A Health Data Science Exchange

Accelerating the appropriate use of data science to close the health equity gap

digital square

Authors

Amanda BenDor
Ashley Bennett
Emily Carnahan
Skye Gilbert
Hallie Goertz
Anna Volbrecht

Advisory Group

The authors of this publication greatly appreciate the contributions, perspectives, and expertise from the Health Data Science Exchange Advisory Group. The Advisory Group was convened in late 2019 to provide guidance and input in the formative stages of this publication. Members of the Advisory Group include:

Fazilah Shaik Allaudin
Padmanabhan Anandan
Caitlin Augustin
Jeff Bernson
Manisha Bhinge
Nejma Cheikh
Marelize Gorgens, *Lead, Digital Health Innovation and Data Science, World Bank Group*
Fred Hersch
Hitesh Hurkchand
Ulla Jasper, *Policy Officer, Fondation Botnar*
Francois Jung
Karin Kallander
Dan Kim
Orin Levine
Tumainiel Macha
Amani M'Bale
Liz McGovern
Michele Monroe
Steven Parkinson, *Senior Advisor, The Global Fund to Fight AIDS, Tuberculosis, and Malaria*
Dan Rosen
Kelly Saldana
English Sall
Virginia Simushi, *Principle ICT Officer, Ministry of Health, Zambia*
Jeffrey Michael Smith
Nana Twum-Danso
Thomas Wiegand
Kate Wilson, *CEO, DIAL*

Suggested citation

Table of Contents

Value Proposition

Background

Countries continually strive to improve the effectiveness, efficiency, and impact of health systems and the health services they provide. Through health system strengthening efforts, countries and their supporting partners are identifying new tools and approaches to accelerate progress. With global targets such as the Sustainable Development Goals, countries are placing a great emphasis on tools and approaches that can contribute to the attainment of health equity and universal health coverage.

Digitalization of health systems has been an important component of health system strengthening in many countries. Digital systems are providing more granular, higher-quality data to support government decision-makers. The data produced by these systems can provide insight into complex health service delivery questions, creating continual opportunities for improvement and progress. As digital technologies become more sophisticated and more deeply integrated into health systems, governments require data science tools and technical capacity to make full use of the data available to them.

Stakeholders, especially those working in low-resource contexts, need to know where to find scalable data science assets with proven ability to improve health system performance, and ultimately health outcomes.

Data science assets are the tools, resources, and approaches that support the intentional and integrated use of data to produce actionable insights that strengthen health systems. There is increasing demand from health system managers, policymakers, and investors to use data science tools and approaches alongside digital technologies for health system strengthening. Ensuring these tools and approaches are accessible and curated for these stakeholders is an important, but absent, resource in the global health sector. A Health Data Science Exchange (HDSE) could provide a solution.

Current status of data science in global health

Health system managers, policymakers, investors, implementing partners, and other stakeholders are investing time and resources to strengthen scalable, interoperable country health information technologies and systems. This investment has led to increased quantity, frequency, and granularity of data.

Despite investments in health information systems and the rapid increase in available data, data science assets continue to be unknown or sparingly used in many low-resource contexts. There are significant missed opportunities to generate insights from data, due in part to:

» **Limited evidence and awareness of which data science investments lead to impact,** in part due to lack of investment in applied research and research methods. There is no established framework for evaluating the impact of data science, contributing to the lack of evidence. For example, there is a limited evidence base to understand the impact of Artificial Intelligence (AI) tools, particularly in low-resource contexts.[1] This results in limited, misaligned, and fragmentated data science utilization in global health, which prevents high-impact innovations from securing the investment needed for replication and scale.

» **Lack of globally accepted criteria for evaluating data science assets,** which creates challenges in comparing assets and limits consensus on how data science can be best used for country contexts.For example, a recent report on AI in global health noted "there is no current definition of acceptable performance standards, accuracy rates, and patient health outcomes against which to measure AI. Further, perspectives vary widely on what standards AI tolls should adhere to in order to be used across LMICs, and for patients and doctors to have trust in them."[2]

1 USAID, The Rockefeller Foundation. Artificial Intelligence in Global Health: Defining a Collective Path Forward. Washington, DC: USAID; 2019. Available from: https://www.usaid.gov/cii/ai-in-global-health

2 USAID, The Rockefeller Foundation. Artificial Intelligence in Global Health: Defining a Collective Path Forward. Washington, DC: USAID; 2019. Available from: https://www.usaid.gov/cii/ai-in-global-health

- **Lack of visibility on available data science assets,** making it challenging for country health leaders to understand the full ecosystem of assets available to them. Lack of visibility also leads to duplication of effort, rather than further utilization and development of existing tools and approaches.
- **Overemphasis on unproven, frontier data science tools or approaches, such as machine learning and artificial intelligence,** leading to uncertainty and misalignment about the utility of these frontier assets. Additionally, this may de-emphasize the importance of existing, proven interventions and foundational prerequisites (i.e., high-quality data) and create confusion when assessing available evidence. Frontier data science tools are essential for the transformation of health systems—but they are applied most effectively in a context of strong existing data systems.
- **Inequitable distribution of data science expertise and capacity worldwide,** which further disadvantages health systems in low-resource settings that may lack the knowledge, skills, and information to design and implement high-quality, sustainable data science interventions.
- **Lack of operational guidance or prerequisites for harnessing data science within health systems,** creating a high technical barrier of entry even when data science assets are available.
- **Lack of clear problem identification or articulation that is understandable to both health and data science professionals,** which limits alignment around what health system challenges data science can help solve.
- **Risk of continued fragmentation as data science is applied within health verticals,** which may increase burden on health system managers with duplicative, fragmented, disconnected data dashboards and tools. A systems approach would better serve health system managers and frontline health workers long-term.

Currently, there is no place for stakeholders to convene to learn about existing data science assets, how data science assets can be harnessed most efficiently to support country health systems, and who supports innovation in this space. A stronger link needs to be made between data science assets and the needs of decision-makers working within low-resource health systems. An HDSE could lead the way in bringing more visibility to how data science, through all phases of a data life cycle, can support countries to reach health system goals.

Data science to improve health outcomes

The intentional and integrated use of data science can produce actionable insights for health system stakeholders by strengthening the use of data at all levels. The application of data science tools and approaches can support improvements at any point, from data capture to implementation of data-led activities (Figure 1).

FIGURE 1. High-level overview of how data science can contribute to overall health system strengthening.

Health systems continually improve core functions through health system strengthening activities.

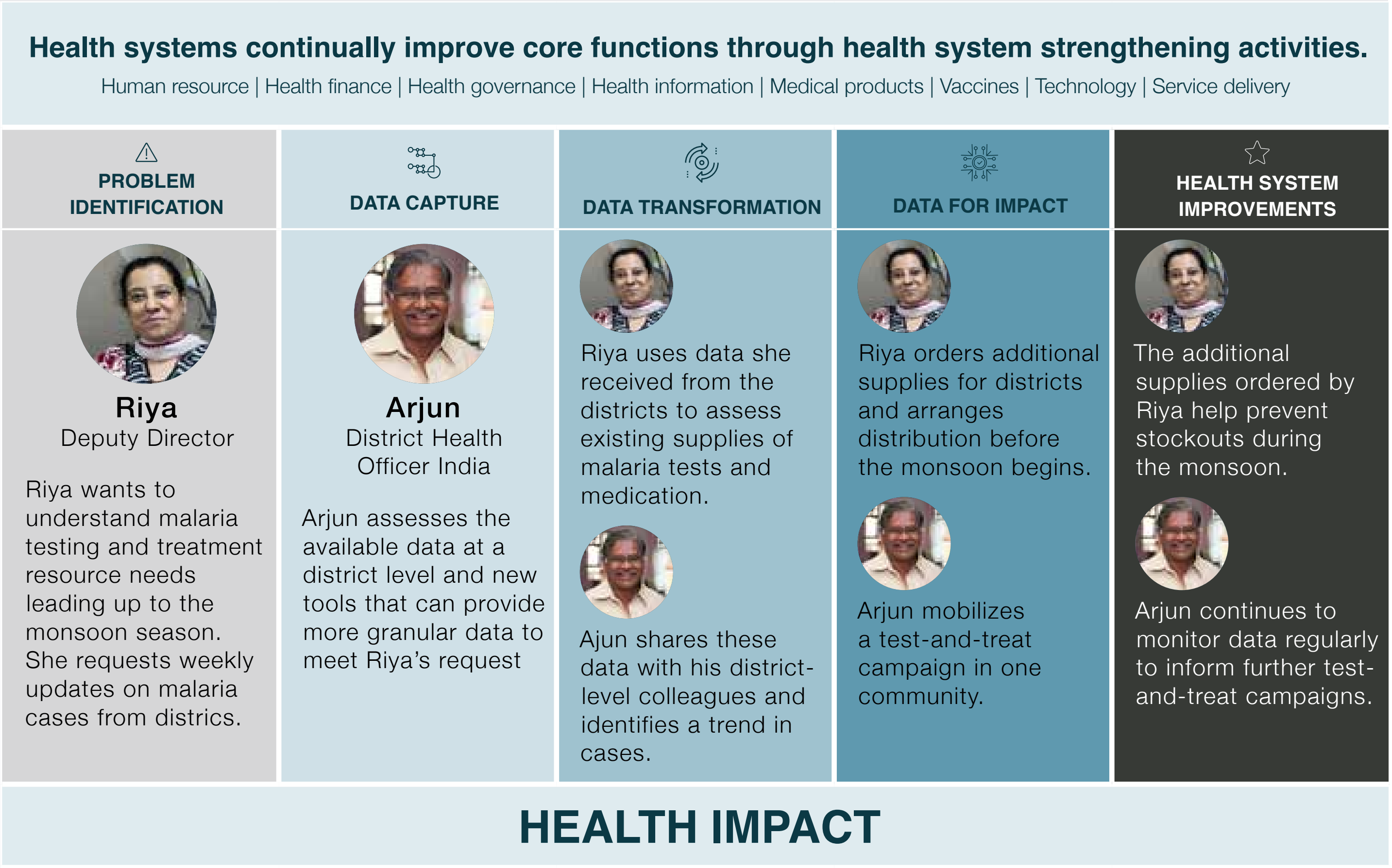
Human resource | Health finance | Health governance | Health information | Medical products | Vaccines | Technology | Service delivery

PROBLEM IDENTIFICATION	DATA CAPTURE	DATA TRANSFORMATION	DATA FOR IMPACT	HEALTH SYSTEM IMPROVEMENTS
Health system managers or other stakeholders identify "data science-able" questions that can be informed using data.	Data are captured through digital health tools by frontline health workers or health system managers. Data are also integrated from non-health-specific data sources.	Data are processed, analyzed, and visualized to create relevant, actionable insights for decision- and policymakers.	Decision-makers take data-led action to increase effectiveness, efficiency, or reach of health system performance.	Improvements in health systems lead to health care that is accountable, affordable, accessible, and reliable.

HEALTH IMPACT

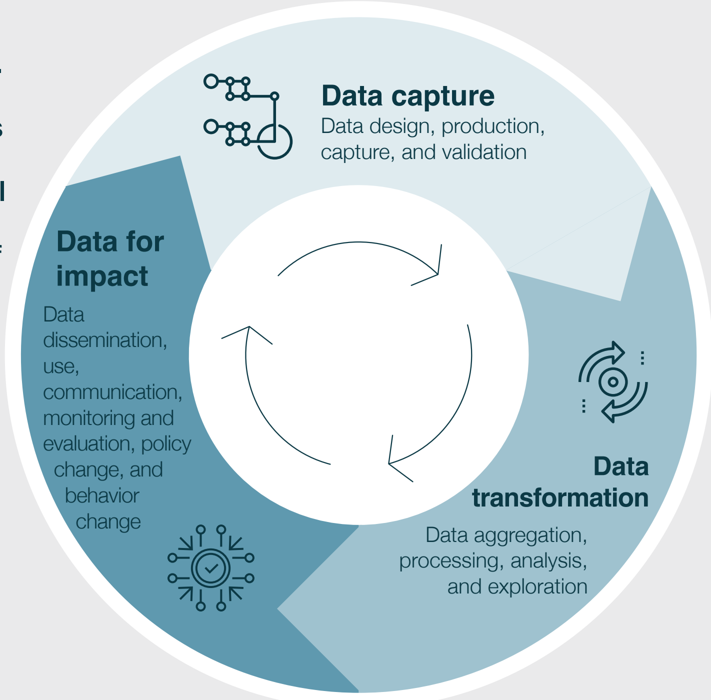
Decision-makers and policymakers working at all levels of the health system interact with data throughout the data lifecycle (Figure 2). By ensuing data science assets are supporting the entire life cycle, data are improved from collection to insight.

FIGURE 2. An example of data use in the health system context.



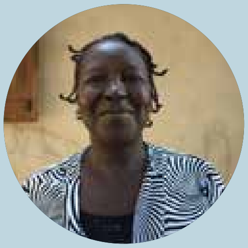
By ensuring data science assets are supportive of the entire life cycle, data are improved from collection to insight. This is not a linear process but rather a cycle of data use that provides continual opportunities to integrate new data sources, new analysis practices, and ongoing decision and action to strengthen health systems (Figure 3).

FIGURE 3.
Data use as a cycle. Rather than a linear process, data use is often cyclical, with a focus on continual assessment and improvement of both health system functions and use of data itself.



Stakeholders within the data science cycle

While government stakeholders, like Riya and Arjun, are the primary audience for many data science interventions, others in the global health ecosystem would benefit from an HDSE, as they engage with health systems and health system strengthening activities.



Researchers
Winnie leads a group of health researchers at Makerere University. Her team wants to collect data on mental health services that are available to HIV patients in Uganda. Some raw data are available. Winnie needs to identify tools to collect additional data and validate the quality of the data her team plans to use.

Civil society and nongovernmental organizations (NGOs)
Jonathan, a digital health lead for an international NGO based in Rwanda, has been asked by the Ministry of Health to support the establishment of a data security system with protocols at the regional data warehouse in Butare. Jonathan wants to identify existing standards, policies, and technologies that can support data security.

Private sector companies
Patricia works for a technology company that supports the Ministry of Health in Vietnam to use artificial intelligence and modeling to predict dengue outbreaks. She has evidence that outbreaks are likely in two regions in the next six months. After promising initial results, Patricia wants to share her work, solicit input from the global community, and find opportunities to apply these models in other settings.

Investors
Allison is a next-generation philanthropist driven to invest in helping countries use artificial intelligence and machine learning to improve health outcomes—specifically, through community prevention of communicable diseases. Allison does not want to duplicate existing investments, so she wants to understand investment gaps in the data science field.

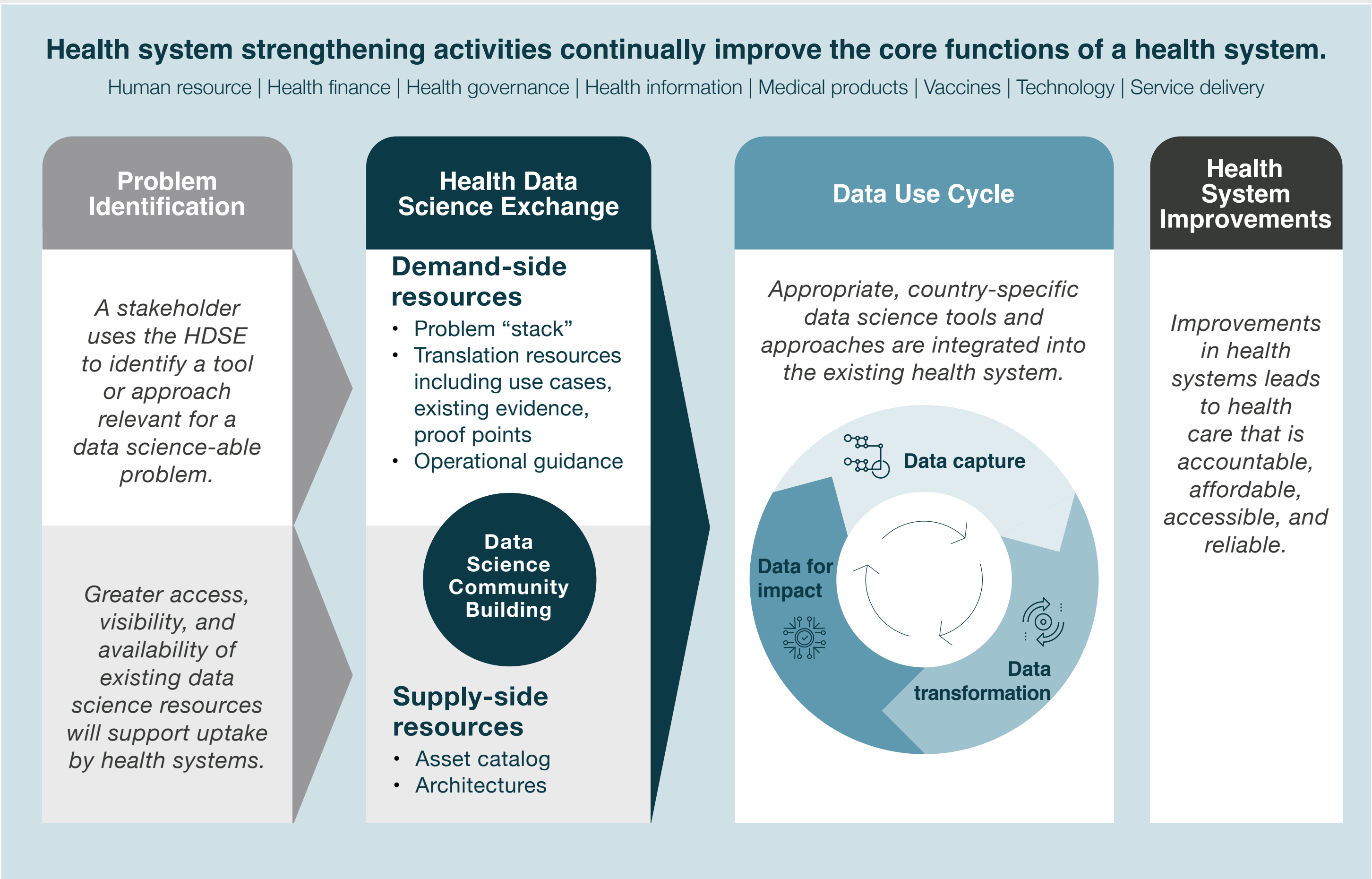
***These personas are illustrative of potential users involved in the HDSE. They were developed to clarify the potential uses and functions of an Exchange.*

The role of a Health Data Science Exchange

To accelerate the appropriate use of data science tools and approaches, stakeholders would benefit from a platform where they can convene to learn about existing data science assets, how data assets can be harnessed most efficiently in countries, and who supports innovation in this space. An exchange is a forum in which assets are made available to stakeholders on a curated platform. Exchanges bring together supply and demand by increasing overall visibility into the availability and quality of tools on the supply side, and increasing understanding of user needs on the demand side. Examples include Digital Square, Global Digital Health Partnership, United Nations Children's Fund (UNICEF) Innovation Fund, United Nations (UN) Global Pulse, Digital Impact Alliance (DIAL), and NetHope. No such exchange exists for data science in global health.

An HDSE would be a virtual space to bring together stakeholders with data science assets, operational guidance, use cases, and other resources that support the strengthening of low-resource health systems. The goal of the HDSE would be to connect country stakeholders to vetted data science assets and information necessary to appropriately deploy data science along the data use life cycle to improve health outcomes. There is significant potential for such an exchange to strengthen health systems and health outcomes through the appropriate scale of data science assets and resources. The creation of this platform and community would bring stakeholders together to solve complex challenges and accelerate the implementation of robust data science assets in low- and middle-income countries.

FIGURE 4. Supporting data use for health system strengthening through an HDSE



Users of the Health Date Science Exchange

The primary users of an HDSE are country government stakeholders across national and subnational levels, including officials from ministries of health, ministries of information and communication technology (ICT), data managers, and health facility staff who interact with data. The impact of a functioning HDSE ultimately benefits how they manage, transform, and use their data to improve the health outcomes of their communities. Because the focus of the HDSE is on the use of data science in public health, the HDSE will be designed specifically to support those working within public health systems or those working on systems used by the health system. The HDSE will be designed for users from a variety of backgrounds—from data and computer scientists to health managers and health area specialists.

An HDSE will ultimately connect users with the best, most efficient, and sustainable resources. Supporting them to effectively put into place and scale data science assets are investors, NGOs, multilateral and bilateral organizations, researchers, and software vendors.

User stories

A functioning HDSE would make innovations in how data science is applied across the data science cycle accessible for all countries to improve health services.

	**Government stakeholders** After talking to colleagues in other districts and some internet searches, Arjun finds the HDSE platform and uses the search function to find tools that could meet his needs. He also requests support in evaluating these tools. A monitoring and evaluation specialist from another district in India responds. They work together to identify the best option, and Arjun shares his experience with other district health officers in Odisha state.
	**Researchers** A colleague suggests that Winnie visit the HDSE platform. Accessing the platform for the first time, Winnie watches the demo video for how to use the Exchange. Using an advanced search function and asset tags, she looks for appropriate data validation tools. With her team, Winnie narrows the options to two tools and uses the HDSE platform to ask if anyone else has experience with these tools. Winnie receives advice and selects a data validation tool. In addition, she finds a training document that provides additional details on the tool's functionality. Users of the HDSE share best practices on data privacy, with specific guidance on mental health data. Winnie incorporates these into her research and receives positive feedback on her strong protection of data. Six months later, Winnie shares her experience on the HDSE by uploading her own user story so that others can learn more.
	**Civil society and nongovernmental organizations** Jonathan's organization and the investor funding the project in Rwanda are part of the HDSE. Jonathan engages with the HDSE community to understand which tools and approaches would best be deployed in Rwanda to meet their needs. After discussions on the HDSE platform, Jonathan has identified several technology solutions as well as a policy document from Indonesia, which he is adapting for his context. In addition, Jonathan has identified a group of other data warehouse managers that he creates a WhatsApp group with to continually share best practices.

Private-sector companies

Using feedback from other members of the HDSE, Patricia improves her models. She works with regional authorities to ensure prevention measures are in place and additional treatment commodities are available. Her models prove correct and an outbreak occurs. However, the outbreak is contained, and patients are treated. Patricia shares the updated models and approaches she deployed, along with data on health impact, so other government ministries and stakeholders can replicate them. Patricia also receives questions from other stakeholders about the dengue-specific incidence data that her team used.

Investors

Allison joins the Advisory Group of the HDSE. She presents on an Advisory Group call to make a case for investments in artificial intelligence for data collected from digital tools used at the community level. Through the HDSE, she learns what tools exist and the existing funding gaps. She is able to identify fellow investors, connect with them on their funding priorities, and determines where to prioritize her own investments. Through an open platform, she routes $1 million to fill gaps, requiring that all activities be shared through the HDSE when completed.

The function of a Health Data Science Exchange

An HDSE would be a virtual, interactive space to coordinate and align health and technology users to support the appropriate scale and use of data science assets through the data life cycle. It will house the tools, resources, approaches, and technologies that support data use and adaptive management in the service of health goals. The platform itself would be a launchpad for engagement and discussion, bringing stakeholders together to facilitate the transparent exchange of services, resources, and information between funders, implementers, practitioners, innovators, and country governments to execute decisions. Supply-side resources embedded in the platform will showcase how data science assets can drive improvement in health services and strengthen health services across disease verticals. These assets will be supported with demand-side resources focused on driving demand and capacity within countries. Additionally, an HDSE would coordinate investors to prevent misalignment and duplication of efforts across this growing field. An example of this is facilitating investments in a transparent manner to encourage alignment.

Through a wide but attainable range of activities, the HDSE will solve the challenges of which data assets are vetted to be scaled in countries and how country stakeholders can leverage the knowledge of experts to harness deep insights from their own data to improve health.

Health Data Science Exchange

Demand-side resources
- Problem "stack"
- Translation resources including use cases, existing evidence, proof points
- Operational guidance

Data Science Community Building

Supply-side resources
- Asset catalog
- Architectures

An HDSE will transparently* perform *the following functions:

- **Curate a problem "stack"** to clarify common or shared health system challenges that could be addressed through the application of data science assets: These problem statements could help clarify data science needs in a way that supports conversations between health system professionals and data scientists.

- **Curate a set of "translation resources"** designed to inform country decisionmakers on the value, utility, and availability of data science assets: These translation resources could take several forms:

  - Use cases: Successful examples of data science projects that include information on specific assets, operational prerequisites, implementation processes, evidence of health system impact, and generalizable lessons learned.

  - Evidence and proof points: A combination of existing evidence resources, mapping gaps in available evidence, and support to generate new evidence in order to clarify the impact of data science on health systems and health outcomes.

- **Provide operational guidance:** Curated operational best practices, including but not limited to data and system prerequisites; data management, security, privacy, and ethics; system integrations; country capacity requirements; and other guidance from normative bodies.

- **Curate an asset and architecture catalog:** Entries on specific data science assets and architectures, including information on possible applications (tied to the problem stack), technical specifications, classifications, and links to additional resources.

- **Build community:** Activities or mechanisms designed to strengthen connections between individuals, organizations, investors, and government institutions working on data science in a public health context.

- **Support capacity strengthening** by including existing training and other skill-building resources in the asset compendium.

- **Support the development, promotion, and uptake** of data science standards, policies, and vetting processes by a representative, technical governing, or normative body

- **Reinforce digital and data standards** developed by normative bodies that ultimately make it easier for those generating datasets to harness clean, secure data for analysis

- **Coordinate investments** in data assets to encourage greater alignment and coordination

The HDSE would coordinate closely with initiatives like the Health Data Collaborative, amplifying and supporting initiatives if and as they increasingly fulfill these listed functions, and focusing efforts wherever gaps remain.

Managing a Health Data Science Exchange

In order for an HDSE to function, a neutral coordinating partner with experience in both health data science and in coordinating a complex community of actors in global digital health would serve as a convener. The coordinating partner would need experience in managing virtual and physical communities, including experience in online forum management. Investment from multiple investors in this Exchange would signal investor alignment and reduce investor duplication in the emerging space.The HDSE would function like some of the existing digital health marketplaces (i.e., Digital Square), adapting successful components from other marketplaces (i.e., Asia eHealth Information Network capacity building trainings).

Recommendations

The HDSE Advisory Group envisions a world in which the appropriate use of data science assets closes the health equity gap. The role data science plays in the global health ecosystem is critical to support stakeholders to use data efficiently and innovatively to directly improve health systems and health services.

The Advisory Group makes the following recommendations:

Establish the Exchange.

The following illustrative activities would support the launch of a human-centered, virtual HDSE:

- Investors adopt Advisory Group recommendations and commit to funding an HDSE over the next five years.
- Advisory Group and investors identify a neutral convener to manage the Exchange and foster community harmonization.
- Neutral convener drafts a roadmap for the Exchange, including a proposed curation protocol and/or vetting process.
- Neutral convener hosts a forum for input and feedback from stakeholders, with emphasis on input from country stakeholders.
- Neutral convener develops a beta version of the Exchange and soft launches the platform with country public health and ICT users.
- Based on soft launch, neutral convener documents where Exchange provides greatest value to users, adapts Exchange to further meet user needs, and develops a costed plan for Exchange maintenance.
- Launch platform and include stories of how the platform has already helped countries solve health challenges.
- Continue to serve country and other stakeholders' needs for visibility by continuously engaging users for feedback and ensuring platform provides valuable information and connections.

Research the impact and scalability of data science assets.

The following illustrative activities would accelerate evidence generation on the value of data science assets:

- Investors partner with neutral convener to launch a public call offering funds to support an impact evaluation of currently deployed data science assets.
- Investors partner with neutral convener to launch a public call offering funds to scale and evaluate drivers of scalability for two to three use cases in which a collection of data science assets has already demonstrated health impact.
- Winning applicants conduct the research, regularly sharing lessons learned to enable the community to adopt and use key insights more quickly.
- Research findings are published and promoted, building alignment and confidence within the global community on what data science assets lead to impact and what supports scalability.

Asset compendium

The HDSE would be a virtual space to bring together stakeholders with data science assets, operational guidance, use cases, and other resources that support the strengthening of low-resource health systems. The role, function, and users of an HDSE are explained in the value proposition. This asset compendium complements the value proposition by illustrating the types of use cases, data science assets, and translation resources that could be included in an HDSE.

The asset compendium comprises three use cases where data science approaches have been applied in low-resource settings to strengthen health systems. Each use case was focused on implementation of data science approaches related to data capture, data transformation, and/or data for impact. From these three use cases, an initial, curated set of data science assets and a classification schema to organize the assets were developed.

Selecting the use cases

The asset compendium is anchored on three uses cases with an established link to health system impact in a low-resource setting. The HDSE Advisory Group assisted with the identification of potential use cases and informed the criteria used to prioritize use cases. Two tiers of criteria were developed (Figure 1): tier 1 criteria outlined the minimum requirements a use case must meet, and tier 2 criteria were considered to prioritize among the use cases that met all tier 1 criteria.

FIGURE 1. Prioritization criteria for use case selection

Tier 1	Tier 2
Identify which use cases meet the minimum criteria to be considered.	***Prioritize among those use cases that meet the minimum criteria.***
+ *Demonstrated health system impact in a low- and middle-income-country setting.* + *Availability of information and/or willingness of the implementing organization to partner on developing the use case.* + *Inclusion of at least two health data science assets.*	+ *Strength of evidence for health system impact.* + *Variety of health data science assets represented within and across use cases.* + *Variety of countries represented across use cases.* + *Variety of organizations represented across use cases.* + *Level of Advisory Group interest.*

Based on an initial landscape review and input from the Advisory Group, 16 potential use cases were identified where data science approaches had been implemented related to a broad range of health focus areas, including malaria, immunization, air pollution, maternal and child health, pneumonia, and tuberculosis, among others. Each use case was assessed based on the first tier of prioritization criteria; 9 of the 16 use cases met the tier 1 criteria. The primary reason some use cases did not meet the tier 1 criteria was limited evidence of health system impact. For example, one potential use case applied machine learning algorithms to immunization data and demonstrated the ability to improve vaccine forecasting by 70 percent as compared to the best models currently in use; however, the algorithms have not yet been used in practice to inform immunization planning in a low-resource setting. In general, many data science interventions have generated excitement but have yet to be applied in low-resource health systems or lack documented health systems impact.

Among the nine use cases that met the tier 1 criteria, the HDSE Advisory Group prioritized three use cases based on the tier 2 criteria. The three selected use cases are:

» **The introduction of ICT tools** to support frontline workers to improve maternal and child health in Bihar, India.

» **The pilot implementation of a data-driven decision support tool** to improve community-based diagnosis and treatment of child health in Kano State, Nigeria.

» **Strengthened data analytics tools and capacity** for malaria elimination in Southern Province, Zambia.

Classifying health data science assets

For an HDSE to be successful, users must be able to locate and learn about the health data science assets they need to support their health systems. In order for the Exchange to efficiently match user needs to asset offerings, and to appropriately organize navigation, the Exchange will classify each asset based on its characteristics and intended use.

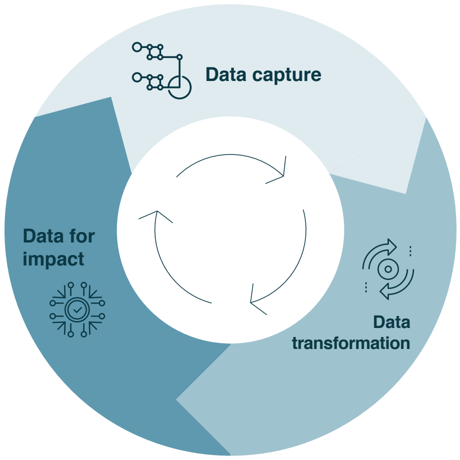
Different users will approach the Exchange with different requirements. To address different levels of detail needed by various types of users, the Exchange will organize assets into two levels of classification: high-level categories and detailed subcategories.

High-level: Applied data science cycle and the 5 Ps

This level of classification will serve the needs of users who are less involved in data science from day to day, and who are most concerned with the capabilities and impact that assets will unlock. This type of user may include investors, clinicians, policymakers, and program leaders.

At a high-level, assets will be organized by the phase(s) of the applied data science cycle that they support. The applied data science cycle was developed based on a landscape review of existing data science approaches and data life cycles, and with input from the HDSE Advisory Group, to support classification into three data science phases:

- **From concept to capture | Data capture assets** including assets related to data design, data production, data location, data capture, and data validation.
- **From data to insights | Data transformation assets** including assets related to data aggregation, data processing, data analysis, and data exploration.
- **From insights to real world impact | Data for impact assets** including assets related to data dissemination, data use, data communication, monitoring and evaluation, policy change, and behavior change



In addition, at the high level, assets are categorized into one of the following categories, referred to as “the 5 Ps” that reflect the type of asset:

- **Products:** The tools, systems, or software that enable users to collect, report, and analyze data. Product assets include, but are not limited to, technology solutions such as data hosting services, data processing algorithms, and data aggregation tools.
- **Policies:** The guidance and documentation that govern and enable data production and use, and which are informed by the data and insights generated by other data science assets. Policies may include data privacy policies, data sharing templates, and regulatory frameworks.
- **Processes:** The management structure and practices surrounding data generation and use. These processes may be formal or informal and include such assets as workflow maps for specific work types, or methodologies for data use.
- **People:** The workforce that supports health data utilization, health system planning, and service delivery at all levels. This category encompasses assets that are intended to strengthen the workforce capacity for data-related activities through training, behavior change, and vendor outsourcing.
- **Proof**: The body of evidence or evidentiary standards used to demonstrate impact or rigor of a data science application. Assets in this category include research protocols, standards for evidence strength, and case studies of health impact.

Detailed level

A more detailed classification will be applied to support data science professionals, innovators, vendors, system architects, and other stakeholders who may have more specific types of assets they are interested in finding. Once assets are classified at a high level by the applied data science cycle and the 5 Ps, they will be further classified into more precise subcategories. The bulleted sub-categories in Table 1 illustrate example subcategories for consideration, although they are not meant to be comprehensive at this stage.

TABLE 1. Health data science asset classification.

	DATA CAPTURE ASSETS	DATA TRANSFORMATION ASSETS	DATA FOR IMPACT ASSETS
PRODUCTS	» Synthetic patient datasets or dataset creation protocols. » Data harvesting/mining/scraping services and tools.	» Open algorithms and code libraries. » Data management assets (cleaning, de-identifying, etc.).	» Methods and approaches for localization of data science assets. » Standardized language translation services across dictionaries.
POLICIES	» Data standards and guidelines. » Data structure requirements.	» Predictive models. » Data privacy and inclusivity requirements.	» Regulatory policies for assets. » Data-enabled policymaking assets. » Pricing and procurement templates.
PROCESSES	» Data flow templates. » Business understanding—model protocols for design (e.g., CRDM).	» Business process/wireframe templates (workflows, decision trees, semi-structured/ accelerator kits). » Visualization templates.	» Crosscutting methods and approaches (strategy, implementation, adaptive management). » Impact assessment approaches.
PEOPLE	» Certified/accredited outsourcing vendors, policymakers, institutions. » Learning content. » Data-enabled behavior change models.		
PROOF	» Research methods and protocols. » Library of evidence tagged to assets. » Guidance on rating the credibility/strength of evidence.*		

**All listed subcategories are representative and have not been formally vetted by the HDSE Advisory Group.*

Classification considerations

The asset classification is structured to give users a simple way to find the tools they need. However, to ensure that the searchability of the Exchange remains flexible, the classification will not be rigid and assets could be classified through a tagging system in multiple categories and subcategories.

Use cases and health data science assets

To demonstrate the variety of translation resources, operational guidance, and health data science assets that would be classified within the Exchange, three illustrative use cases are included here. Each includes an introduction to the public health challenge addressed by the data science approach, key characteristics of the "problem stack" that the use case assets address, and implementation and scale considerations tailored to each use case. The health data science assets included in the use cases will not automatically be included in the Exchange and will be subject to the same vetting process as all other assets once that process is identified.

USE CASE

Information Communication Technology tools for frontline workers to improve maternal and child health in Bihar, India

The challenge

In Bihar, India, frontline workers (FLWs) are deployed by the government to provide family health services to women and children; these services are aimed at reducing local rates of maternal, newborn, and child mortality; fertility; and undernutrition.

FLWs who provide these continuum of care services (CCS) can mobilize communities to seek and use care services, but they are not always equipped with appropriate data to identify and monitor clients, make referrals, provide counseling, or identify danger signs.

Accredited social health activists and anganwadi workers are the FLWs that provide health services at the community level in Bihar. Without ICT tools, the expected practice is to use paper-based home visit registers to track visits and interactions with beneficiaries. FLWs without ICT tools therefore lack a way to seek real-time information that they need to do their jobs and serve their beneficiaries effectively.

The solution

CARE International introduced an ICT CSS intervention in Bihar, India, beginning in mid-2012 to increase the coverage and quality of services that FLWs provide, enhance their communication with beneficiaries, and facilitate supervision of FLWs.

The ICT-CCS intervention was implemented as part of the ongoing Ananya program created by the Bill & Melinda Gates Foundation to improve maternal and child health (which began in 2011). The intervention included introduction of a mobile phone tool for FLWs, along with training and technical support.

The ICT-CCS tool combines registration of beneficiaries, scheduling of home visits, and guided protocols, along with audiovisual job aids. FLWs enter client informationinto the tool, including registration and subsequent visits, that is processed by a back-end server that manages the scheduling of home visits for each pregnant woman and mother with young children in each FLW's coverage area and provides FLWs with reminders about the timing of home visits. The tool also includes checklists and videos that are intended to support the FLWs in communicating health-related information to beneficiaries during these visits.

Project dates
2012 to 2015

Location
Bihar, India

Health focus area
Maternal and child health

Partners

Dimagi

CARE International

BBC Media Action

Grameen Foundation

MOTECH

Mathematica Policy Research

Bill & Melinda Gates Foundation

Further information

Borkum E, Sivasankaran A, Sridharan S, et al. Evaluation of the Information and Communication Technology (ICT) Continuum of Care Services (CCS) Intervention in Bihar. Princeton, NJ: Mathematica Policy Research; 2015. https://www.mathematica.org/our-publications-and-findings/publications/evaluation-of-the-information-and-communication-technology-ict-continuum-of-care-services-ccs

Impact

Figure 1 shows the logic model for the ICT-CCS intervention. The expectation was that if FLWs understood and used the tools, they would improve their service provision, which in turn would result in improved health behaviors and lead to positive health impact.

Mathematica Policy Research evaluated the pilot intervention through a randomized control trial after two years of implementation. The study measured the value-add of the ICT-CCS tool by examining whether the introduction of the tool let to changes in how FLWs provided services and in beneficiary behavior and practices.[1] Seventy health subcenters were divided into treatment and control groups (35 subcenters in each); the treatment group received the ICT-CCS intervention and the control group did not. The study used qualitative and quantitative data collection methods to understand differences between the treatment and control groups.

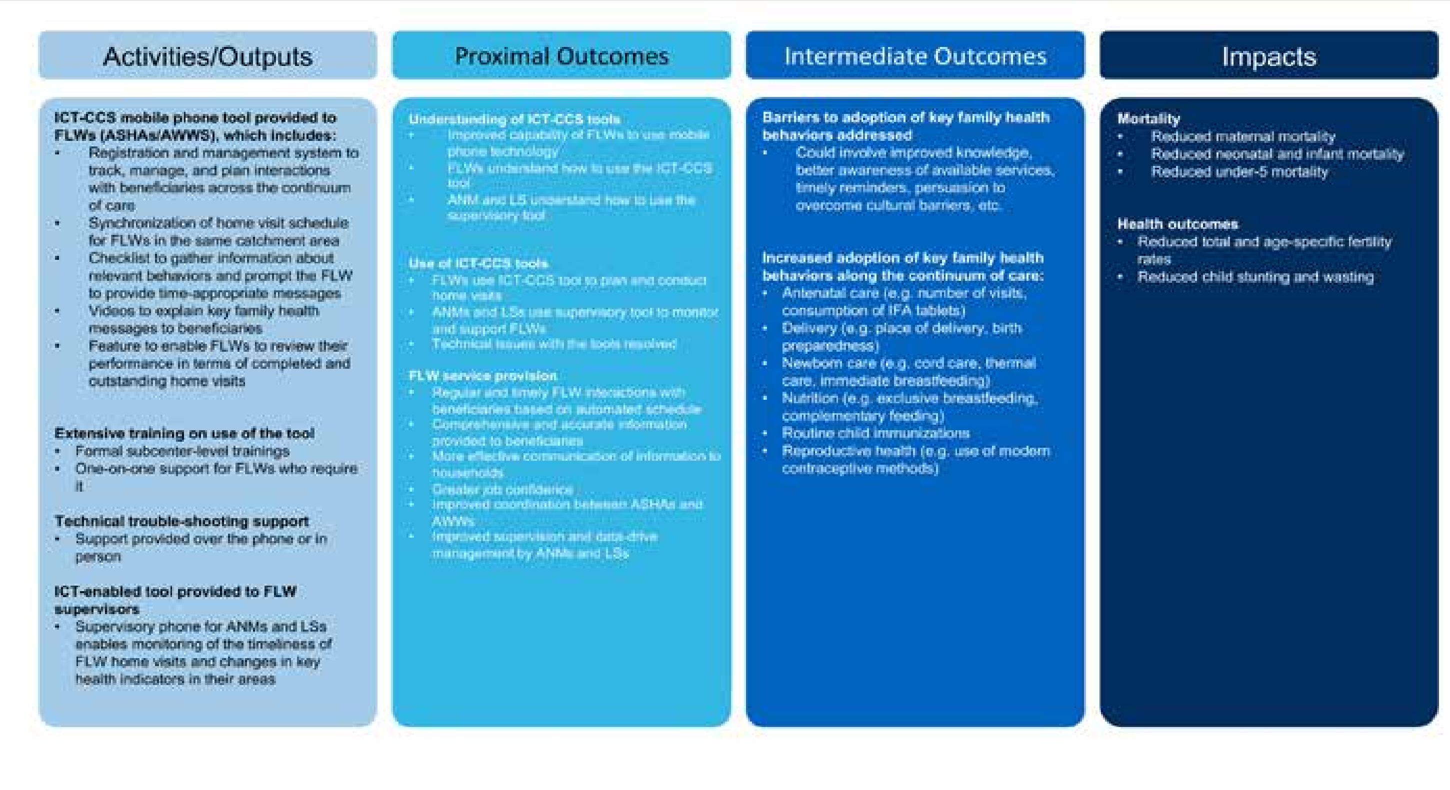
The study found that CARE's intensive training on the tool was effective in increasing FLW understanding of and use of the tool, although use of the different ICT-CCS features varied by user. The tool led to some improvement in coordination of FLW home visits and increased job confidence. There was a statistically significant increase in FLWs asking other FLWs to conduct a home visit when they were unable to do so (p-value = 0.018). Other measures of coordination increased in the treatment group compared to the control but not at a statistically significant level. Increase in job confidence was shown by a statistically significant increase in percent of treatment group FLWs who reported that they thought they had the skills required for their jobs (p-value = 0.039). However, the intervention was less successful at improving supervision practices. There was no statistically significant difference between treatment and control groups for frequency of meetings between FLWs and their supervisor outside of subcenter meetings. There was also no difference in frequency of supervisors joining FLWs on home visits or percent of supervisors reviewing home visit registers or work phones at subcenter meetings.

There was also evidence that the intervention improved the frequency of interactions between FLWs and beneficiaries—a goal of the intervention package—and improved the quality of the interactions, albeit to a lesser extent. In the treatment group, more beneficiaries received at least two FLW home visits in the final trimester, home visits within one week and one month of delivery, and a complementary feeding home visit. Beneficiaries in the treatment group were more likely to receive advice on exclusive breastfeeding and complementary feeding but not on other topics like antenatal and newborn care. There was also a statistically significant increase in beneficiaries reporting that a FLW used job tools during home visits, but there were no differences in average duration of visits or whether the FLW spoke to other family members during the visit.

The intervention also improved health behaviors in the areas of antenatal care, child nutrition, and reproductive health, as well as some impacts on delivery and newborn care. For antenatal care, the intervention significantly increased the proportion of beneficiaries receiving at least three antenatal care visits, the proportion of beneficiaries using iron supplementation during pregnancy, and birth preparedness practices. For delivery and newborn care, there was a significant increase in the proportion of mothers breastfeeding immediately after birth. For child nutrition, there was a significant increase in reports of children aged 6 months or older eating solid or semisolid food and that the child started eating solid or semisolid food by 6 months of age. For reproductive health, there was a significant increase in the use of permanent contraception methods and ever using temporary modern contraception methods (no difference in current use of temporary contraception methods).

1 Borkum E, Sivasankaran A, Sridharan S, et al. *Evaluation of the Information and Communication Technology (ICT) Continuum of Care Services (CCS) Intervention in Bihar*. Princeton, NJ: Mathematica Policy Research; 2015. https://www.mathematica.org/our-publications-and-findings/publications/evaluation-of-the-information-and-communication-technology-ict-continuum-of-care-services-ccs

FIGURE 1. Logic model for the ICT-CCS intervention in Bihar, India



Note: ANM, auxiliary nurse midwives; ASHA, accredited social health activist; AWW, anganwadi worker; CCS, continuum of care services; FLW, frontline worker; ICT, information and communication technology; IFA, iron-folic acid; LS, lady supervisors.

Scale and future health system applications

In India, the ICT-CCS intervention piloted in Bihar laid the groundwork for a larger scale-up of mobile health (mHealth) tools nationwide. The Integrated Child Development Scheme, launched in 1975, is one of India's national flagship programs to support the health, nutrition, and development needs of children and women. This support takes place through a network of anganwadi centers that provide services to pregnant women, children, and their mothers. Following the 2015 results of the ICT-CCS intervention, the Government of India began to strengthen the Integrated Child Development Scheme in seven states using an mHealth intervention called Common Application Software, which was installed on smartphones with accompanying multilevel data dashboards. This system is intended to be a job aid for FLWs, supervisors, and managers. It aims to ensure better service delivery and supervision by enabling real-time monitoring and data-based decision-making.[2] Scale-up is continuing, as of February 2020, more than 625,000 anganwadi workers across 28 states have been using the application for service delivery.[3]

Future scale-up of the ICT-CCS intervention to other geographies and for health applications outside of maternal, newborn, and reproductive health is certainly possible. To our knowledge, the full ICT-CCS intervention package as described in this use case has not been applied in other settings; however, individual data science assets that are part of this use case have been widely applied in other contexts.

CommCare was the mobile data collection and service delivery platform that was implemented as part of this use case to register and monitor beneficiaries, track immunizations, schedule home visits, provide health information directly to beneficiaries (through checklists and videos), and monitor performance of FLWs. CommCare is used by more than 2,000 projects across 80 countries for mobile data collection and reporting.[4] For example:

+ **In Tanzania, FLWs use CommCare to assist with danger sign identification and referrals.**
+ **In Uttar Pradesh, India, FLWs use CommCare for real-time guidance on counseling and decision-making, as well as for time-sensitive alerts.**
+ **In Guatemala, the CommCare platform is used to receive continuous training and to perform community health promotion and prevention activities.**
+ **In Zanzibar, traditional birth attendants use CommCare to record permissions for emergency transport from family members and to facilitate payment for transportation.**[5]

Most recently, CommCare has been used to support the response to the COVID-19 pandemic. Dimagi created a template application to implement the World Health Organization's First Few X (FFX) cases protocol.[6] As of March 30, 2020, 500 organizations have downloaded the application.[7] At the time of publication, Dimagi is also working on additional template applications, trying to secure free messaging to support self-reporting workflows for positive COVID-19 cases, and pursuing WhatsApp integration.

2 Nimmagadda S, et al. (2019). Effects of an mHealth intervention for community health workers on maternal and child nutrition and health service delivery in India: Protocol for a quasi-experimental mixed-methods evaluation. BMJ Open. 9. e025774. 10.1136/bmjopen-2018-025774. https://www.researchgate.net/publication/332039883_Effects_of_an_mHealth_intervention_for_community_health_workers_on_maternal_and_child_nutrition_and_health_service_delivery_in_India_Protocol_for_a_quasi-experimental_mixed-methods_evaluation#pfa

3 "Anganwadi Workers Were given Smart Phone for ICDS-CAS." Fresherslive. Fresherslive Current Affairs, March 13, 2020. https://www.fresherslive.com/current-affairs/articles/anganwadi-workers-were-given-smart-phone-for-icds-cas-24927.

4 https://www.dimagi.com/commcare/

5 CommCare. "The CommCare Evidence Base for Frontline Workers Overview." August 2019. https://www.dimagi.com/toolkits/commcare-evidence-base/

6 https://confluence.dimagi.com/display/commcarepublic/COVID-19+Template+App%3A+WHO+FFX+Protocol

7 Digital Square Webinar: Global Goods Adaptation for COVID-19 Response. March 30, 2020. https://path.zoom.us/rec/play/vpJ4cLqop243HILDuASDAvYrW9S7K6usgScb8_UEzRm8VXAGNlamNeBBMbbewJNnd8VplC77G3m4hu0O?autoplay=true&startTime=1585575432000

Implementation considerations

Connectivity and mobile phone usage

During implementation of the ICT-CCS intervention, FLWs did report some technical issues such as poor internet connectivity, which limited ability to synchronize records with the main server, as well as logistical challenges, including delays in receiving funds to cover internet charges. Potential users of the intervention package should ensure that internet bandwidth will be sufficient in the local area and that there are backup systems in place in case of outage. To resolve the challenge of refunding FLWs for internet costs, users should consider implementing a mobile payment scheme to reduce lag time.

Although not specific to the ICT-CCS implementation in Bihar, other programs using CommCare have reported issues with using mobile phones if there is no convenient way to charge or fix them if they break.

Health worker burden

FLWs who were part of this intervention were required to fill out paper registers, while learning to use the new ICT tool and implementing it. Because of this, some FLWs reported that the tool added to their workload. Users who consider introducing this tool should consider workload burden on health workers and build in staging or other mitigation strategies to reduce the drain on FLW time. It should be noted, that if successful, the intervention is intended to provide a net reduction in FLW workload with the efficiencies that would result from ideal use of the tool.

Training and support

Results from the pilot evaluation highlight the importance of the training approach that provided intensive support to familiarize FLWs with the ICT-CCS tool. The training approach implemented by CARE International is included as a data science asset within this use case. Future applications of this use case should ensure that sufficient training and support—both formal and informal—are provided for end users. The training and support offered should be tailored to the end users' starting comfort level with technology. In the case of the Bihar implementation, many FLWs had little to no experience with mobile phones beyond making phone calls.

Data science assets

PRODUCT

PROCESS

CommCare mobile health platform

Data Collection

Data Transformation

Data for Impact

CommCare is an offline-capable mobile data collection and service delivery platform designed for everything from simple surveys to comprehensive longitudinal data tracking.

A straightforward application builder allows for easy digitization of surveys and forms, as well as the integration of decision support, notifications, and SMS (short message service, or text) messaging. Programs can be scaled from the community to the national level, thanks to simple device deployment and translation features.

In the Bihar ICT-CCS system, CommCare served as the phone-based interface for FLWs. The interface includes forms, checklists, videos, and data collection tools. Beneficiary information entered in CommCare is sent to a central cloud server. MOTECH is a back-end server that integrates beneficiary data that have been entered in CommCare and manages schedules for each beneficiary, which get updated back to the FLWs' phones.

The CommCare ICT-CCS tool includes features to register and monitor beneficiaries, track immunizations, schedule home visits, provide health information directly to beneficiaries (through checklists and videos), and monitor performance of FLWs. These features include:

- **Beneficiary registration form.**
- **Home visit scheduler.**
- **Immunization scheduler.**
- **Beneficiary management form.**
- **Interactive checklists to guide FLW-beneficiary interactions.**
- **Videos for education / behavior change communication.**
- **FLW performance report.**

Mobile data collection tool

Data management / business processes

Decision support

For more information

https://dimagi.com/commcare/

PRODUCT

PROCESS

MOTECH open source enterprise software

Data Collection

Data Transformation

Data for Impact

Hosting services

Dimagi's MOTECH is a CommCare-based interface that supports the integration of scalable mobile services and health information systems.

MOTECH implements the OpenHIE (open source health information exchange) standards, which are emerging as the global standards for interoperability of health information systems and registries. MOTECH is designed to enable integration with a set of self-service features, which enables the sharing of data between systems to be configured without software developers or code changes. MOTECH supports integration with District Health Information Software 2 (DHIS2) and Open Medical Record System (OpenMRS).

In the Bihar ICT-CCS system, CommCare served as the phone-based interface for FLWs. The interface includes forms, checklists, videos, and data collection tools. Beneficiary information entered in CommCare is sent to a central cloud server. MOTECH is a back-end server that integrates beneficiary data that have been entered in CommCare and manages schedules for each beneficiary, which get updated back to the FLWs' phones.

For more information

https://motechproject.org/

PROOF

Randomized controlled trial to evaluate the ICT-CCS intervention:

Study design, research protocol, and data collection tools

Data Collection

Data Transformation

Data for Impact

Research methods

Research protocols

Mathematica Policy Research conducted a randomized controlled trial to evaluate the impacts of the ICT-CCS intervention in the Saharsa district of Bihar.

The research protocol and study design included evaluation questions, specifications on data collection processes, sample selection and size, and analysis methods. The evaluation used a mixed-methods approach that included quantitative data from surveys with FLWs and beneficiaries and qualitative data from field visits and senistructured interviews with implementing partner staff, FLWs, and beneficiaries.

The evaluation was designed to answer the following questions:

» **What was the ICT-CCS intervention, and how was it implemented? To what extent did FLWs understand how to use the new ICT-based tools? What were the practical challenges or barriers to using the tools?**

» **What was the impact of the ICT-CCS intervention on FLW-household interactions? Did ICT-based tools lead to an improvement in the quantity and quality of FLW-household interactions?**

» **Did the intervention lead to improvements in maternal and child health outcomes among beneficiaries? Did ICT-based tools lead to improvements in key health outcomes across the family health continuum? If so, were these improvements larger for certain subgroups of the population than for others?**

For more information

Borkum E, Sivasankaran A, Sridharan S, et al. Evaluation of the Information and Communication Technology (ICT) Continuum of Care Services (CCS) Intervention in Bihar. Princeton, NJ: Mathematica Policy Research; 2015. https://www.mathematica.org/our-publications-and-findings/publications/evaluation-of-the-information-and-communication-technology-ict-continuum-of-care-services-ccs

PEOPLE

Training approach for frontline workers

Data Collection

Data Transformation

Data for Impact

Methods and approaches for learning

FLWs' understanding of the ICT-CCS tool increased over time, the result of an intensive training effort by CARE International.

Dimagi (the firm that built the CommCare mobile data collection platform) initially trained one CARE ICT coordinator at each subcenter, who then conducted trainings for the FLWs. The training approach used in the ICT-CCS intervention was much longer and more intensive than the common week-long, one-time trainings on mHealth tools. The ICT tool was a new way of working for the FLWs, which required a longer training, with gradual learning over time. The training was built specifically to address the modules included in the ICT platform and was implemented in waves to improve retention.

One critical aspect of the training approach was the use of a practice mode and follow-on learning. The practice mode allowed FLWs to try out simulated beneficiary interactions as needed without the pressure of real-life data collection. Follow-on learning included access to a suite of video and audio resources. FLW supervisors were also able to see the data that FLWs in their supervisory group were collecting and inputting. Finally, the approach prioritized the eventual transfer of supervision to national health system staff in the interest of scale-up and sustainability.

Based on interviews and demonstrations with FLWs two years after the ICT-CCS intervention began, there was a high level of understanding on how to use the home visit scheduling feature and the checklist function. The high understanding was attributed to CARE's effort to provide ongoing training and mentoring opportunities inside and outside of formal meetings. Formal meetings consisted of 16 three-hour long training sessions over eight weeks for FLWs to learn the basics of the tool. CARE staff were able to connect FLWs who would benefit from additional instruction to informal mentoring opportunities before the next training session. This helped prevent the FLW from becoming increasingly behind their peers and allowed them to join the next formal learning session at the same level of understanding as the rest of the group.

For more information

Borkum E, Sivasankaran A, Sridharan S, et al. Evaluation of the Information and Communication Technology (ICT) Continuum of Care Services (CCS) Intervention in Bihar. Princeton, NJ: Mathematica Policy Research; 2015. https://www.mathematica.org/our-publications-and-findings/publications/evaluation-of-the-information-and-communication-technology-ict-continuum-of-care-services-ccs

USE CASE

Data-driven decision support to improve community-based diagnosis and treatment

The challenge

Treatable conditions (such as pneumonia, diarrhea, and malaria) result in millions of deaths each year among children living in low- and middle-income countries.

More than half of these deaths could be prevented with early clinical assessments and appropriate treatment. A primary way to achieve this is through access to health care professionals who can perform high-quality integrated assessment that will identify those children who are sick and require immediate intervention. However, many children in these settings do not have access to high-quality health care. Some programs have expanded access through community health workers (CHWs), also known as FLWs, but FLWs do not always have the time, resources, or ability to diagnose and treat patients with high-quality care.

The World Health Organization and UNICEF have developed a strategy called the Integrated Management of Childhood Illness (IMCI) to improve case management, overall health systems, and family and community health practices to ensure children receive high-quality care. A multicountry evaluation of the IMCI strategy has shown that it improves health worker performance and quality of care, and can improve child health outcomes if implemented well.[1] However, other studies have shown that achieving adherence to IMCI guidelines can be challenging. Most FLWs lack fast assessment support that aligns with the WHO's IMCI guidelines and enables FLWs to effectively use the guidelines to assess, treat, and diagnose their patients.[2]

Project dates
2018

Location
Nigeria

Health focus area
Child health

Partners

THINKMD

eHealth Africa

Kano State Primary Health care Management Board

Further information

http://www.thinkmd.org/

https://www.ehealthafrica.org/medsinc

Finette BA, McLaughlin M, Scarpino SV, et al. Development and Initial Validation of a Frontline Health Worker mHealth Assessment Platform (MEDSINC®) for Children 2–60 Months of Age. The American Journal of Tropical Medicine and Hygiene. 2019;100(6):1556–1565. https://doi.org/10.4269/ajtmh.18-0869.

1 Organization. Integrated Management of Childhood Illness. https://www.who.int/maternal_child_adolescent/topics/child/imci/en/

2 Mulaudzi MC, 2015. Adherence to case management guidelines of Integrated Management of Childhood Illness (IMCI) by healthcare workers in Tshwane, South Africa. S Afr J Child Health 9: 89–92.

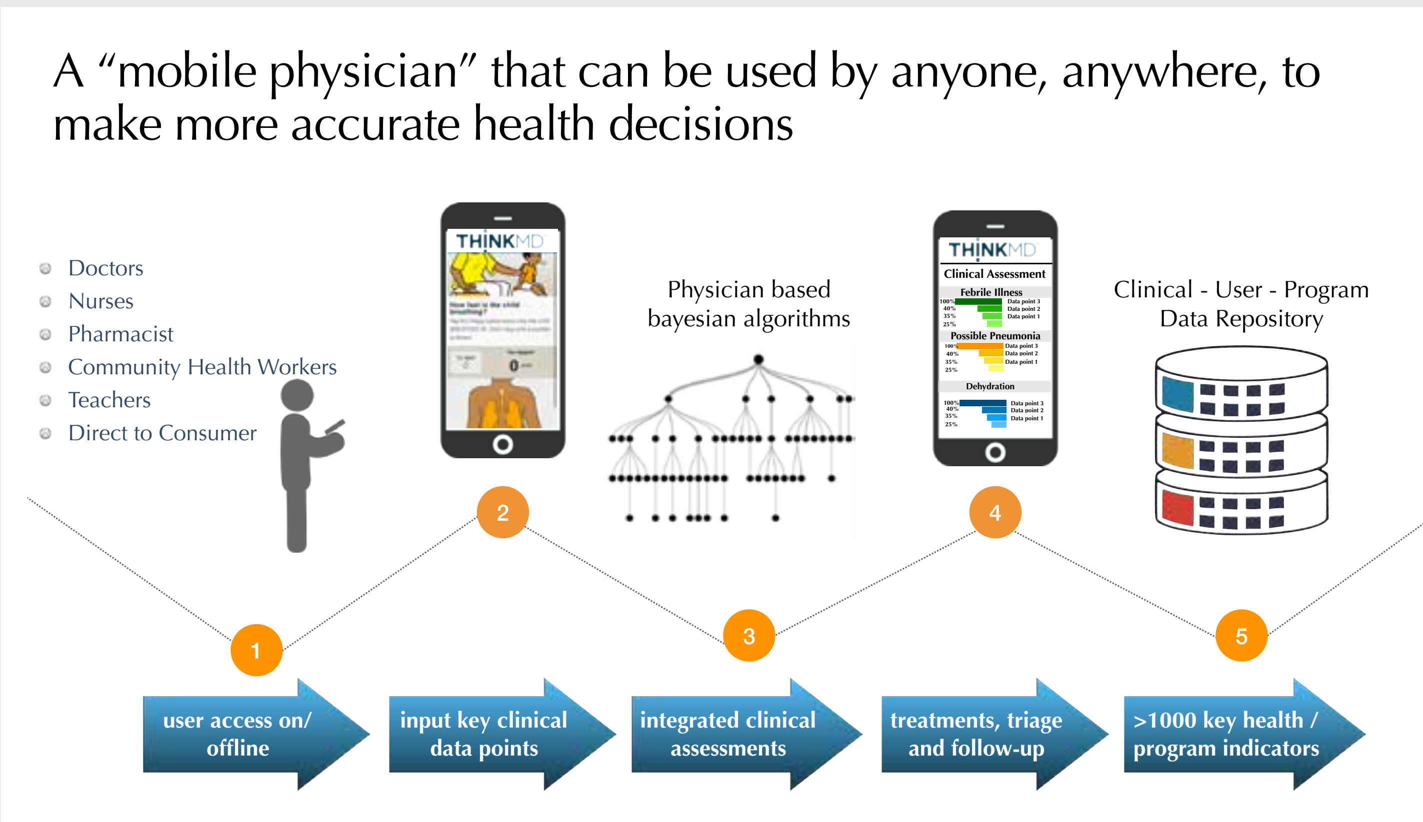
The solution

THINKMD developed a digital health point-of-care tool to support FLWs to assess, diagnose, and treat patients.

To use it, FLWs gather a series of data points about the patient's past and current history, symptoms, vital signs, and physical examination findings. The THINKMD technology relies on underlying algorithms to weight these data points compared to clinical best practices and global guidelines to generate a clinical risk assessment and recommendations for triage, treatment, and follow-up (Figure 1). In this way, the THINKMD technology can help any user to make a data-driven decision to identify how sick a person is, what illness they may have, and what next steps to take—thereby expanding access to high-quality health care.

The THINKMD technology was introduced in 15 health facilities in Kano State, Nigeria, in April to August 2018, to support FLWs to diagnose and treat patients and improve adherence to the WHO IMCI guidelines.

FIGURE 1. **The THINKMD technology processes acquired clinical data points through data algorithms to generate patient-specific triage, treatment, and follow-up recommendations**



Impact

In Kano State, Nigeria, the pilot of the THINKMD technology showed a 41 percent observable and 50 percent actual improvement in the ability of FLWs to appropriately diagnose sick children, according to the WHO-IMCI guidelines.[3] This was based on a pre-/post-observational analysis of 28 FLWs during the use of their current paper-based approach ("pre" n = 384 child assessments) compared to during their use of the THINKMD technology ("post" n = 384 assessments). Based on usability and acceptability surveys:

+ **93 percent of FLWs said the technology was very easy to learn and use.**
+ **100 percent of FLWs said the technology was a valuable training tool.**
+ **100 percent of FLWs said the technology helped them identify sick children.**
+ **100 percent of FLWs said they were extremely likely (75 percent) or likely (25 percent) to recommend use of the technology to other FLWs.**

This pilot built on prior research that was conducted to validate the THINKMD technology in three countries: Burkina Faso, Ecuador, and Bangladesh.[4] The validation study indicated a high specificity correlation between the THINKMD technology diagnosis and the "gold standard" of diagnosis by a local health care provider (LHP). Of the four key clinical conditions that the technology assessed—respiratory distress, dehydration, sepsis (systemic inflammatory response syndrome), and acute malnutrition—the correlation between the THINKMD and LHP assessment ranged from 55 percent to 97 percent. The 55 percent correlation between THINKMD and LHP assessment was in acute malnutrition in Bangladesh and was an outlier, with the next lowest correlation at 82 percent. Without that outlier, the comparative specificity of the four risk assessments was between 85 percent and 100 percent (the outlier had a 75 percent specificity for acute malnutrition in Bangladesh). As expected, there was a wide sensitivity range (0.12 to 0.81) between THINKMD and LHP assessment due to the low prevalence of these conditions in the test population. The validation study also asked the FLWs to provide input on how easy the THINKMD platform was to learn, how easy it was to use, and if it allowed them to do their job better. The results showed positive usability and feasibility responses from FLWs and LHPs related to the ease of use, learning, and job performance associated with the THINKMD technology (Figure 2).

While the initial validation study of the THINKMD technology showed promising results, a key limitation was the use of physician-based assessments as the benchmark or "gold standard" for comparison. Health practitioners are not error-proof and there can be large variability in clinical assessment results across practitioners. Another limitation raised by stakeholders was the lack of transparency into the back-end algorithms that merged internationally recognized WHO guidelines with a novel physician-based logic.[5] However, this is often the case for machine learning- and artificial intelligence-based algorithms used in digital health technologies and disclosure limitations of clinical algorithms regularly occur in peer-reviewed medical literature.[6] A final limitation was that the THINKMD technology had not been independently evaluated, although the validation testing approach was approved by institutional review boards and published in a peer-reviewed journal.[7] To ensure the quality of clinical assessments by mHealth platforms, more studies will be necessary to continue to validate and build the evidence base.

3 eHealth Africa, THINKMD, Kano State Ministry of Health. Community Health Worker (CHW) Adherence to the Integrated Management of Childhood Illness (IMCI) Protocol in Kano State, Nigeria before, during, and after use of the digital health platform MEDSINC. MEDSINC Implementation Field Report. 2018. https://www.ehealthafrica.org/medsinc

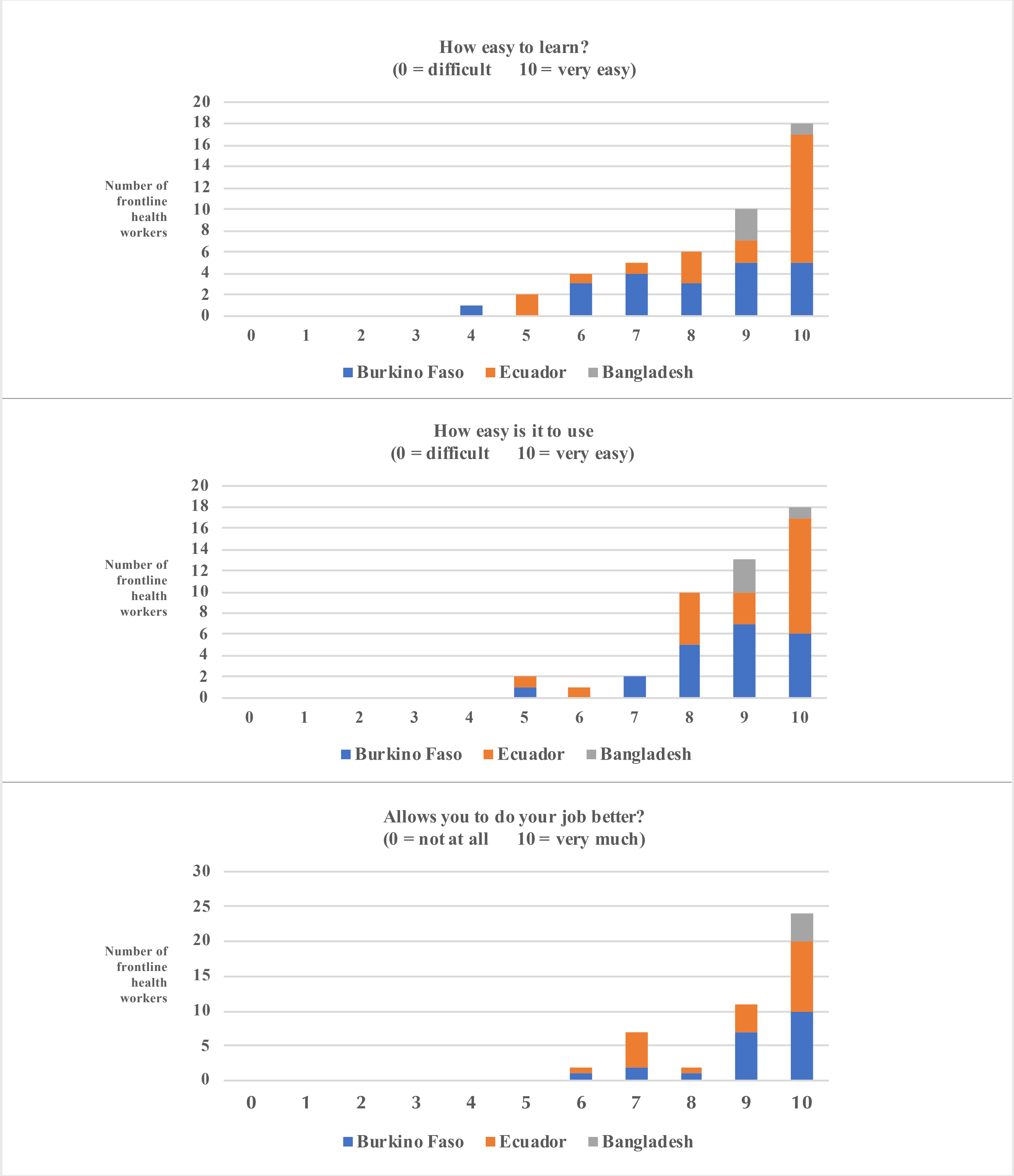
4 Finette BA, McLaughlin M, Scarpino SV, et al. Development and Initial Validation of a Frontline Health Worker mHealth Assessment Platform (MEDSINC®) for Children 2-60 Months of Age. Am J Trop Med Hyg. 2019;100(6):1556-1565. doi:10.4269/ajtmh.18-0869

5 Ansermino JM, Wiens MO, Kissoon N. Letter to the Editor: Evidence and Transparency are Needed to Develop a Frontline Health Worker mHealth Assessment Platform. Am J Trop Med Hyg. 101(4), 2019, p. 948. doi:10.4269/ajtmh.19-0411a

6 Finette BA, McLaughlin M, Scarpino SV, et al. Authors' Response. Am J Trop Med Hyg. 101(4), 2019, p 949. doi:10.4269/ajtmh.19-0411b

7 Finette BA, McLaughlin M, Scarpino SV, et al. Authors' Response. Am J Trop Med Hyg. 101(4), 2019, p 949. doi:10.4269/ajtmh.19-0411b

FIGURE 2. Usability and acceptability of the THINKMD technology among frontline workers



Scale and future health system applications

The data science assets in this use case include the THINKMD technology itself, the associated training approach, as well as the study design for the initial validation study of the technology and a field-based pilot that followed. The THINKMD technology could be expanded to other settings to support FLWs to improve their ability to appropriately diagnose and treat sick children. At the time of publication, the THINKMD technology has been used to assess 300,000 patients across more than ten countries.[8] For example, in Zambia, Healthy Learners[9] is training teachers to be school health workers as part of an integrated school health model. The THINKMD technology is currently used in 45 schools to increase quality of care and health outcomes of over 110,000 school-aged children, with plans to continue to scale the approach in partnership with the Ministry of Health and Ministry of General Education of the Government of Zambia.[10]

The data science assets related to the validation study and field-based pilot of the THINKMD technology could be adapted and used to inform the study design for other digital health tools. The data themselves could also be used for other applications, including disease surveillance and project evaluation. For instance, the platform gathers significant user- and program-specific information that could be continuously analyzed for project evaluation purposes or to complete key performance indicators. The data could also be used to illustrate or unearth health system barriers or challenges in program delivery. For example, one district's data showed an unusually high instance of malnutrition. The data generated by the tool enabled the project managers to quickly isolate which two clinics had high rates of malnutrition, and they were able to determine that the facilities' scales used to weigh children were broken.

The THINKMD technology can also be adapted for emergency response. In response to the COVID-19 pandemic, THINKMD has released a COVID-19 screening and educational tool. The tool guides users through a self-risk assessment to indicate whether they have symptoms associated with COVID-19. The algorithm is being dynamically updated based on the latest clinical data and guidelines. Based on individual results, the tool provides information on seeking medical care and recommended response measures. Additionally, the data can be aggregated for emergence and monitoring of the pandemic and building predictive disease models to inform population-level responses to the pandemic.[11]

8 https://thinkmd.org/

9 https://www.healthylearners.org/

10 THINKMD. "THINKMD Partners with Healthy Learners to Scale School-based Health Program in Zambia." 2020. https://thinkmd.org/thinkmd-partners-with-healthy-learners-to-scale-school-based-health-program-in-zambia/

11 https://thinkmd.org/covid-19/

Implementation considerations

User acceptability

A mobile health platform that is validated also needs to be useable by FLWs to reach scale. The challenge with many mobile health platforms is that the FLW is making subjective decisions that require specific skills to do correctly. Learning and maintaining those skills require training and supervision, which in turn require a significant amount of funds. The THINKMD platform seeks to guide FLWs through collecting the data, leading to more data collection for less effort.

Adaptability

The THINKMD technology can be configured to change the language, user interface, or underlying guidelines/treatment protocols.

Health system linkages

The THINKMD technology can be linked with the existing country health information system to ensure the captured data are not only used at the point-of-care for patient-specific diagnosis and treatment, but also to inform population health interventions. For example, if the data were aggregated and available to district health officials, they could use the information to better target resources to specific geographic areas based on disease burden. Although this functionality was not included in the Kano State pilot in Nigeria, THINKMD has partnered with Ona to integrate the functionality of THINKMD's FLW platform with Open Smart Register Platform (OpenSRP), an mHealth platform that connects FLWs to national health systems.

Cost

Potential implementers and program managers should consider the cost-benefit analysis of adopting THINKMD. There are cost requirements to acquire THINKMD and train cadres of healthcare workers in its operation. However, the platform is intentionally designed to enable sustainable scale over time. This would be the result of cost savings achieved through reductions in requirements for supervision and a less expensive training at the front end. For example, the evaluation of the pilot in Nigeria estimated a more than 50 percent reduction in training cost when the THINKMD technology training was implemented using a distance learning approach.[12,13]

12 eHealth Africa, THINKMD, Kano State Ministry of Health. Community Health Worker (CHW) Adherence to the Integrated Management of Childhood Illness (IMCI) Protocol in Kano State, Nigeria before, during, and after use of the digital health platform MEDSINC. MEDSINC Implementation Field Report. 2018. https://www.ehealthafrica.org/medsinc

13 https://www.businesswire.com/news/home/20180130005503/en

Data science assets

PRODUCT

THINKMD technology for data-driven diagnosis and treatment

Data for Management

Data Transformation

Data for Impact

Mobile data collection tool

Decision support tool

THINKMD developed a digital health point-of-care tool to support FLWs to assess, diagnose, and treat patients.

To use it, the FLWs gather a series of data points about the patient's past and current history, symptoms, vital signs, and physical examination findings.

The THINKMD technology relies on an underlying algorithm to weight these data points compared to clinical best practices and global guidelines to generate a clinical risk assessment and recommendations for triage, treatment, and follow-up. This back-end technology design uses clinical knowledge and practice data points to inform the technology's algorithm, which then learns from data generated by users of the technology platform. Essentially, the technology translates how physicians would analyze clinical data points and creates an algorithm to replicate that process. The THINKMD technology uses an ever-growing set of data to continuously improve and modify its core algorithm for diagnosis.

For more information

http://www.thinkmd.org/medsinc

Finette BA, McLaughlin M, Scarpino SV, et al. Development and initial validation of a frontline health worker mHealth assessment platform (MEDSINC®) for children 2 -60 months of age. The American Journal of Tropical Medicine and Hygiene. 2019;100(6):1556 -1565. doi:10.4269/ajtmh.18-0869

PROOF

Design of the field-based pilot of the THINKMD technology:

Study design, research protocols, and data collection tools

Data Collection

Data Transformation

Data for Impact

Research methods

Research protocols

The THINKMD suite of tools was evaluated through a two month pilot study among CHWs in Kano State, Nigeria.

The field-based pilot study was designed to understand the feasibility, acceptability, and usability of the THINKMD digital clinical assessment platform and how its use could improve adherence to IMCI clinical guidelines.

The study design, research protocol, and data collection tools were co-designed by eHealth Africa, THINKMD, and Kano State Primary Healthcare Management Board. The research protocol and study design included the research questions, an operational definition of IMCI clinical adherence, specifications on data collection processes, and sample selection and size. The data collection tools included usability and acceptability surveys as well as tools to capture observational data on CHW adherence to the IMCI guidelines.

The research protocol included the following research questions:

1. Do the CHWs find the THINKMD clinical technology usable? Do they find it an acceptable technology to perform clinical assessments and to treat sick children or to refer them to seek care with a trained clinician?
2. Does the MEDSINC platform improve the healthcare capacity and clinical assessment consistency of CHWs to accurately assess a child compared to currently used clinical assessment, triage, and treatment tools?
3. Does the THINKMD analytics and visualization platform improve programmatic and national-/state-level health monitoring, evaluation, and development compared to currently used information monitoring systems?
4. What is the potential health and economic impact of a THINKMD implementation based on assessments and interventions (preventative and curative) provided by CHWs?

For more information

eHealth Africa is a Nigerian nonprofit organization with the mission to "build stronger health systems through the design and implementation of data-driven solutions that respond to local needs and provide underserved communities with tools to lead healthier lives." https://www.ehealthafrica.org/medsinc

PROOF

Design of the validation study of the THINKMD technology:

Study design, research protocol, and data collection tools

Data Collection

Data Transformation

Data for Impact

Research methods

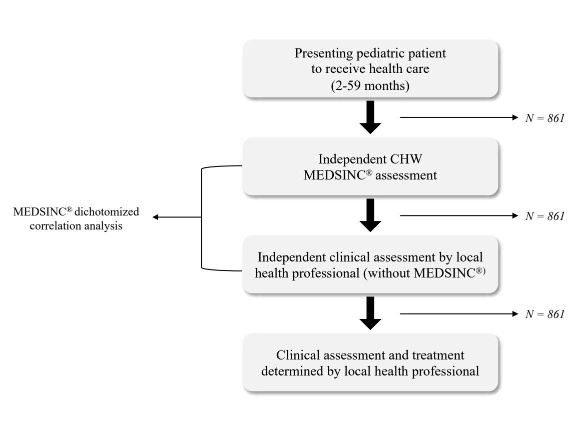
Research protocols

Validation testing was done to gain insights into the clinical logic accuracy of the THINKMD technology.

Field-based testing was conducted in Burkina Faso, Ecuador, and Bangladesh. The validation testing compared clinical assessments that were generated from FLWs using the THINKMD technology to clinical assessments by local health care professionals for the same children. UNICEF, local ministries of health, icddr,b and other local research organizations were engaged as co-leads of the study.

The validation study design specified the study sites, sample selection and enrollment process (Figure 1), and the statistical methods for analyzing results. The validation study also included questionnaire surveys to assess the usability of the THINKMD technology from stakeholders, including health care workers at national, regional, local, and community levels as well as child caregivers.

FIGURE 1. Validation study design and recruitment of subjects



For more information

Finette BA, McLaughlin M, Scarpino SV, et al. Development and initial validation of a frontline health worker mHealth assessment platform (MEDSINC®) for children 2 -60 months of age. The American Journal of Tropical Medicine and Hygiene. 2019;100(6):1556 -1565. https://doi.org/10.4269/ajtmh.18-0869.

PEOPLE

Training approach for use of the THINKMD technology

Methods and approaches for learning

To introduce end users to the THINKMD technology, a train-the-trainer approach was taken as part of the field-based testing in Burkina Faso, Educador, and Bangladesh.

This entailed initial training occurred of LHPs who then trained the FLWs in groups of two to four participants. Average traning time for each group was four to six hours. The training content was the same at each level (LHP and FLW) and included a presentation on the background, functionality, and full use of the THINKMD platform. The participants also worked through test cases, including practice on taking heart and respitory rates using a metronome, and using colleagues as example patients. For each site, there were training modifications based on available technology, number of participants, experience of participants, and other requests from the collaborators.

The THINKMD user interface also facilitates additional follow-on training in an interactive manner. The technology includes the ability to teach the user how to evaluate patients and how to use the tool itself thorugh visuals, GIFs, and videos that teach and reinforce how users should inform the data points collected through the input system. Users also can play these visual aids after they have added data to the platform, in order to ensure that they performed the diagnostic evaluation properly.

In Nigeria, the training-of-trainers model was conducted remotely due to financial constraints. This distance learning approach resulted in a more than 50 percent reduction in training cost. Of the CHWs trained in Nigeria, 93 percent said the THINKMD technology was very easy to learn and use. Additionally, 100 percent of CHWs said that the THINKMD technology was a valuable training tool.

For more information

Finette BA, McLaughlin M, Scarpino SV, et al. Development and initial validation of a frontline health worker mHealth assessment platform (MEDSINC®) for children 2 -60 months of age. The American Journal of Tropical Medicine and Hygiene. 2019;100(6):1556 -1565. https://doi.org/10.4269/ajtmh.18-0869.

eHealth Africa; THINKMD; State Ministry of Health, Kano State, Nigeria. Community Health Worker (CHW) Adherence to the Integrated Management of Childhood Illness (IMCI) Protocol in Kano State, Nigeria Before, During, and After Use of the Digital Health Platform MEDSINC. MEDSINC Implementation Field Report. [Publisher location: Publisher]; 2018. https://www.ehealthafrica.org/medsinc.

USE CASE

Data analytics for malaria elimination in Zambia

The challenge

Malaria cases persist in Zambia despite a focus on control and elimination. Health system managers do not have timely visibility into when and where cases occur (or are forecasted to occur) to inform where to allocate scarce prevention and treatment resources.

At the start of the VNM project in Zambia in 2014, health workers supporting the malaria control and elimination program faced several interrelated data challenges—a "problem stack." They were not able to pull as much data as they needed from the existing DHIS2 system. They were also unable to integrate data from other sources; for example, data on medical supply stocks could not be integrated with information on malaria cases. It was not possible to generate the data visualizations needed to answer the complicated analytical questions they needed to answer. From both a technological and public health perspective, the existing data system was insufficient to meet the goal of eliminating malaria-related deaths by 2030.

The solution

PATH, the Tableau Foundation, and the Government of Zambia launched a partnership in 2015 to integrate new tools and systems for data use to support the Zambian Ministry of Health in its effort to bring the malaria incidence rate down to zero.

The VNM platform drew on a novel approach to tackle these complex analytical questions, which require many different types of data and the tools to turn those data into insights and actionable intelligence; this approach can be applied to many other types of use cases. Operational dashboards were introduced to help district health personnel and frontline health workers track, treat, and report on the disease. Much of the work to inform and design the platform was done by health workers in Zambia to ensure the final system met the country's needs.

In order to scale its early successes, the VNM partnership expanded its technology solutions to support health workers in using data to understand where, when, and how to intervene to control and eliminate malaria. As of 2020, the partnership includes eight technology and service companies that support satellite imagery and location insights, automated workflow for data analytics, cloud-based communication, database solutions, and storage services (Figure 1).

Project dates
2015 to present

Location
Zambia

Partners

PATH
Tableau Foundation
Government of Zambia
Alteryx
DataBlick
Mapbox, Slalom
DigitalGlobe
Exasol
Twilio

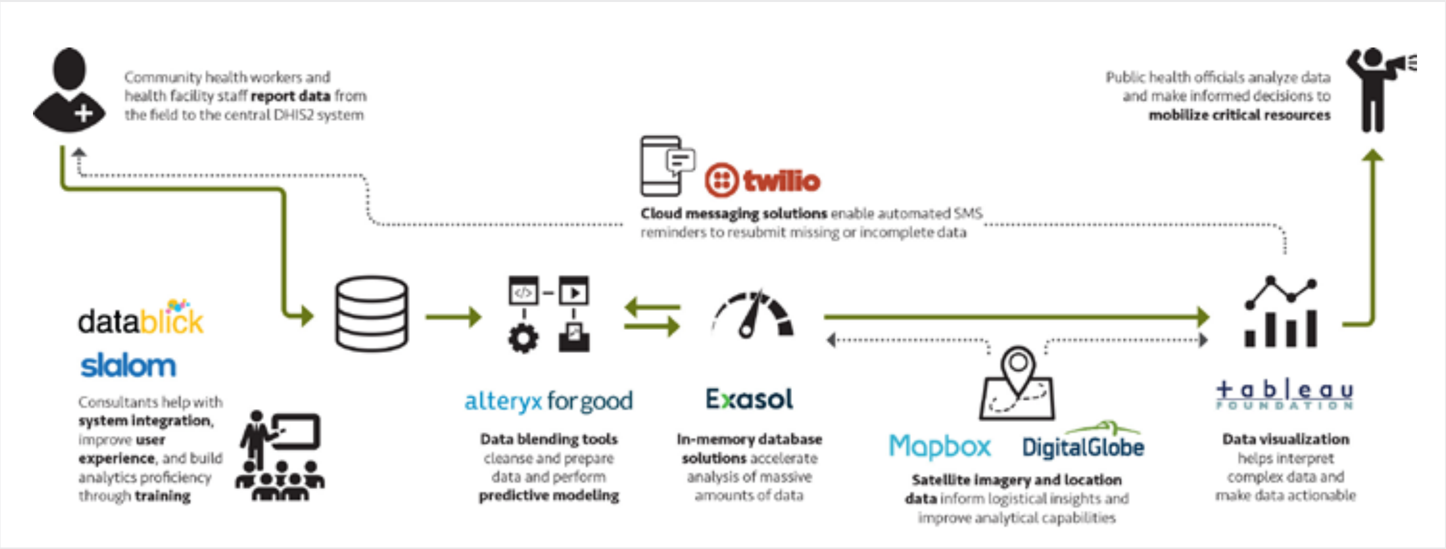
Further information

Visualize No Malaria website @ PATH

Visualize No Malaria – Data for Action Webinar

Using Technology to Advance Global Health: Proceedings of a Workshop. Country-level Digital Health Strategies

FIGURE 1. The Visualize No Malaria (VNM) "stack" of technology solutions and partners

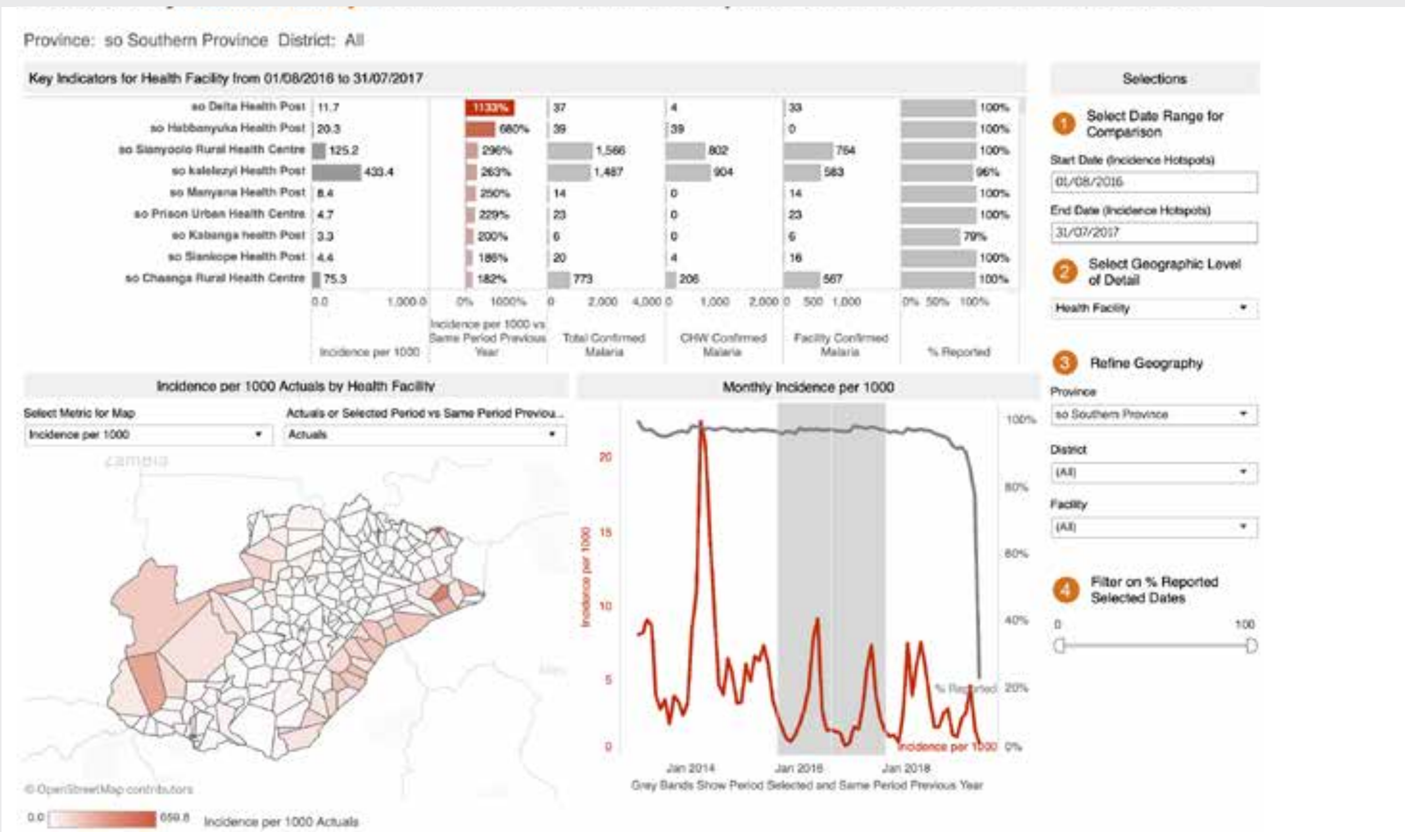


Impact

As part of the VNM project in Zambia, insights from the data visualizations informed how district health managers pursued gaps in surveillance information, deployed health workforce, and adjusted the supplies of malaria drugs and diagnostic tests across Southern Province and eventually across the nation. The VNM platform has contributed to increased malaria surveillance reporting rates and timeliness, improved reporting accuracy, and enabled more data-driven intervention targeting and response.[1] Armed with data literacy and new tools, more than 8,000 frontline health workers now provide high-quality community care and serve as the foundation of Zambia's malaria surveillance system. Users report that the platform has revolutionized the quality and reliability of data and reports and has sped up the accessibility of data and insights to health workers (Figure 2).

The VNM project—together with a complementary set of ongoing interventions, including vector control initiatives like indoor residual spraying and malaria bednet distribution, a mass drug administration program, and intensive CHW training in disease surveillance—contributed to an overall 92 percent decline in malaria-related deaths in Southern Province from 2014 to 2017.[2]

FIGURE 2. Visualize No Malaria incidence dashboard



1 https://www.path.org/articles/real-time-malaria-dashboards/

2 https://www.path.org/visualize-no-malaria/

Scale and future health system applications

Within Zambia, the technology stack used to support the VNM platform is already being adapted to address other health system challenges beyond malaria control and elimination. The platform has also been deployed to support malaria control in Senegal and Ethiopia and has been used to support essential health services in several other countries. In Tanzania, the Alteryx and Tableau assets were used to replace existing data preparation tools used by Tanzania's BID Initiative on childhood immunization. This change improved the data visualizations and significantly accelerated data processing time. Through the US President's Malaria Initiative VectorLink Project in Zambia, Mozambique, Madagascar, Ethiopia, and Mali, the full VNM stack is being used to improve use of malaria case data and improve the supply chain for indoor residual spraying projects.

Each time the VNM platform was adapted for a new environment, the time requirements for scale have significantly decreased, drastically reducing the start-up costs. This acceleration was made possible in part by introducing commercial technology solutions, but also by rigorously prioritizing the use of tools that were fast, were easy to use, and required a minimum of additional training. Further, the VNM platform is adaptable—it is designed for data, not for a particular disease area. Because significant work was done during the Zambia VNM project to test and revise the platform components, less work was required for subsequent applications.

Applications for emergency response

Several components, or data assets, that are part of the VNM platform were leveraged in the response to the Ebola outbreaks in 2014 and 2015 in Guinea, and in other parts of West Africa in 2016 and 2018.[3] In Guinea, the business intelligence asset from Tableau[4] was used in tandem with a CommCare module from Dimagi to devise a mobile-based patient and contact tracking and tracing system. This digital approach did not replace the existing paper-based system, but added the capability of real-time reporting through the use of smartphones. The assets could be applied to disease surveillance for any infectious disease, as well as many noncommunicable diseases, to provide greater intelligence to local, regional, and national program managers, ministry of health officials, and policymakers about financing and resource needs, to respond quickly to outbreaks, and to help spot small outbreaks before they become large epidemics.

From late 2019 to 2020, data assets that are components of the VNM stack have also been adapted to respond to the spread of COVID-19. As of this publication, Tableau's data visualization software and Mapbox mapping assets have been used to create the COVID-19 Data Hub,[5] which aggregates several vetted datasets and presents a visualization of the disease's spread in every country. The Alteryx data management asset has been used to estimate and forecast hospitalization rates[6] in US counties to predict health workforce and supply needs. Twilio's cloud communications platform is being used to boost telehealth capabilities.

3 https://solutionscenter.nethope.org/assets/collaterals/Glob_Health_Sci_Pract-2015-Sacks-GHSP-D-15-00207.pdf

4 http://flipthemedia.com/2015/11/tableau-foundation-building-better-world-data/

5 https://www.tableau.com/covid-19-coronavirus-data-resources

6 https://community.alteryx.com/t5/Alteryx-Use-Cases/Modeling-the-impact-of-COVID-19-on-healthcare-systems-with/ta-p/553437

Implementation considerations

Before adopting the health data science assets that are part of the VNM platform, potential users should consider several factors that may determine whether implementation will be successful.

Health system readiness

The health system should be ready to adopt a sophisticated technology stack; implementation will be faster and more successful if data are captured in electronic systems versus a paper-based system or spreadsheet software such as Microsoft Excel. The optimal environment to successfully implement the VNM assets would already have an existing data pipeline that automatically reports from multiple sources to a central data warehouse. In many countries, health information systems comprise many disparate data sources that require investment of time and skill to clean, combine, and consolidate the data to create dashboards.

Countries should consider data needs and data quality requirements based on high-priority analytical questions. Solutions can start by using the existing data instead of focusing on “perfect” data, as combining data sources and improving their quality can take significant time and resources to achieve. In the case of VNM in Zambia, the Malaria Control and Elimination Partnership in Africa (MACEPA) had worked with the Ministry of Health to develop and establish the community-based reporting tools, systems, and processes of the surveillance system and other necessary inputs to the platform.

When building strong data and reporting pipelines, it is important to remember health workforce requirements. The automatic and manual platform and component updates to run the VNM platform require certain skills, and that will affect what project staff are required to successfully use the platform. Potential users of the VNM platform should consider whether they can employ staff who are experienced with automated reporting systems on the implementation team. There may also be significant data preparation expertise required when the data quality is low, or if data sources are not linked and interoperable.

Additional considerations are required for data visualization, particularly when including maps of geospatial information. Two data characteristics have been important to the success of VNM: (1) whether the source data include coordinates and/or links to geospatial information, and (2) whether the administrative boundaries are up to date. These factors influence how effectively the mapping layers of computation, interactvity, and display will be set up and managed. Potential users of the VNM platform should consider whether their existing data provide datasets with these characteristics.

User-centered learning culture

VNM put the users—health workers in Zambia’s malaria control and elimination program—at the center of the system development process. The VNM team created a learning framework based on four principles (expose, explore, explain, and empower[7]) to create a data culture of learning and informed action. Additionally, VNM intentionally chose tools that were designed with end users in mind; that is, tools that could be adapted to the existing skill levels, fast implementation, and high usability. The data products (dashboards) themselves were also designed and created in partnership with end users. As described in the asset classification below, the user-centered design approach coupled with a willingness to pursue non-integrated data sources and move beyond existing operational systems were both important factors in the success of the VNM implementation in Zambia.

7 https://www.path.org/articles/inspiring-data-insight-four-guiding-principles/

Cost and capability

The VNM platform includes certain data science tools that are proprietary, or not open source, which may or may not fit the needs, policies, and resources of the user's health system. Modular systems like the VNM platform are very flexible and adaptable. However, they can be expensive, especially in terms of upfront costs like purchasing software licenses and training needs. They also require time for training and hiring. However, as a result of the VNM partnership, modular packages, reduced pricing, and licensing grants are now available, which can make the costs more acceptable. Cost must also be balanced with capability. Analytical questions often outpace the capabilities of operational systems used for routine data collection, and having the tools and human resources is necessary to achieve greater impact.

When it comes to data visualization, it may not make sense for a national program to use a software platform like VNM if the goal is to serve a simpler purpose, such as to generate traditional reports. VNM's use of Tableau would be an expensive option if the software is not going to be used to its full capability.

Partnerships and support

The VNM project was supported by many partners with technical expertise who were able to contribute in-kind donations of staff time to configure the required platforms. The use of volunteers to support the Zambia VNM platform was an integral part of the user-centered design process. Particularly with modular systems that involve several different technology owners, the relationships between key stakeholders are also critical to the success of such large-scale platforms. For VNM, trust between the private sector, NGO partners, and the Government of Zambia was an important facilitator of the project's success.

Data science assets

PRODUCT

PROCESS

Alteryx automated data workflows and data management

Data Collection

Data Transformation

Data for Impact

Automated workflow

Data blending tools

Cloud-based data services

Alteryx provides tools that automate workflows to quickly and efficiently process thousands of laboratory results, notify health workers, and track interventions.

The software allows users to rapidly curate, process, and archive large amounts of data to enable near real-time understanding of insights and trends. Instead of waiting months to translate data into action, users can process results daily with automated outputs sent to the end user to inform timely decision-making. Alteryx's data-blending tools lend additional capability to users to cleanse and prepare data and perform predictive modeling.

In Zambia, the Alteryx automated workflow integrates geospatial data and case surveillance data in a data pipeline that informs the data visualization dashboards used by district-level health managers. The data are sourced from the DHIS2 server and local data sets and are integrated using a cloud-based interface. The automated workflows enable faster processing of lab results. This means that positive malaria cases are communicated to health workers on a daily basis instead of monthly; more rapid evaluation, validation, analysis, and action based on survey data as it is collected; and weekly summaries of predicted malaria cases.

For more information

Alteryx is a software company headquartered in Irvine, California, with a development center in Broomfield, Colorado. The company's products are used for data blending and advanced data analytics. Alteryx has a stated goal of enabling advanced analytics to be performed by non-specialists. https://www.alteryx.com/

PATH + Alteryx Join Forces to Fight Malaria in Africa. Alteryx Community articles. June 19, 2019. https://community.alteryx.com/t5/Alteryx-Use-Cases/PATH-Alteryx-Join-Forces-to-Fight-Malaria-in-Africa/ta-p/182839

PRODUCT | PEOPLE

Mapbox satellite imagery and location data

Data Collection | Data Transformation

Mapbox provides mapping resources, including satellite imagery and location data. It recruits volunteers around the world to crowdsource data about human settlements and structures that can be overlaid with existing data on infrastructure, waterways, and weather.

In Zambia, the location data were combined with surveillance data to create algorithms that predict where malaria cases emerge. These predictive models enable community health workers to track and treat incidents faster, meaning new cases can be prevented and lives can be saved. Maps and satellite imagery allowed the project team to create more accurate geospatial-based dashboards to visualize case distribution and clustering, and to identify issues with underlying data on the locations of facilities and CHWs.

Location data

Satellite imagery

Crowdsourcing input approach

For more information

Mapbox is the location data platform for mobile and web applications. Mapbox's apps reach more than 300 million people each month. https://www.mapbox.com/

"Supporting malaria elimination for 60 million people" [blog post]. Mapbox Blog. April 25, 2019. https://blog.mapbox.com/supporting-malaria-elimination-for-60-million-people-96053bbd8a0a

PRODUCT | PROCESS

Supplemental data from nonroutine systems

Data Collection

To power the VNM analytics platform, surveillance data are harvested from DHIS2. However, additional datasets are also necessary, such as administrative boundaries, health facility data, additional population data, satellite imagery, and crosswalks and manually generated extension datasets that provide merged and cleansed data for use within automated processes.

Obtaining these data can be a complex and time-consuming effort due to local policies; where there have not been resources put toward data governance, additional effort is required to cleanse and link data so it can be useful.

Data

For more information

Supporting malaria elimination for 60 million people [blog post]. Mapbox Blog. April 25, 2019. https://blog.mapbox.com/supporting-malaria-elimination-for-60-million-people-96053bbd8a0a

PRODUCT

PEOPLE

Twilio cloud communication platform

Data Collection

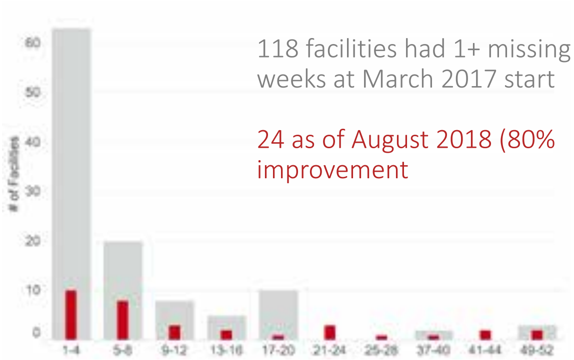
Automated SMS communication

Cloud-based data services

Health worker communication tools

Twilio provides a cloud communication platform that allows users to exchange valuable insights and enables automated collection of missing and incomplete data from CHWs via text message.

In Zambia, use of Twilio to communicate with health workers via text resulted in a rapid increase in reporting completeness and timeliness, including back-entry of missing data up to two years prior. Through the VNM partnership, Twilio created a custom cloud-based feedback loop that uses SMS to share data back with frontline health workers. This application has the capability to evolve to include voice and video in addition to text messaging. The bar chart below shows the decrease in the number of facilities with missing reports over the timeline of the implementation of the SMS feedback loop in Southern Province, Zambia.



Viz Alerts, theTwilio communications platform, is an open source Python script designed to identify data gaps, alert the local health workers of those gaps, and integrate missing data when provided by health workers. When the Viz Alerts algorithm detects an incomplete dataset, the script prompts an SMS message via Twilio to the appropriate health worker to submit the missing report; the health worker can quickly reply by submitting the missing data over SMS. Viz Alerts incorporates the missing data automatically, and users can access the updated dataset through Tableau visualizations.

From Marianne Soko, the malaria focal point person in Itezhi-Tezhi district, Zambia: “This is really great because the alert allows me to follow up on the missing reports, and they can be updated both at facility and community level.”

For more information

Twilio powers the future of business communications. Enabling phones, VoIP, and messaging to be embedded into web, desktop, and mobile software. https://www.twilio.com/

Talbot J. PATH uses SMS data alerts to help end malaria in Zambia [blog post]. Twilio Blog. June 14, 2017. https://www.twilio.com/blog/2017/06/path-uses-sms-data-alerts-to-help-end-malaria-in-zambia.html.

PROCESS | POLICY

Data Transformation
Data for Impact

Predictive modeling algorithms

Custom algorithms

Predictive models

Predictive modeling uses data and stastical methods to predict or forecast outcomes.

Alteryx software is capable of leveraging many different kinds of data to inform these predictions. In the case of malaria, these data could include population mobility, case surveillance, geospatial data, meterological or climate data, and others to create algorithms that inform predictions of where malaria cases may emerge.

Thus far, predictive modeling has not been fully used in support of the VNM platform in Zambia and its application is at the proof-of-concept stage. However, even as Zambia has seen significant progress in its fight against malaria, it will still be challenging to address some of the persistent drivers of the epidemic. Continued innovation and new tools to are essential to reach elimination. For example, malaria parasites that reside in asymptomatic individuals represent one of these epidemic drivers. In the future, predictive models may enable CHWs to determine who might be carrying the malaria parasite and to track and treat cases faster, meaning new cases can be prevented and lives can be saved.

For more information

Walker A. PATH and Tableau Foundation #VisualizeNoMalaria Project: Development of an Automated Predictive and Forecasting Malaria Cases Capability. Seattle: Tableau; [YEAR]. https://www.tableau.com/sites/default/files/whitepapers/path_and_tableau.pdf.

PRODUCT | PEOPLE

Data Collection
Data Transformation
Data for Impact

User-centered design approach

Data visualization

Collaborative design

Data use tools

User-centered design is an iterative process that focuses on the end users and their needs to inform the design process of a tool or solution.

For the VNM platform in Zambia, this methodology focused on understanding who the tools are meant to serve, for what functions, and in service of what goals. This approach also supports the creation or strengthening of a data use culture.

To develop the integrated platform used for VNM, particularly the dashboards used for data visualization and their backend analytics, PATH used a design approach that centered on the users in Zambia at the national, district, and facility levels. PATH and other VNM partners held workshops with local health officials to determine system requirements and data use goals. The resulting platform was iteratively designed to meet the local needs of Zambia's health system.

For more information

PATH is an international health organization driving transformative innovation to save lives. Working in partnership with national governments, PATH is leading the way toward a malaria-free world by focusing on new vaccines, treatments, diagnostics, and approaches. https://www.path.org/visualize-no-malaria/

PRODUCT | PROCESS | PEOPLE

Tableau data visualization

Data Transformation

Data visualization

Data insights

User-centered/ People-centered design

Tableau provides a powerful data visualization engine that allows end users to create user-friendly dashboards.

In Zambia, Tableau dashboards were used to analyze the malaria situation in real time. Expert users from DataBlick and Slalom traveled to Zambia and worked closely with district health workers and the Ministry of Health to iteratively develop the dashboards. Other assets from the VNM partnership were able to connect into the data pipeline that feeds the visualizations created using Tableau software.

Before the VNM partnership begn, district health managers generated simple spreadsheets of data in a time-consuming process that could only occur quarterly or annually. Now, with the capabilities of the dashboards, health workers are able to have daily and weekly information at their fingertips to enable them to make decisions on health workforce and supply allocation, track down missing information, and control local outbreaks. Visual data preparation provided a faster method for system design than writing code and also served as a way to work around the lack of data warehouses. In Senegal, local staff were able to quickly learn from the Zambia experience and develop dashboards tailored to local needs.

The designs of a number of the dashboards built as part of VNM are sufficiently useful that they have been re-used for the US President's Malaria Initiative's Malaria Data Integration and Visualization for Eradication (M-DIVE) project that is providing malaria data to countries supported by the US President's Malaria Initiative.

For more information

Tableau Software is a software company headquartered in Seattle, Washington, which produces interactive data visualization and analytics products. Tableau's mission is to help people see and understand their data. https://www.tableau.com/

Myrick, Neal. "Visualize No Malaria crosses new frontiers in fight against disease." [blog post]. Tableau Blog. April 25,2019. https://www.tableau.com/about/blog/2019/4/visualize-no-malaria-crosses-new-frontiers-fight-against-disease.

PEOPLE | PROCESS

DataBlick and Slalom analytics training

Data Transformation

Analytics training materials

Data use culture approach

Volunteer engagement

Analytics training for end users can build their capacity to understand and interpret their data, derive meaning from data visualizations, and understand how to use data insights to drive decisions and actions

In Zambia, volunteer expert users from DataBlick and Slalom helped to build analytics proficiency through training programs for users in Zambia. DataBlick also built a capability to issue reports to district-level managers who need to know when health facilities are not reporting data into the surveillance system. The goal of this effort is to enable the 100+ district-level health information officers and malaria focal point staff to manage and use data effectively in the fight against malaria. Training materials were developed with two aims: to build local capacity to use the data visualization dashboards built through the VNM partnership and to improve overall data literacy.

For more information

DataBlick is a boutique consulting firm that transforms data into actionable art. https://www.datablick.com/

Slalom is a modern consulting firm focused on strategy, technology, and business transformation. https://www.slalom.com/home

PRODUCT

Exasol analytics database

Data Collection

Database management system

In-memory analytics

Exasol is a high-performance in-memory analytics database that can work with data on premises, in the cloud, or both.

The architecture enables users to perform advanced analytical tasks on large volumes of data within the database itself. It can analyze large volumes of data in real-time to accelerate data insights and reporting.

In Zambia, Exasol is used alongside Alteryx to analyze data from the DHIS2 combined with satellite imagery data. Exasol's in-memory analytics database compresses massive amounts of data and serves them up fast, which is essential in low-resource, lower-bandwidth settings like Zambia. An Exasol database was developed to store the input variables and the output data. It connects both to Alteryx as an in-memory data source for the automated workflow and as a live data source to Tableau for the data visualization dashboards. .

For more information

Exasol is an analytic database management software company. Its product EXASolution is an in-memory, column-oriented, relational database management system. https://www.exasol.com/en/products/database-architecture/

PRODUCT

DHIS2 health management information system

Data Collection

Database

Data management

District Health Information Software 2 (DHIS2) is an open source, web-based health management information system platform.

The core DHIS2 software development is managed by the Health Information Systems Programme (HISP) at the University of Oslo. HISP is a global network composed of 11 in-country and regional organizations, providing day-in, day-out direct support to ministries of health and local implementers of DHIS2.

Zambia uses DHIS2 for data collection from health facilities countrywide and as a mobile reporting tool for the country's Malaria Control Program. CHWs and health facility staff report data from the field to the central DHIS2 system. In Zambia, DHIS2 has been integrated directly with Tableau, which allows for embedding Tableau visualizations within DHIS2 and directly pulling data from DHIS2 into Tableau.

For more information

https://www.dhis2.org/

Digital Square brings partners together to improve how the global community designs, uses, and pays for digital health tools and approaches. By strengthening the coordination among government officials, technological innovators, donor and implementation partners, and others across borders and boundaries, Digital Square reorients the market to better match tools and approaches to the needs of countries and communities. Digital Square is an initiative housed at PATH and is funded by the United States Agency for International Development, the Bill & Melinda Gates Foundation, and a consortium of other donors.

This publication was prepared by PATH staff and funded by The Rockefeller Foundation.Health Data Science Exchange Advisory Group members participated in an individual capacity, and their participation does not confer organizational endorsement of this work.

digital square