



HSS Evaluation Collaborative

Learnings for HSS Evaluations

July 2022

Working Group 1

Contents

Acronyms	ii
Executive summary	iv
Purpose and objectives of this review	iv
Methodology and limitations	iv
Findings	iv
GHI evaluation objectives, approaches, and methodologies	iv
Challenges of HSS evaluations	vii
Use and uptake of HSS evaluation findings	vii
Strengths, weaknesses and lessons learned	viii
Approach and methods	viii
Utilization of findings	viii
Future considerations	viii
1 Background	1
2 Methods and limitations	1
3 Findings	2
3.1 HSS strategic focus and priorities	2
3.2 GHI evaluation objectives, approaches, methodologies	4
3.3 Challenges of HSS evaluations	9
3.4 Use and uptake of HSS evaluation findings	10
3.5 Strengths, weaknesses and lessons learned from evaluation approaches methods and utilization	11
Approach and methods	11
Utilization	12
4 Future considerations	13
Addressing challenges and gaps in HSS evaluation	13
Annex 1	14
Annex 2: FA2 Review of HSS evaluations: Review of Findings (WIP) October 21 2021	21

Acronyms

APL	Adaptable program loan
CBHI	Community-based health insurance
FCE	Full country evaluations
GF	The Global Fund to Fight AIDS, Tuberculosis and Malaria
GHI	Global Health Institutions
HSIS	Health System & Immunisation Strengthening
HSS	Health System Strengthening
HSSE	Health System Strengthening Evaluation
HSSEC	Health System Strengthening Evaluation Collaborative
IDIQ	Indefinite Delivery, Indefinite Quantity contract
IEG	Independent Evaluation Group
KII	Key informant interview
LHSS	Local health system sustainability
MEL	Monitoring, evaluation and learning
MTAPS	Medicines, Technologies and Pharmaceutical Services
OIG	Office of Inspector General
PBF	Performance-based financing
PCE	Prospective Country Evaluations
PPAR	Project Performance Assessment Report
RSSH	Resilient Sustainable Systems for Health
TERG	Thematic Evaluation Reference Group

TOC	Theory of Change
TRP	Technical Review Panel
UHC	Universal health coverage
USAID	The United States Agency for International Development
WBG	The World Bank Group
WHO	World Health Organization

Executive summary

Purpose and objectives of this review

Health systems strengthening (HSS) is widely understood to be key to achieving universal health coverage and to ensuring robust responses to health emergencies. In recent decades, global health investors have put more attention and investment towards HSS, leading to accelerated efforts to evaluate HSS policies and programs initiated by those investments. Yet, a common definition and framework for how to evaluate HSS interventions remains elusive, hampering efforts to strengthen, coordinate and amplify HSS programs.

The Health Systems Strengthening Evaluation Collaborative (HSSEC) brings together key global and national stakeholders to suggest ways to strengthen the quality of evaluations of health systems strengthening (HSS) investments in LMICs and to improve coordination across stakeholders in this space.

This paper aims to inform learning and collaboration to improve the understanding, performance and quality of Health Systems Strengthening (HSS) evaluation.

The paper provides an overview of HSS evaluations commissioned by Global Health Institutions¹ (GHI) from 2010 to the present. It describes the GHI's strategic focus, priorities, objectives and methodological approaches to HSS evaluation and identifies challenges and lessons learned. The paper also describes how evaluation findings have been used and highlights opportunities for improvement.

Methodology and limitations

Methods used for developing the review were:

- A focused literature review of GHI-related HSS evaluations based on GHI website searches and recommendations from members of Health System Strengthening Evaluation Collaborative (HSSEC) Working Group 1 (WG1). The evaluation samples from each GHI were selected based on evaluation type, objectives and methodologies, and data availability and quality. Documents included country evaluation reports, global evaluation and synthesis reports, meta-analysis reports, position papers and strategy documents, and annual project report documents.
- The literature review was supplemented by five key informant interviews (KIIs) with GHI stakeholders and consultant evaluators to explore the utilization of findings.

Relevant data were extracted, reviewed and synthesized to generate findings for this report. Given the heterogeneity of the evaluations, there was limited scope for comparing GHI evaluation performance; instead, the paper seeks to highlight some key challenges and lessons learned.

Findings

GHI evaluation objectives, approaches, and methodologies

- **The conceptualization of current HSS evaluations falls into two main categories: evaluation which reflects systems thinking, and evaluation focused on systems related to specific diseases or national immunization programs.** World Bank Group (WBG) and USAID's thinking, learning and evaluation practices reflect the understanding of health

¹ Gavi, the Vaccine Alliance (Gavi), the Global Fund to Fight AIDS, Tuberculosis and Malaria (GF), the United States Agency for International Development (USAID) and the World Bank Group (WBG) – referred to in this paper as Global Health Institutions (GHIs).

systems as a comprehensive set of interconnected parts that must function together to be effective and that changes in one area will have repercussions elsewhere in the system. For Gavi and the Global Fund (GF), evaluation approaches largely reflect the primary focus of support which is strengthening immunization and disease-specific delivery systems, with limited evaluation of wider health system benefits and impacts.

- **Three broad categories of evaluation can be identified** from the literature:
 - **Evaluations of GHI HSS-specific policies and related programming.** For example, GF review of support to Resilient Sustainable Systems for Health (RSSH) through RSSH-specific and disease grants; Gavi country and global evaluations and meta-analysis of HSS grants; USAID's HSS programs such as the Local Health System Sustainability Program (LHSS); WBG's evaluation of support to Health System Modernisation in Armenia.
 - **Evaluations of HSS-specific interventions.** For example, the evaluation of USAID's Medicines, Technologies and Pharmaceutical Services (MTAPS) program, which aims to strengthen a country's pharmaceutical systems and supply chain management.
 - **Evaluations of HSS interventions as part of an evaluation of a wider portfolio of GHI support or strategy.** For example, GF Strategic Reviews and Prospective Country Evaluations (PCE) and Gavi Full Country Evaluations (FCE), which included some assessments of RSSH and HSS interventions as part of a wider portfolio analysis of GF and Gavi support, and WBG health sector evaluations which assess some aspects of HSS in relation to evaluating WBG's contribution to primary health care programs.
- **Evaluations vary in objectives, instruments and independence.** For example, USAID and WBG conduct self-evaluations of HSS closely linked to decision-making and organization learning objectives; WBG, Gavi and GF conduct independent HSS evaluations with an emphasis on accountability and learning at the strategy/Board level.
- **Most types of evaluations employ mixed methods to address the multiple objectives and evaluation questions of HSS evaluations.**
- **Evaluation approaches have evolved to assess the complexity of programs and related HSS evaluation questions through a shift towards prospective evaluation, which aim to understand not only 'what happened' but, crucially, why change occurred.** This is evidenced through Gavi and GF investments in FCE and PCE and USAID's HSS programs, which have moved to prospective monitoring and learning throughout program implementation. The idea of longitudinal work has benefits, but the implementation of prospective work can be considerably improved.
- **Theory-based HSS evaluations are common, with varying degrees of use of Theories of Change (TOC).** TOCs are often used for GHI organizational strategies and some monitoring, evaluation and learning (MEL) plans. There is evidence of TOCs being used for some HSS evaluations, such as the WBG's impact evaluations. Still, in general, few evaluations sufficiently elaborate on the theory and underlying assumptions to be evaluated or discuss how evaluation findings validated or modified the assumptions of change. In the case of the GF, there is no TOC for how GF RSSH and/or community systems investments are expected to strengthen the health system.
- **There is limited discussion of the relevance and suitability of evaluation methods and approaches used for the HSS evaluations in the evaluation reports reviewed.** Although there is some evidence of Independent Evaluation Group (IEG) evaluations ranking the relevance of an evaluation design and methods, notable across the literature was the

limited discussion or reflection on the suitability of evaluation methods and data sources for the questions being evaluated, making it difficult to draw conclusions about the appropriateness of methods and value of evaluations to GHIs/countries. For example, few evaluations elaborate on the suitability of methods used to evaluate the contribution of GHI HSS support in relation to program outcomes.

Challenges of HSS evaluations

Evidence from the review highlights common challenges with HSS evaluations:

- **A lack of common definitions, approaches and guidance** across and within some GHIs on what HSS is, what it is trying to achieve, and how to do it. This makes evaluation design challenging in terms of identifying what is important to measure and track and determining the appropriate choice of evaluation methods.
- **The unclear scope of HSS grants/interventions or unclear TOCs** limits understanding of what HSS interventions are trying to achieve and the complex causal links and pathways to achieve intended outcomes/impact. These challenges are less likely to be reported in evaluations of specific HSS interventions with more clearly defined objectives and evaluation questions compared to evaluations of HSS interventions that were part of an evaluation of a wider portfolio of GHI support or strategy.
- **Methodological challenges concerning the use of indicators and availability of data, and difficulties in establishing a counterfactual.**
- **Inadequate time and resources to observe behavioral change within health systems** with program/grant cycle logic and timelines widely perceived to be too short (usually three-year cycles and often assessed annually), affecting the feasibility of evaluations to demonstrate change/impact.
- **The lack of an overarching strategy for undertaking HSS evaluations within GHIs.** The evidence suggests it is not always clear how HSS evaluations fit within broader organizational strategic frameworks and plans, raising questions regarding purpose, choice of evaluation, target audience, and rationale for multiple evaluations (either of the same grant/intervention over time or of different grants/intervention) and timing.

Use and uptake of HSS evaluation findings

A number of interrelated approaches broadly characterize the use of evaluation findings:

- **'Within-program' learning** was a key objective of monitoring and evaluation work in prospective evaluations (Gavi FCEs, GF's PCEs, USAID's annual reports of LHSS) and WBG project-level self-evaluations.
- **Summarizing program achievements and results and identifying lessons learned** were a key focus of retrospective evaluations, USAID HSS end-of-program reports, and some Gavi country evaluations of HSS.
- **Informing policies and strategic directions**, as is the intention of IEG thematic evaluations such as the review of WBG's Support to Health Services, and also the GF Thematic Review of RSSH, which aimed to generate evidence on the implementation, successes and challenges related to RSSH support, to inform future Strategy Committee/Board decisions.
- **There was limited insight from the literature regarding how evaluation learnings and findings were used and their value and influence at global and country levels.** For example, there was some limited evidence and examples in PCE synthesis reports for how PCE findings were informing programmatic changes at country level and enabling changes in business model policies and practices at Secretariat level. The FCE real-time data collection allowed for early identification of issues at country level, but it was not evident how often action was taken as a result of FCE findings and what this achieved.
- Gavi carried out multiple process and strategy reforms; however, the challenges flagged by evaluations over time (e.g., difficulties with understanding the expectations of investments and use case of the evaluation; time and data constraints presenting a

technical challenge for evaluation) have not all been addressed to date. That said, there is evidence that Gavi has used findings and recommendations from HSS evaluations to shift the focus of HSS activities away from supporting any HSS activities to only funding HSS activities related to immunization.

Strengths, weaknesses and lessons learned

Approach and methods

- Of key importance, emerging from the review and KIIs is the necessity within GHIs to develop and agree on common HSS evaluation guidance, a set of appropriate HSS indicators within GHIs, and a common approach to asking evaluation questions across a range of contexts, to enable the build-up of country comparisons over time.
- More structured approaches to HSS evaluation, including using TOCs, conceptual frameworks, data gathering tools and standardization of evidence for synthesis, have been found to promote better comparability of findings across countries. However, these approaches need to ensure that context-specific evaluation needs are addressed using local data, for example, by providing global level templates, evaluation tools and guidance for evaluation plans and encouraging local adaptation of templates and tools.

Utilization of findings

- Review evidence points to the critical importance of joint collaboration between GHIs and country stakeholders in the design (inception phase) of evaluations to ensure country buy-in from an early stage and clarity regarding evaluation scope, what the HSS interventions are expected to achieve, and generate interest in learning and use of findings.
- Best practices emphasize tailoring dissemination strategies to stakeholder needs and processes and providing knowledge translation support. Lessons from GF PCE suggest this was limited by resources and capacities, and as a result, dissemination processes were largely focused on the annual report and dissemination meeting, which did not always align with program implementation timelines or decision-making windows. A key lesson for future evaluations is to ensure resources are available for multiple modes and more timely dissemination of findings.

Future considerations

This review identified published evaluations carried out by the GHIs and synthesized key findings, challenges and lessons learned. Future considerations based on the review of evaluations come from recommendations in the documentation and an analysis of gaps and good practices identified in the literature and by the review team.

- **GHI guidance for evaluators should be standardized and coordinated across countries** to promote improved quality and use of findings and should be consistently used by evaluators to support the comparability of findings across contexts.
- **Common definitions of HSS and an appropriate set of agreed indicators should be developed** to systematically track HSS interventions and enable a robust HSS evidence base to develop, supporting country comparisons over time.
- **Develop and/or strengthen theoretical frameworks and TOCs to clearly articulate expectations for HSS support** and elaborate on assumptions for how interventions will bring about change. Ensure GHI technical guidance on HSS is congruent with the TOC and other related organizational systems and monitoring, evaluation and learning thinking.

- **Continue to learn and build on the experience of assessing GHI contribution to outcomes** and impact through greater consideration and reflection on the suitability of methodologies to achieve this (e.g., HSS-specific prospective and retrospective evaluation approaches and through a combination of established and experimental mixed methods).
- **Ensure resources (technical, human, time) are available to support the close involvement of GHIs and country stakeholders** in the design and implementation of evaluation and learning efforts and ensure adequate feedback mechanisms are in place to promote the uptake of evaluation findings.
- **Strengthen cooperation, coordination of HSS approaches, investments, evaluations, and technical support across GHIs**, including at country level. This should be done to improve complementarity and identify niche areas for how each GHI can support HSS and wider health goals.
- **Where appropriate, establish organization-wide MEL plans and structures** that support rationale, planning and organization of HSS evaluations, linked to wider strategy, and continue to provide and strengthen learning platforms for collecting and disseminating evidence and findings.
- **Continue to strengthen and utilize learning components of HSS evaluation** to advance understanding, practice and lessons learned on systems thinking, methods and approaches appropriate to evaluation purpose/objectives.

1 Background

The paper was undertaken by the Health Systems Strengthening Evaluation Collaborative (HSSEC) Working Group 1 (WG1). It focuses on HSS evaluations commissioned by Gavi, the Vaccine Alliance (Gavi), the Global Fund to Fight AIDS, Tuberculosis and Malaria (GF), the United States Agency for International Development (USAID) and the World Bank Group (WBG) – referred to in the paper as Global Health Institutions (GHIs).

The paper intends to inform learning and collaboration and serve as a reference for discussion, aiming to improve understanding, performance and quality of Health System Strengthening (HSS) evaluation. Specifically, the paper aims to:

- Provide an overview of GHI-commissioned HSS evaluations from 2010 to the present day.
- Describe the strategic focus and priorities for HSS, objectives and methodological approaches of the GHI-commissioned HSS evaluations.
- Identify challenges, strengths, and weaknesses of the GHI-commissioned HSS evaluations, from structuring evaluations to disseminating and utilizing evaluation findings.
- Describe how evaluative approaches have been applied to answer HSS evaluation questions and highlight opportunities for improvement.

2 Methods and limitations

Six research questions were developed and refined during data collection to achieve the objectives outlined above. Additionally, a process-oriented assessment framework was devised to provide an overarching understanding of organizational performance and assessment in HSS evaluation and to identify categories of data to be extracted from documents for analysis and synthesis.

A focused literature review of GHI-related HSS evaluations was undertaken based on GHI website searches and recommendations from members of HSSEC WG1. The evaluation samples from each GHI enabled a variety of evaluations to be reviewed based on evaluation type, evaluation objectives and methodologies, and availability and quality of data.

The documents reviewed included country evaluation reports, global evaluation synthesis and meta-analysis reports, global HSS thematic reviews, GHI position papers, and annual project reports. The literature review was supplemented by five key informant interviews (KIIs) with GHIs stakeholders and consultant evaluators to explore the utilization of findings.

Data from the literature review and KIIs were extracted, analyzed and synthesized to generate the findings for this report.

Methodological limitations identified include:

- Secondary analysis was undertaken based on publicly available evaluation reports. As such, the findings of this paper are dependent on the content, quality and robustness of those existing evaluation reports.
- The diversity of HSS evaluations undertaken across GHIs and limited data availability for some areas of assessment has made the generalization of findings challenging and limited scope for synthesis. Results should be treated with caution.
- Given the heterogeneity of the evaluations and data constraints, there was limited scope for comparing GHI evaluation performance; instead, this paper seeks to highlight key challenges and lessons learned.

A full listing of evaluation documents reviewed can be found in **Annex 1**, and a review of findings (work in progress slide set) can be found in **Annex 2**.

3 Findings

3.1 HSS strategic focus and priorities

GHIs have different definitions, ideas, strategic focus and priorities of health systems and the means of strengthening health systems are similarly varied. A brief overview of GHI focus and priorities for HSS follows.

Gavi: Gavi has increased HSS commitments over time and supports a range of activities under the HSIS framework launched in 2017. The HSS window has adjusted its stated purpose and guidance, and investments now target four strategic areas - data, supply chain, leadership and coordination, and demand promotion and community engagement. This strategic focus is reported to have helped move the HSS window towards more targeted, outcome-oriented support to countries². HSS funding is used to address health system bottlenecks which hinder progress towards improved immunization demand and coverage. Despite increased funding for HSS (from US\$200m disbursed in 2007-8 to US\$1.3bn approved for HSS support in 2016-2020³), HSS support is small in scale (most HSS grants are under US\$5m per annum) and it is unclear whether these grants are catalytic, as envisioned⁴.

Gavi has more than a decade's experience in HSS evaluation and has established an evaluation system specific to its HSS investments. Some evaluations are managed by countries and others by the Gavi Secretariat (see Table 1). Retrospective and prospective evaluations, country and global level HSS and meta-reviews have successfully identified evaluation learning and improved HSS grant management and implementation. However, the evaluations highlight the continued challenge of generating good quality evidence linking HSS investments to improved coverage rates or stronger health systems.

Global Fund: The GF has increased commitments and investment in Resilient Sustainable Systems for Health (RSSH). As of 2019, it was the largest multilateral investor in grants for systems for health (over US\$1bn a year) with funds targeting seven operational objectives, including strengthening laboratory and data systems and integrated service delivery⁵.

The GF has managed RSSH evaluations largely through the Thematic Evaluation Reference Group (TERG), which has commissioned broader Prospective Country Evaluations (PCEs) and Strategic Reviews as well as RSSH-specific reviews, and to a lesser extent through the Technical Review Panel (TRP) and Office of Inspector General (OIG).

Evaluation reports highlight strategic issues such as a significant proportion of investments (66%) focusing on short-term 'systems support' such as salaries or equipment, rather than longer-term changes in policies, structures and behaviors, which could sustain change, weak performance metrics for RSSH and the need to better define and clarify expectations of RSSH support. The latest strategy (2023 to 2028) further defined RSSH priorities, emphasizing integrated people-centered services and community systems.

Additional issues raised by evaluations include the necessity to resolve 'tensions' regarding the GF's approach to RSSH investments. These include the prioritization and operationalization of RSSH funding vis-à-vis disease-specific grants, which can complicate evaluation as RSSH investments are more difficult to track when programmed through disease grants.

² Kenney C, Glassman A (2019) Gavi's approach to health systems strengthening, Centre for Global Development, CDG Note.

³ Ibid., 9

⁴ Ibid., 8

⁵ The Global Fund (2019) 41st Board Meeting Report of the Technical Evaluation Reference Group for Board Information GF/B41/11 Geneva 15-16 May 2019 https://www.theglobalfund.org/media/8529/bm41_11-terg_report_en.pdf

USAID: USAID launched their *Vision for HSS 2030*⁶ in 2021, which adopts of 'whole of society approach' to addressing health system challenges and prioritizes the achievement of three interrelated outcome goals – equity, quality, and resource optimization. The Vision builds on USAID's previous focus on health system functions (the six building blocks). It recognizes that by orienting support towards achieving health system outcomes, most activities will integrate across more than one building block. The Vision is grounded in a TOC and complemented by an HSS Learning Agenda⁷, guided by six HSS learning questions on systems thinking, measurement and sustainability. It supports program improvement by generating, synthesizing and disseminating evidence from progress in implementing USAID's HSS programs.

USAID evaluations of HSS programs are commonly presented in the form of regular or final program reporting. There is not one standardized approach to evaluation required across USAID-funded HSS programming. TOCs are not usually presented in reporting; however, there has been a demonstrable shift to incorporating results frameworks in evaluations in recent years.

USAID's HSS investments are implemented through multicountry, multi-year (six to eight years), large-scale programs such as the Local Health System Sustainability project (LHSS) and Health Systems 20/20. These programs often contain smaller-scale pilot interventions and evaluations at country levels, e.g. Community-Based Health Insurance (CBHI) Pilot Schemes in Ethiopia. HSS programs have had a ceiling of approximately US\$210m over the course of the program.

World Bank: Since 2007, WBG's health policy focus has shifted to a health systems approach implemented through mechanisms such as the Health Results Innovation Trust Fund, the lending instrument 'Program-for-Results', and the Global Financing Facility, established in 2015. WBG's health sector investments and research are focused on helping countries achieve universal health coverage (UHC) by 2030 through stronger primary health systems and the provision of quality and affordable access to health services for all. Within its focus area of strengthening health systems and health financing, WBG invests in four priority areas:

- Ramping up investments in affordable, quality primary healthcare
- Engaging the private sector and unlocking new models for health financing and delivery
- Going beyond health to improve health outcomes and support communities
- Changing health financing by catalyzing domestic resources to build sustainable national health systems⁸

Evaluations are delivered through three main modalities, which have strong learning objectives: independent evaluation conducted by the Independent Evaluation Group (IEG), mandatory self-evaluation, and demand-driven self-evaluation. IEG evaluations relevant to HSS are structured as multi-level, multi-project 'Major Evaluations', which may comprise corporate, thematic and sectoral evaluations, and 'Meso Evaluations', such as country and project level evaluations.⁹ Over the past ten years, the International Development Association¹⁰ has provided US\$13.5 billion to fund essential health interventions. However, this is identified as insufficient, with the financing gap to achieve UHC in the poorest 54 countries estimated at US\$176 billion annually.¹¹

⁶ USAID, *USAID Vision For Health System Strengthening 2030*, 2021. <https://www.usaid.gov/global-health/health-systems-innovation/health-systems/Vision-HSS-2030-download>

⁷ USAID, *Health Systems Strengthening Learning Agenda*, 2021. <https://www.usaid.gov/global-health/health-systems-innovation/health-systems/resources/learning-agenda>

⁸ World Bank, *Universal Health Coverage*, 2022. <https://www.worldbank.org/en/topic/universalhealthcoverage#2>

⁹ World Bank Group, *Evaluations*, 2022. <https://ieg.worldbankgroup.org/evaluations>

¹⁰ Part of the WBG, which aims to reduce poverty by providing zero to low-interest loans (<https://ida.worldbank.org/en/about>)

¹¹ <https://www.worldbank.org/en/news/speech/2019/09/23/world-bank-group-president-david-malpass-remarks-at-the-2019-un-high-level-meeting-on-universal-health-coverage-uhc>

3.2 GHI evaluation objectives, approaches, methodologies

The following table illustrates the characteristics of a sample of global and country-level evaluations reviewed for this assignment.

Table 1: Characteristics of HSS evaluations from selected examples of global and country level evaluations reviewed for this assignment

Evaluation	Type	Objectives	Approach and Methods
Gavi			
Multi-country evaluations <ul style="list-style-type: none"> HSS support in 20 countries (2015, 2018). 	<ul style="list-style-type: none"> Retrospective end of grant evaluations commissioned by Gavi. Implemented at country level through local institutions with support from Gavi. 	<ul style="list-style-type: none"> To understand Gavi HSS implementation efficiency, effectiveness, and results, including contribution of Gavi to epidemiological coverage. To provide lessons learned and recommendations for program improvement. 	<ul style="list-style-type: none"> Approach: Evaluation framework covering inputs - outcome level results. Variable use of TOCs. Mixed methods: Literature review, key informant interviews, site visits. Surveys and secondary analysis of program data.
Global Review <ul style="list-style-type: none"> Review of HSS support (2019). 	<ul style="list-style-type: none"> Retrospective evaluation of strategy review commissioned by Gavi. Implemented externally by international public health organization (Swiss TPH). 	<ul style="list-style-type: none"> To assess design, implementation, and results of HSS support. To identify challenges and learn from recent experience to guide further development of HSS support, including the relevance/design of an HSIS evaluation. 	<ul style="list-style-type: none"> Approach: Theory-based using TOC from FCE. Social determinants of health framework informed approach to contribution. Strong focus on results and measurement – inputs to outcomes. Mixed methods: Quantitative analysis to model contribution to outcomes. Document review, key informant interviews, qualitative review of 16 countries.
Global Fund			
Prospective Country Evaluations (PCE) <ul style="list-style-type: none"> RSSH and non-RSSH in eight countries (2018-2021). 	<ul style="list-style-type: none"> Prospective process evaluation of GF portfolio commissioned and managed by TERG, of which RSSH was a component. Implemented externally through two international consortia of global and national research and consulting firms (IHME/PATH; Euro Health Group). 	<ul style="list-style-type: none"> To improve national program performance and understand how GF business model works in practice over time, including helping/hindering progress to results. To support progress towards data-based estimates of outcomes and impact. Provide lessons learned and inform approaches to evaluation. Build in-country evaluation capacity. 	<ul style="list-style-type: none"> Approach: Theory-based with multiple TOCs. M&E framework covered inputs to impact. Mixed methods: Contribution analysis, root cause analysis, grant tracking. Statistical correlation of results chain and health system modelling. Impact evaluation through measurement of outputs (geospatial analysis, coverage indicators, disease burden metrics).

<p>Global evaluations of GF RSSH support</p> <ul style="list-style-type: none"> 2019 TERG RSSH Thematic Review. 	<ul style="list-style-type: none"> Retrospective thematic review related to GF Strategy implementation commissioned and managed by TERG. Implemented externally through international consulting firm. 	<ul style="list-style-type: none"> To assess progress on implementation of RSSH Strategic Objective at the country level, focusing on effectiveness of RSSH investments and improvement of impact of RSSH support. 	<ul style="list-style-type: none"> Approach: Use of analytical framework based on key areas of analysis and evaluation questions. No TOC in use. Mixed methods: Document review, key informant interviews, 11 country case studies, data analysis and synthesis.
USAID			
<p>Country impact evaluation</p> <ul style="list-style-type: none"> Evaluation of Community-Based Health Insurance (CBHI) Pilot Schemes in Ethiopia. 	<ul style="list-style-type: none"> Retrospective end of project impact evaluation. Commissioned and led by USAID and Health Sector Financing Reform project with Ethiopia's Health Insurance Agency. 	<ul style="list-style-type: none"> To evaluate the impact of the CBHI pilot schemes on equitable access to sustainable quality health care, increased financial protection, and effective social inclusion among the communities. 	<ul style="list-style-type: none"> Approach: Literature review component described theoretical/ conceptual framework for CBHI. Overarching approach not detailed. Mixed methods: Document review, key informant interviews and focus group discussions. Quasi-experimental methods (comparing before and after, and pilot participants vs control group), quantitative surveys.
<p>Global Annual Report</p> <ul style="list-style-type: none"> The Local Health System Sustainability Project (LHSS) under the USAID Integrated Health Systems IDIQ. October 2020. August 29, 2019 – September 30, 2020. 	<ul style="list-style-type: none"> Annual progress report developed by the USAID Local Health System Sustainability Project. Project is implemented through USAID cooperating agency, Abt Assoc. 	<ul style="list-style-type: none"> Project aims to support USAID's HSS priorities to expand access to UHC (increase financial protection and population and service coverage of essential services. Summary report with progress tracking and reflections on early lessons and successes. 	<ul style="list-style-type: none"> Approach: Theory-based results framework linking activities to objectives with indicators and learning plan/activities. Approach enables continuous learning cycle of identifying and solving systemic HSS issues. Mixed methods: quantitative and qualitative reporting on status of deliverables, qualitative reporting of lessons learned, project-wide 'pause and reflect' sessions and learning workshops.
World Bank			
<p>Country Program Evaluation</p> <ul style="list-style-type: none"> Armenia: Improving Health Care Utilization, Evaluation of WB Support to Health System Modernization (2019) 	<ul style="list-style-type: none"> Retrospective Project Performance Assessment Report (PPAR). Commissioned and undertaken by WBG's Independent Evaluation Group (IEG). 	<ul style="list-style-type: none"> To shed light on the effectiveness, consolidation, and sustainability of the whole adaptable program loan (APL) program investments for the improvement of health care services. 	<ul style="list-style-type: none"> Approach: Theory-based with TOC. No specific HSS framework. Results framework was designed and used to monitor progress throughout APL implementation. Mixed methods: Literature review, health sector portfolio analysis, key informant interviews. Secondary data analysis, site visits to secondary care hospitals.

<p>Global evaluation</p> <ul style="list-style-type: none"> Support to Health Services Achievements and Challenges (2018). 	<ul style="list-style-type: none"> Major (Sectoral) Evaluation: Retrospective thematic and sector evaluation of WBG support, commissioned and undertaken by IEG. 	<ul style="list-style-type: none"> To assess the roles and contribution of the WBG support to health services in client countries and assess whether WBG goals were achieved. To provide lessons and recommendations for achieving greater development effectiveness in future support to health services. 	<ul style="list-style-type: none"> Approach: Theory-based with overarching evaluation intervention logic and use of specific intervention-centric TOCs. Mixed methods: portfolio analysis of project features, outcomes, indicators and drivers of success or failure. Intervention and country case studies, social network analyses.
--	---	--	--

The following observations on GHI HSS evaluation objectives, approaches, and methodologies can be drawn from the literature review and key informant interviews.

The conceptualization of current HSS evaluations falls into two main categories: evaluation which reflects systems thinking, and evaluation focused on systems as they relate to specific disease or national immunization programs. WBG and USAID's evaluation thinking, learning and evaluation practices reflect an understanding of health systems as a comprehensive set of interconnected parts that must function together to be effective and that changes in one area will have repercussions elsewhere in the system. For Gavi and the GF, evaluation approaches largely reflect the primary focus of their support which is strengthening immunization and disease-specific delivery systems, with more limited evaluation of wider health system benefits and impacts.

Three broad categories of evaluation can be identified from the literature:

- **Evaluations of GHI HSS-specific policies and related programming.** For example, GF RSSH thematic review; Gavi country and global evaluations and meta-analysis of HSS support; annual progress reporting for USAID's different HSS programs such as the Local Health System Sustainability Program (LHSS); WBG's evaluation of WB support to Health System Modernization in Armenia.
- **Evaluations of HSS-specific interventions.** For example, the evaluation of USAID's Medicines, Technologies and Pharmaceutical Services (MTAPS) program, which aims to strengthen a country's pharmaceutical systems and supply chain management.
- **Evaluations of HSS interventions as part of an evaluation of wider GHI portfolio support or strategy.** For example, GF strategic reviews and Prospective Country Evaluations (PCE) and Gavi Full Country Evaluations (FCE), which include some assessment of RSSH/HSS interventions as part of a wider portfolio assessment of GF and Gavi support to countries; and WBG health sector evaluations which assess some aspects of HSS but the primary focus is on evaluating WBG contribution to access, coverage and quality of services. A specific HSS evaluation framework is less likely to be used in all these cases.

Evaluations vary in their objectives, instruments and independence. For example, USAID and WBG conduct self-evaluations of HSS or related programs which are undertaken by operational and management teams, closely linked to decision-making and organization learning objectives. Independent HSS evaluation is conducted by WBG, Gavi and GF with an emphasis on accountability and learning at the strategy/board level.

Most types of evaluations employ mixed methods to address the multiple objectives and evaluation questions of HSS evaluations. For example, WBG has a health portfolio of impact evaluations using Randomised Control Trials (RCTs), which are often complemented by process evaluation methodologies to add depth to 'how' and 'why'. RCTs are less evident in other GHI evaluation approaches, partly reflecting the difficulty of establishing, specifying, or measuring a control group to compare investments/interventions. In some GF and Gavi process evaluations, attempts have been made to model impact using experimental quantitative methods, with mixed results. For example, GF PCE Health Systems Modelling proved less methodologically robust than anticipated and was advised not to be used as a stand-alone method for explaining the linkages between inputs and outcomes. The Gavi Swiss Tropical and Public Health Institute (TPH) evaluation experienced challenges with the variables used to model Gavi's impact across a portfolio of over 70 countries with very different socio-economic and demographic characteristics.

Evaluation approaches have evolved to assess the complexity of programs and related evaluation questions through a shift towards prospective evaluation. Evaluation approaches have evolved in recognition of the need to understand not only 'what happened' because of an intervention but, crucially, why change occurred.

The Gavi FCE and GF PCE used process evaluation and implementation research to understand the complex systems in which Gavi and GF investments operated, including, but not exclusively, HSS/RSSH. These evaluations adopted a combination of methods, data and analytical approaches to evaluate health systems and generate recommendations throughout the grant cycle, not just at the end of the evaluation. For example, the PCE adapted the GF Technical Review Panel (TRP) '4S' framework to assess the nature of RSSH investments¹² and used deep dives and root cause analysis to identify and resolve health system-related gaps and bottlenecks.

USAID HSS programs have also moved from retrospective reporting to more prospective monitoring and learning throughout program implementation. While Health Systems 2020, Health Financing and Governance, and USAID Applying Science to Strengthen and Improve Systems (ASSIST) programs published final reports which identified lessons learned for future activities, LHSS has been publishing annual reports which identify lessons learned from ongoing implementation and course correction.

Theory-based HSS evaluations are common, with varying degrees of use of Theories of Change. For example, WBG's impact evaluations conducted by the IEG enable testing hypotheses set out in TOCs through experimental methodologies that aim to assess impact using a counterfactual. Evaluation efforts of USAID HSS evaluations vary by program and are generally summarized in program and annual reports rather than published as formal evaluations. TOCs are not generally presented in these reports (though more recent program reports present a 'results framework' that might fulfil similar objectives). For Gavi, the use of TOCs is an integral part of the monitoring, evaluation and learning culture and are applied regularly to evaluations. In the case of the GF, there is no TOC for how investments in RSSH and community systems strengthening are expected to strengthen the health system.

Even when used, few evaluations explain the theory in detail or go beyond the TOC graphic to outline the assumptions of change to be evaluated. Furthermore, there is little discussion around how the TOC is used during the evaluation process and how the evaluation findings validate or modify the assumptions of change which are important for informing program course correction.

There is limited discussion of the relevance and suitability of evaluation methods and approaches used in the literature reviewed. There is evidence that WBG IEG evaluations rank the relevance of an evaluation design and methods. Still, across GHIs, there was limited discussion or reflection on the suitability of evaluation methods and data sources, making it difficult to draw conclusions about the appropriateness of methods used to answer evaluation questions and the value of evaluations to GHIs/countries. Some commonly cited design issues for all evaluation categories related to limited available data and/or too few relevant indicators to adequately evaluate progress. For example, the GF PCE struggled to access recent programmatic data, which significantly limited the prospective nature of the evaluation and the ability to feedback 'real-time' programmatic progress.

There has been a notable shift in the culture of HSS evaluations with an increased emphasis on learning. For example, while there was little mention of the use case in early Gavi-commissioned evaluations of HSS (such as the Gavi Country evaluations of HSS), Gavi has come to recognize the importance of identifying the purpose of the evaluation and the audience that is expected to take action on findings. Key informants highlighted that these considerations are critical in determining whether an evaluation is needed and the nature/approach of the evaluation. Evidence from USAID also supports the emphasis on learning, with their new HSS Vision being complemented by an HSS Learning Agenda, which is guided by six HSS learning questions on systems.

¹² The 4S framework devised by the Global Fund Technical Review Panel refers to a framework that assessed RSSH investments across the 'health system development continuum', to determine whether RSSH investments were principally targeting 'system support' as opposed to 'system strengthening.' A modified version of the 4S framework was used by the PCE to assess the nature of RSSH investments in GF grants in PCE countries.

3.3 Challenges of HSS evaluations

The review identified common challenges which can impact the ability to measure intended and unintended effects of HSS investments and make evaluation design, implementation and relevance more problematic:

There is a lack of common definitions, approaches and guidance across and within some GHIs on what HSS is, what it is trying to achieve, and how to do it.¹³ This makes evaluation design challenging in terms of knowing what is important to measure and track, the choice of methods to do so, and the ability to develop comparable approaches to evaluating HSS interventions and outcomes.

Unclear scope of HSS grants/interventions or unclear TOCs limits understanding of what HSS interventions are trying to achieve and the causal links and pathways to achieve intended outcomes/impact. Evaluations have drawn attention to the lack of explanatory TOCs, and/or overly simplified and linear frameworks.

These challenges were less likely to be reported in evaluations of specific HSS interventions, which had more clearly defined objectives and evaluation questions. For example, evaluations of Gavi's performance-based financing (PBF) scheme or the USAID's Community-Based Health Insurance (CBHI) Pilot Schemes in Ethiopia (see Table 2 below) were more easily able to link the intervention to its intended outcomes and impact through the evaluation questions compared to evaluations of broader HSS efforts or strategies. The GF's Thematic Review of RSSH asked broader evaluation questions relevant to understanding the impact of the intervention: "What have RSSH investments delivered in country, and how can investments be better focused for improvements in program quality/ efficiency?"¹⁴ The evaluation identified that it was not clear to many stakeholders how RSSH investments are intended to improve health outcomes and strengthen health systems. This is partly due to the absence of an articulated theory of change but also to difficulty differentiating RSSH investments from disease grants and variability in how they are recorded across countries, impeding measurement of results¹⁴.

Table 2: Intended effects of evaluated interventions on HSS

Intervention	Intended to strengthen the HS by:
USAID CBHI Pilot Scheme	Improving financial access, quality, resource mobilization and strengthening community participation
Gavi PBF	Improving coverage and equity
GFATM RSSH	Unspecified

Inadequate time and resources to observe behavioural change within health systems with program/grant cycle logic and timelines widely perceived to be too short (usually three-year cycles but often assessed annually), affecting the feasibility of evaluations to demonstrate change/impact. This has been observed as a challenge in WBG, Gavi and GF evaluations. Additionally, limited resources available for HSS and, in some cases (Gavi and GF), the relatively small size of HSS investments also makes it difficult to demonstrate change/impact.

Methodological challenges concerning the use of indicators and availability of data and difficulties in establishing counterfactual HSS metrics/performance indicators can be inappropriate or incomplete, and this can complicate evaluation. Further, lack of baseline and/or ongoing programmatic and financial data presents challenges when evaluating the technical soundness, strategic focus, and value for money of HSS interventions, as well as impacting ongoing monitoring of HSS efforts and course correction. As well, evaluating the contribution of GHI health system investments and interventions to outcomes is problematic due to difficulties in establishing a

¹³ See also HSSEC brief "Boundaries and definitions of Health System Strengthening: implications for evaluation" March 2022.

¹⁴ TERG Position Paper, RSSH, July 2019

https://www.theglobalfund.org/media/8793/terg_resilientsustainablehealthreview_paper_en.pdf?u=637319004264830000

counterfactual or a control/intervention design. In the example of the Gavi FCE and GF PCE, contribution was assessed largely through qualitative and descriptive methods rather than experimental methods. Some examples from the evaluations are provided below.

A WBG evaluation of Support to Health Services project noted that projects aiming to improve health outcomes faced limited achievement against objectives, "partially due to the use of indicators such as mortality rate, that may not be sufficiently sensitive over the project life span" (World Bank Group, 2018).

The evaluation of Gavi's PBF component of its HSS support stated that "increases in coverage are difficult to achieve within one year and, even more importantly, current measurement systems are not precise enough to measure such changes" (Gavi, 2018).

GF strategic key performance indicators monitor progress on four of seven RSSH sub-objectives, accounting for 35% of total RSSH investments. The remaining sub-objectives are intended to be measured through evaluation studies; however, it is not clear if these areas are being tracked at country level at all (OIG report, 2019).

Lack of overarching strategy for undertaking HSS evaluations and/or evaluations not planned strategically or with a consistent approach. The evidence suggests it was not always clear how HSS evaluations fitted within broader organizational strategic frameworks and plans, raising questions regarding purpose and use, target audience, the rationale for multiple evaluations (either of the same grant/intervention over time or of different grants/intervention) and timing. For example, Gavi FCE set out to evaluate the impact of HSS on vaccine coverage, but too few HSS grant activities were implemented in FCE countries to allow measurement, at least partly due to Gavi rules or processes¹⁵. Whilst this underscores the value of approaches that assess implementation, it also highlights challenges with timing and ensuring the 'right conditions' are in place to ensure evaluation objectives can be met.

3.4 Use and uptake of HSS evaluation findings

A number of interrelated approaches characterize the use of evaluation findings:

- **'Within-program' learning was emphasized** as a key objective of monitoring and evaluation work in prospective evaluations (Gavi FCEs, GF's PCEs, USAID's annual reports of LHSS) and WBG project level self-evaluations.
- **Summarising program achievements and results and identifying lessons learned** were a key focus of retrospective evaluations and USAID HSS end-of-program reports. Gavi country evaluations of HSS also aimed to assess results and lessons learned but had limited and variable information about case use. For example, the Gavi-commissioned evaluation in Somalia (2016) set out that the evaluation would provide input into the next phase of the Gavi HSS application process. However, the evaluation report for Burkina Faso does not set out in its mandate how evaluation findings are expected to be used.
- **Informing policies and strategic directions** is the intention of WBG IEG thematic evaluations. Examples include the review of WBG's support to Health Services and the retrospective GF Thematic Review of RSSH, which aimed to generate evidence on the implementation, successes and challenges related to RSSH investments and to inform future Strategy Committee decisions.

There is limited insight from the literature on how evaluation learnings and findings were used and their value and influence at global and country level. For example, there was limited evidence

¹⁵ Soi et al (2020) How to evaluate the implementation of complex health programmes in low-income settings: the approach of the Gavi Full Country Evaluations in Health Policy and Planning, 35, 2020 ii35-ii46.

in PCE synthesis reports for how PCE findings were informing programmatic changes at country level and enabling changes in business model policies and practices at Secretariat level. FCE real-time data collection allowed for early identification of issues at country level. However, it was not evident how many times action was taken early on as a result of FCE findings and what this achieved.

Gavi has carried out multiple process and strategy reforms. However, the challenges flagged by evaluations over time (e.g. difficulties with understanding the expectations of investments and use case of the evaluation; time and data constraints presenting a technical challenge for evaluation) have not all been taken up or addressed to date. There is evidence that evaluation findings which recommended Gavi increase its HSS focus to be more aligned to their strengths and major areas of work has led to a shift in the scope of HSS activities funded, i.e., supporting more specific HSS activities related to immunization. There is evidence of evaluation findings influencing the greater use of TOCs.

3.5 Strengths, weaknesses and lessons learned from evaluation approaches methods and utilization

Approach and methods

- Of key importance emerging from the review and key informant interviews is the necessity within GHIs to develop and agree on common HSS evaluation guidance, a set of appropriate HSS indicators within GHIs, and an approach to asking evaluation questions across a range of contexts, to enable the build-up of country comparisons over time.
- Findings from the WBG, Gavi and GF found that structured approaches to HSS evaluation, including using TOCs, conceptual frameworks, data gathering tools and standardization of evidence for synthesis, enabled comparability across countries. However, these approaches need to ensure that context-specific evaluation needs are addressed using local data, for example, by providing global generated templates, evaluation tools and guidance for evaluation plans and by encouraging local adaptation of evaluation templates and tools. In the case of the GF PCE, more deductive evaluation designs for country PCE teams to implement led to improved synthesis, with a clear trade-off between synthesis of generalizable findings across all eight countries and more inductive, flexible, and ‘country owned’ approaches.
- Gavi and GF experience highlighted the strengths of prospective evaluations for managing and assessing complex interventions and for using flexible, innovative approaches and multiple methods concurrently designed and implemented to assess HSS issues. For example, as the scope and evaluation questions became increasingly targeted, methods and tools became more analytical, evidenced through deep dives, which were considered valuable for exploring a problematic health system issue and drilling down on specific pathways of change.
- However, compared to ‘regular’ evaluations, the cost of the prospective approach was significantly higher, and the results were unclear. Lessons indicate that both the PCE and FCE were successful in building country evaluation capacity over time but striking a balance between global and country stakeholder interests, priorities, and evaluation questions was challenging and added complexity.
- Additionally, across FCEs and PCEs, balancing breadth versus depth of evaluation scope affected the degree to which HSS/RSSH was systematically examined. Evaluators highlighted this issue “due to the wide scope of the FCE, there is a limited ability to examine all issues in detail”, with FCEs taking a wider approach to considering HSS, among other outcomes. This contrasts with HSS-specific evaluations, which were able to

draw clearer links between the intervention and elements of HSS it intended to affect, enabling more direct identification of lessons learned for implementing HSS-specific evaluations.

- Lessons learned suggest that retrospective HSS evaluations should, where possible, adopt mixed methods and innovative analytical techniques which combine qualitative and quantitative data. Further, to understand HSS impact and return on investment, evaluations should consider analyzing the full results chain from national decision-making to community outcomes.
- A lesson learned for assessing HSS outcomes is that detailed, robust TOCs and mixed methods are required. Additionally, the use of TOCs can strengthen MEL organizational culture by considering and supporting the use case, including whether it is necessary to demonstrate contribution and/or impact, what is it that is being evaluated, who is the evaluation for, how do evaluations link to strategic KPIs and the commissioning of evaluations.
- Related to the above point, having an overarching project or MEL framework can help organize evaluation and learning work. WBG's experience of using such a framework for a large nutrition evaluation suggests this helps coordinate actors in countries to address what's missing, is a good way of gathering evidence and can help develop more systematic learning about whether the right things have been measured.

Utilization

- The review highlighted the importance and benefits of close collaboration between GHIs and country stakeholders in the design (inception phase) and implementation of evaluations to ensure buy-in and clarity on intentions, scope and feasibility, what HSS interventions are expected to achieve, and to generate interest in learning and use of findings.
- Findings from the review indicate that countries need greater support for understanding evaluations, for example, capacitating ministries of health to improve data sources and availability and strengthen analysis and use of data for decision making and accountability. The importance of developing standardized datasets and making microdata available widely was raised by stakeholders across the GHIs. A key lesson for future evaluations is to ensure adequate resources and time are available to support countries through evaluation processes, including the joint development of frameworks, training and adaptation of evaluation templates and tools for country use, and joint analysis, discussion and use of findings.
- Best practices emphasize tailoring dissemination strategies to stakeholders and providing knowledge translation support. Lessons from GF PCE suggest this was limited by resources and capacities, and as a result, dissemination processes were largely focused on the annual report and dissemination meeting, which did not always align with program implementation timelines or decision-making windows. A key lesson for future evaluations is to ensure resources are available for multiple modes of timely dissemination of findings.
- Discussions with KIIs highlighted the importance of having clear champions and mechanisms in place to ensure findings are applied in decision-making processes. For example, key informants highlighted the role of the GF Strategy Committee in ensuring that recommendations from the 2015 and 2020 Strategic Reviews were taken up to inform future strategies and plans. WBG IEG takes up findings of thematic evaluations using standard feedback mechanisms to inform future WBG policies and strategic

direction; for self-evaluations, there are fewer direct mechanisms for ensuring that lessons from these evaluations are applied more broadly.

4 Future considerations

This review identified published evaluations carried out by the GHIs and synthesized key findings, challenges and lessons learned. Future considerations based on the review of evaluations come from recommendations in the reviewed evaluation documentation and an analysis of gaps and good practices identified in the literature and by the review team.

Addressing challenges and gaps in HSS evaluation

- **GHI guidance for evaluators should be standardized and coordinated across countries** to promote improved quality and use of findings and should be consistently used by evaluators to support the comparability of findings across contexts.
- **Common definitions of HSS and an appropriate set of agreed indicators should be developed** to systematically track HSS interventions and enable a robust HSS evidence base to develop, which supports country comparisons over time.
- **Develop and/or strengthen theoretical frameworks and TOCs to clearly articulate expectations for HSS support** and elaborate on assumptions for how interventions will bring about change. Ensure GHI technical guidance on HSS is congruent with the TOC and other related organizational systems and monitoring, evaluation and learning thinking.
- **Continue to learn and build on the experience of assessing GHI contribution to outcomes** and impact through greater consideration and reflection on the suitability of methodologies to achieve this (e.g., HSS-specific prospective and retrospective evaluation approaches and through a combination of established and experimental mixed methods).
- **Ensure resources (technical, human, time) are available to support the close involvement of GHIs and country stakeholders** in the design and implementation of evaluations and learning efforts, and ensure adequate feedback mechanisms are in place to promote the uptake of evaluation findings.
- **Strengthen cooperation and coordination of HSS approaches, investments, evaluations, and technical support across GHIs**, including at country level. This should be done to improve complementarity and identify niche areas for how each GHI can support HSS and wider health goals.
- **Where appropriate, establish organization-wide MEL plans and structures** that support rationale, planning and organization of HSS evaluations, linked to wider strategy, and continue to provide and strengthen learning platforms for collecting and disseminating evidence and findings.
- **Continue to strengthen and utilize learning components of HSS evaluation** to advance understanding, practice and lessons learned on systems thinking, methods and approaches appropriate to evaluation purpose/objectives.

Annex 1

FA2 Repository of GHI documents and evaluations including those reviewed.

The Bill and Melinda Gates Foundation

- HSS Evaluation Lessons Learned Final.pptx
- Organization Deep Dives.pptx
- Summary of literature read on Health Systems Strengthening.docx
- Worksheet from Organization Deep Dives.xlsx

Gavi

Background (Reviewed)

- Board Committees and Advisory Committees
- evaluation-advisory-committee-terms-of-reference.pdf
- Gavi Alliance evaluation of Independent Review Committee-managerial response.pdf
- Gavi HSS 2000-2015
- HSIS - Data
- HSIS and Measurement
- HSIS Supply Chain
- HSIS support framework overview
- HSIS Support Framework.pdf
- HSS evaluations
- hss-policy-timelinepdf.pdf
- Review of Independent Review Committees

Full Country Evaluations (FCEs)

- FCEs Summary.pdf

Country-level

Bangladesh (Reviewed)

- 2015-fce-bangladesh-briefpdf.pdf
- 2015-full-country-evaluations-report---bangladeshpdf.pdf
- 2016-fce-bangladesh-briefpdf.pdf
- 2016-full-country-evaluations-report---bangladeshpdf.pdf

Mozambique

- 2015-fce-mozambique-briefpdf.pdf
- 2015-full-country-evaluations-report---mozambiquepdf.pdf
- 2016-fce-mozambique-briefpdf.pdf
- 2016-full-country-evaluations-report---mozambiquepdf.pdf

Global-level (Reviewed)

- 2015 FCE Cross Country Brief.pdf
- 2015 FCE Cross Country Report.pdf
- 2016 FCE Cross country Brief.pdf
- 2016 FCE Cross country Report.pdf
- FCEs Management Response 2016.pdf

Peer-reviewed Papers

- Carnahan et al. (2020) - Lessons Learned From Implementing Prospective, Multicounty Mixed-Methods Evaluations for Gavi and the Global Fund.pdf
- Soi et al. (2020) - How to evaluate the implementation of complex health programs in low-income settings - the approach of the Gavi Full Country Evaluations.pdf

Global HSS Reviews (Reviewed)

- 01-Swiss TPH Review of HSS Support.pdf
- 02-CEPA meta-analysis 2016.pdf
- 03-Meta-review 2018 update.pdf
- 04-Five Priorities for the Next Five Years.pdf
- HSS Review Management Response 2019.pdf

HSS Country Evaluations (Reviewed)

- Gavi (Country v Gavi commissioned comparison) .xlsx
- Gavi Commissioners & Evaluators 2.xlsx

Selected for review

- 05-Burkina Faso Gavi Eval.pdf
- 05a-Summary Burkina Faso Gavi Eval.pdf
- 06-Country Eval Afghanistan.pdf
- 06a-Summary Country Eval Afghanistan.pdf
- 09-Nepal Gavi Eval.pdf.pdf
- 09a-Nepal Gavi Eval.pdf.pdf
- 10-Ethiopia Gavi Eval.pdf.pdf
- 10a-Ethiopia Gavi Eval.pdf.pdf
- 11-Yemen Gavi Eval.pdf.pdf
- 11a-Yemen Gavi Eval.pdf.pdf
- 12-Ghana Gavi Eval.pdf.pdf
- 12a-Ghana Gavi Eval.pdf.pdf

Additional

- 08-Eval HSS Chad and Cameroon.pdf

- 08a-Advantages and disadvantages Chad and Cameroon.pdf
- 08b-Resource allocation Chad and Cameroon.pdf
- 13- Madagascar Gavi Eval.pdf
- 13a- Madagascar Gavi HSS Eval.pdf
- 14 - Chad Gavi Eval.pdf
- 14a- Chad Gavi Eval.pdf

Global Fund

Background (Reviewed)

- GF Guidance Note - Epidemiological and Impact Analysis.pdf
- GF Modular Framework Handbook (2019).pdf
- GF RSSH Focus.pdf
- GF Strategic Framework for Data Use for Action and Improvement at Country Level, 2017-2022.pdf
- Monitoring & Evaluation Framework - Monitoring & Evaluation - The Global Fund to Fight AIDS, Tuberculosis and Malaria
- RSSH 2S Budget Analysis (extracted)
- RSSH Indicator Guidance Sheet.xlsx
- TERG (2020) - Thematic Review on Sustainability, Transition and Co-financing (STC) Policy.pdf

Global HSS Reviews (Reviewed)

- GF Board Meeting (2019) - TERG Report.pdf
- GF Information Note on Building Resilient and Sustainable Systems for Health (RSSH).pdf
- OIG Report (2019) - Managing investments in Resilient and Sustainable Systems for Health.pdf.pdf
- Review of Global Fund Investments in Resilient and Sustainable Systems for Health in Challenging Operating Environments (2016).pdf
- TERG Thematic Review of RSSH (2019).pdf
- terg_strategicreview2020position_paper_en.pdf
- TRP report (2018).pdf

Prospective Country Evaluations (PCEs)

Country-level

- About the PCE | IHME
- PCE Evaluation Activities | IHME
- PCE Evaluation Design | IHME
- PCEs - Euro Health Group A/S

PATH/IHME Consortium

- Guide to Reports.docx

DRC (PATH/IHME)

- 2018 DRC_PCE_2018_Annual_country_Report_EN.PDF
- 2019 Prospective_Country_Evaluation_DRC_annual_report_2019.pdf
- 2020 DRC_PCE_Final_annual_report_2020_English.pdf
- 2021 DRC_PCE_Annual_Country_Report_final_2021_English.pdf
- 2021 Extension_Report_DRC_Final_30_June_2021_English.pdf

Guatemala (PATH/IHME/CIESAR)

- 2018 guatemala_PCE_2018_Annual_Country_report_final_English.pdf
- 2019 Guatemala_PCE_annual_country_report_2019_English.pdf
- 2020 Guatemala_PCE_2020_Annual_Report_ENGLISH_Final.pdf
- 2021 Guatemala_2020-21_Annual_Country_Report_final_31_March_2021.pdf

Senegal (PATH/IHME/ISED/UCAD)

- 2020 Senegal_PCE_annual_report_2019-20_English_Final.pdf
- 2021 Extension_Report_SEN_Final_Draft_Oct2021_-_ENGLISH.pdf
- 2021 Senegal_PCE_2020-21_Annual_Country_Report_final_31_March_2021_Eng.pdf

Uganda (PATH/IHME/IDRC)

- 2018 PCE_synthesis_infographic_Uganda_final_CcJDDvy.pdf
- 2018 Uganda_PCE_2018_Annual_Country_Report_final.pdf
- 2019 Uganda_2019_Annual_Country_Report_April_2019_Final_pdf.pdf
- 2020 Uganda_PCE_2019-2020_Annual_Report.pdf
- 2021 Extension_Report_UGA_Final_August_2021.pdf
- 2021 Uganda_PCE_Annual_Country_Report_Final_2021.pdf

UCSF/EHG/Itad Consortium

Global-level (Reviewed)

- PCE 2018 terg_2018-pce-synthesis_report_en.pdf
- PCE 2019 terg_2019-pce-synthesis_report_en.pdf
- PCE 2020 terg_2020-pce-synthesis_report_en.pdf
- PCE 2021 terg_2021-pce-synthesis_report_en.pdf
- PCE MGMT RESPONSE 2021 terg_secretariat-management-response-terg-pce-synthesis-2021_report_en.pdf

Peer-reviewed Papers

- Carnahan et al. (2020) - Lessons Learned From Implementing Prospective, Multicountry Mixed-Methods Evaluations for Gavi and the Global Fund.pdf
- Phillips et al. (2020) - Bringing a health systems modelling approach to complex evaluations - multicountry applications in HIV, TB and malaria.pdf
- terg_position-paper-pce-synthesis-report-2021_paper_en.pdf

- [terg_secretariat-management-response-terg-pce-synthesis-2021_report_en.pdf](#)

Health Systems Modeling documents (Reviewed)

- [Chang et al. \(2017\) - Dynamic modeling approaches to characterize the functioning of health systems - A systematic review of the literature.pdf](#)
- [DRC_PCE_Final_annual_report_2020_English.pdf](#)
- [Phillips et al. \(2020\) - Bringing a health systems modelling approach to complex evaluations - multicountry applications in HIV, TB and malaria.pdf](#)
- [Verguet et al. \(2019\) - Health system modelling research - towards a whole-health-system perspective for identifying good value for money investments in health system strengthening.pdf](#)

Global Financing Facility

- [GFF-Annual-Report-2019-2020.pdf](#)
- [GFF-Results-Monitoring-Framework.pdf](#)
- [ME_frameworks_for_GFF_countries.xlsx](#)

Other

- [Cordaid-3G Analysis - HSS_March-2021.pdf](#)
- [Resources shared by WG1 \(included in repository\).docx](#)

HS Frameworks

- [Bennett & Peters \(2014\) - Assessing National Health Systems - Why and How.pdf](#)
- [HSS_MandE_framework_Nov_2009.pdf](#)
- [Sacks et al. \(2019\) - Beyond the building blocks - integrating community roles into health systems frameworks to achieve health for all.pdf](#)
- [van Olmen \(2012\) - Health systems frameworks in their political context - framing divergent agendas.pdf](#)
- [WHO FW for HS Performance Assessment.pdf](#)

Systems thinking

- [Adam \(2014\) - Advancing the application of systems thinking in health.pdf](#)
- [Cassidy et al \(2019\) - Mathematical modelling for health systems research - a systematic review of system dynamics and agent-based models.pdf](#)
- [Editorial \(AJPH, 2006\) - Systems Thinking and Modeling for Public Health Practice.pdf](#)
- [Peters \(2014\) - The application of systems thinking in health - why use systems thinking.pdf](#)
- [WHO \(2009\) - Systems thinking for health systems strengthening.pdf](#)

UNICEF

- [The UNICEF Health Systems Strengthening Approach.pdf](#)

USAID

- [1-MEL Guide.pdf](#)
- [2-compendium of indicators.pdf](#)

- 3-LHSS Y1 Report.pdf
- 3a-LHSS Y1 Report annexes.pdf
- 4-GSAM Final Evaluation.pdf
- 5-MTAPS 2020 Annual Report.pdf
- 5a-MTAPS 2019 Q3 report.pdf
- 6-PMQ+ y1 Annual report.pdf
- 7- DigitalSquare Strategy June 2020.pdf
- 8-HRH2030 Philippines Final Report.pdf
- 9-USAID GH Strategic Framework.pdf
- 10-Marshalling the evidence.pdf
- 11-HGF Annual Report Y6.pdf
- 12-ASSIST Final Report.pdf
- 13-HS2020 Final Report.pdf
- 14-Framework and Guideline for the Assessment and Evaluation of Health Systems Strengthening Program.pdf
- 15-Partners for Health.pdf

WBG

- What Difference Does Good Monitoring & Evaluation Make to World Bank Project Performance?.pdf

Background (Reviewed)

- IEG Methodology | Independent Evaluation Group
- IEG Org Chart.pdf
- Task Team Leader Toolkit.pdf

HRITF / RBF Health

- Summaries of RBF Health IEs.csv

Impact Evaluations (Reviewed)

- Afghanistan-cluster-randomized-trial.pdf
- Haiti-Zeng et al 2013.pdf
- India-Quasi-Experimental-Evaluation-of-Insurance-and-Health-Outcomes.pdf
- Zambia - Impact Evaluation of Zambia's Health Results-Based Financing Pilot Project.pdf
- Zimbabwe-Impact-Evaluation-Overview-Report.pdf

Other Impact Evaluations (Not reviewed)

- 101505-BRI-Box393260B-PUBLIC-PUBDATE-12-01-15-DOI-10-1596-K8566.pdf
- Burkina-Faso-Impact-Evaluation-Results-Report.pdf
- Burkina-Faso-Qualitative-Research-Impact-evaluation-report.pdf
- Financial risk protection from social health insurance-INDIA.pdf
- Haiti-Eilcher et al.pdf
- Impact Evaluation on PBF in Haut Katanga District - DRC.pdf
- Nigeria-NSHIP-IE-Report.pdf

- PBF for Health in Haut Katanga_IE Executive Summary_0.pdf
- Rewarding provider performance to enable a healthy start to life - evide..._0.pdf
- w19046.pdf
- WPS8059.pdf
- WPS8060.pdf
- WPS8162.pdf
- Zimbabwe-Impact-Evaluation-Overview-Report.pdf

Toolkits

- Cost Effectiveness Analysis of RBF Programs - A Toolkit.pdf
- HRITF Qualitative Research Synthesis Report.pdf
- HRITF Qualitative Research Tool.pdf
- IE_Toolkit_2012.08.21_ENG.pdf
- Impact Evaluation in Practice.pdf
- PBF Toolkit.pdf
- Toolkit Descriptions.docx

IEG - Independent Evaluation Group

- Behind the Mirror - A Report on the Self-Evaluation Systems of the World Bank Group.pdf
- Evaluation Capacity Development | Independent Evaluation Group

Guidance & Toolkits (Reviewed)

- Designing-a-Results-Framework-for-Achieving-Results-A-How-to-Guide.pdf
- Impact Evaluations & Development.pdf
- Managing Evaluations - a how to guide for managers and commissioners of evaluation.pdf
- MethodsSourceBook.pdf
- World Bank Evaluation Principles

Major Evaluations (Reviewed)

- World-Bank-Group-Support-to-Health-Services-Achievements-and-Challenges-An-Independent-Evaluation Summary.pdf
- World-Bank-Group-Support-to-Health-Services-Achievements-and-Challenges-An-Independent-Evaluation.pdf

Other publications

- Changing_Landscape_of_Dev_Eval_Training.pdf
- Evaluation2030-final2.pdf
- How the Bank learns 0- learning_results_eval.pdf
- M&E System - case of Chile 1990-2014.pdf

PPARs (Reviewed)

- ppar_39118 Ghana.pdf
- ppar_armeniaimprovehealthcare.pdf
- ppar_boliviahealthsector.pdf
- ppar_niger_0716.pdf

- ppar-philippineshealth-10062017_0.pdf
- wbp_gavi_alliance2.pdf

SIEF

- SIEF Evaluations Summaries - Health

Health Evaluations

- Impact of Social Accountability Interventions on Healthcare Delivery and Health Outcomes in Uttar Pradesh, India.pdf
- Improving Supply Chain for Essential Drugs in Low Income Countries Results from a Large Scale Randomized Experiment in Zambia.pdf
- Strengthening malaria service delivery through supportive supervision and community mobilization in an endemic Indian setting - an evaluation of nested delivery models.pdf

Annex 2: FA2 Review of HSS evaluations: Review of Findings (WIP) October 21 2021

<https://www.itad.com/wp-content/uploads/2022/07/2020-078-HSSEC-WG1-FA2-Annex-1.pptx>