



Working paper 574

Review of public financial management diagnostics for the health sector

Sierd Hadley, Tom Hart and Bryn Welham

February 2020

Key messages

- Rising interest in public financial management (PFM) in the health sector is increasing demand for relevant diagnostic tools.
- This paper identifies eight diagnostic tools that provide information on PFM in relation to public health services.
- These diagnostic tools differ considerably in their methodology, areas of focus, cost and timeframe for implementation.
- It is critical to carefully agree the 'problem' that a tool is intended to analyse, since no single tool covers all aspects of financial management in the health sector.
- A tool that secures strong interest and encourages collaborative working between stakeholders is likely to yield most benefits.

Readers are encouraged to reproduce material for their own publications, as long as they are not being sold commercially. ODI requests due acknowledgement and a copy of the publication. For online use, we ask readers to link to the original resource on the ODI website. The views presented in this paper are those of the author(s) and do not necessarily represent the views of ODI or our partners.

This work is licensed under CC BY-NC-ND 4.0.

Acknowledgements

This document is based on research jointly produced by the Overseas Development Institute (ODI) and the United Nations Children’s Emergency Fund (UNICEF) in order to inform the development of an approach for UNICEF country offices to analyse and address public financial management bottlenecks in the health sector. The original analysis was conducted by Tom Hart, Sierd Hadley, Gundula Loeffler and Bryn Welham of ODI and overseen by Jennifer Asman, Ulla Griffiths and David Hipgrave of UNICEF. Richard Hughes provided programming support and Elizabeth Tribone oversaw production of the final paper. This final paper was edited for a general audience and was funded by the Bill & Melinda Gates Foundation and UNICEF.

The authors are grateful to Jennifer Asman, Hélène Barroy, Ulla Griffiths, Srinivas Gurazada, Jason Lakin, Jo Kemp and Mark Miller for providing peer review comments. The conclusions of this document – including any errors and omissions – remain the responsibility of the authors.

Contents

Acknowledgements	3
List of boxes and tables	5
Acronyms	6
Executive summary	7
1 Public financial management and health diagnostic tools	10
1.1 Health and public financial management diagnostics – a growth industry	10
1.2 The challenges of governance diagnostics	12
1.3 Identifying and describing existing diagnostic tools	16
1.4 Key lessons from the comparison of the tools	21
1.5 Concluding points	23
2 Annex: description of the tools	28
2.1 Comparing the types of tool	28
2.2 Five public financial management and health tools reviewed	30
2.3 Three allied tools reviewed	37
References	41

List of boxes and tables

Boxes

Box 1	The governance of public financial management indicators	13
Box 2	Diagnostic tools that link public financial management and health services	17
Box 3	An embarrassment of riches?	23
Box 4	Problem-driven approaches to institutional reforms	25

Tables

Table 1	Summary of the methodology and focus of the diagnostic tools	8
Table 2	Pros and cons of different methodological approaches	15
Table 3	The diagnostic and ‘allied’ tools that relate to public financial management and health	18
Table 4	Approach used in the five PFM and health diagnostic tools and three ‘allied tools’	19
Table 5	The main methodological approach used	19
Table 6	Areas of focus across different types of tools	21
Table 7	Diagnostic characteristics of Public Expenditure Tracking Survey	30
Table 8	Diagnostic characteristics of WHO PFM and health financing process guide	32
Table 9	Diagnostic characteristics of World Bank FinHealth PFM-in-health toolkit	33
Table 10	Diagnostic characteristics of PEFA for health	34
Table 11	Diagnostic characteristics of OECD Budgeting Practices for Health Survey database	36
Table 12	Diagnostic characteristics of UNICEF (modified) Tanahashi analysis	37
Table 13	Diagnostic characteristics of WHO Health Financing Country Diagnostic	39
Table 14	Diagnostic characteristics of World Bank Public Expenditure Review	40

Acronyms

HSS	health systems strengthening
OECD	Organisation for Economic Co-operation and Development
PEFA	Public Expenditure and Financial Accountability
PETS	public expenditure tracking surveys
PFM	public financial management
UHC	universal health coverage
WHO	World Health Organization

Executive summary

This working paper reviews existing diagnostic tools that seek to provide information on the effectiveness of public financial management (PFM) systems in relation to publicly delivered health services in low- and middle-income countries. It aims to provide donors, government and other practitioners with a comparative ‘menu of options’ for considering what – if any – PFM and health diagnostic tools might support their particular objectives.

Interest in the relationship between the management of public funds and the delivery of health services has increased in recent years. The number of tools that seek to understand and assess PFM institutions, practices and outputs in health has also grown significantly. From the broad set of tools that are available across different disciplines, this review identifies five tools that look at the relationship between PFM and health directly and three ‘allied’ tools that look at aspects of PFM among a wider range of other concerns over health service delivery. These tools are summarised in Table 1 below.

The tools vary significantly in their methodological approach, cost to deliver, time to implement and frequency of application to date. Some tools such as Public Expenditure Tracking Surveys (PETS) also vary considerably in their own right, with different studies varying in their methodology, cost and areas of coverage. Chapter 2 of this paper provides a profile for each tool to allow a detailed comparison. There are, however, some broad and important differences to highlight in the methodological approach and the analytical basis for investigating PFM challenges in the health sector.

The main methodological split is between tools that use a benchmarking or narrative-based approach. Benchmarking tools like the PEFA for health compare PFM systems to an external standard. They either score or describe the degree of compliance with these standards. On the other

hand, narrative approaches such as PETS rely more on expert assessment or research. Each of these approaches has strengths and weaknesses – many of which reflect the same ongoing issues, challenges and debates that surround any attempt to measure, assess and/or quantify the quality of public institutions in general or public finance processes specifically.

In general, narrative-based tools are easier to tailor to the specific context of a country, while benchmarking tools offer greater comparability across countries. Benchmarking tools are also arguably more ‘actionable’ because they tend to compare specific parts of the PFM system of a country with a desirable end state or international standard. Achieving that end state may later be used to guide reforms. Narrative tools may not provide such a clear direction of travel, but when they are used well, they can provide a more tailored analysis that can be used to guide domestic PFM reforms. While PFM diagnostics are increasingly using benchmarking approaches (often with a scoring system), the tools used in the health sector offer a broad range of approaches, possibly because it is less clear what a ‘good’ PFM system for the health sector should look like.

The tools take different views (if only implicitly) on the links between PFM and health service delivery – and what ‘good’ PFM means. Most of the tools review procedures, practices and outputs that do not have an obvious direct link to service delivery outcomes. Indeed, in many cases the link between the PFM issues identified by the tool and actual improvements in health services need to be inferred. However, one set of tools asks how PFM bottlenecks can be removed to support better health services, while another set asks whether PFM systems can be more aligned to health financing objectives of resourcing, pooling and purchasing. It is also notable that some tools look exclusively at the

Table 1 Summary of the methodology and focus of the diagnostic tools

Name of tool	Sponsoring organisation	Methodology	Area of focus	Guiding question being answered
Five diagnostic tools that focus on PFM and health				
Public Expenditure Tracking Surveys (PETS)	Originally developed by World Bank but also used by others	Narrative tool (based on quantitative data)	Health service delivery	To what extent are budgeted resources reaching front-line facilities?
PFM and health financing process guide	WHO	Benchmarking tool (descriptive)	Health financing policy and processes	How well is the PFM system aligned with the health financing objectives of revenue-raising, pooling, and purchasing?
PEFA for Health	Adapted separately by DANIDA and USAID	Benchmarking tool (ordinal scoring)	General PFM in the health sector	How closely do PFM systems in the health sector align with international good practices for PFM?
FinHealth: PFM-in-health toolkit	World Bank	Narrative tool (qualitative)	General PFM in the health sector	What are the key PFM and health finance bottlenecks or enablers that constrain or support service delivery results at the provider level? How do these relate to PFM systems?
Budgeting Practices for Health Survey	OECD	Descriptive tool with limited analysis	General PFM in the health sector	What are the institutional arrangements for preparing the health budget, and what is the role of the finance ministry?
Three allied tools with information relevant to the PFM system				
Modified Tanahashi Bottleneck Analysis	Originally developed by WHO but used mostly by UNICEF	Narrative tool (based on quantitative data)	Health service delivery	What are the supply and demand side constraints to the delivery of subnational health services?
Health Financing Country Diagnostic tool	WHO	Benchmarking tool (descriptive, drawing on quantitative data)	Health financing policy and processes	How well does the health financing system support the health financing goals of revenue-raising, pooling and purchasing?
Public Expenditure Review (PER)	World Bank	Narrative (based on qualitative and quantitative data)	Health financing policy and processes	Is the budget for health adequate, fairly allocated, and used effectively; and what impact does this have on health outcomes?

PFM system, while others like the Tanahashi bottleneck assessment includes PFM as one of a number of supply-side influences on health services.

Given that these tools aim to do different things according to different methodologies, and with different resource implications, selecting the right tool to match the agreed ‘problem’ or gap in knowledge is an important part of the decision process. Indeed, it is important to note that no single tool will be able to cover all themes of interest and all parts of the PFM system to provide a complete PFM assessment. There are also some notable gaps in the tools reviewed. Critically, none of the tools place a great deal of focus on diagnosing or investigating the non-technical issues affecting service delivery (e.g. political economy or incentive issues). In practice, however, a well-delivered PFM diagnostic is

likely to uncover many of these in the course of its work. Most also suffer from a relatively weak evidence base and conceptual framework linking improved PFM and improved health services.

Finally, there are widespread views that suggest the way the tool is used is a major factor behind supporting institutional change. Evidence of policy impact from the different tools is rarely documented and would be a useful issue for further research. However, a clear message from several key informants (practitioners who have used one or more of the tools reviewed) is that securing genuine government interest and engagement in the diagnostic exercise is important for success; and, in terms of real-world reform impact, the technical design of the tool is less important than the ability of the diagnostic process to bring relevant stakeholders together in a spirit of solving problems.

1 Public financial management and health diagnostic tools

This paper reviews eight diagnostic tools that aim to assess the strengths and weaknesses of PFM in relation to publicly delivered health services in low- and middle-income countries. The intention of the paper is to help government officials and their development partners to make informed choices over which tool – if any – might suit their particular needs. Instead of providing a ranking of effectiveness or identifying the best tool, this review outlines the available eight options and highlights the advantages and disadvantages of particular tools. This is done in the context of a brief discussion of the challenges and opportunities of PFM measurement and diagnostics in general.

This review comes at a time of particular interest in links between PFM and health services. Sustainable Development Goal 3 – ‘Ensure healthy lives and promote well-being for all at all ages’ – sets a target for achieving universal healthcare (UHC) by 2030. While donor funding for health has increased substantially over recent decades, from around \$7 billion in 1990 to more than \$30 billion in 2016 (IHME, 2017; Xu et al., 2018), the overwhelming majority of funds to achieve UHC in the long run, as countries develop, are expected to come from, and flow through, national financial systems (Fan and Savedoff, 2014). The international debate currently places a strong emphasis on the ‘transition’ out of aid dependence, heightened by administrative cut-off points for receiving international aid (Silverman, 2018; Engen and Prizzon, 2019). This puts a premium on efficiency

and effectiveness of the PFM systems that support domestic spending.

The paper is divided into two sections. Chapter 1 is the main body of the review, setting out an analysis of the key issues related to PFM and health diagnostics. Chapter 2 examines each of the tools in detail using a common framework and provides the underlying evidence on which the discussion in section 1 draws.

The rest of chapter 1 is structured as follows: section 1.1 outlines the growing interest in PFM and health; section 1.2 sets out the general challenges associated with measuring the quality of PFM systems in general, which also apply to any attempt to do this in the health sector; section 1.3 provides a summary comparison of eight diagnostic tools; and section 1.4 draws out the key lessons from this comparison of the tools. Section 1.5 concludes with points for potential users of the tools to consider.

1.1 Health and public financial management diagnostics – a growth industry

1.1.1 The emerging interest in public financial management from the health sector

This review is being conducted against the backdrop of a growing interest in the links between PFM and health services among development agencies. This interest stems from the confluence of three broad trends (e.g. Welham et al., 2017). Firstly, over recent decades and in line with a broader shift in development policy

towards questions of good governance, health sector practitioners have increasingly considered questions of system-level organisation, governance and management as a determinant of service delivery, rather than simply focusing on the overall quantity of funding available or the technical correctness of health policies. This is perhaps most clearly expressed in the health systems strengthening (HSS) agenda, of which consideration of financial systems are often a part (Hafner and Shiffman, 2013).

Secondly, as already described, there is an increased awareness that the bulk of resources for financing an increase in health coverage will come from domestic resources, managed by national PFM processes, rather than through external development assistance. This is reinforced by formal ‘transition points’ where external funding to support certain health programmes (e.g. vaccination programmes supported through Gavi) is expected to rapidly decline in a number of middle-income countries. This puts a premium on making sure that national finance processes are able to support effective service delivery.

Thirdly, the shift by most donors away from general budget support towards sector budget support or sector programming has brought into clearer focus the importance of PFM in specific social sectors. As donors withdraw from general budget support, the lens through which they look at PFM and fiduciary risks is increasingly sector-specific. There have also been increasing calls for PFM reforms to be more explicitly linked to real-world improvements in development outcomes: sometimes summarised by the phrase ‘PFM for what?’ (Welham et al., 2013). This agenda recognises the increasing pressure for PFM reforms to demonstrate that they are delivering *something* in terms of better development outcomes, and not simply improving PFM systems as ends in themselves.

Taken together, these trends promote the idea that PFM systems and processes are not just a matter for a finance ministry and multilateral organisations, but are also of relevance to a ministry of health and its partners (Barroy et al., 2019). This move has resulted in new communities of practice for health financing,

such as the World Bank ‘PFM for service delivery’ group and the World Health Organization (WHO) agenda on fiscal space, PFM, health financing and UHC.

1.1.2 The emerging research on public financial management in the health sector

A key assumption underpinning this agenda is that the effectiveness of PFM systems is important for supporting effective and efficient health services delivery. At first glance this would seem intuitively obvious. From the health side of the debate, health services require numerous inputs to be brought together at the right place and the right time; many of these inputs require financing; and it is the role of PFM systems to ensure that this financing is available. From the PFM side of the debate, one of the three core objectives of a PFM system is efficient operation of public services, alongside maintaining a sustainable fiscal position and effective allocation of resources (see for example Schick, 2013).

There is some emerging evidence for this view, although it is not extensive. Quantitative, cross-country research has found some association between quality of PFM systems (as measured by the Public Expenditure and Financial Accountability (PEFA) framework) and the effectiveness of delivery of certain key health outputs and outcomes (Welham et al., 2017; Piatti-Fünfkirchen and Smets, 2019). However, this kind of cross-country analysis has significant limitations, and the studies do not identify the mechanism (or mechanisms) by which financial management processes support service delivery or how these work across different contexts. So, while countries with better PFM may also be able to deliver more effective health services, it is difficult to pinpoint the exact contribution that PFM plays in the chain of causality.

Researchers studying PFM and researchers from a health financing tradition have so far presented different explanations for why PFM matters. Welham et al. (2017) adopt more of a PFM perspective, concluding that the strongest evidence points to a need for countries to ‘get the basics of PFM right’ in terms of providing reliable flows of inputs to facilities; and alongside this, building accountability structures, which

may be outside the PFM system, to oversee the use of these resources. Another prominent view is that weak PFM systems can undermine the health financing goals of pooling and strategic purchasing (Barroy and Kabaniha, 2018).

A recent position paper published by WHO has adopted elements of both arguments (Barroy et al., 2019). However, existing research would struggle to predict which PFM reforms would have the greatest positive impact on health service delivery.

Lastly, it should be noted that all of the studies have emphasised another intuitive (and perhaps obvious) point expressed in the HSS agenda: finance and effective PFM systems are not the only, or even main, drivers of effective health service delivery. A wide range of non-PFM related technical and non-technical factors will also affect the quality of health services delivery.

1.1.3 The development of new diagnostic tools

As well as prompting further research on PFM and health, the new agenda has encouraged another important development: the design of new diagnostic tools. As will be discussed in section 1.2, measuring institutional quality has become an increasingly important feature of international development research and practice. There are numerous examples of diagnostic tools for PFM, but also for other areas of governance. Agencies working on health systems strengthening are no exception. There are a range of tools and measurement approaches available for assessing performance in the six building blocks of health systems strengthening: (1) service delivery; (2) health workforce; (3) health information systems; (4) access to essential medicines; (5) financing; and (6) leadership/governance (WHO, 2007; 2010). Indeed, internet searches conducted for this research found around a dozen tools for assessing the quality of health information systems alone.

At the time of writing (November 2019), a number of international agencies were in the process of developing diagnostic tools that would support analysis of PFM systems in the health sector. These agencies include UNICEF, WHO and the World Bank. Indeed, the background material for this paper was initially prepared

to support the process for designing a new approach for UNICEF country teams – which is excluded from the analysis below because it is still in the early stages of development. It was also suggested in the peer review process that other agencies such as Gavi may also be considering new diagnostic approaches. These new tools build on an already considerable base of PFM and health systems tools that provide an extensive menu for practitioners and government officials to draw on in their work. The rest of this paper considers some key issues that will be important for selecting the right tools for the right purpose at the right time.

1.2 The challenges of governance diagnostics

PFM diagnostic tools use some form of standardised assessment to make a judgement as to whether, and to what degree, PFM systems are operating effectively in a particular country. These diagnostic tools are widely seen as useful, at least as judged by the growth in the number of tools and in the number of times they have been applied across the world (see PEFA, 2018, for a list of more than 40 PFM diagnostic tools; the PEFA website lists more than 600 individual PEFA assessments).

The use of PFM diagnostic tools has, however, not been without criticism. At first glance, the aim seems straightforward: to measure how well PFM institutions, processes and systems are operating. However, the history of measuring the quality of public institutions, processes and systems *in general* – and PFM institutions *in particular* – has proved both difficult and contentious. The discussion that follows provides a brief overview of the challenges and opportunities of using different forms of diagnostic tool to measure institutional effectiveness in order to situate the health-focused PFM tools in this broader debate.

1.2.1 A brief history and typology of PFM diagnostics

The use of diagnostics to judge the quality of public institutions so as to inform policy decisions emerged clearly in the 1990s (see Box 1). Governance and the effectiveness

of state institutions were increasingly seen as key determinants of economic and social development (Lateef, 2016). Development institutions began to use institutional analysis both to understand varying levels of economic and social development across different countries and to inform their programming. In part, this was as a way to explain the varying results achieved by very similar development policies applied to different countries and as a way of addressing the question of how to build new state institutions in transition economies after the collapse of the Soviet Union (Knack et al., 2003).

The interest in PFM institutions in particular gained greater prominence in the early 2000s. This was reinforced by the need to assess how

well debt relief under the Highly Indebted Poor Country Initiative would be managed by national public finance systems in eligible countries and by the emergence of general budget support as a mainstream aid modality. The institutional failings that contributed to the 1997 Asian financial crisis also increased the attention given by the International Monetary Fund (IMF) and World Bank to the role of PFM systems in supporting fiscal transparency. By the early 2000s, a wide range of tools was available for assessing the quality of PFM systems, and this number has continued to grow (Allen et al., 2004; Mackie, 2011; PEFA Secretariat, 2018).

The current range of tools available allows for different levels and areas of assessment,

Box 1 The governance of public financial management indicators

The governance of indicators themselves has been an important part of the debate over PFM diagnostic tools. In the 1990s, many donors developed their own tools to assess fiduciary risks and systems of public expenditure management, including the United Nations Development Programme (UNDP), the IMF, the European Union (EU) and the World Bank. Recipient countries were subject to overlapping missions and an overwhelming number of externally driven and inconsistent recommendations (Allen et al., 2004). Multiple reform plans were also designed to provide financial support to implement recommended changes. This all imposed heavy transaction costs on governments.

These concerns gave rise to a ‘strengthened approach’ to supporting PFM reform. This approach aimed to support a country-led agenda, coordinated programme of support and shared information pool. There was to be a concerted effort to rationalise the number of diagnostic tools used in countries, which gave birth to the PEFA programme – a widely used indicator set managed by a Secretariat under the supervision of a Steering Committee of different donors. While PEFA has proved a hugely successful and influential tool on almost any measure, broader aims of the strengthened approach have arguably fallen away (Hadley and Miller, 2016).

Today, there are new challenges. The power of the PEFA diagnostic approach has led to the emergence of a growing number of new tools under different custodians, often borrowing from a similar methodology. The move away from comprehensive PFM reform programmes to underpin general budget support appears to be fuelling this trend. The current landscape looks much more like it did in the 1990s. Indeed, a government official from a sub-Saharan African country recently complained to one of the authors of this report of the never-ending flow of new diagnostics being implemented in the finance ministry of his country. He suggested it was a case of ‘death by diagnostics’.

While there are benefits to the profusion of tools from which a practitioner or government official can choose, there are also clear risks. For readers of this paper, it may be useful to reflect on the way that new diagnostic tools for PFM and health are developed and piloted across agencies, how existing tools will be used and coordinated in country, and how development agencies can put governments in the driving seat of the process.

with some tools providing a picture of the system as a whole, and others focusing on a specific sub-system such as accounting. There are also differences in methodology: some tools *benchmark* country systems against international best practices; some collect information in a *database* for cross-country comparison; and some rely mainly on a *narrative assessment* of how a country's PFM systems work, often based on 'expert opinion'. Benchmarking assessments that provide clear scores against objective and universal criteria have become the dominant approach in PFM diagnostics (e.g. PEFA, the IMF Fiscal Transparency Code), while more descriptive assessment approaches to PFM that do not produce clear 'scores' (such as Public Expenditure Tracking Surveys, PETS) have taken more of a backseat in recent years (Hood et al., forthcoming).

1.2.2 Emerging critiques of institutional diagnostics and how they are used

The rise of institutional assessments (including of PFM), and of quantitative benchmarking tools in particular, has experienced something of a backlash (Grindle, 2007; Andrews, 2008). There is a contested debate on the usefulness and robustness of PFM diagnostic approaches, particularly benchmarking approaches that use ordinal scoring methodologies (Andrews, 2007; Andrews et al., 2014; Hadley and Miller, 2016; Hood et al., forthcoming).

These criticisms should be seen in the context of historical debates over the best way to measure the quality of institutions. The initial effort to develop quantitative indicators of institutional performance was seen as necessary to complement narrative approaches and to support more systematic cross-country comparison (Knack et al., 2003). The first generation of indicators included the Worldwide Governance Indicators, which ranked countries based on expert assessments or perceptions of broad institutional challenges such as corruption or the rule of law. The second generation of

governance indicators emerged in the early 2000s and are often referred to as actionable indicators.¹ These measured specific processes and outputs of particular parts of the governance system (e.g. anti-corruption laws) rather than the broader outcome (e.g. perceptions of corruption). This was done in a way that could be replicated across countries – often by benchmarking practices against an established international standard or norm. This change in approach has clearly shaped the methodology used in many PFM diagnostic tools, including the PEFA framework.

Commentators have argued that current benchmarking approaches are overly determinative in that they explicitly set out a narrow, normative view of what 'good' PFM systems should look like (often informed primarily by developed country experience) and then inappropriately rank developing country performance against this (explicitly through quantitative/ordinal scoring or implicitly through narrative description). Others have argued that quantitative/ordinal PFM diagnostics focus too much on 'form' (i.e. the existence of a formal system) rather than 'function' (i.e. the actual results produced). There are also concerns that while quantitative/ordinal PFM diagnostics may be useful in themselves, a 'cookie cutter' approach to turning their conclusions into very similar-looking PFM reform plans tends to undermine their positive benefit. In contrast, a narrative/descriptive approach might allow for greater responsiveness to local context, but this comes at the risk of producing conclusions that are less actionable and results that are harder to compare across countries. Some of the general strengths and weaknesses of each approach are summarised in Table 2.

To illustrate some of these differences, it is possible to compare findings from a PEFA assessment (a benchmarking tool) with PETS (a narrative tool). Tracking surveys have been used to identify 'leakages' in budget execution, but these same studies do not necessarily

1 Actionable indicators had four criteria: (1) they were generated through an objective process that could be validated and replicated; (2) they were available for a large number of countries; (3) they could be measured accurately and consistently across countries; and (4) they should focus on specific governance systems or outputs, and not be unduly affected by exogenous influences outside that system.

Table 2 Pros and cons of different methodological approaches

Approach		Common strengths and weaknesses
Benchmarking (e.g. PEFA framework)	Ordinal (scoring/ranking) or descriptive	<ul style="list-style-type: none"> Standardised assessment methodologies allow for easier comparison between countries and over time Findings are usually specific and actionable but not necessarily prioritised Assessment of performance is against an externally set benchmark and therefore not always sensitive to complexities of the specific context
Narrative evaluation (e.g. PETS)	Usually either predominantly qualitative or quantitative	<ul style="list-style-type: none"> Assessment methodologies are less standardised making it harder to compare across countries and over time Findings are less specific and actionable but offer more scope for prioritisation More tailored to country context
Database (e.g. budget practices survey)	Descriptive with both qualitative and quantitative elements	<ul style="list-style-type: none"> Highly standardised results allow for straightforward cross-country comparison Strongly descriptive approach means few (if any) clear actionable or prioritised actions Standardised approach to questions and responses reduces scope for detailed discussion of context

provide a clear picture of the systems that need to be strengthened in order to reduce wastage and corruption. On the other hand, a PEFA assessment does not identify the scale of leakages, but will identify areas of internal audit, expenditure controls and other PFM systems that do not comply with good practice standards, on the assumption that increasing adherence to good practices will reduce the likelihood of resource leakages.

Database tools and benchmarking tools are more similar, but database tools are usually more descriptive. The Organisation for Economic Co-operation and Development (OECD) budget practices survey (a database tool) and PEFA (a benchmarking tool) both have questions related to the use of reserve funds. PEFA gives a score of ‘A’ if the reserve funds used are on average under 3% of the national budget, while the OECD budget survey records only the level of the reserve fund and what it was used for, without explicitly setting a threshold for ‘good’ or ‘bad’ performance for a particular country. While both frameworks take a normative judgement that reserve funds are an important part of the PFM system, PEFA also sets out a series of performance benchmarks with which countries

should expect to comply in order to meet ‘good practice’ levels of performance.

Limitations with the diagnostic tools are most problematic when they are used inappropriately. Norway is the only advanced economy to have conducted and then published a national PEFA assessment. The assessment returned mainly high ‘A’ scores, but eight indicators were rated ‘C’ or ‘D’. Reviewing these results, the finance ministry agreed that the assessment had revealed weaknesses in areas such as procurement and the follow-up of external audit reports. However, the ministry dismissed other low scores as being less important – either because the scores reflected the decentralisation of primary service provision to municipalities or because current systems were considered appropriate for the Norwegian state.² This illustrates that a diagnostic tool must be filtered appropriately for the local context.

1.2.3 The implications for health and PFM diagnostics

This review cannot fully explore or resolve the various arguments about the usefulness and/or robustness of PFM diagnostic tools of different kinds. It is, perhaps, enough to note that while PFM diagnostic tools remain popular, they face numerous criticisms regarding: (1) their ability

2 For the full response and report see: <https://norad.no/en/toolspublications/publications/2009/public-financial-management-performance-report---norway/>

to fully identify, describe and capture the totality of PFM institution and system functioning; and (2) the way they are used in practice to inform policy decisions. Recent critiques are aimed at quantitative/ordinal benchmarking approaches in particular, but the debate on earlier approaches highlights that all diagnostics will face limitations. This has prompted calls for donor agencies to improve the way that these measurements are used (Hadley and Miller, 2016; PEFA Secretariat, forthcoming). The health and PFM diagnostic tools presented in the following sections should therefore be situated within the ongoing debate between an increasing demand for more indicators and diagnostics on the one hand and concerns about how accurate such assessments are and how wisely they are used in practice other.

1.3 Identifying and describing existing diagnostic tools

This review has identified eight PFM and health-related diagnostic tools through two main avenues. Firstly, the authors have previously undertaken research into the relationship between PFM and healthcare (Welham et al., 2017) and a number of these tools were raised, referred to, or otherwise identified in the course of that work. Secondly, the authors proactively raised the issue of health and PFM diagnostics with other institutions, agencies and individuals working in the field and asked them to share their knowledge. It should be noted, therefore, that this list is not necessarily entirely exhaustive and there may be other financial management and health diagnostic tools available that have not been identified through these methods.³ The following discussion outlines

how the eight identified tools can be categorised, conceptualised and considered across a number of dimensions. This discussion will help potential users of these tools to select the most useful tool for their particular purposes.

1.3.1 Identifying diagnostic tools for understanding PFM in the health sector

PFM diagnostic tools can usefully be placed into three categories in terms of their particular focus on, and relation to, the health sector (summarised in Box 2). The main focus of this review is on the second area of focus: PFM diagnostic tools that specifically look at the health sector. These tools most clearly represent the intersection of health and PFM considerations. Tools relevant to the third area are also reviewed (albeit in less detail), given that in some cases they represent well-known, longstanding and frequently used tools that consider some elements of PFM amid a wider discussion of the health sector. These are mostly drawn from the field of health financing, which has been the locus for recent WHO work on PFM in the health sector (Barroy and Kabaniha, 2018). Those in the first area of focus are not covered on the grounds that they take a general view on the strength of the PFM system rather than providing information that is of specific interest to the health sector. A summary of this kind of tool can be found in the stocktake conducted for the PEFA Secretariat (2018).

It should also be noted that this paper does not discuss tools that provide: (1) a fiduciary risk review of health PFM systems; or (2) financial audits of healthcare institutions or health providers. This is because reviews of fiduciary risk typically have a narrow focus on assessing the strengths and weaknesses of the system in correctly managing donor aid flows, rather

3 It should also be noted that this review considers only ‘public’ healthcare provision. In reality, each country will have a variety of public, private, charitable and/or social insurance systems to raise financing for health, and this money may be spent through a mixture of public, private and/or non-profit/charitable mechanisms. For the purposes of this review, the diagnostic tools under discussion are most suitable for those parts of the health system that are publicly financed (i.e. funds are disbursed by government – even if these funds have been originally raised through a variety of tax and/or insurance mechanisms) and publicly delivered (i.e. health facilities are public sector institutions themselves, or are non-public sector but are so heavily regulated by government and under public authority that they operate as quasi-public sector institutions). The key determinant for the appropriateness of use of these diagnostic tools is the degree to which health services and their financing are delivered by and through government/public systems. In many countries, there may not be a clear dividing line between different types of healthcare financing and delivery. It should also be noted that in some countries publicly delivered and publicly funded healthcare may be a minority of overall healthcare provision.

Box 2 Diagnostic tools that link public financial management and health services

PFM diagnostic tools can be placed into three categories:

1. PFM diagnostic tools that do not focus on the health sector in particular (e.g. Fiscal Transparency Evaluation; PEFA)
2. PFM diagnostic tools that look at the health sector in particular (PEFA for health; PETS)
3. Health system diagnostic tools that look at a number of areas that can include PFM, but where PFM is not necessarily the particular focus (Public Expenditure Reviews, modified Tanahashi framework)

than focusing on the effectiveness of the PFM system as a whole. Some approaches do consider elements of ‘value for money’ in the management of funds; however, as the PEFA stocktake (2018: 6) concludes:

...[fiduciary risk assessment tools] are always applied or commissioned by a development partner and are often a mandatory part of their procedures. For governments considering tools to further their reform agendas, the fiduciary risk assessment tools will rarely be sufficient.

Such tools therefore serve a different purpose from the diagnostic tools covered in this note. Regarding financial audits, these typically consider the correctness, probity and/or efficiency of resource use in terms of compliance with financial rules within *a particular institution* or *part* of the health sector. While practitioners may want to draw on the information produced by such reports to inform their assessments, they are not considered further in this paper.

On this basis, five diagnostic tools have been identified that most clearly fit into the second area of focus in Box 2. In addition, the third area of focus contains three ‘allied’ tools that consider PFM in health to some degree as part of a consideration of other elements of health service delivery or health sector structure. These are listed in Table 3, with the institution that created, sponsored, funded or was otherwise involved in its development.

1.3.2 Summary of the approach used in the eight tools

The eight tools take different approaches to considering issues of PFM and health. Indeed, reading a PETS study and an OECD health systems characteristics survey makes very clear how different the approaches are. Table 4 provides a few lines of summary explaining what each tool does.

1.3.3 Comparison of methodologies used

The identified tools use different methodological approaches to deliver their assessments (see Table 5). Of the eight, three have adopted a benchmarking approach, meaning that they explicitly assess performance against a stated external standard of what good performance looks like. One of these has a prominent quantitative/ordinal methodology (PEFA for health) while the others take a more descriptive approach of comparing performance against a stated level (the two WHO tools). A further four tools take a narrative evaluation approach, meaning that they assess performance, but this is not done against a single external standard. Of this subset, two use a largely qualitative descriptive/narrative approach to discuss performance, supported by quantitative data in places (World Bank FinHealth tool and Public Expenditure Reviews (PERs)). In contrast PETS and the Tanahashi analysis have a strong quantitative basis to support the narrative evaluation. Finally, one tool – the OECD budget practices for health – takes a database approach, meaning that it assembles a common set of data for a number of countries, but does not use this to make any explicit judgement of performance against an external standard.

Table 3 The diagnostic and ‘allied’ tools that relate to public financial management and health

Name of tool	Institution or sponsoring entity	Sources for methodology or tool guides
Diagnostic tools that focus on PFM and health		
Public Expenditure Tracking Surveys (PETS)	World Bank, as original developers, although there are now numerous methodologies used by different institutions	<ul style="list-style-type: none"> • World Bank guidance on using PETS • 2010 Stocktake paper on experience with PETS • Paper summarising different PETS experiences • Paper summarising methodologies for using PETS to assess small-scale projects
PFM and health financing process guide	WHO	<ul style="list-style-type: none"> • WHO guide
PEFA for health	Two separate applications were identified: <ul style="list-style-type: none"> • a ‘home-grown’ approach developed by the Danish International Development Agency (DANIDA) office in Mozambique, and also used in Tanzania (not formally published) • a methodology, funded by the United States Agency for International Development (USAID), that applies the PEFA framework to the health sector 	<ul style="list-style-type: none"> • Mozambique case study • ‘Guided Self-Assessment of PFM Performance – A Toolkit for Health Sector Managers’ (USAID-funded guidance note)
FinHealth: PFM-in-health toolkit	World Bank	<ul style="list-style-type: none"> • Unpublished, forthcoming 2020¹
OECD Budgeting Practices for Health Survey	OECD	<ul style="list-style-type: none"> • Webpages covering methodology and output
‘Allied tools’ with information relevant to the PFM system		
Modified Tanahashi Bottleneck Analysis	Originally developed by WHO but most extensively used by UNICEF	<ul style="list-style-type: none"> • Original Tanahashi paper • UNICEF paper providing an overview of the approach
Health Financing Country Diagnostic tool	WHO	<ul style="list-style-type: none"> • WHO guide
PER	World Bank	<ul style="list-style-type: none"> • The original World Bank guidance for evaluating public expenditure • Online repository of published World Bank Public Expenditure Reviews

¹This review was based on an unpublished draft of the toolkit.

As already explained in more general terms, these methodologies have different limitations. The PEFA for health tool may provide some information specific to the health sector that is not captured in a general PEFA assessment, but it will also be open to the same criticisms as any other PEFA assessment. Most notably, a PEFA assessment does not spell out if a low score is problematic, or if a high score means systems are working well in practice. The benchmarking approach used in the two WHO diagnostics is not actionable in the same way as a PEFA assessment but is still guided by normative views

of what constitutes a good health financing system. If these views do not hold universally, then practitioners will need to judge whether the benchmarks are appropriate for the context in which they work. Similar concerns would also apply to any benchmarking study that draws on the results of the OECD Budgeting Practices for Health Survey. On the other hand, the narrative evaluations used in the FinHealth toolkit, PERs, PETS and the modified Tanahashi framework may offer more context-specific information but all are likely to require further work to determine how best to respond to the assessment findings.

Table 4 Approach used in the five PFM and health diagnostic tools and three ‘allied tools’

Name of tool	Summary of approach
PFM and health tools	
PETS	Tracking flows of funds within the health system, from the budgets proposed by central government down to the funds actually received by service delivery units. In some studies this is complemented by service delivery or service quality assessments. Some studies also look beyond cash and consider the flow of other kinds of resources (e.g. drugs, medical equipment). PETS focus predominantly on downstream issues of whether funds reach service delivery units and how these are accounted for.
PFM and health financing process guide	Comparing health sector PFM practices with best practice. The guide sets out key elements of what makes good PFM practice and health financing in the health sector and provides a method for determining the degree to which national health systems accord with these practices, and – relatedly – where they are misaligned. This analysis can then form the basis of a reform plan for government.
PEFA for health	Assessing the performance of predominantly upstream planning, budgeting and reporting systems for the health sector using a modified version of the ‘standard’ PEFA assessment. Scoring is ordinal against externally set indicators with benchmarks that allow for a judgement of basic, good or advanced performance. Scores can be compared between countries and over time.
FinHealth: PFM-in-health toolkit	An assessment of both upstream and downstream processes in managing finance and other inputs in the health sector covering key parts of the budget cycle, the relationship of financial inputs to best practice in health financing, and asking a range of questions about how these inputs combine to deliver health services on the ground.
OECD Budgeting Practices for Health Survey	A top-down high-level description of OECD member country budgeting practices for health that show the different institutional frameworks and budgeting instruments used to manage healthcare expenditure. The resulting information is almost entirely descriptive (i.e. what is the nature of specific features of the health budgeting system at the present moment) rather than analytical (i.e. how well does the system operate against a particular standard).
‘Allied tools’	
Modified Tanahashi Bottleneck Analysis	A bottom-up tool that encourages service users, service providers and other ground-level stakeholders to come together and identify blockages and challenges to effective health service delivery, including those related to finance and flow of funds, with a view to agreeing reforms to remove them.
Health Financing Country Diagnostic tool	A top-down, upstream budgeting tool that encourages comparison between best practice in financing healthcare at a national level and actual national practice, with a view to developing a reform plan.
PER	A top-down analytical tool that compares overall expenditure in the health sector with overall service delivery outputs along a number of dimensions and at a number of levels. This is analysed over time and in comparison with other countries.

Table 5 The main methodological approach used

Approach	Tool	
Benchmarking	Ordinal (scoring/ranking)	<ul style="list-style-type: none"> • PEFA for health (DANIDA/USAID)
	Descriptive	<ul style="list-style-type: none"> • Health Financing Country Diagnostic (WHO) • PFM and health financing process guide (WHO)
Narrative evaluation	Predominantly qualitative	<ul style="list-style-type: none"> • FinHealth: PFM-in-health toolkit (World Bank) • PER (World Bank)
	Predominantly quantitative	<ul style="list-style-type: none"> • PETS (World Bank) • Modified Tanahashi (UNICEF)
Database	Descriptive with both qualitative and quantitative elements	<ul style="list-style-type: none"> • Budgeting Practices for Health Survey (OECD)

It is also important to point out that there is variation in the degree to which the methodology of the tools is ‘standardised’. Some tools (e.g. the two WHO diagnostics) offer a single standard, methodology paper, whereas some offer a wide variety of different implementation options (e.g. PETS and PERs). In other cases, the tools represent modifications of existing methodologies and are not standardised or necessarily ‘approved’ by development agencies (e.g. PEFA for health). Potential users of these tools will need to bear in mind, therefore, that some tools are ‘ready to go’ in terms of a clear methodology for implementation and a large number of case studies, whereas others will require some groundwork to finalise the approach in the absence of an extensive case study library.

1.3.4 Comparison of the areas of focus

As well as the methodology, practitioners will be keen to understand what areas of PFM or the health sector a diagnostic tool can cover. As illustrated in Table 6, the eight diagnostic tools can be grouped into three broad areas of focus:

- **General PFM systems in the health sector.** This structures the discussion around the ‘standard’ high-level stages of the budget or PFM cycle, such as budget formulation; approval; execution; accounting and reporting; and audit. It tends to ask to what degree the health sector is complying with the overall rules and procedures of the central government PFM cycle and focuses at an aggregate level.
- **Health financing as it relates to PFM.** This approach tends to focus on the appropriateness, fairness and sustainability of the specific methods and policies by which healthcare is financed and delivered in a country, and on the degree to which this is in accordance with what the health sector is deemed to need. The focus is more on the ‘intermediate’ level between issues of compliance with central government PFM systems (as in the previous category) and the experience of service delivery on the ground (as in the next category).

- **Service delivery and PFM.** These questions are much more concerned with the experience of service delivery units, and (sometimes) their service users. It tends to ask whether cash or other inputs are actually reaching service delivery units and, if so, how these resources are being used and accounted for.

This categorisation can further be divided into those tools that look upstream at planning, budgeting and reporting at a central level, and those that look ‘downstream’ at budget execution, flow of funds and cash utilisation at facility level. Tools in the first two categories tend to look at upstream issues and in the third very clearly at downstream issues of the reality of budget execution.

Inevitably, this grouping will not be perfect. The three categories outlined are not fully discrete and issues can cut across them: as a result, the tools do not always fit neatly into these groups. The PETS framework, for example, has been put in the service delivery category, but there are significant differences in the methodology used for individual studies and some applications would be closer to the first category (PFM systems). Likewise, although most of the areas for investigation in the World Bank PFM-in-health toolkit concern upstream policy, finance and budgeting, the toolkit is sufficiently broad that parts of it would cut across all three categories. Nevertheless, this provides a further method of categorising the eight tools against different types of PFM and health relationships to aid tool selection.

1.3.5 Frequency of application

The tools presented here also vary substantially in the frequency of their application. At one extreme, there are several hundred PERs (not only in health) and numerous different forms of PETS that have been applied in a wide variety of contexts over 20 years (although less so in recent years). This is also true for the modified Tanahashi framework, which has its roots in an approach pioneered at WHO in the 1970s, and has since been modified and applied widely, increasingly by UNICEF country offices. At the other extreme, some tools have been completed only a handful of times (PEFA for health).

Guidelines or examples of these tools are not all easily accessible through a web search and there are no evaluation or survey articles, especially for the tools that have been used less often or are still emerging. The description in chapter 2 of each tool discusses this in more detail.

1.4 Key lessons from the comparison of the tools

The review of the eight tools highlights some key points that might be useful for potential users of the tools. There are limitations in what different methodologies measure; differences in the way that PFM challenges are linked to health service delivery performance; general gaps relating to non-technical factors like the political economy; and important variations in the cost and resources needed to execute the different tools well. These issues are elaborated in turn.

1.4.1 Variation in the overall aim, methodology and approach of the tools

One clear lesson is that – despite all working in the broad area of PFM and health – the tools aim to do very different things at different levels of analysis. This is perhaps unsurprising given the sheer number of institutions, processes and activities that could be considered within the category of ‘PFM and health’. Given these differences, the tools offer very different kinds of information. Some tools provide analysis of upstream PFM systems at the whole-of-sector level (e.g. PEFA for health), while others provide information about the flow of resources to frontline service delivery units (e.g. PETS). Some tools identify areas where sector PFM processes are aligned or misaligned with global best practice (e.g. the WHO tools), while others capture PFM issues as part of a holistic analysis of broader health service delivery challenges (e.g. modified Tanahashi). The OECD budget practices database explicitly aims to provide

Table 6 Areas of focus across different types of tools

Type of diagnostic	Areas of focus	Diagnostics
PFM systems in the health sector	<ul style="list-style-type: none"> Strengths and weaknesses of PFM systems across the budget cycle in the health sector Budget preparation, budget execution and reporting; and sub-systems such as procurement and payroll management in the health sector An ‘upstream’ view of predominantly central government-led PFM issues 	<ul style="list-style-type: none"> PEFA for health FinHealth: PFM-in-health toolkit
Health financing policy and processes	<ul style="list-style-type: none"> Performance of national systems to finance healthcare across key criteria How far these financing functions are contributing to the UHC goals of (1) equity in the use of health services; (2) quality of care; and (3) financial protection How the higher-level national PFM system is affecting delivery of these key goals ‘Upstream’ view looking at overall systems outputs compared to inputs, and how decisions on health financing are made at a central government level and delivered through national systems 	<ul style="list-style-type: none"> WHO health financing country diagnostic WHO PFM and health financing process guide OECD Budgeting Practices for Health Survey Public Expenditure Review
Health service delivery	<ul style="list-style-type: none"> The real-world availability and flow of funds (and sometimes other inputs) to support key inputs for service delivery (i.e. staff, medicines, facilities) Service coverage and effectiveness at the facility level Downstream PFM issues such as regularity of transfers to providers, reporting on resource use, payments of wages, delivery of priority medical inputs etc. 	<ul style="list-style-type: none"> Modified Tanahashi PETS

cross-country comparison, which most of the other tools do not.

The tools also vary in the degree to which they consider PFM problems as part of the wider health delivery system challenges (e.g. Tanahashi), or as a set of problems to be investigated in their own right (e.g. PETS). A narrower focus may support more actionable PFM-related recommendations, but this also risks overstating the likelihood that improvements in PFM will ‘fix’ more general challenges with governance and accountability in the health system. On the other hand, a broader approach may give a good sense of the major PFM challenges facing a sector but provide insufficient depth to understand which parts of the PFM system really matter and should therefore be strengthened as a matter of priority. Critically, no single tool is able to capture and distil all information that potential users might wish to know across the entire health sector and PFM system.

1.4.2 Explicit links to service delivery

A key challenge for all of the diagnostic tools is to link PFM constraints to actual service delivery challenges. This is an important observation since the ultimate objective of all these tools is to use their analysis of PFM problems to support better health services. Tanahashi is the most focused on measuring the effective availability of the key inputs (staff, medical supplies and infrastructure) that are needed to provide health services, but it has no formal methodology for seeking to understand the PFM constraints that might be leading to shortages. For most of the other tools, the link between failures in PFM systems in the health sector and failures in health service delivery as a result must either be inferred through other sources of information or assumed based on prior knowledge.

For the upstream tools in particular – such as PEFA for health or the WHO instruments – establishing a clear link between PFM weaknesses identified by the tools and actual on-the-ground service delivery failures will almost certainly require (potentially significant) additional analysis. The tool that comes closest to linking PFM constraints and service delivery outcomes is PETS, where it can be reasonably

straightforward to link, for example, how flow of funds (or lack thereof) is affecting service delivery. However, these studies are often constrained by data limitations, and many PETS have not clearly identified which element of the PFM system – if any – is causing the bottleneck. For this, more hybrid studies like the PETS applied in Timor Leste would be needed (Nixon and Bredenkamp, 2014).

1.4.3 Non-technical factors affecting service delivery

None of the tools reviewed provides a thorough method of analysing political economy (or other non-technical) factors that might underlie dysfunctions in the allocation, flow and use of resources in the health system. This matters since it is unlikely that health service delivery challenges in developing countries can be solved solely by fixing technical financing issues. Indeed, the typical assumption underlying the reform of public institutions is that ‘softer’ issues of leadership, management and behavioural change are crucial for success, and this requires political will and senior-level championing.

It could be argued that investigating these issues is simply not the role of a PFM-focused diagnostic tool, and other tools should be deployed to deal with this issue. Alternatively, it may be reasonable to expect that, through the process of gathering PFM-related data about health service delivery challenges, a team of researchers will inevitably access a large amount of tacit and informal knowledge regarding political economy factors affecting health service delivery. Nevertheless, it should be noted that diagnostic tools such as these will not systematically review the broader context of what might be driving these limitations in the first place. This inevitably puts more emphasis on the need to carefully judge the way in which the diagnostics are used to guide reforms in a particular context.

1.4.4 Cost and resources needed

Potential users of these tools will need to consider carefully the resource cost of the studies, which can vary considerably. A desk-based diagnostic, such as the WHO tools, that draws predominantly on existing government data

and limited qualitative discussion with central government staff could be relatively inexpensive and swift. On the other hand, a study that requires large amounts of primary data collection from health facilities (e.g. collection of quantitative data or semi-structured interviews, as is the case for PETS or Tanahashi) would cost significantly more. Relatedly, the team and skills mix needed for diagnostic tools need to be carefully considered with the data requirements in mind. As an example, one key informant suggested that attempts to produce ‘integrated’ PERs joining together expertise on sector policy and PFM administration tended to fail because of challenges in bringing together a team of experts with the right skills mix to cover both areas. Potential users of these tools will need to consider how much money they have to spend on this exercise and the availability of the right skills before selecting a tool.

1.5 Concluding points

Drawing conclusions from reviewing such a wide range of diagnostic tools is not straightforward. The purpose of this review has been to summarise and outline the tools available to potential users across a number of dimensions, rather than rank or rate their individual

effectiveness (see Box 3 for an update of the tools available). Indeed, the preceding points have highlighted the differences between the tools that make such direct comparison difficult. Taking this as a starting point, some key points on the use (and misuse) of these tools for potential users can be drawn from this exercise.

1.5.1 The need to carefully select the right tool(s)

As explained earlier, there is no single tool that can answer all the relevant questions about PFM and health. There is no comprehensive or ‘meta’ assessment framework that looks at issues of high-level, sector-wide upstream planning and budgeting all the way down to downstream issues of funds flow and PFM bottlenecks at frontline facilities. The World Bank FinHealth toolkit comes closest to attempting this by asking a large range of questions, grouped across 24 key health functions, that collectively cover a wide range of both upstream and downstream issues. However, the sheer number of, and potential depth of response to, the questions in the toolkit would represent a significant practical challenge in terms of the volume of data collection and analysis required. In practice, users will either have to be selective about what they need to know or consider commissioning more than one

Box 3 An embarrassment of riches?

This review has identified eight tools for consideration – five that focus specifically on PFM and health service delivery and three that consider PFM among other factors. At the time of writing (November 2019), at least two other PFM and health-related diagnostic analysis tools were being developed or piloted by institutions working in the field. This raises the issue of whether there are already ‘enough’ tools available. Indeed, some stakeholders have expressed the view that further tools are unnecessary and will simply confuse national governments and their donor partners.

It is difficult to answer comprehensively the question of whether there are already enough tools. On the one hand, national PFM systems affecting health are complex and vary significantly from country to country, and the problems generated by their poor performance will be similarly varied. This might suggest that a wide range of different tools would be useful so as to allow users to pick precisely the right one. On the other hand, a review of some of the conclusions from the most widely used tools (e.g. PETS, as summarised in Welham et al., 2017) suggests that the same issues recur across countries. This would suggest that a limited number of tools is sufficient in order to identify the PFM challenges in most contexts.

diagnostic tool if their knowledge gaps about health and PFM span a wide range of issues. Given the large number of diagnostics already being rolled out in low- and middle-income countries, it seems fair to recommend that donor agencies opt for greater selectivity rather than number.

When choosing the right tool, users will need to ask where they think the knowledge gap lies: are there concerns about how funds move in the health system (which might suggest a PETS-style funds tracker)? Or is the issue about improving the overall amounts of financing for the system and where it comes from (wherein the WHO health financing tools might be most useful)? Or is it about the efficiency of resource use within the sector (which would suggest more of a PER-style policy review)? Being clear on the gap in knowledge and how a particular tool will fill that gap is critical. Indeed, if the gap in knowledge is very specific and clearly agreed, a full ‘diagnostic’ of the kind discussed here may be less useful than an investigative review of the particular issue at hand.

An increasingly popular way of identifying the knowledge gap would be to focus carefully on the most pressing ‘problem(s)’ in the health sector that need to be solved (Andrews, 2013). Stakeholders would invest time and effort in jointly understanding what the general issue is, breaking this down into smaller problems and/or refining the specific problems, and using this to determine if a lack of knowledge or gap in understanding about health systems PFM is part of the identified problem(s). Based on the understanding of this knowledge gap, the right diagnostic tool or alternative study approach can be selected. This would mean being clear in particular about the ‘level’ of the knowledge gap – for example, is it about national-level planning, budgeting and financing across the health sector, lack of funds transfer to service delivery units, or how funds are actually used at local level? This will directly affect the choice of tool for investigating the gap in knowledge in this area.

1.5.2 It’s not what you do, it’s the way that you do it

One related lesson that was drawn from interviews asking about the way that the tools were used *in practice* is that they are most effective when embedded in broader reforms in which the government is genuinely invested. These tools typically offer a ‘one-off’ knowledge product as their output – but this needs to be absorbed and used if it is ultimately to have a positive impact. Key informants who had used these tools repeatedly highlighted the importance of government engagement with the results of the studies. As with any form of technical assistance, a strong interest by government in the findings and subsequent policy recommendations of the tool’s conclusions will give these PFM and health diagnostics the best chance of having a positive impact.

As part of this point, it is worth noting again that where practitioners were interviewed regarding the use of these tools, they typically noted that, in reality, the formal methodology of the assessment tool was often less important than how it was used to bring together actors from across government to jointly share information, analyse problems and identify solutions. In particular, it was noted in several cases that the mere fact that officials from the finance and health ministries were working together to gather and analyse health and PFM financial data encouraged a new sense of joint understanding and joint working between these institutions. The data gathering and analysis process allowed for an exchange of information, positions and views on mutual challenges that otherwise had no particular forum for discussion. Using the tool to *start or support a conversation* between stakeholders about delivering positive change, based on relevant data was frequently cited as being more important than the technical perfection (or otherwise) of the tool itself.

This resonates strongly with the recent push for ‘problem-driven’ approaches to institutional reform that move away from encouraging ‘best practices’ towards ‘best-fit’ solutions, working more closely with the grain of the domestic political economy (Levy, 2014). More information is provided on these approaches

in Box 4, but the overriding conclusion for practitioners is to position any diagnostic as part of a broader strategy that will help to facilitate or broker change over time, and to ensure that this process both informs the choice of diagnostic and provides an avenue for the legitimisation and uptake of its results. For many, this will be a departure from the stereotype of a diagnostic study – conducted by an independent expert for the purposes of a donor seeking to drive change

from the outside or design the next phase of its programme of support.

1.5.3 Do these tools add anything to the existing HSS agenda?

This review raises significant questions about whether there is indeed a discrete and recognisable thing as ‘health and PFM’ that is clearly separate from either regular health systems strengthening (HSS) or standard PFM systems improvement. HSS approaches typically

Box 4 Problem-driven approaches to institutional reforms

Over recent years a ‘new consensus’ has emerged on how institutional change can and should be delivered in developing countries. This new consensus puts an emphasis on a bottom-up iterative searching for solutions that will address real-world problems that have been diagnosed locally. This reacts to the perceived longstanding failures of ‘traditional’ models of institutional reform which are seen to have focused on diagnosing problems as failures to meet international best practice, and which typically recommended that countries adopt a single ideal-type solution that has worked well in other often more institutionally advanced environments. A number of sources put forward different formulations of this approach (e.g. ‘Thinking and Working Politically’; ‘Doing Development Differently’; ‘Going with the Grain’ and delivering ‘Change in Challenging Contexts’), but perhaps the best known is the ‘problem-driven iterative adaptation’ (PDIA) approach (see Andrews, 2012; Andrews et al., 2017).

PDIA seeks to achieve strong buy-in from the organisations that are developing the reforms by solving real-world problems that resonate with officials. A problem-driven approach changes the nature of external support for institutional reform (including those funded by donors). Rather than approaching poorly functioning institutions as a problem that external actors can diagnose, prescribe and treat, it suggests a different role. While external actors may be well placed to share international examples and outline what best practice looks like in other countries, this approach moves their primary role to a more relationship-based ‘softer’ function of facilitating, convening and ‘dot-joining’ among institutions and their staff so that they can agree on what the problems are, and decide how they can be solved.

While this represents a new consensus on delivering institutional reform, it naturally brings its own challenges. The pace of reform may be slow if institutional change can only happen at the speed set by the (poorly performing) institutions themselves. Reform efforts may look quite low-level, basic and unimpressive (at least at first) as institutions begin to solve basic and easy problems first. It also pre-supposes (from the view of external actors) a genuine desire to improve performance within institutions that are held back primarily by a lack of expertise, and/or inability to effectively collaborate to problem-solve within themselves or with other institutions.

Nevertheless, the PDIA approach continues to provide the basis of a new consensus about delivering institutional reform. In many ways it is an approach that empowers smaller institutions that often have less money to spend as it puts a premium on the ability to adopt the institutions’ perspectives, being responsive to local needs and building good relationships with key stakeholders.

look at the range of finance, consumables, staff, policy, leadership and management factors that need to come together to deliver health services.⁴ Does it help to separate out ‘PFM’ from these other issues for particular focus? Or is there sufficient information on PFM embedded within the existing ‘building blocks’ of HSS?

The tools implicitly take different views on this. The approach of the Tanahashi method situates PFM issues firmly within the range of other ground-level bottlenecks to service delivery. Once a bottleneck has been identified, discussions can then be used to determine whether the supply of inputs has been undermined by weaknesses in the PFM system or by other factors. The WHO tools separate out the analysis of PFM, but explicitly consider whether these impede broader health financing goals (e.g. is there *risk-pooling*?). A PEFA for health study, on the other hand, provides no clear linkage with the HSS approach. So, a wide spectrum of view is available within the tools reviewed here, at least implicitly.

Underlying these differences is also another question: whether improved PFM performance *in general* is sufficient to support improved health services, or if the health sector has *particular needs*. The WHO tools make the strongest claim that there is something special about PFM in the health sector, which means that specific aspects of upstream PFM in the sector need particular attention (again linked to the ability to pool risk and support strategic purchasing). This is not necessarily a universally shared view. The PEFA for health guidance material produced by the Health Finance and Governance programme, for example, makes no mention of these sector-specific objectives and instead focuses on the PEFA indicators that are likely to be important to a line ministry in general. Further, the absence of ordinal scoring tools in this area, as compared to the broader landscape of PFM diagnostics,

suggests that there is much weaker agreement on what a benchmark for ‘good PFM for health services’ looks like. Indeed, there are examples of areas where the views of PFM experts and health financing experts differ substantially. While health financing practices are increasingly encouraging the use of quasi-contracting arrangements in the public sector as part of efforts to introduce more ‘strategic purchasing’, the longstanding advice on PFM reforms has argued that countries should avoid contracting in the public sector before there is a strong practice of enforcing contracts in the private sector (Schick, 1998).

Resolving these debates is beyond the scope of this review, but it is important to note that these disagreements are unsurprising given the contested conceptual basis for these tools. As discussed, there is a lack of clear consensus in the research literature about the mechanisms through which PFM will most affect service delivery, and therefore which mechanisms should receive most attention in diagnosing weaknesses and formulating reforms. Any attempt to formulate an improved diagnostic tool therefore needs to be done in close dialogue with the growing body of evidence looking at PFM in the health sector. But without research progress in specifying how PFM reforms affect health service delivery and which reforms are likely to have the greatest positive impact in different contexts, many of these diagnostics will be based on shaky conceptual foundations.

1.5.4 There is a need for more evidence of policy impact

The question of ‘and so what?’ in terms of impact as a result of these diagnostic studies remains to be answered. This document has provided a survey-style overview and description of tools with some discussion of practical experience

4 It is worth noting that ‘health systems strengthening’ is itself a contested term. Indeed, a recent survey of HSS interventions noted that the term lacks a clear operationally useful definition (Witter et al., 2019). Within typical conceptions of HSS, the issues of ‘PFM in healthcare’ might fall into two of the main areas of (1) ‘financing’ (overall health systems financing from public/private sources, or ‘financing plus’ interventions such as performance-based financing, purchasing and contracting reforms that combine financing and governance reform); and (2) ‘leadership and governance’, which deals with systemic governance issues. It could be that by not making ‘PFM’ a separate topic in healthcare delivery, the HSS more accurately captures the cross-cutting nature of PFM systems and the outcomes they support (Welham et al., 2017).

in their use. However, it has not attempted to answer the (very important) question as to whether these tools (individually or collectively) have actually made a positive difference to health service delivery in developing countries – and if so, how has that happened

Potential users will rightly want to know if a certain tool tends to have greater policy impact than others and, if so, how the chances of policy impact can be maximised through particular use

of that tool. This is a question that merits further research. In the meantime, potential users of these tools should assure themselves that there is indeed a specific *PFM* problem that requires investigation – and not just a service delivery failure caused by other factors. Developing strong relationships with country partners and a deeper understanding of the national context will be key to using a tool successfully.

2 Annex: description of the tools

2.1 Comparing the types of tool

The main section of this document has outlined how the tools were identified and explained how they fit into various classifications. This section describes each of the eight tools in detail against a number of common criteria.

As noted in the main discussion, the tools are divided into two categories:

- five diagnostic tools that focus specifically on the relationship between health and PFM (PETS; PEFA for health; World Bank FinHealth toolkit; WHO health financing process guide; OECD Budgeting Practices for Health Survey);
- three ‘allied tools’ that look at other topics in the health sector but which touch upon PFM issues (WHO health financing guide; World Bank PER; modified Tanahashi).

The focus of the discussion is on the first category of tools – those that deal directly with the relationship between PFM and health services delivery.

2.1.1 Comparing the PFM and health tools

The tools are reviewed against the common characteristics outlined below.

Key objective and type of assessment

This provides a brief description of the tool’s objectives and focus within the area of PFM and health.

Example of application and number of applications

This discusses how many cases and/or applications of the methodology have taken place. Potential users will be interested to note the diversity in frequency of applications – or, at least, in the number of publicly available applications.

Methodology

This briefly explains the general approach to gathering and analysing information within the tool. It sets out whether the diagnostic tool attempts to deliver: (1) a quantitative assessment (i.e. a quantitative or ordinal score against an externally defined standard of performance); (2) a ‘narrative’ assessment (i.e. a qualitative description of performance that may be set against an expected standard – either externally set or compared to the ambitions of the country itself); or (3) a ‘database’ assessment where the same information is collected across a number of countries and made available for secondary research.

Information requirements

This discusses how much information is required to deliver an application of the diagnostic tool and what kind of information is required (e.g. detailed facility-level financial data; or high-level, whole-of-system data).

Timeframe

This sets out an estimate for the time required to deliver the methodology, building on available

case study evidence. Again, some of the tools assessed here have extensive case studies from which to draw such an estimate; others are much more limited in their implementation.

Inputs required

Each diagnostic assessed requires substantial inputs, including the cost of implementing studies and the expertise needed to execute a diagnostic assessment effectively. This section outlines, where available, the likely resource cost and expertise required to deliver this kind of study.

Access to information

This outlines whether the available case studies or other real-world applications of these

diagnostic tools found any particular challenges in accessing the key information required.

Stakeholder engagement

To make the diagnostic a success, most studies require a degree of engagement with government officials, given the nature of the information required to deliver the study. This section outlines any particular observations on the kind of engagement required for a successful diagnostic.

A brief discussion provides additional information or views on the relative strengths and weaknesses of the tool. This is added in a final section in each table under ‘Selected commentary’.

2.2 Five public financial management and health tools reviewed

Table 7 Diagnostic characteristics of Public Expenditure Tracking Surveys

Diagnostic characteristic	PETS
Key objectives	<p>A bottom-up quantitative tool focused on the facility level that typically: (1) tracks the flow of resources (usually cash, but also 'in-kind' resources in some cases) through different levels of administration in the health sector, from central government down to the service delivery unit; and/or (2) identifies problems, challenges and leakages in this flow of cash and in-kind resources, as well as (in some cases) problems with the deployment of human resources at facility level.</p> <p>It should be noted that PETS can vary considerably in scope. Some restrict themselves to simply following the flow of cash from central government to local level; others expand their coverage to include flows and deployment of non-cash resources (e.g. staff and medicines); others go further still to include analysis of service delivery performance (sometimes called a 'Quality of Service Delivery Survey').</p>
Example of application	<p>Previous research (Welham et al., 2017) identified 27 PETS covering the health sector published between 2000 and 2016.</p>
Number of times used (estimate)	<p>As above, 27 published studies identified though many more are unpublished and/or not easily accessible. Judging from this subset of studies, there has been a general trend away from using PETS, with the number of published studies falling from 21 between 2000 and 2004, to just 4 between 2010 and 2014 (Welham et al., 2017).</p>
Methodology	<p>There is no fully standardised approach to conducting PETS. Variation in case study approach is certainly apparent from a review of the publicly available case studies despite some 'standardised' guidance documents being available for those conducting PETS (e.g. Koziol and Tolmie, 2010). This has raised some discussion over which aspects of PETS could or should be standardised in future (e.g. Gauthier and Reinnika, 2007; Gauthier, 2010).</p> <p>Methodologies are predominantly quantitative, supported (in some cases) by qualitative evidence. A quantitative approach to tracking the flow of funds is, to some degree, intrinsic to the core of a PETS. Thus, the 'basic' PETS approach involves gathering primary financial data on the movement and availability of funds at each administrative level of the health system compared to expectation. This can then be combined with qualitative interviews with officials to understand how resources are used and/or with qualitative and quantitative data about actual service use.</p>
Information requirements	<p>Information requirements for PETS are significant. Indeed, the requirement to gather significant amounts of quantitative data at different levels of government has proven to be a real challenge in the context of weak PFM systems where budgets, cash flows and expenditure are not routinely recorded.</p> <p>Effective PETS require a good understanding of the way resources are moved between different levels of administration in order to begin designing the study. They would therefore often begin with a mapping exercise to understand how money, goods and/or people are supposed to move through the health system. In practice, these systems are often highly complex and in several studies, PETS teams report that they have not always managed to clearly identify how systems are supposed to work.</p> <p>Financial data need to be available, reasonably accurate at each level of administration and shared willingly or the work of the PETS will be undermined. Pre-survey data-gathering tests can help to assess realistically what is likely to be available. Again, there are examples of PETS that have been severely limited either by poor record-keeping, absence of records entirely, or because information is not available in sufficient detail.</p> <p>A typical PETS will survey more than 100 facilities across multiple districts/regions, and gather financial information for the previous two to three years for these facilities. This will then be used to create a picture of how money and sometimes other resources have moved between different levels over time. That said, the balance between qualitative and quantitative information gathered through PETS can vary depending on the objective of the study. For example, the PETS in Timor Leste focused quantitative data-gathering on far fewer facilities, but made greater use of qualitative data-gathering as a means of understanding how and why funds and other goods move through the system (Nixon and Bredenkamp, 2014).</p>

Table 7 Diagnostic characteristics of Public Expenditure Tracking Surveys (contd)

Diagnostic characteristic	PETS
Timeframe	<p>The time needed to deliver a PETS will depend on the scope of ambition and the availability of key information. As noted, PETS can vary considerably in their design, and the local context will vary in terms of ability to provide the necessary information.</p> <p>As examples of the variation in timeframe, a PETS in Brazil took more than a year: three to four months for design and pre-testing, field work of five months and seven months of analysis and publication. A PETS conducted in Timor Leste took more than 18 months to complete.</p>
Inputs required	<p>Again, this will vary depending on the scope of the PETS and the ability of government at different levels to provide information.</p> <p>It is recognised that 'costs and time demands have limited the application of PETS' (Gurkan et al., 2009). Some have estimated the costs for a single sector at between \$75,000 and \$200,000, though others have claimed PETS have exceeded \$1 million in some cases. In the Timor Leste case, interviews suggest that the study was 'under \$100,000', though this may not have included World Bank staff costs.</p> <p>Teams conducting PETS are typically interdisciplinary. The required expertise usually covers: survey design and methodology; extensive knowledge of the sector being reviewed; extensive understanding of national and local PFM systems; and an accountancy or other financial background to support tracking of funds.</p> <p>Depending on the approach taken, a PETS may also require recruitment, training and deployment of large numbers of temporary enumerators (data-gathering staff) for the initial stages. This will typically add significantly to the cost.</p>
Access to information	<p>PETS require significant access to financial (and sometimes non-financial) information. As noted, there are often numerous challenges with data collection in countries with weak PFM systems. Indeed, many PETS studies report a failure to produce intended results due to the low quality and/or lack of information.</p>
Stakeholder engagement	<p>Support from government (at different levels) is typically critical to the success of a PETS. The quantitative data requirements are often significant, and much of this data will need to come from public records that can likely only be accessed with government assistance. In some cases, PETS were conducted by (or jointly with) government officials.</p>
Selected commentary	<p>It is difficult to provide a single conclusion on the usefulness of PETS approaches given the large variation in study approaches, the methodologies applied and the quality of execution. Broadly, the most effective health sector PETS have been relatively successful at identifying PFM-related problems in the health sector, such as delays in funding, bureaucratic capture, leakages and staff absenteeism. It should be noted, however, that there are numerous examples within the PETS case studies where poor specification of the study and/or limitations with available data mean that the data gathered could not be interpreted in a meaningful way. As a result, successful PETS studies require careful specification of the scope of work and an awareness of the realistic availability of data to support the study.</p> <p>PETS can be an effective tool for judging how and to what degree the identified financial bottlenecks impact on actual health system outputs, particularly when paired with service delivery data (although gathering both types of information will naturally add to costs). PETS have generally been weaker tools for understanding why these constraints exist, given their typical focus on 'following the money' and tracking flow of funds. One way of addressing this issue is to complement the survey approach with more in-depth qualitative interviews to understand the underlying causes of the constraints, although such an approach would move beyond a strict focus on expenditure tracking. Given the downstream budget execution focus of PETS, they have relatively little to say about the upstream processes of budget preparation.</p>

Table 8 Diagnostic characteristics of WHO PFM and health financing process guide

Diagnostic characteristic	WHO PFM and health financing process guide
Key objectives	A health financing-focused tool looking top-down that: (1) provides a framework to assess how well health budgeting practices and financial rules are aligning with international best practice in health financing; and (2) provides guidance to help policy-makers diagnose misalignments and obstacles in regards to matching health financing objectives and PFM systems, with a view to formulating a reform plan.
Example of application	Full or partial applications of the guide have occurred in Malawi, Myanmar and Tanzania.
Number of times used (estimate)	As above; seemingly a small number of countries.
Methodology	<p>Qualitative description and analysis (supported in some cases by specific pieces of quantitative data) analysed with reference to an externally defined standard.</p> <p>The guide is arranged around into six assessment modules that correspond in part to well-known steps in the budget cycle:</p> <ul style="list-style-type: none"> • laying the groundwork • health budget formulation • health budget execution and provider payment • budget accounting and reporting • fiscal sustainability • options for achieving better alignment between PFM and health <p>Each module has a set of questions, culminating in an assessment table that asks for qualitative comparison between the current situation and best practice.</p>
Information requirements	The assessment requires extensive data on the government health financing system; however, much of this can usually be drawn from existing systems rather than requiring new research. There is relatively little required in terms of quantitative data, but much more in terms of qualitative institutional detail that can only be obtained through interviews and meetings.
Timeframe	The guide estimates that completing the assessment process would take three to four months.
Inputs required	The assessment is designed to be carried out by a technical working group of officials from the ministry of finance and ministry of health (plus other relevant agencies). While government officials may in theory be able to complete the work itself, the tool recognises that they may benefit from external support in the form of PFM and health financing experts to coordinate the work and undertake initial analysis of the data.
Access to information	Not found. Some of the key data relating to high-level issues of health financing are likely to be publicly available. However, given the qualitative descriptive/analytical approach to considering issues of institutional process, there is likely to be a need for 'insider' information as to how PFM and health financing systems work.
Stakeholder engagement	The assessment is designed to be carried out by government officials, supported (in practice) by external consultants. As noted, the kind of qualitative information required is likely to mean significant government official involvement in the study.

Table 8 Diagnostic characteristics of WHO PFM and health financing process guide (contd)

Diagnostic characteristic	WHO PFM and health financing process guide
Selected commentary	<p>This tool provides a framework to assess how well health budgeting practices and country financial rules are aligned with health financing best practice. This is done by outlining the best practices that PFM systems should support in the areas of: budget formulation; budget execution and purchasing/provider payment; budget monitoring; and fiscal sustainability. The tool provides qualitative guidance for health and finance policy-makers to diagnose misalignments between health financing policies and these PFM good practices as a means to identify the main obstacles and subsequently actions that can improve alignment.</p> <p>Like PEFA for health, the focus of this tool is a top-down, high-level view of predominantly upstream PFM processes, rather than a focus on downstream facility-level finance challenges. The tool certainly sets out a clear model for what good health financing PFM practices look like. However, the extent to which this effectively diagnoses the most important PFM constraints on the health system in a particular country depends on prior agreement that departures from 'best practices' are indeed the upstream PFM-related drivers of weaknesses in service delivery. The framework is therefore vulnerable to the criticisms that there is relatively little empirical evidence in favour of these PFM 'best practices' as a means to deliver effective healthcare services. For example, the framework assumes, rather than demonstrates, that adoption of Medium-Term Expenditure Frameworks and use of programme-based budgeting will lead to better PFM practices in the health sector and therefore better health results.</p> <p>In a similar critique to those of some PEFA indicators, this approach of measuring the gap between current PFM practices and 'best practice' does not directly measure the actual functioning (or otherwise) of the health system itself. In terms of identifying non-PFM drivers of health services delivery challenges, the tool would provide some space to discuss why and to what degree non-PFM factors are driving poor performance in the identified upstream PFM areas, although this is not the particular focus of this tool.</p>

Table 9 Diagnostic characteristics of World Bank FinHealth PFM-in-health toolkit

Diagnostic characteristic	World Bank FinHealth PFM-in-health toolkit
Key objectives	A framework to help World Bank country teams identify key challenges and opportunities with PFM arrangements in country health systems, linking together service delivery bottlenecks, PFM systems and health financing objectives. Undertaking a diagnostic using the toolkit will result in options for strengthening PFM arrangements to support improved health service delivery.
Example of application	None publicly available. This review was based on early pilot studies from two countries.
Number of times used (estimate)	Pilot studies to inform the design of the tool took place in Kyrgyzstan, Guinea-Bissau, Laos, Armenia and Myanmar. Forthcoming studies are expected in India, Azerbaijan, Peru and Brazil. At the time of writing, none of these were publicly available.
Methodology	Qualitative description and analysis across 24 specific areas for investigation, supported in many cases by specific pieces of quantitative data.
Information requirements	Significant. The 24 areas for investigation cover a wide range of both upstream and downstream issues in health sector PFM and service delivery. While some information required can be gathered from a desk review of pre-existing sources of information (e.g. PEFA, PERs, other health sector reviews), others will require original data collection. The tool asks specifically for a sample survey of health service delivery units and key informant interviews with representatives from key institutions.
Timeframe	This is not explicitly stated. It would depend in part on the scale and depth of the data collection undertaken to provide evidence to answer each of the 24 areas. A report aiming to respond to all areas outlined in the tool would likely take at least 9 to 12 months to move from inception to final publication.
Inputs required	The World Bank suggests that the cost of a FinHealth analysis would range between \$100,000 and \$300,000, depending on the coverage of the study and the size of the country.

Table 9 Diagnostic characteristics of World Bank FinHealth PFM-in-health toolkit (contd)

Diagnostic characteristic	World Bank FinHealth PFM-in-health toolkit
Access to information	Fully responding to all 24 areas would require access to significant information. Some would be publicly available and could be drawn from existing sources (e.g. previous PEFA or PER documents); however, the toolkit foresees a significant role for original data collection in terms of sample surveys of health facilities and key informant interviews.
Stakeholder engagement	Government engagement at various levels (central ministry, subnational government, local service delivery unit) would be crucial for a successful exercise. A great deal of the upstream planning and budgeting information regarding PFM systems will come from government sources. The requirement to sample service delivery units and to undertake key informant interviews with representatives of government institutions would also require good engagement with public sector counterparts.
Selected commentary	<p>The framework aims to provide a ‘whole-of-system’ view, from the upstream planning and budgeting to the downstream budget execution process grouped across 24 health functions, derived from a framework that brings together supply-side service delivery issues, key PFM systems and health financing best practice. Using this framework, the tool aims to make clear the linkages between identified PFM problems, health financing in general, and service delivery. In this way, it is different to – for example – a PEFA for health approach, which focuses on the operation of PFM systems but with no specific focus as to how these affect service delivery.</p> <p>Unlike PEFA, the FinHealth toolkit does not set clear external benchmarks against which performance can be quantified and/or assessed. Instead, users of the tool work across questions grouped into 24 pre-defined health functions to draw their own narrative judgements based on the data collected, which are predominantly qualitative with some quantitative elements. Many of the questions within each identified function are in the form of ‘To what extent...’ and ‘To what degree...’, offering an open-ended response where users will have to determine for themselves at what point a sufficient response has been provided. A few of the health function question areas are directly linked to PEFA indicators. It includes a number of questions and areas of focus that could be considered outside a strict definition of PFM system functioning (e.g. payroll, procurement).</p> <p>The tool is certainly broad in that it puts forward a wide-ranging framework for understanding the interrelated issues of delivery bottlenecks, PFM challenges and health financing best practice. A full response to all its questions would provide a great deal of information on health service delivery and its PFM challenges. However, this also represents its major challenge: answering every question within the 24 health function areas – and collecting all the qualitative and quantitative data required to do it – may prove a significant logistical and resource challenge to carry out in its entirety.</p>

Table 10 Diagnostic characteristics of PEFA for health

Diagnostic characteristic	PEFA for health
Key objectives	A top-down approach that uses a modified version of the PEFA framework to provide a judgement of the strength and weaknesses of PFM systems in the health sector using standard PEFA indicators, modified as necessary.
Example of application	Mozambique, 2015 (two rounds of application) and Tanzania preliminary report (incomplete and in draft), 2016, both funded by DANIDA. The UK Department for International Development (DFID) has reportedly used a similar approach in Nepal. However, these documents are not all publicly available through a web search. A USAID-funded document sets out a suggested approach for using the PEFA framework to assess PFM performance in the health sector. The Mozambique and Tanzania cases do not explicitly refer to this guide so it is unclear whether they were developed based on this approach or developed independently.
Number of times used (estimate)	As above; seemingly used in three to four different circumstances with around three applications through the DANIDA work and a potential application in Nepal with DFID. It is not known how many (if any) applications of the USAID-funded documents approach have been undertaken.

Table 10 Diagnostic characteristics of PEFA for health (contd)

Diagnostic characteristic	PEFA for health
Methodology	<p>Quantitative measurements (with some qualitative assessment) of performance across a range of modified 'standard' PFM indicators scored against benchmarks to provide ordinal marks (i.e. A, B, C).</p> <p>There is no official or standard PEFA-for-health framework. The approach in Mozambique and Tanzania was relatively straightforward: to take a PEFA indicator that applies to the whole of central government then restrict its scope to only the health sector (or health ministry). Where the PEFA indicators were not appropriate for use at a sector level, they were not used. The Mozambique and Tanzania studies cover most of the standard pre-2016 PEFA 'pillars' and use modified versions of most of the standard PEFA indicators. The USAID-funded methodology document takes a similar approach but restricts the sector assessment to 12 of the 31 pre-2016 PEFA assessment framework indicators. It provides detail on the mechanics of a scoring methodology for each indicator based on the PEFA A to D approach.</p> <p>Given PEFA's nature as a central government assessment tool that looks predominantly at high-level upstream systems, the approach does not assess in detail the performance of PFM at the lower/district level responsible for frontline delivery of health services.</p>
Information requirements	<p>The PEFA methodology is well established and the data requirements for each PEFA indicator are clear (although the documents here use the pre-2016 assessment framework). The case studies reviewed here intentionally stayed close to the original PEFA indicator methodology, even as they modified it to restrict the focus to the health sector. Identifying the data required to respond to PEFA indicators based on the health sector is therefore relatively straightforward. However, as with a conventional PEFA exercise, the actual data-gathering effort will depend on positive engagement by government.</p>
Timeframe	<p>The PEFA for health work in Mozambique and Tanzania was estimated to have taken around one year to complete, based on informal discussion with those involved in the study. As noted, examples of use of the USAID-funded tool have not been found.</p>
Inputs required	<p>Estimates are around \$120,000 for the Mozambique study and \$90,000 for the Tanzania study, reflecting the easier travel logistics in the latter compared to Mozambique and the availability of national consultants working for lower fees. As noted, examples of use of the USAID-funded tool have not been found.</p>
Access to information	<p>For Mozambique and Tanzania, the government was cooperative in these studies and willing to share information. One advantage of a PEFA-style assessment is that the PEFA data requirements are already known by government. Furthermore, using the 'PEFA brand' was thought to be useful as it carried some form of official status and was already well recognised by government.</p>
Stakeholder engagement	<p>The managers of the PEFA for health studies in Mozambique and Tanzania reported good engagement from government.</p> <p>Crucially, they emphasised that in practice the collaborative data collection process involving government, donors and the research team was of great benefit in itself in bringing together stakeholders (e.g. finance ministry, health ministry, districts) to identify their problems and discuss solutions. Indeed, this process of working together using the same data sources was considered as beneficial as the delivery of the conclusions of the report.</p>

Table 10 Diagnostic characteristics of PEFA for health (contd)

Diagnostic characteristic	PEFA for health
Selected commentary	<p>At present, there is no ‘official’ PEFA for health. The standard PEFA framework is designed to apply to central government, not to specific sectors. The PEFA for health approaches reviewed here are therefore not ‘formal’ PEFA products, and are not endorsed or formally supported by the PEFA Secretariat. They are not subject to an official ‘PEFA check’ like a standard PEFA assessment. The Mozambique and Tanzania studies are ‘home-grown’ and represent a decision made by a development agency (DANIDA) country office to experiment with applying the framework in the health sector in a particular context (initially in Mozambique, and then a further application through different means in Tanzania). The USAID-funded guidance sets out one possible ‘standardised’ approach to using PEFA indicators in the health sector, although examples of its use could not be found.</p> <p>As noted in Part 1 of this paper, PEFA studies use predominantly quantitative data to compare performance against an externally defined standardised benchmark. PEFAs in the health sector therefore bring the same strengths and weaknesses as a standard PEFA assessment: a focus on high-level legal/institutional form at the risk of missing out on consideration of actual ground-level outcomes; a focus on central government rather than lower-level governments and/or health facilities; and a standardised scoring that allows for comparison over time and between countries, but which may miss subtle nuances and differences between countries.</p> <p>The focus on PFM systems – not health service outputs – means that the level of analysis is too high to easily understand how identified PFM problems actually create ground-level bottlenecks that affect service delivery. This relationship would need to be inferred or the subject of further investigation. In addition, a PEFA approach means that there is no systematic review of the non-PFM issues affecting health services delivery. However, recent reforms to the standard framework have emphasised the importance of discussion of the non-PFM country context in a PEFA assessment; and it is likely that the data collection process required to complete a PEFA for health would in any case offer the chance to explore relevant non-PFM issues beyond the strict requirements of the assessment. PEFAs for health would be of particular value in exploring if key upstream PFM processes are working better, worse or the same in the health sector as in central government as a whole.</p>

Table 11 Diagnostic characteristics of OECD Budgeting Practices for Health Survey database

Diagnostic characteristic	OECD Budgeting Practices for Health Survey
Key objectives	<p>To collect information on the key institutional characteristics of national health budgeting procedures for comparative purposes.</p> <p>Note: this tool is best described as a ‘database’ approach to health and PFM issues. It is not, strictly speaking, a ‘diagnostic’ tool. Country responses to standard questions (in some cases using a fixed response methodology) are made available in one place, but further activity would be required from researchers to actually take the information and use it as part of an investigative study. It is included in this study as the information contained in the database is likely to be highly relevant to researchers looking to investigate health and PFM issues.</p>
Example of application	<p>The database containing all survey responses from 26 countries is available on the relevant OECD webpage. (https://qdd.oecd.org/subject.aspx?Subject=0EF6332A-E1A0-4340-A8F6-E8692B1BBA1A).</p>
Number of times used (estimate)	<p>One application of the survey in 2013. There is no publicly available information as to whether further rounds are planned.</p>

Table 11 Diagnostic characteristics of OECD Budgeting Practices for Health Survey database (contd)

Diagnostic characteristic	OECD Budgeting Practices for Health Survey
Methodology	A mixture of qualitative and quantitative questions answered by national governments covering three broad topics: (1) the role of health in the budgeting process; (2) policies used by budget agencies to influence health spending; (3) decision-making by budget agencies; and (4) the challenges of budgeting for health in decentralised contexts. The answers are then summarised into an accessible database through which researchers can search for specific information. This can form the basis of secondary analysis in comparing and contrasting countries.
Information requirements	Significant; although national governments are the respondents to this survey – not individual healthcare providers – and are therefore well placed to answer the questionnaire. For researchers using the database, there are no information requirements as the summary data is provided by the government itself.
Timeframe	One application of the survey in 2013, with a follow-up comparative summary document published in 2015.
Inputs required	The survey is managed centrally by OECD liaising directly with national governments. In contrast to other approaches discussed here, this survey is not delivered by individual researchers. The survey itself is extensive and complex, but since it is completed by national governments themselves (most likely ministries of health) they are well placed to answer these questions.
Access to information	The survey is completed by national governments (most likely ministries of health), which will already have had access to the data required, or most likely be able to find it.
Stakeholder engagement	Not explicitly stated. Most of the questions posed in the questionnaire would be answerable by central government ministries of health using existing data. Where data is not available (e.g. on very specific questions) the ministry of health would be well placed to engage with other stakeholders to obtain it.
Selected commentary	This survey provides a wide range of information on various aspects of national practices in budgeting for health across a set of common questions. The survey asks for a range of information (often quite detailed) grouped into the four areas described, and receives this in both qualitative and quantitative forms. The survey results can be searched digitally to facilitate easy comparison between country responses. The approach does not attempt to score, mark, rank or assess effectiveness of national health systems a result of this survey; instead, the information provided by national governments is made available for secondary review by interested parties. In this way the survey aims to provide a database of standardised information that could be used for diagnostic purposes and is therefore somewhat different to the other tools in this review.

2.3 Three allied tools reviewed

Table 12 Diagnostic characteristics of UNICEF (modified) Tanahashi analysis

Diagnostic characteristic	Modified Tanahashi analysis
Key objectives	This is a bottom-up service delivery-focused tool that aims to determine the bottlenecks to health service delivery on both the supply side and the demand side. It uses quantitative and qualitative data to provide an assessment that is broadly narrative/descriptive in nature, supported in many places by quantitative data. Finance-related blockages are assessed within the broader consideration of supply- and demand-side bottlenecks.
Example of application	Henriksson et al. (2017) 'Bottleneck analysis at district level to illustrate gaps within the district health system in Uganda'; UNICEF application laid out in O'Connell and Sharkey (2013).
Number of times used (estimate)	This tool is widely used by UNICEF as part of its Health Systems Strengthening approach. The Tanahashi framework has been used in Nigeria, Uganda, Ghana, Kenya, Tanzania and India, to name a few countries.

Table 12 Diagnostic characteristics of UNICEF (modified) Tanahashi analysis (contd)

Diagnostic characteristic	Modified Tanahashi analysis
Methodology	<p>A mixed quantitative and qualitative approach. The focus of quantitative measurement is on the flow and availability of key inputs for service delivery (particularly staff and drugs) in primary care facilities.</p> <p>The Tanahashi tool identifies six determinants of effective health coverage: three supply side, two demand side, plus quality:</p> <ul style="list-style-type: none"> • Supply side: availability of essential health commodities; availability of human resources; accessibility – physical access of services • Demand-side: initial utilisation – first contact of multiple contact services; adequate coverage – continuity/completion of care • Quality: effective coverage – quality/impact on health outcomes. <p>The gap in coverage for each determinant reflects the size of the bottleneck to health system effectiveness. The study is usually done by focusing on a specific set of medical services, or by using ‘tracer’ interventions that serve as proxies for the broader functioning of the health system.</p>
Information requirements	<p>The quantitative and qualitative data requirements are typically extensive.</p> <p>As an example, the data in a Uganda study was collected as part of a larger donor project through repeated household and health facility surveys. This was done across two districts over 2.5 years from November 2011 to April 2014. The household survey used continuous cluster surveys that covered 6,513 women who were pregnant 12 months prior to the survey. The facility survey covered 50 facilities surveyed every four months. This represents a significant investment in data-gathering, which will likely have use beyond the original Tanahashi study that collected it.</p>
Timeframe	<p>Undertaking large-scale panel/cluster surveys over time in multiple locations will require a long timescale. Survey design and analysis/publication are also likely to be considerable.</p> <p>The mentioned Uganda study had a significant data collection process carried out quarterly over a 2.5-year period.</p>
Inputs required	<p>A full Tanahashi study is likely to be costly, given the scale of the data collection it requires.</p> <p>Total costs for the Uganda project cited are not available. For a Tanzania case study, the estimated cost per district of carrying out the survey was around \$20,000. This included: training district staff in data collection; three rounds of surveys (the challenge of which will vary depending on the size and population of the district); and holding workshops with district staff to analyse the results. The cost of analysis, write-up, publication and dissemination would be additional.</p>
Access to information	<p>This study approach requires collection of both existing secondary data (from government financial and/or service delivery data systems) as well as primary data collection (predominantly through qualitative surveys or focus groups). As with PETS studies noted before, a lack of financial information and/or poor quality financial information from government sources is likely to undermine the robustness of the study.</p>
Stakeholder engagement	<p>Given the scope and intensity of data collection, engagement from government at different levels (national, district, facility) is required.</p>
Selected commentary	<p>The modified Tanahashi analysis aims to determine bottlenecks to health service delivery on both the supply side and the demand side. It is not solely PFM-focused, but instead looks broadly at service-level delivery bottlenecks and challenges (of which PFM-related issues will be only a part). As the analysis is only done at the district (or other lower-tier) level of delivery and is mostly quantitative, it is less able to identify the high-level PFM constraints that might be creating the bottlenecks. For example, having identified staff shortages as a blockage, the survey does not fully investigate the higher-level reasons behind staffing shortages (e.g. is it poor distribution of staff across the country, an overall shortage of financing for staff etc.)</p> <p>The tool is able to effectively identify non-PFM constraints to service delivery, including factors such as health-user awareness and health-provider compliance with protocols. Indeed, the broad nature of the tool encourages this wide approach and some of the studies reviewed made extensive use of qualitative interview or focus group information to provide a great deal of information on the non-PFM drivers of (non-)delivery.</p>

Table 13 Diagnostic characteristics of WHO Health Financing Country Diagnostic

Diagnostic characteristic	WHO Health Financing Country Diagnostic
Key objectives	A top-down narrative/descriptive tool focused on health financing. It aims to: (1) assess a country's health financing system relative to the goal of universal health coverage (UHC); and (2) analyse the challenges faced in moving towards UHC. This is intended to inform the development of a health financing reform strategy.
Example of application	Global Network for Health Equity (GNHE) report for Uganda in 2014 (Zikusooka et al., 2014).
Number of times used (estimate)	GNHE have applied the tool in at least 18 countries.
Methodology	<p>The approach is predominantly qualitative with some supporting quantitative data.</p> <p>The diagnostic carries out:</p> <ul style="list-style-type: none"> • a quantitative assessment of levels, trends and the composition of health system spending covering high-level issues like total health expenditure, government expenditure, private expenditure etc. • a descriptive qualitative analysis of three pre-identified key health financing sub-systems with a view to identifying areas of under-performance. The three areas of health financing focus are revenue-raising mechanisms, pooled funding arrangements and purchasing • a descriptive qualitative analysis of progress towards the UHC goals (equity, quality, financial protection) and associated intermediary objectives (equity in resource distribution, efficiency and transparency and accountability) with a view to reaching plausible conclusions as to the likely reasons why the system is under-performing.
Information requirements	<p>The quantitative data required are usually available from government sources or pre-existing databases such as WHO's Global Health Expenditure Database.</p> <p>Qualitative descriptive data on institutional structures and processes can be collected using a series of suggested questions. The framework suggests various dimensions, categories and question areas that can help to structure an assessment of the data gathered.</p> <p>Some aspects of the qualitative assessment may need to draw on previous analysis to fully consider certain complex issues such as financial protection (including estimates of the incidence of catastrophic and/or impoverishing health expenditure) and the overall benefit incidence of health expenditure.</p>
Timeframe	Not found.
Inputs required	Completing the diagnostic requires a team of health economists familiar with financing issues.
Access to information	Not found. Some of the quantitative data could be found in public documents (e.g. national health accounts, government expenditure data). However, the qualitative information would require access to government officials.
Stakeholder engagement	Not found; although as noted the requirement of qualitative interview data would suggest that government engagement would be required in order to make the data gathering a success.
Selected commentary	<p>This approach aims to qualitatively assess a country's health financing system relative to the goal of UHC so as to inform the development of a high-level health financing reform strategy. To do this, data are compiled on health financing and the degree to which the UHC goals (equitable coverage, quality and financial protection) are currently being reached.</p> <p>This tool provides a high-level and descriptive approach to considering health financing and PFM in a national health system. There are few externally defined benchmarks for what counts as achieving the UHC goals and as a result, the extent to which the health financing system is missing and/or achieving these goals is partly a matter of judgement. The linking of PFM issues to ground-level service delivery is not the core focus of the tool and has to be assumed. The resulting diagnostic may point to the general presence of underlying PFM constraints to service delivery insofar as they affect high-level progress towards UHC, but does not necessarily explicitly discuss them. The tool does identify non-PFM blockages to service delivery to some extent in that it encourages a discussion of the broader reform environment in the context of moving to UHC. The tool puts a great deal of emphasis on analysing context, yet leaves it up to the user to interpret the specific relevance of local service delivery contexts to the health financing system.</p>

Table 14 Diagnostic characteristics of World Bank Public Expenditure Review

Diagnostic characteristic	World Bank Public Expenditure Reviews (PERs)
Key objectives	To assess the relationship between the public expenditure going into a particular sector, and outputs and outcomes achieved by that sector over time, with a view to assessing overall efficiency and effectiveness of public spending.
Example of application	A large number of PERs have been completed in developing country health sectors (among many other sectors). Examples can be found on the World Bank website.
Number of times used (estimate)	There are 589 PERs available on the World Bank website, of which a great many deal with health. Often health is examined in combination with other topics (e.g. health and nutrition; health and population).
Methodology	A mixture of qualitative and quantitative methods are used to assess the efficiency and effectiveness in public spending. The core PER guidance (dating from 1996) sets out some common elements of a PER approach; however, in practice, individual PERs adapt and amend the core guidance to address areas of specific concern. Common characteristics of PERs are both: (1) comparisons over time so as to assess relative change in sector spending and outcomes; and (2) comparisons between similar countries so as to benchmark performance against comparators.
Information requirements	Significant. Detailed information on sector and sub-sector spending over time is required; along with information on outputs and outcomes achieved by sector and sub-sector.
Timeframe	Not explicitly stated; but PERs can take up to two years to deliver from inception through to publication.
Inputs required	A multi-disciplinary team of sector specialists is required. PERs typically require staff with expertise in econometric and statistical analysis.
Access to information	Most PERs (or at least, those that make it to publication) appear to be successful in gathering the information they require from government.
Stakeholder engagement	Significant engagement with government would be required. Government is likely to hold the key financial and performance information required to undertake the core work of a PER. The core of a PER is on 'whole-of-sector' inputs and outputs from a top-down perspective at various levels, but some also undertake original research at a lower level, for example through a sample of facility-level outputs or engagement with service providers and/or users (although this is not universally done).
Selected commentary	<p>PERs are a well-known and longstanding policy diagnosis tool that touches on aspects of public expenditure management in a particular sector. There are now more than 400 PER studies available in English on the World Bank website. These studies can take a number of forms, but broadly they seek to analyse expenditure compared to outcomes over time in selected sectors, compare performance to peer countries and set out a range of recommendations for reform to improve effectiveness of government expenditure overall. They rarely consider PFM issues in depth; typically they are more concerned with issues of expenditure, the value for money of sector policy choices and overall outcomes. PERs may consider the flow of funds through the sector, but broader PFM institution-related issues are not usually a particular focus of these studies.</p> <p>Recently the World Bank has begun using an 'integrated PER' approach that actively considers both PFM institutional strength and performance, as well as policy issues within a particular sector. An integrated approach can involve an assessment of PFM institution effectiveness across the sector in areas of: (1) the legal framework; (2) budget planning and preparation; (3) budget execution and reporting; and (4) compliance and review. This move towards integration comes, in part, from a recognition that many of the recommendations of a PER cannot be effectively implemented without improvements and reform to the underlying PFM institutions that manage finance in a particular sector (personal communication, 2019). More closely joining up analysis of the effectiveness of sector policy with the effectiveness of the PFM systems that support service delivery could therefore lead to more useful and impactful policy recommendations.</p>

References

- Allen, R., Schiavo-Campo, S. and Garrity, T.C. (2004) *Assessing and reforming public financial management, a new approach*. Washington DC: World Bank
- Andrews, M. (2007) 'What would an ideal public finance management system look like?' in A. Shah (ed.) *Budgeting and Budgetary Institutions*: 359–383 Washington DC: World Bank
- Andrews, M. (2008) *Good government means different things in different countries*. Working Paper RWP08-068. Cambridge MA: Harvard Kennedy School of Government
- Andrews, M. (2013) *The limits of institutional reform in development: changing rules for realistic solutions*. Cambridge: Cambridge University Press
- Andrews, M., Pritchett, L. and Woolcock, M. (2012) *Escaping capability traps through problem-driven iterative adaptation (PDIA)*. CFGD Working Paper 299. Washington DC: Centre for Global Development
- Andrews, M., Cangiano, M., Cole, N., de Renzio, P., Krause, P. and Seligmann, R. (2014) *This is PFM*. Cambridge MA: Center for International Development at Harvard University
- Andrews, M., Pritchett, L. and Woolcock, M. (2017) *Building state capability: evidence, analysis, action*. Oxford: Oxford University Press
- Barroy, H. and Kabaniha, G. (2018) 'Building strong public financial management systems towards universal health coverage: key bottlenecks and lessons learnt from country reforms in Africa.' Draft technical background report for the regional workshop on public financial management for sustainable financing for health in Africa, 25–28 September. Nairobi: World Health Organization
- Barroy, H., Kabaniha, G., Boudreaux, C., Cammack, T. and Bain, N. (2019) *Leveraging public financial management for better health in Africa: key bottlenecks and opportunities for reform*. Health Financing Working Paper No. 14. Geneva: World Health Organization
- Cashin, C., Bloom, D., Sparkes, S., Barroy, H., Kutzin, J. and O'Dougherty, S. (2017) *Aligning public financial management and health financing: sustaining progress toward universal health coverage*. Geneva: World Health Organization
- Engen, L. and Prizzon, A. (2019) *Exit from aid: an analysis of country experiences*. London: Overseas Development Institute
- Fan, V.Y. and Savedoff, W.D. (2014) 'The health financing transition: a conceptual framework and empirical evidence' *Social Science and Medicine* 105: 112–121
- Gauthier, B. (2010) *PETS-QSDS in sub-Saharan Africa: a stocktaking study*. Washington DC: World Bank.
- Gauthier, B. and Reinnika, R. (2007) *Methodological approaches to the study of institutions and service delivery: a review of PETS, QSDS and CRCS*. Washington DC: World Bank
- Grindle, M.S. (2007) 'Good enough governance revisited' *Development Policy Review*, 25(5): 553–574
- Gurkan, A., Kaiser, K. and Voorbraak, D. (2009) 'Implementing public expenditure tracking surveys for results: lessons from a decade of global experience.' PREM Notes No. 145. Washington DC: World Bank
- Hadley, S. and Miller, M. (2016) *PEFA: what is it good for? The role of PEFA assessments in public financial management reform*. London: Overseas Development Institute
- Hafner, T. and Shiffman, J. (2013) 'The emergence of global attention to health systems strengthening' *Health Policy and Planning*, 28(1): 41–50. (doi:10.1093/heapol/czs023)
- Henriksson, D.K., Fredriksson, M., Waiswa, P., Selling, K. and Swartling Peterson, S. (2017) 'Bottleneck analysis at district level to illustrate gaps within the district health system in Uganda' *Global Health Action* 10(1)

-
- Hood, C., King, M. and Piotrowska, B. (forthcoming) ‘Evaluating the efficacy of expenditure control over time: a new standard approach?’ Unpublished conference paper
- IHME – Institute for Health Metrics and Evaluation (2017) *Financing global health 2016: development assistance, public and private health spending for the pursuit of universal health coverage*. Seattle WA: IHME
- Knack, S., Kugler, M. and Manning, N. (2003) ‘Second-generation governance indicators’ *International Review of Administrative Sciences* 69: 345–364
- Koziol, M. and Tolmie, C. (2010) *Using public expenditure tracking surveys to monitor small projects and small-scale programs: a guidebook*. Washington DC: World Bank
- Lateef, K.S. (2016) *Evolution of the World Bank’s thinking on governance*. Background paper for 2017 World Development Report. Washington DC: World Bank
- Levy, B. (2014) *Working with the grain: integrating governance and growth in development strategies*. Oxford: Oxford University Press
- Mackie, A. (2011) *Stocktaking study of PFM diagnostic instruments volume I – main report*. 4th High Level Forum on Aid Effectiveness
- Nixon, H. and Bredenkamp, C. (2014) *Timor-Leste health resource tracking study*. Washington DC: World Bank
- O’Connell, T. and Sharkey, A. (2013) *Reaching universal health coverage through district health system strengthening: using a modified tanahashi model sub-nationally to attain equitable and effective coverage*. UNICEF Maternal, Newborn and Child Health Working Paper, December 2013
- Odaga, J., Henriksson, D.K., Nkolo, C., Tibeihaho, H., Musabe, R., Katusiime, M., ... and Valadez, J.J. (2016) ‘Empowering districts to target priorities for improving child health service in Uganda using change management and rapid assessment methods’ *Global Health Action* 9(1)
- PEFA Secretariat (2018) *Stocktake of PFM diagnostic tools 2016*. Washington DC: PEFA Secretariat
- PEFA Secretariat (forthcoming) ‘Using PEFA to support PFM improvement: volume IV of the PEFA handbook.’ Washington DC: PEFA Secretariat ([https://pefa.org/sites/default/files/3Using%20PEFA%20to%20support%20PFM%20Reform%20\(Vol%20IV\).pdf](https://pefa.org/sites/default/files/3Using%20PEFA%20to%20support%20PFM%20Reform%20(Vol%20IV).pdf))
- Piatti-Fünfkirchen, M. and Smets, L. (2019) *Public financial management, health financing and under-5 mortality*. IDB Working Paper Series IDB-WP-976. Washington DC: Inter-American Development Bank
- Schick, A. (1998) ‘Why most development countries should not try New Zealand’s reforms’ *World Bank Research Observer* 13(1): 123
- Schick, A. (2013) ‘Reflections on two decades of public financial management reforms’, in M. Cangiano, T. Curristine and M. Lazare (eds.) *Public financial management and its emerging architecture: 21–76* Washington DC: International Monetary Fund
- Silverman, R. (2018) *Projected health financing transitions: timeline and magnitude*. CGD Working Paper 488. Washington DC: Centre for Global Development
- Welham, B., Krause, P. and Hedger, E. (2013) *Linking public financial management dimensions to development priorities*. London: Overseas Development Institute
- Welham, B., Hart, T., Mustapha, S. and Hadley, S. (2017) *Public financial management and health services delivery*. London: Overseas Development Institute.
- WHO – World Health Organization (2007) *Everybody’s business: strengthening health systems to improve health outcomes: WHO’s framework for action*. Geneva: World Health Organization
- WHO (2010) *Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies*. Geneva: World Health Organization (http://www.who.int/healthinfo/systems/WHO_MBHSS_2010_full_web.pdf)
- Witter, S., Palmer, N., Balabanova, D., Mounier-Jack, S., Martineau, T., Klicpera, A., ... and Gilson, L. (2019) ‘Health system strengthening—reflections on its meaning, assessment, and our state of knowledge’ *International Journal of Health Planning and Management* 34(4): e1980–e1989

Xu, K., Soucat, A., Kutzin, J., Brindley, C., Maele, N.V., Touré, H., ... and Siroka, A. (2018) *Public spending on health: a closer look at global trends*. Geneva: World Health Organization

Zikusooka, C.M., Kwesiga, B., Lagony, S. and Abewe, C. (2014). *Universal Health Coverage Assessment Uganda*. The Global Network for Health Equity (http://gnhe.org/blog/wp-content/uploads/2015/05/GNHE-UHC-assessment_Uganda-1.pdf)



**Evidence.
Ideas.
Change.**

ODI is the UK's leading independent think tank on international development and humanitarian issues.

Our mission is to inspire and inform policy and practice which lead to the reduction of poverty, the alleviation of suffering and the achievement of sustainable livelihoods.

We do this by locking together high-quality applied research, practical policy advice and policy-focused dissemination and debate.

We work with partners in the public and private sectors, in both developing and developed countries.

ODI
203 Blackfriars Road
London SE1 8NJ

+44 (0)20 7922 0300
info@odi.org

odi.org
odi.org/facebook
odi.org/twitter