily and analyzed monthly, quarterly, and/or annually.

Suggested comparator: Across hospitals, type of hospitals.



Source: Ministry of Health, Malaysia (2017)

Length of stay and by extension bed occupancy rates can be a reflection of the efficiency of hospital services as a shorter stay will reduce costs and shift the care from inpatient to outpatient settings. However, it is difficult to compare this measure across countries as treatment protocols may vary. ALOS may also reflect the readiness and/or quality of outpatient/primary health care. Therefore comparison between hospitals within-country is recommended. As the quality of claims data improves it is recommended to compare ALOS by case-mix.

Hospital re-admission rate

Simple definition: Percentage of unplanned and unexpected hospital readmissions for conditions that are tracked through the health system like acute myocardial infarction, pneumonia, asthma and diabetes.

Frequency of data collection: Usually analyzed monthly or annually.

Suggested comparator: Across facilities/hospitals.

Potential source of data:
Administrative data, claims data, medical records.

30-day Hospital Readmission rate (%), May 2017



Source: Administrative hospital data, Country X, 2017

Benchmarking guidelines
No established benchmark.
Suggest benchmarking against past performance or average facility.

This indicator is tricky to measure across countries as it is influenced by factors such as treatment protocols, provider incentives, quality of care, and the availability of outpatient and community care. For these reasons, it is suggested to look for outlier facilities or providers within countries or regions. If doing cross-country comparison, it is important to check how readmissions are defined – for what conditions and within what time period. Common time periods are 30-day and 90-day readmissions and common conditions include chronic obstructive pulmonary disease, acute myocardial infarction, pneumonia, asthma, diabetes, hip and knee replacements. Analysis is also dependent on the quality of data, good medical record keeping, coding practices, and reporting compliance. This indicator also requires unique patient identifiers to be able to track patients in the healthcare system.

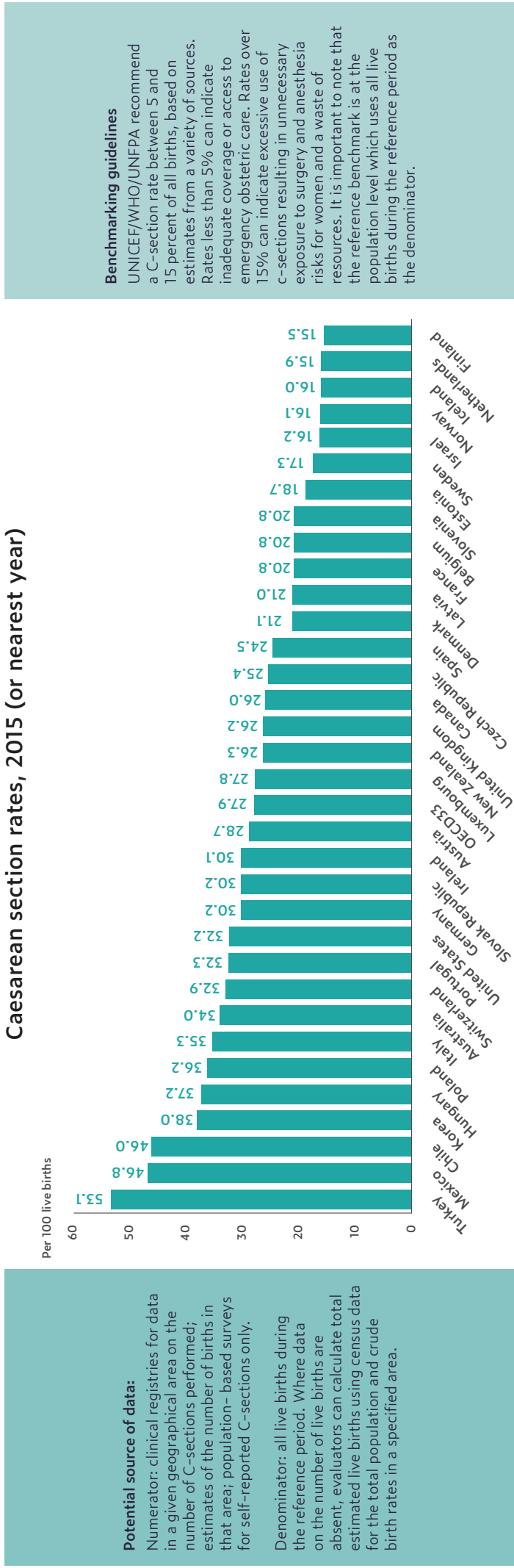
Cesarean section (C-section) rates

Simple definition: The percent of pregnant women who have a C-section in a specific geographical area and reference period.

Frequency of data collection: Administrative data is usually analyzed monthly, quarterly, and/or annually. Survey data is usually every 5 years.

Suggested comparator: Across countries; across facilities/providers; past performance/trend data; by urban/rural; by public/private facility.

Caesarean section rates, 2015 (or nearest year)



As treatment protocols may vary across countries, comparison within country is also recommended. For within country comparison, an alternative indicator is the proportion of facility deliveries that are C-sections. Here, the source of data can be from facility or claims data. The denominator is live births at the facility not all live births within a catchment area. This indicator may be significantly higher than c-section rates at the population level because it does not adjust for higher risk pregnancies and referrals. While specifying an appropriate range of target percentages within a facility is impractical, it might still be interesting to benchmark against average facilities/providers of similar type to look for outliers. The Robson Scale has been applied in some settings to help monitor the rates of c-section amongst subgroups of pregnant women (e.g. high risk versus low risk) however this will likely require access to patient history and medical records.

(Link to Robson Scale:<https://apps.who.int/iris/bitstream/handle/10665/259512/9789241513197-eng.pdf;jsessionid=A048ECA988C32D049AD25B2E2D07E68A?sequence=1>)

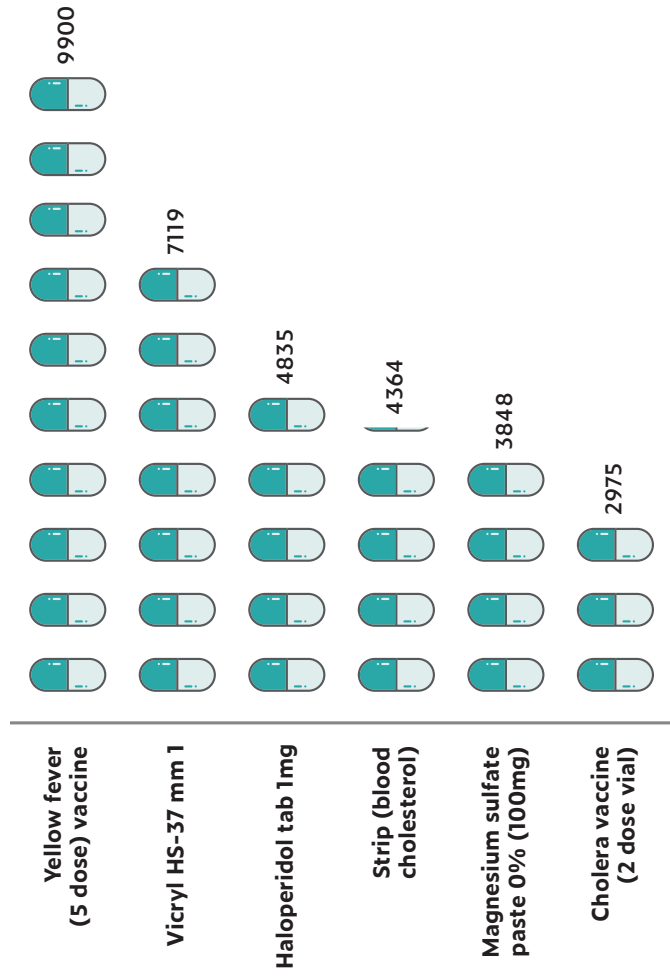
Expired drugs

Simple definition: Proportion of products past their use-by date in stock over the total amount of product in stock. Can also be measured in monetary value if the quantity of discarded expired drug is recorded along with the market price of the drug.

Frequency of data collection: Administrative data could be collected monthly, yearly. Facility surveys are usually ad hoc when funding is available. Inventory or logistics management systems can produce daily/real-time reports.

Suggested comparator: Across facilities, pharmacies, medical storage centers.

Loss due to expired drugs in Hospital A, Top 6 items in USD, 2017



Potential source of data:

Facility based surveys, administrative data, Service Provision Assessments. If available, Logistics Management and Information System (LMIS) or inventory management system.

Benchmarking guidelines

Ideally zero. Suggest benchmarking against past performance or average facility.

Source: *Inventory management system, Country X, 2017*

Expired drugs could be a reflection of weaknesses throughout the entire supply chain from procurement to distribution to inventory management and forecasting at the facility level. For example, long procurement and delivery lead times may incentivize facilities to stockpile drugs; storage and inventory management practices such as a first-in, first-out policy may not be in place; and facilities may overestimate demand for certain products. Other useful indicators to look at might include the order turnaround time (i.e. the number of days between when the order was received and delivered), the frequency of orders and the prescription rate or utilization of commonly expired products.

Drug stockouts

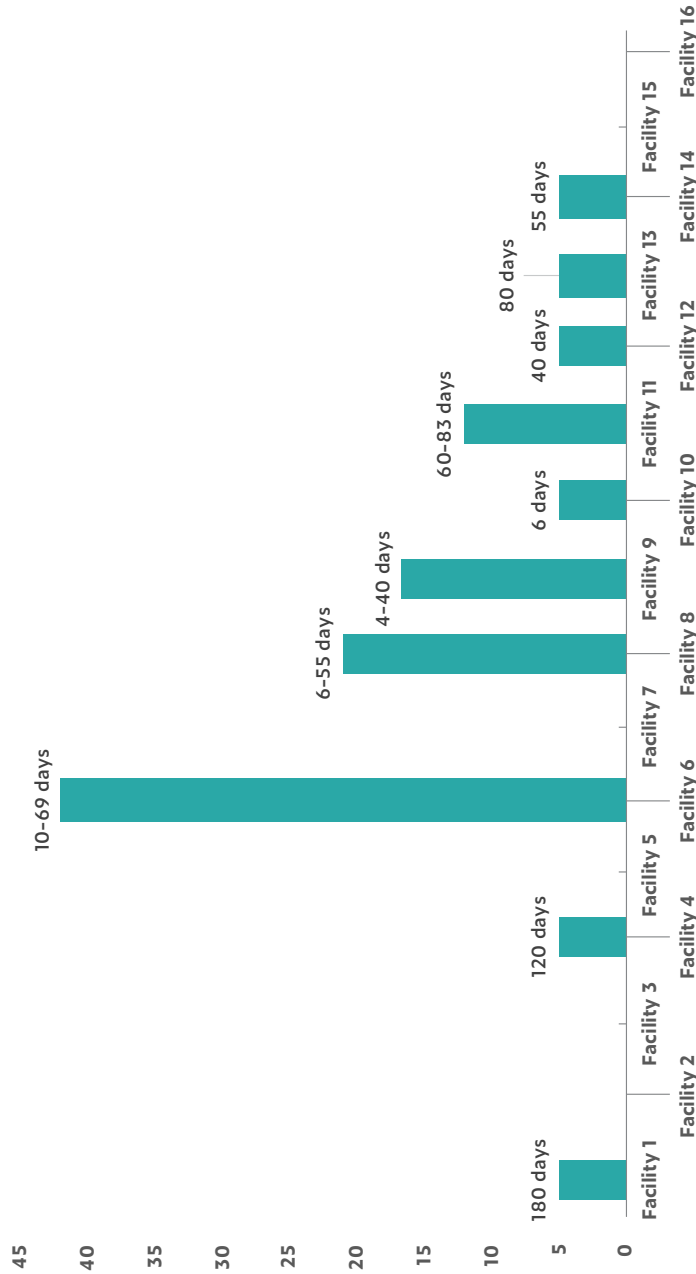
Simple definition: Number or percent of service delivery units that experience a stockout of a particular product that should be in stock over a defined period of time. It could also be defined as number, or percent of products that are unavailable at a particular service delivery unit.

Frequency of data collection: Administrative data could be collected monthly, yearly. Facility surveys are usually ad hoc when funding is available. Inventory or logistics management systems can produce daily/real-time reports.

Suggested comparator: Across facilities, pharmacies, medical storage centers.

Potential source of data:
Facility based surveys, administrative data, Service Provision Assessments. If available, Logistics Management and Information System (LMIS) or inventory management system.

Stock-out rate (%) and number of days items were not available, across facilities, 2017



Benchmarking guidelines
Ideally zero. Suggest benchmarking against past performance or average facility.

Source: *Inventory management system, Country X, 2017*

Just like expired products, drug stockouts could also be a reflection of weaknesses throughout the supply chain. Lack of funds and/or complicated and lengthy procedures to access funds and order drugs could be an issue. Here too, indicators such as the proportion of expired or ruined drugs, the order turnaround time, the frequency of orders, and the prescription rate or utilization of stocked-out items might help facilities/pharmacies/central medical stores better forecast and manage drug inventories.

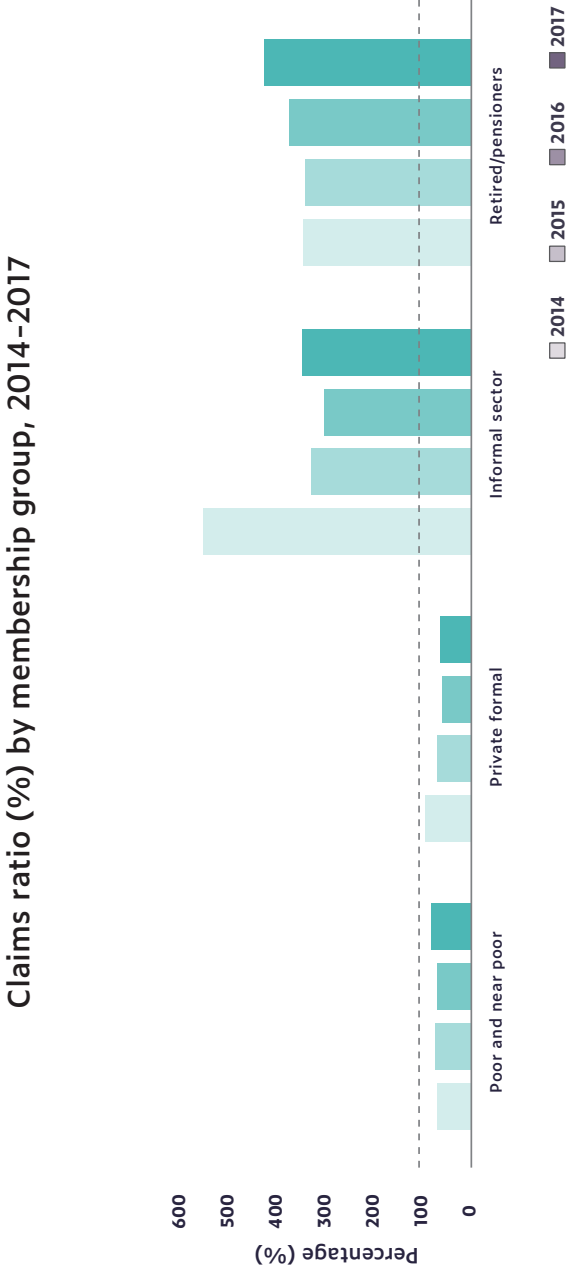
Claims ratio

Simple definition: A ratio calculated as accrued claims divided by accrued premiums. This is sometimes also referred to as the loss ratio.

Frequency of data collection: Monthly, quarterly, annually.

Suggested comparator: By membership type, against past performance. If more than one scheme by insurance scheme.

Claims ratio (%) by membership group, 2014-2017



Source: *National health insurance agency claims and premium database (2019)*

Benchmarking guidelines
The "expected" claims ratio should be 100% minus a margin for administrative expenses and a reasonable level of claim reserves. For a private insurance company, a margin for profit would also be subtracted.
For mature national health insurance systems, one would expect the ratio to be about 90%.

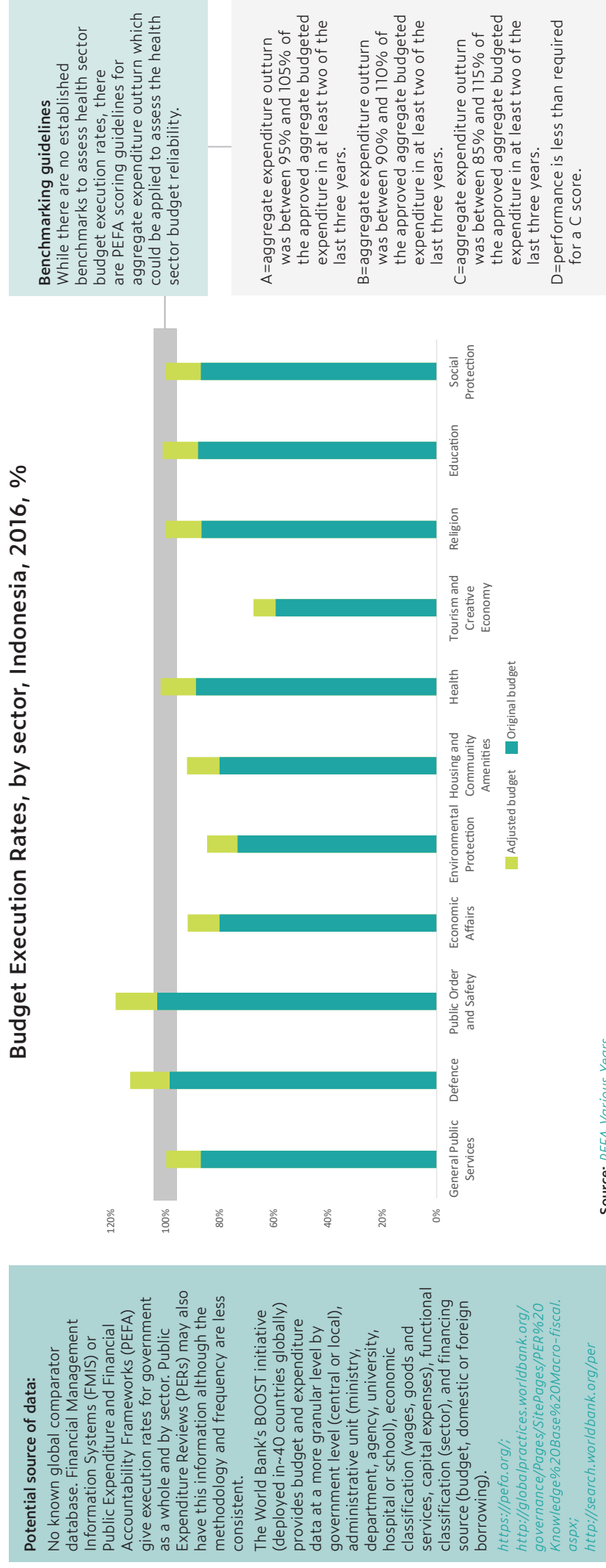
If the claims ratio is regularly over 100% over an extended number of years it means either the premiums are not being properly calculated using sound actuarial principals and/or there are issues on the claims payment and administrative expense side. If comparing across insurance schemes, it is important to account for differences in premiums, membership types, coverage rates, and cost containment features.

Budget execution rates (BERs)

Simple definition: The percentage of the approved budget for health in a given fiscal year that was actually executed (*In the public financial management world this is referred to as budget reliability and measured by aggregate expenditure outturn*).

Frequency of data collection: Budget documents and Financial Management Information System available monthly and annually; Public expenditure and financial accountability frameworks and Public expenditure reviews are less frequent.

Suggested comparator: Across countries; Other sectors/line ministries; by expenditure category (e.g. capital vs recurrent); trend data.



Source: PEFA Various Years

Many of the reasons for poor BERs are generally outside the control of the MOH (e.g. weak capacity to forecast revenues, delays in the release of funds, in-year budget adjustments, or rigidities on how funds can be spent across programs). For this reason while BERs may be a poor measure of the sector performance, it is still a helpful indicator in budget negotiation discussions if health sector policy makers understand the rules and processes that influence BERs. A greater granular assessment of execution rates is also extremely helpful. For instance, Ministries that are very human resource intense (like education and health) are likely to have a better execution rate than others because wages are quasi statutory. For this reason, it might make more sense to compare the BER of wages and salaries, non-wage recurrent expenditures, and investment expenditures in the sector separately.

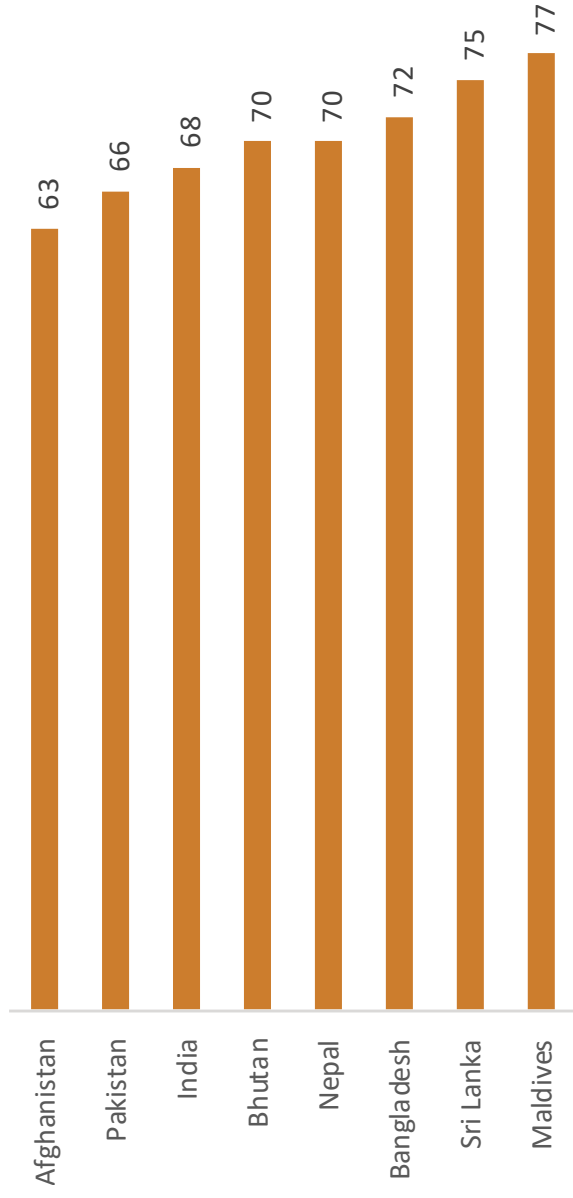
Life expectancy at birth

Simple definition: Average number of years that a newborn is expected to live if current mortality rates continue to apply.

Frequency of data collection: Vital registry data should be annual. Survey data usually every 5 years.

Suggested comparator: The relative straightforward and standardized measurability of these indicators allows for comparability across countries. It may also be of interest to policy makers to compare by socioeconomic status (e.g. wealth quintile, level of education, gender).

Life expectancy (in years), 2017



Potential source of data:

Vital registry data. Household surveys such as the Demographic and Health Survey (DHS), the UNICEF Multiple Indicator Cluster Survey (MICS).

<https://dhsprogram.com>

<https://mics.unicef.org>

Benchmarking guidelines

No established benchmark the higher the better. Suggest also benchmarking against past performance or relevant regional/income country comparators/averages.

Source: *World Bank (2019). World Development Indicators. (<https://databank.worldbank.org/source/world-development-indicators>)*

As a high level outcome indicator, it will be important to choose relevant health system input, output, and process measures in developing a results chain narrative.

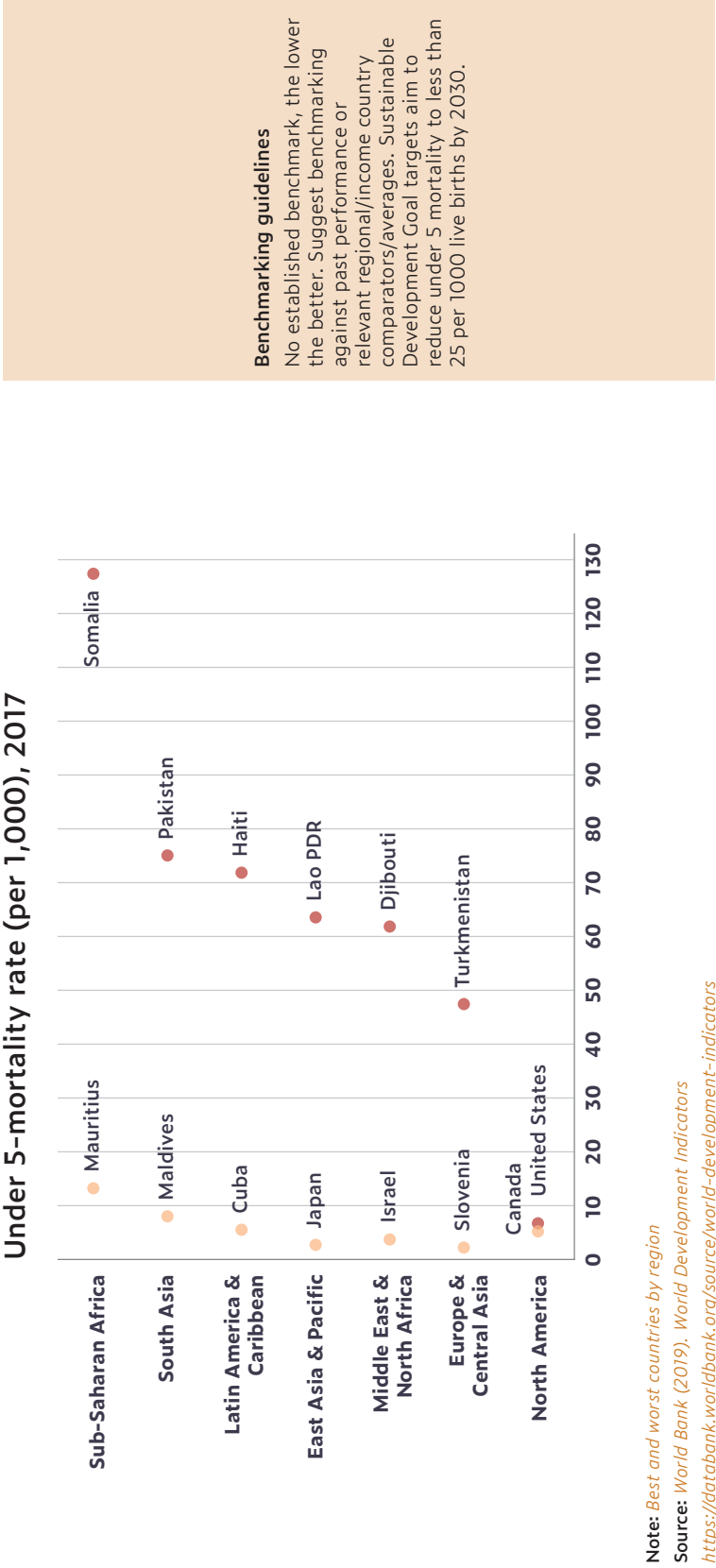
Under-five mortality

Simple definition: Under-five mortality (U5M) rate is the probability that a child will die before reaching age five.

Frequency of data collection: Vital registry data should be annual. Survey data usually every 5 years.

Suggested comparator: The relative straightforward and standardized measurability of these indicators allows for comparability across countries. It may also be of interest to policy makers to compare by region (e.g. province, urban/rural), or socioeconomic status (e.g. wealth quintile, level of education, gender).

Potential source of data:
Vital registry data. Household surveys such as the Demographic and Health Survey (DHS), the UNICEF Multiple Indicator Cluster Survey (MICS).
<https://dhsprogram.com>
<https://mics.unicef.org>



Neonatal (less than 28 days) and infant (less than 1 year) mortality rates are important parts of the overall under-5 mortality rate. Disaggregated analyses can help identify vulnerable populations for better targeting. As most U5M deaths are preventable, it will be important to choose relevant input, output and process indicators in developing results chain narrative. For example, number of nurses, postnatal coverage, fully-immunized children, relevant child-health specific readiness measures (SALA (https://www.who.int/healthinfo/systems/sara_introduction/en/) or SDI (<https://www.sdindicators.org/>), and others such as diagnostic accuracy of common childhood illnesses like diarrhea or pneumonia, and adherence to clinical guidelines.

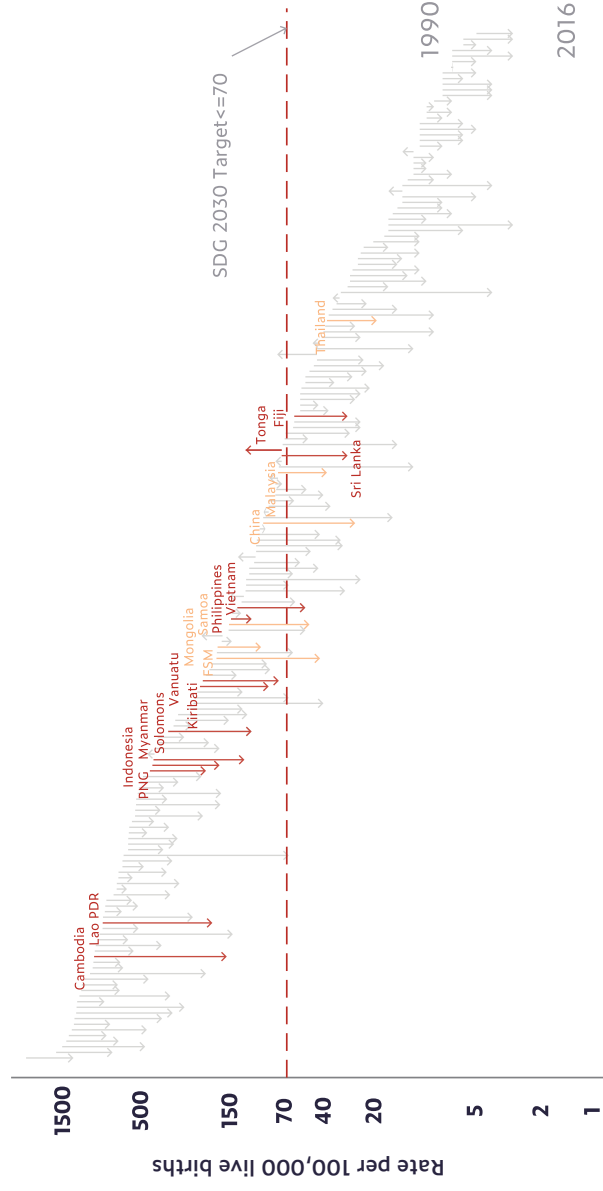
Maternal mortality rate

Simple definition:The annual number of maternal deaths from any cause related to or aggravated by pregnancy or its management, relative to the total number of live births. The indicator is reported as deaths per 100,000 live births.

Frequency of data collection: Vital registry data should be annual. Survey data occurs usually every 5 years.

Suggested comparator: The standardized measurability of these indicators allows for comparability across countries. It may also be of interest to policy makers to compare by region (e.g. province, urban/rural), or socioeconomic status (e.g. wealth quintile, level of education, gender).

Maternal mortality per 100,000 live births (1990-2016)



Source: World Bank (2019). World Development Indicators. (<https://databank.worldbank.org/source/world-development-indicators>)

Potential source of data:

Vital registry data, health services records, census. Household surveys such as the Demographic and Health Survey (DHS), the UNICEF Multiple Indicator Cluster Survey (MICS).

<https://dhsprogram.com>

<https://mics.unicef.org>

Benchmarking guidelines

No established benchmark, the lower the better. Suggest benchmarking against past performance or relevant regional/ income country comparators/ averages. Sustainable Development Goal targets aim to reduce maternal mortality to less than 70 per 100,000 live births by 2030.

As most maternal deaths are preventable, it will be important to choose relevant input, output and process indicators in developing the results chain narrative. For example, number of midwives, antenatal care coverage, access to skilled care during childbirth, postnatal coverage, relevant service-specific readiness measures SALA (https://www.who.int/healthinfo/systems/sara_introduction/en/) or SDI (<https://www.sdiindicators.org>), and others such as diagnostic accuracy, adherence to clinical guidelines. Maternal death audits are also an important additional source of information to explain causes of maternal mortality.

TB burden

Simple definition: The burden of TB disease can be measured in terms of: i) Incidence – the number of new and relapse cases of TB arising in a given time period, usually 1 year; ii) Prevalence – the number of cases of TB at a given point in time; and iii) Mortality – the number of deaths caused by TB in a given time period, usually 1 year.

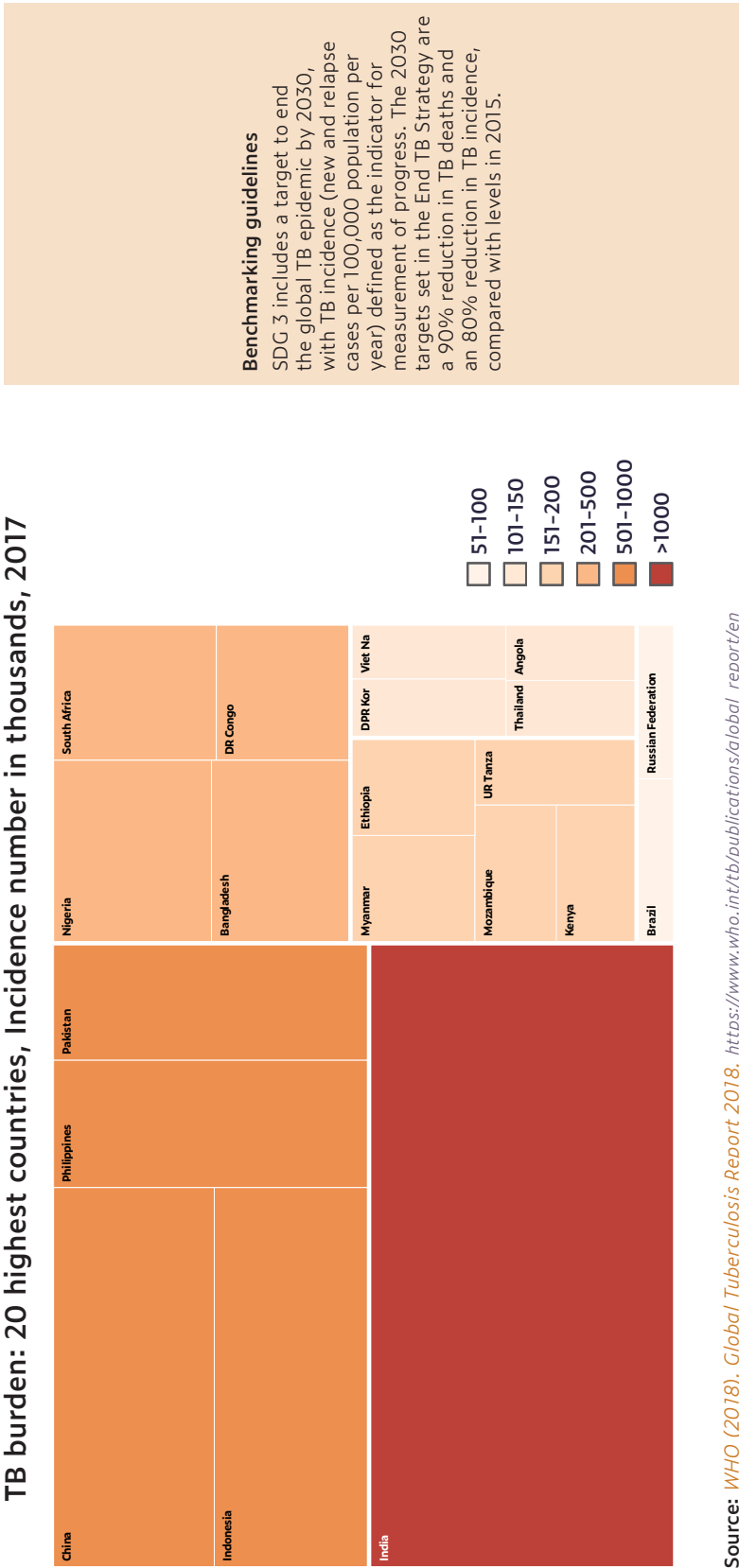
Frequency of data collection: TB notifications/surveillance data are ideally captured in real-time and reviewed weekly or monthly. Prevalence surveys are at least 5 years apart.

Suggested comparator: Across countries, by region, high risk populations (e.g. people living with HIV, injecting drug users, prisons, nursing homes, healthcare workers).

Potential source of data:

The WHO produces a yearly Global TB report. There are also i) TB prevalence surveys – a method used if notification data from routine surveillance are thought to be inaccurate or incomplete and when there is an estimated relatively high prevalence of TB (more than 100 per 100,000 population) as they are expensive to conduct; ii) TB notifications – a method used by 144 countries with low levels of underreporting (generally high income and some selected middle-income countries); and iii) National inventory studies – these measure the level of underreporting of detected TB cases (method used in 6 countries); and iv) Case notification data combined with expert opinion about case-detection gaps – this method is the least preferred and it is relied upon only if one of the other three methods cannot be used (currently used in 43 countries).

TB burden: 20 highest countries, Incidence number in thousands, 2017



Source: WHO (2018). *Global Tuberculosis Report 2018*. https://www.who.int/tb/publications/global_report/en

The ultimate goal is to directly measure TB burden from TB notifications in all countries. Notifications are considered a good proxy of incidence in countries that have high-performance surveillance systems and good access to quality care that result in few undiagnosed cases. As the surveillance system improves and the number of cases decreases, countries no longer undertake expensive TB surveys that would require larger and larger sample sizes. To assess progress made due to a program or intervention it may be more informative to also look at indicators along the entire TB cascade including other TB indicators such as TB case detection, TB notification, TB treatment success rates, TB-service specific readiness measures (SARA https://www.who.int/healthinfo/systems/sara_reports/en/ and SDI <https://www.sdiindicators.org/indicators>) and diagnostic accuracy and adherence to clinical guidelines.

Catastrophic health expenditure

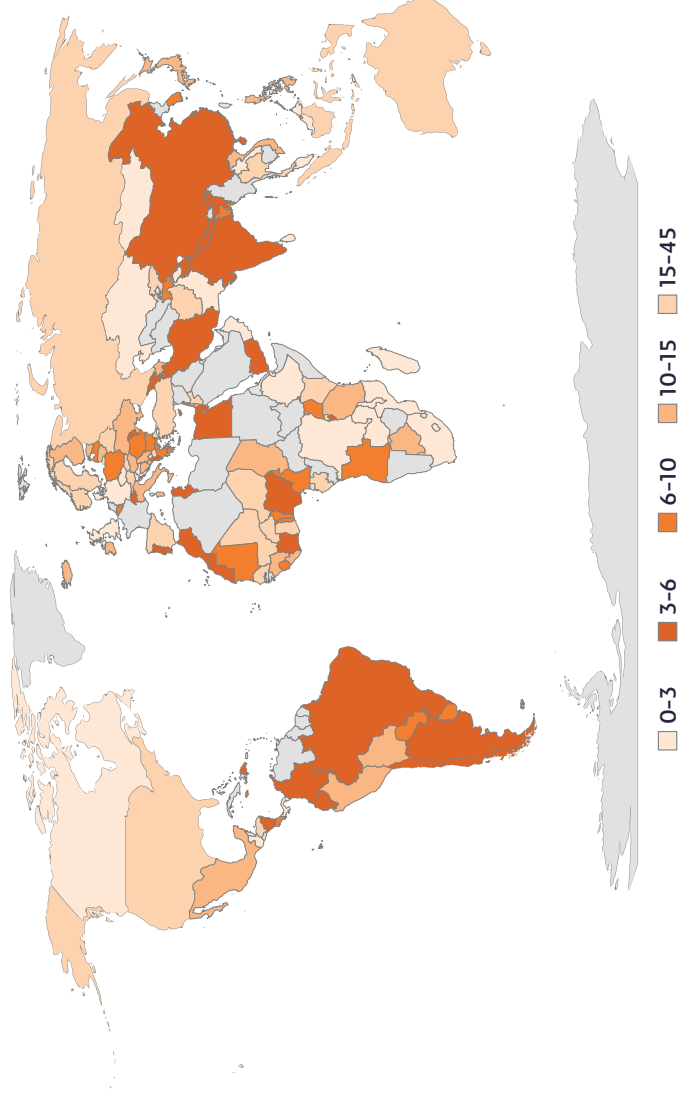
Simple definition: Percentage of households who spend more than 10% of their total household consumption on health expenditures.

Frequency of data collection: Usually every 5 years however some regions/countries collect data on an annual basis. The Health Equity and Financial Protection Indicators database will be updated annually.

Suggested comparator: Across countries. Some surveys also produce representative data at the subnational level allowing within country comparisons. It may also be of interest to policy makers to compare by socioeconomic status (e.g. wealth quintile or income groups).

Potential source of data:
Household surveys. More recently the World Bank has started to produce regular estimates through its Health Equity and Financial Protection database (HEFPI).
<http://datatopics.worldbank.org/health-equity-and-financial-protection/>

**Incidence of catastrophic health expenditure, % of population
(Latest year, 10 Percent threshold)**



Note: Grey = no data

Source: WB (2019) HEFPI database

Benchmarking guidelines

No established benchmark, the lower the better. Suggest benchmarking against past performance or relevant regional/income country comparators/averages.

While this indicator is commonly used as a measure of financial protection it must be interpreted with caution as it cannot distinguish between i) households that draw upon savings to finance often unexpected medical care and those that have no such reserves and ii) individuals who forgo treatment because they either do not have access to health care or cannot afford it. It also frequently only captures the direct cost of treatment and medication while ignoring cost of transportation, time costs, informal payments, etc.

Catastrophic expenditure should be interpreted together with service coverage and impoverishing expenditure. The working paper has much more detail on how to calculate and interpret financial protection.

The World Bank's support to the Joint Learning Network for UHC
is made possible with financial contributions from the following partners:



BILL & MELINDA
GATES *foundation*



THE
ROCKEFELLER
FOUNDATION



JOINT
LEARNING
NETWORK

For Universal Health Coverage

www.jointlearningnetwork.org

