Policy Research Working Paper 10407

A Framework for the Economic Evaluation of Digital Health Interventions

Thomas Wilkinson
Mengxiao Wang
Jed Friedman
Marelize Görgens

WORLD BANK GROUP
Health, Nutrition and Population Global Practice &
Development Research Group
April 2023

Policy Research Working Paper 10407

Abstract

Numerous digital health interventions have been piloted in response to the health care challenges low- and middle-income countries face. Because the opportunity cost of investing in digital health interventions can be large, countries must make choices about which interventions to scale up. To make good investment decisions about digital health interventions, there is a need to define and establish their value, and to use economic evaluation to make informed decisions, however DHIs present methodological challenges for economic evaluation. To address these challenges, this paper first creates a 'gap map' of digital health intervention evidence which reveals a dearth of economic evaluation evidence about digital health interventions; this lack can limit decisions on policy, programming, and appropriate scale-up of digital health interventions. To advance work in this field, this paper then develops an economic evaluation framework that can be used when determining the economic value of digital health interventions. Such a standardized approach, alongside guidance to assist the conduct and use of economic evidence, can improve decision making and investments in DHI under constrained health budgets. The resulting digital health intervention economic evaluation framework consists of 5 steps: (1) determine the context, (2) determine the intervention type, (3) establish the level of complexity, (4) set the analytic principles, and (5) represent the value proposition. Users of the framework should attempt to adhere to its steps and principles, but where this is not feasible or appropriate, they should provide justification for the methodological choice. The framework should facilitate methodological transparency, thereby improving the overall usefulness of economic evaluations of digital health interventions.

This paper is a product of the Health, Nutrition and Population Global Practice and the Development Research Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at mwang8@worldbank.org, jfriedman@worldbank.org and mgorgens@worldbank.org.

A Framework for the Economic Evaluation of Digital Health Interventions

Thomas Wilkinson
Mengxiao Wang
Jed Friedman
Marelize Görgens

Keywords: digital health interventions; economic evaluation; low- and middle- income countries.

JEL classifications: I10, I15, O33

This paper is a product of the Digital Health Flagship Program at the Health, Nutrition and Population Global Practice and the Development Research Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at mwang8@worldbank.org, jfriedman@worldbank.org and mgorgens@worldbank.org.

The writing team thanks the insightful contributions of those who attended the Technical Working Group meetings in July 2019 and October 2020 (in person and virtually) and/or who participated in subsequent consultation and discussion on development of the framework. We also acknowledge the strategic advice of David Wilson, Program Director, Health Nutrition and Population Global Practice, World Bank.

---

1. Introduction

1.1 Background of the Digital Health Intervention Economic Evaluation Framework

The number of digital health interventions (DHIs) is growing rapidly, with the promise of enormous benefits to health systems and the people who use them. While DHIs have the potential to fundamentally alter the practice of health care and wellness, the opportunity cost of investments is also large and often sunk. This is particularly relevant for developing country contexts where the potential gains from harnessing digital technology can be transformative, bypassing existing infrastructure and process bottlenecks to address substantial unmet health needs. However, the opportunity costs of investing in DHIs that are simply ineffective or impractical, or whose costs outweigh the benefits that could have been achieved had the money been invested elsewhere, is particularly acute in resource-constrained contexts.

One of the most effective tools to generate information to inform health policy decision making is economic evaluation – the comparative analysis of one or more interventions in terms of costs and consequences [1]. Economic evaluation can be used to reflect the local value of DHI investments, facilitating adoption and efficient use of limited resources, realizing the promise of DHIs to rapidly improve global health. However, economic evaluation applied inconsistently or only partially can hinder the decision-making process and even lead to sub-optimal decisions. An economic evaluation framework will provide guidance and structure to facilitate the practice of economic evaluation to optimize its use in good decision making.

1.2 Objective of the DHI Economic Evaluation Framework

The framework aims to provide a consistent mechanism for representing the value of DHIs in context while adhering to the central concept that an economic evaluation is a comparative analysis of alternative courses of action in terms of both their costs and consequences.

It achieves this by recommending a standardized and coherent methodological approach that can be used prospectively or retrospectively by a range of users in the planning, conduct and reporting of economic evaluations of DHIs. The framework builds substantially on existing methodological improvement initiatives in the economic evaluation field [2] [3] [4] and draws on methodological advancements in evidence generation, synthesis and evaluation in digital health [5] [6] [7]. It is not constrained to a singular analytic approach or decision-making construct but utilizes elements of established methodology to accurately represent costs and consequences in context, reflect the nature of the intervention, and complexity of the analytic question. In this way it finds a pragmatic balance between consistency, validity and comprehensiveness and seeks to improve the evidence available for decision making while accommodating the unique attributes of DHIs.

Because cost-effectiveness analysis (CEA) is a predominant method of economic evaluation used to inform investments in health technologies, this framework focuses on adapting it for use in evaluating DHIs. The framework also discusses benefit-cost analysis (BCA), which is more frequently used when evaluating policies implemented in other sectors, including public health interventions implemented outside the health care system.

The framework does not attempt to provide methods to advocate for DHIs over other potential health system interventions. It seeks objective assessment and reflection of all relevant costs and consequences and their distribution, including the opportunity cost of spending of limited health resources. The framework aims to be used to assist in generating economic evidence to improve health in a digital world [8] rather than viewing DHIs as isolated health system investments.

The framework is targeted towards DHIs as defined by the World Health Organization's (WHO) classification of DHIs [9], with additional consideration for economic evaluation of DHIs that are enabled by predictive analytics, and non-digital health interventions that are enabled by predictive analytics as part of a wholistic intervention. The framework is also intended to be used within the larger guidance and technical support ecosystem for investments in digital health, represented in the Digital Implementation Investment Guide (DIIG), a global initiative coordinated by the WHO to outline the steps for country investment in digital health.

1.3 Problems That the DHI Economic Evaluation Framework Addresses

While the development and implementation of DHIs is rapidly expanding, the generation of economic evidence for DHIs is currently disparate, inconsistent and of variable quality, particularly for DHIs containing AI components. The framework will signal to those generating and synthesizing evidence the nature and type of evidence that will be useful for informing purchasing and investment decisions for DHI.

Some DHIs in use or being developed may have demonstrated and measurable outputs but limited meaningful impact to the wider health system or people. Conversely, DHIs may offer extensive indirect or wider health or societal benefits but do not deliver immediate outcomes in terms of health or process improvements that may be expected in a traditional economic evaluation. The framework will provide a signal of the components of value that are important for those investing as well as the intended beneficiaries from DHI implementation.

There is a critical need for DHI development to meaningfully involve the patients and communities they impact, particularly related to applications of AI [10]. A standardized framework has the potential to facilitate the democratization of DHIs by enabling a comprehensive understanding of value and the methods required for representation.

DHIs are frequently complex and have the potential to disrupt existing processes and pathways of care. This creates challenges in terms of common evaluation principles such as identification of comparators and identifying intended and unintended consequences. The framework provides a mechanism for identifying complexity and appropriately tailoring method choice, enabling context specific analysis that seeks to appropriately represent the uncertainty that complexity creates.

2. Scope: Digital Health Interventions and Artificial Intelligence

2.1 What Are Digital Health Interventions?

DHIs are a broad and diverse grouping. The World Health Organization adopts a user-focused categorization of DHIs[1] into interventions [9] for:

- Clients, e.g., text message targeted alerts or medication reminders to patients
- Health care providers, e.g., clinical decision support or diagnostic assistive interventions
- Health system managers, e.g., digital interventions to monitor or organize health care providers
- Data services, e.g., digital interventions for data synthesis and visualization or to parse unstructured data into structured data.

These categories incorporate interventions whose intermediate and final outcomes are likely to be health user experience related (such as improved medication adherence leading to better treatment outcomes) and outcomes that relate to systems efficiency (such as an improved clinical coding data management system leading to reduced administration costs for a provider).

To meet the range of available DHIs, the scope of this framework must also incorporate technologies enabled by artificial Intelligence (AI) processes and systems. As an emerging technology and branch of computer science, AI has the potential to exploit data, extract useful information, provide predictive capabilities and support evidence-based decision making to optimize health system performance [11]. Although no common definition of AI has been universally accepted, AI holds a natural character of being able to make machines utilize information to act intelligently. This report adopts the understanding as proposed by others [10] [12] that something "acts intelligently" when 1) what it does is appropriate for its circumstances and its goals; 2) is flexible to a changing environment and changing goals; and 3) learns from experience and makes appropriate choices given its perceptual and computational limitations. The current applications of AI in health are largely limited to "narrow intelligence", which is the use of data for a specific task, rather than generalized intelligence, which is human-like behavior across a range of activities. Agrawal, Gans, Goldfarb [13] describe an AI application in terms of the anatomy of a task (figure 1), where the existing human approach to a task involves intake of information, with prediction and judgement acting together to result in an action and outcome and an associated information feedback loop where a human learns from each task undertaken. Conceptualization of narrow AI involves separating the prediction element of a specific task while maintaining human judgment and action, ideally resulting in an improved outcome. The use of AI to enhance prediction within a human-operated task is underpinned by the availability and flow of data creating unique considerations for the treatment and role of evidence, with a distinction made between input and feedback data and the initial training data used to develop the AI prediction functionality.

[1] The WHO Classification of Digital Health Interventions is currently being updated (in 2023). Once the revised version is available, this paragraph will be updated with the revised classification.

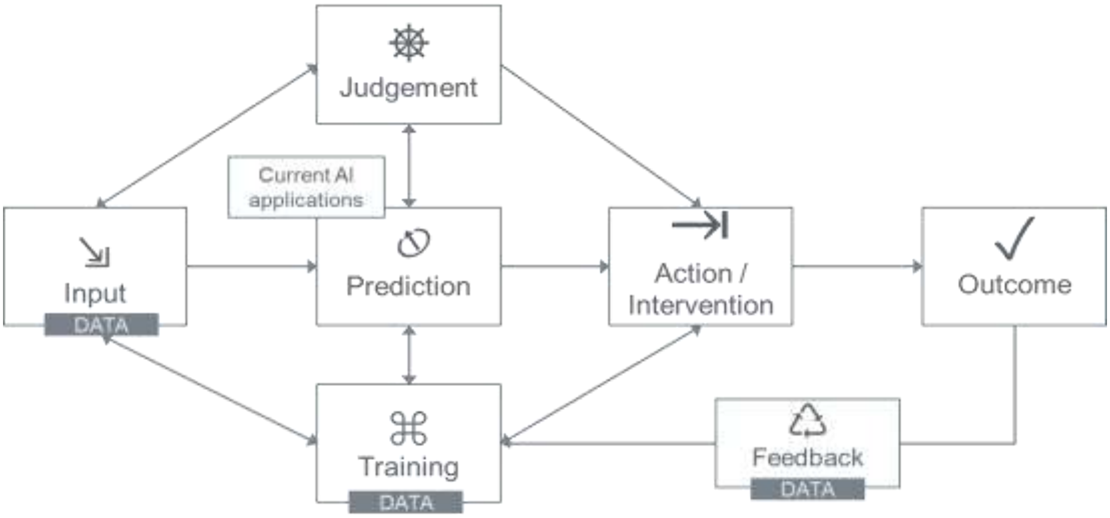

Figure 1 AI and the Anatomy of a Task

Source: Adapted from Agrawal, Gans, Goldfarb. 2018 [13] .

The categories of AI applications in health are described in table 1 below. As the digital health and AI field is constantly evolving, it is not intended to be an exhaustive categorization but a representation of common types of AI.

Table 1 Types of Artificial Intelligence Application in Health

TYPE OF AI	DESCRIPTION
Expert systems (or knowledge-based system)	Applying expert level competence in solving specific problems. It is often based on a series of complex rules (e.g., 'if-then' statements). Advanced by the development of fuzzy logic—a set of mathematical principles for knowledge representation based on probability and uncertainty.
Machine learning	A method for automating data analysis by using algorithms that iteratively identify patterns in data and learn from them. Machine learning applications are generally classified into three broad categories: (1) supervised learning, (2) unsupervised learning and (3) reinforcement learning. Supervised learning uses patterns already identified in data (i.e., training data).
Natural language processing	Determine the meaning of text by using algorithms that allow machines to identify key words and phrases in natural language corpora (ie, unstructured written text). Topic modelling is an approach to NLP that seeks to automatically identify the topics covered in documents by inferring relationships among prominently featured words.
Automated planning and scheduling (or AI planning)	Focused on organizing and prioritizing the activities and managing complex interdependent constraints to achieve a desired goal.
Image and signal processing	Processing large amounts of data from images and signals (i.e., information about the attributes of a particular physical phenomenon). Steps in image and signal processing algorithms typically include signal feature analysis and data classification using tools such as artificial neural networks (ANNs).

Source: Adapted from Wahl 2018 [10]

The advancement of digital health and the large volume of data being generated through digitizing health information and development in mobile health applications has promoted AI as an enabler of health in a digitized world [10].

However, AI technologies may face the same challenges and barriers of implementation and scaling up as with the broader digital health solutions in addition to unique risks and trade-offs. A USAID report on AI in global health identified AI use cases with high potential for impact on global health and current challenges to accelerating AI investment. The report established that all efforts to drive AI forward must be aligned with the general principles and best practices for digital health solutions [14]. The intersection between the user-classification of DHIs and application of AI in relation to the scope of this initiative is described in the next section.

2.2 Applying the Framework to the DHI Landscape

Incorporating the WHO user focused DHI classification and AI functionality results in three main categories of intervention within the DHI economic evaluation framework (in figure 2):

1. Digital Health Interventions without AI enabling functionality – e.g., simple text messaging patient reminders (area 1)
2. Digital Health Interventions with AI enabling functionality – e.g., machine learning radiology diagnostics (area 2)
3. Non-digital health interventions enabled by the use of AI technology – e.g., use of demographic and clinical information for identification of at-risk patients for community health worker deployment (area 3).

These categories are nested within the digital health system context that consists of information systems and the digital health architecture. An important element that is captured within this framework is the traditional "non digital" (or analogue) interventions that are enabled by predictive analytics and may require specific methodological approaches beyond those used traditionally in economic evaluation. Generalized AI development (e.g., AI analytical teams/capacity development within research institutions or development of "investment cases" for AI) without specifically linked health interventions will be outside the scope of the framework.

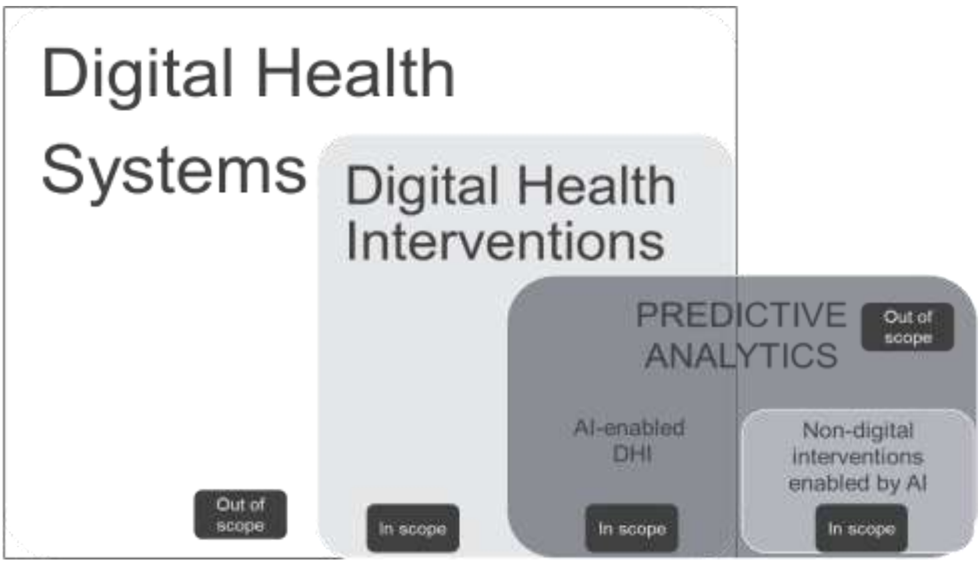

Figure 2 Representation of the Scope of the DHI Economic Evaluation Framework

Source: Authors, World Bank.

2.3 Usefulness to Whom? The Intended Audience

A fundamental aim of the framework is that it should be "useful" to those planning, conducting, and using economic evaluations of DHIs. There would be limited merit in attempting to develop methodology for only some stakeholders related to economic evaluation of DHIs as there needs to be a consistent understanding and representation through the chain of evidence generation and use.

As a World Bank product, the framework has an immediate intended use for country governments supported by World Bank country and global teams. The broader target audience however can be defined into three major groupings:

- Investor: aiming to invest in the highest value interventions
- Product developer: aiming to develop high-value interventions
- Researcher: aiming to analytically assess and represent value.

Investors are those who are determining whether to fund or reimburse a digital health intervention in a health system. This includes country governments, development partners, and private providers. This category also includes those supporting purchasing decisions in digital health, such as development partners providing technical assistance.

Product developers are those who are producing or creating DHIs. This includes developers in private and public research institutions and start-ups, reflecting the multiple different DHI types and use cases. This also incorporates those who are directly funding development of DHIs, noting that it is particularly important that the representation of value at the product development stage should align closely to the representation of value at implementation.

Researchers and analysts are those who conduct economic evaluations of DHIs within research institutes or other public and private organizations including consultancies. This category also includes research funding organizations that may be providing funding for economic evaluations of DHIs, and researchers who may be using economic evaluations in associated research in digital health such as policy analysis or systematic reviews of the literature. This is a particularly important group as it will undertake the application and production of economic evidence and develop methodological improvements and innovation, building on the research agenda promoted by the framework.

3. A Review of the Literature: Evaluating Digital Health Interventions

To determine the utility of an Economic Evaluation framework for DHIs, it is necessary to have a good understanding of the existing evidence base. The World Bank and 3ie developed an evidence gap map (EGM) to investigate the available published literature on impact evaluations of DHIs [15]. The broad goal of this EGM was to facilitate decision-making regarding investment in research to assess the effects and costs of DHIs and AI globally. A more specific objective of the EGM was to support the development of the economic evaluation framework for DHIs, and key findings are summarized below.

3.1 Main Findings and Gaps

For the EGM, 63,014 records were retrieved, and 17,142 duplicates were excluded. Most of the records (n=40,374) were excluded during the title and abstract screening. After the full-text screening, the EGM included 632 Impact Evaluations (IEs) (436 completed and 196 ongoing) and 97 Systematic Reviews (SRs) (89 completed and 8 ongoing).

The EGM reveals two major types of evidence gaps in the literature: 'absolute' gaps, where no or few primary studies have been conducted, and 'synthesis' gaps, where no SR exists despite a cluster of IEs or the quality of the SRs is substandard or the existing SRs are dated. The major gaps are presented below.

3.1.1 The geographic and economic context

The evidence base is skewed towards high-income and Western countries with more limited evidence from low- and middle-income countries (LMICs). In terms of location, there is limited evidence from the WHO regions of Africa, South-East Asia, and Eastern Mediterranean as most of the studies (66%) were conducted in the two regions of America and Europe. Similarly, very few studies were conducted in LMICs and there exists a huge gap in the volume of studies conducted in rich and poor countries. For instance, over three-quarters of the studies were conducted in high-income countries, with the United States alone contributing one-third of the included IE studies.

3.1.2 The intervention coverage

The intervention framework comprises four DHI categories divided further into 28 specific types of DHIs. Across the four intervention categories, almost all studies evaluated interventions for clients and health care providers. Significant evidence gaps were found for the other two categories, which were together covered by five studies only. See Figure 3 for a survey of coverage.

Notwithstanding the large volume of IEs covering clients and providers' DHIs, the evidence base within these intervention categories is unevenly distributed. For example, there is a limited literature on DHIs that empower clients to take the lead on health issues affecting them through, for example, citizen reporting or client-client communication. The focus for most studies is largely on top-down or externally led interventions such as targeted communication interventions. For the health care providers, there is a heavy focus on DHIs concerned with service delivery (e.g., telemedicine or decision support). Interventions that support providers in planning or coordination of health services (e.g., referral or activity scheduling) have received relatively low attention. There is limited evidence of DHIs for health system or resource managers and data services.

3.1.3 Frequency of AI use

Despite the potential role of AI in strengthening health systems and improving the quality of health care, a shallow evidence base was found in the published literature for AI. The number of studies that evaluated AI-powered interventions is extremely low. In total, only 13% (n=83) of the studies covered interventions that incorporated AI. Possible reasons include slow adoption of AI in the health sector, limited use of peer-reviewed publications as a mechanism of measuring impact, or a long and rigorous vetting processes that may be required in the medical field. This could also be due to the inclusion criteria of the EGM, which was focused on IEs that utilize counterfactual analysis. This may not be a common approach in AI-related evaluations.

3.1.4 Health domains

Not many studies covered critical global health domains such as maternal health, nutrition, HIV/AIDS, and family planning/reproductive health. No more than 45 studies covered each of these domains compared to the 343 studies that covered NCDs. While the majority of IEs were concerned with NCDs, there is also a substantial focus on mental health, infectious diseases, and child health. The plausible explanation for this evidence gap is that most of the infrequently covered domains predominantly concern LMICs, where very few DHI studies have been conducted.

3.1.5 Health outcomes

The impact of DHIs on summary or ultimate health outcomes such as mortality, quality of life and health care has been under-explored. The focus of most studies concerned short-term outputs and directly measurable health outcomes (e.g., number of clinic visits). In fact, there is a huge gap in the number of studies that either report intermediate or summary health outcomes.

The health outcomes reported by the studies are organized along the DHI theory of change (TOC) casual pathway from outputs through impact. For this analysis, outputs comprised process outcome (non-therapeutic) as well as knowledge and beliefs. Intermediate outcomes include behavior change, client/provider satisfaction, health status (natural units), process outcome (therapeutic) and health care utilization. The final outcomes consist of health status (aggregated/summary units) and quality of care. Based on this classification, most outcomes reported by the included studies are intermediate (n=2,306 outcomes), followed by impact (n=342) and outputs (n=296).

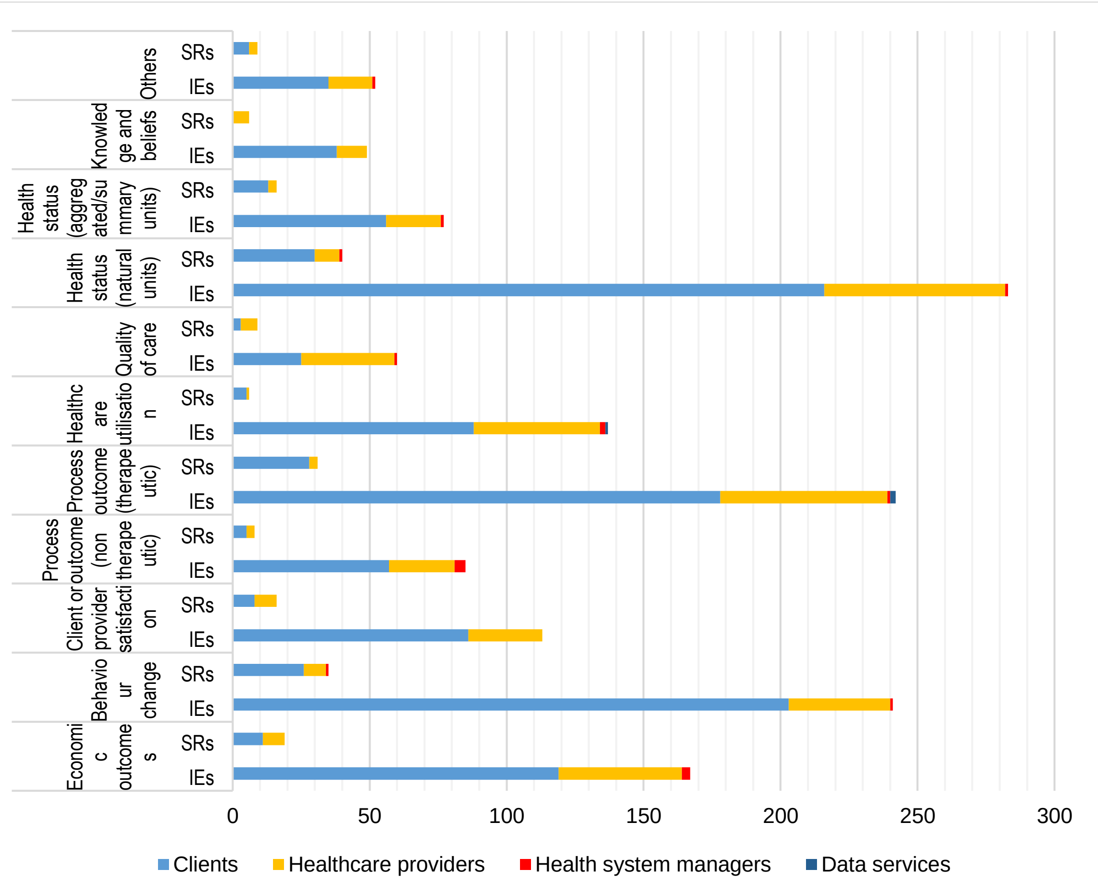

Figure 3 Outcomes reporting from the DHI Evidence base by intervention type

3.1.6 Economic outcomes

Cost data for the interventions are infrequently reported. About 80% (n=632) of the studies did not report any cost data and out of those that reported more than half simply provided cost data without performing further analysis such cost-effectiveness analysis. A majority of the studies that did include CEA covered DHIs for clients (n=119 IEs and n=11 SRs) and health care providers (n=45 IEs and n=8 SRs). Only three IEs and no SR covering DHIs for health system managers reported economic outcomes. Moreover, none of the studies on DHIs for data services included any cost data (See Figure 4).

The economic outcomes reported by the studies are further organized into simple, intermediate and summary/impact based on the DHI's TOC causal pathway. The simple economic outcomes refer to reporting of cost data without any reference to health or well-being outcomes (e.g., cost incurred/saved). However, some refer to outputs such as cost per user or persons reached. Studies reporting simple economic outcomes are generally variations on forms of costing analyses. Intermediate and summary economic outcomes link intervention cost data to select health outcomes. The difference between the two is the level of aggregation and finality of the health outcome. Intermediate economic outcomes focus on immediate or natural outcomes such as cost per infection averted whereas the summary outcomes invoke ultimate health outcomes such as cost per life saved or cost per quality adjusted life-year gained. Highly aggregated outcomes, such as net benefits calculated under a cost-benefit analysis framework, would also be classified as a summary outcome, however no cost-benefit analyses were identified in the included studies.

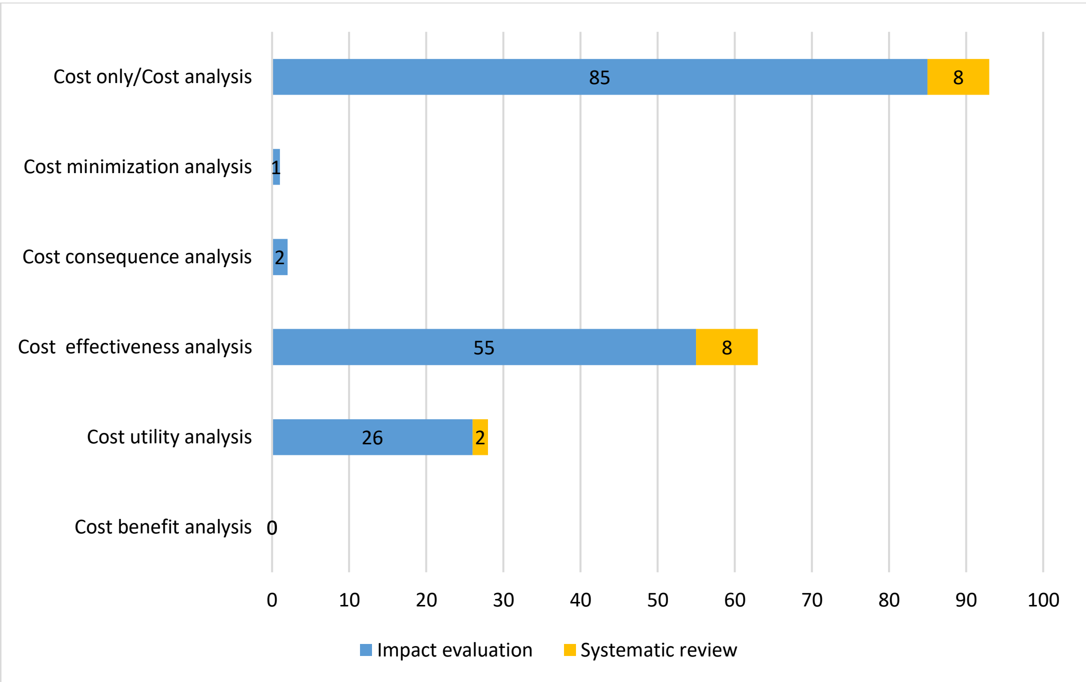

Figure 4 Economic Outcomes reporting from the DHI Evidence base

3.1.7 Evaluation methods

Most studies (n=583, 91%) evaluated the interventions through an RCT framework. The remainder of the studies used quasi-experimental designs. Among the studies with an explicit economic evaluation framework, the majority did not adopt any specific economic evaluation method but instead simply reported the cost incurred or saved. Cost-effectiveness analysis was the most frequently used economic evaluation method (n=63), followed by cost utility analysis (n=28). Cost minimization and cost consequence analyses were reported in one and two studies, respectively. None of the included studies used cost-benefit analysis.

3.1.8 Synthesis evidence gaps

The EGM found significant gaps in evidence syntheses for multiple DHI types. Most of the SRs covered process outcome (therapeutic), economic outcomes, health care utilization, health status (natural units), client/provider satisfaction and behavior change outcomes.

There are opportunities for possible syntheses relating to the effects of client health records and health care decision support system interventions. A substantial number of studies have explored the effects of client health records on health care utilization, quality of care and health status (natural units), but no high-quality SR exists. Another area for potential synthesis is the link between practitioner decision support systems and each of the following outcomes: health care utilization, process outcomes (either therapeutic or non-therapeutic), and quality of care.

The effects of telemedicine on process outcome (therapeutic), health care utilization, health status (aggregated/summary units), health status (natural units), knowledge and beliefs and process outcome (non-therapeutic) present additional synthesis gaps for future reviews. Telemedicine is one of the most common interventions with a large body of evidence across all the outcomes of interest in this EGM. Most of the studies have measured outcomes of health status (natural units), process outcome (therapeutic) and health care utilization. A cluster of completed and ongoing reviews on telemedicine were identified, but only two of them (under the outcome categories of client/provider satisfaction, economic outcomes, and behavior change) are rated as high confidence.

Synthesis gaps also exist among targeted digital health communication interventions as most of the included SR are of substandard quality. This was the most covered type of DHI in the included literature as multiple studies have explored the link between targeted digital health communication and health care utilization (n=56), process outcome (non-therapeutic) (n=47), health status (natural units) (n=134) and client/provider satisfaction (n=67).

3.2 Conclusions and Implications Related to Existing Evidence

This comprehensive assessment of the digital health intervention impact evaluation literature identified substantive gaps for key types of interventions and in the type of outcomes reported.

Based on intervention coverage analysis, there is a need to explore the availability and possible barriers to assessing the health system management and data service interventions. EGM found few IEs that assessed DHI containing an artificial intelligence component. This is one of the areas with substantive evidence gaps that could be addressed by future studies. In terms of health domains analysis, it is evident from the findings that most DHIs were used to address NCDs. Evidence of the effects in other priority global health issues such as HIV, maternal health and nutrition is very limited. These health domains should be considered in future studies to improve the availability of evidence in this field.

The EGM proposed that future studies on DHIs should consider measuring long-term health and economic outcomes, which are relatively neglected in the current literature. These outcomes include those relating to the quality of care, quality of life, survival rate and summary economic outcomes. There is a limited economic evaluation evidence base for DHIs, which can hinder decisions on policy, programming and scale-up.

4. The Case for a DHI Economic Evaluation Framework

4.1 Methods Frameworks for Economic Evaluation in Health

A methods framework guides the planning, conduct and reporting of economic evaluation and assists in maintaining a consistent approach allowing comparison of analytical results over time, and context. An economic evaluation framework builds on the concept of a reference case but is less prescriptive and accommodates a wider range of methodological options.

The World Bank has published substantively in applied benefit-cost analysis with methodological recommendations developed as early as the 1970s with subsequent methods initiatives in 1992 and 2010 [16]. The concept of a reference case was first used by the 1st US panel of cost and cost effectiveness in health care in 1996 [17]. The overarching aim was to establish a standardized methodology for the conduct of economic evaluation (in the case of the US panel, cost effectiveness analysis) as a mechanism to improve the usefulness of economic evaluation in decision making. The US panel members recognized that improving the quality of economic evaluation through methodological advancements and comprehensive guidance was insufficient, and that for economic evaluation to find a place in policy, there needed to be a degree of consistency in the analytic approach for decision makers to be able to have confidence in evidence produced and be able to compare and weigh the evidence for competing investments. The 2000s saw a proliferation of national priority setting agencies (also termed Health Technology Assessment (HTA) Agencies) that built on the reference case concept for the specification of economic evaluation that would inform investment decisions for local national health systems. A high-profile national methodology developed by the National Institute for Health and Care Excellence in England and Wales defined a series of methodological specifications for economic evaluation, including how analytical comparators were to be chosen, instruments for health-related quality of life measurement, and the management of equity and time horizon and discount rates. This enabled a level of consistency in the analytical approach used to inform investment decisions. The establishment of the Health Intervention and Technology Assessment Program (HITAP) in Thailand that was tasked with conducting economic evaluation to inform investments in the country's health benefit package and essential drugs list highlighted the importance of rigorous economic evaluation methodology and processes as a component of a sustainable Universal Health Coverage (UHC) system. HITAP also demonstrated the applicability of this approach for a middle-income country context. The World Health Assembly Resolution 67.23 in 2014 specified the integral role of health technology assessment for countries journeying towards UHC and established a direct link between consistent methods and process when considering evidence of economic evaluation and achievement of UHC. A growing number of countries now have country-specific guidance, reference cases or frameworks for economic evaluation to inform national decision making. While these initiatives are a significant stimulus for economic evaluation methods innovation, the direct global application of these frameworks is limited, as by definition, these are developed from the perspective of a local decision maker concerned with national budgets and national population health in a particular country.

4.2 Existing Economic Evaluation Frameworks for Generalized Interventions

To build on the progress of national-level priority setting and economic evaluation methodology to provide a more globally focused guidance, a Bill and Melinda Gates Foundation funded workstream led by the International Decision Support Initiative (iDSI) created a generalized principle-based reference case that would have applicability to multiple contexts, intervention types and decision makers. The iDSI reference case identified common or “universal principles” of economic evaluation that should be considered for generating economic evidence intended to inform decision making in a UHC context. Further initiatives extended the iDSI reference case, including the Global Health Costing Collaborative [18] that expanded substantially on the “costs” principle, and the Benefit-Cost Analysis Reference Case, that refined multiple principles including the valuation of mortality and morbidity.

The WHO guidance on Generalized Cost Effectiveness Analysis detailed an approach to conducting cost effectiveness analysis that would have applicability at a global, regional, and national level and could be utilized to inform local decision making [19]. Components of the GCEA methodology is used to inform practice of economic evaluation at country and global levels.

Table 2 details four initiatives with an international scope that developed methodological frameworks for the conduct of cost effectiveness analysis and benefit-cost analysis.

A further initiative targeted specifically towards reporting of economic evaluation was the development of the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement in 2013 [20]. The CHEERS statement built on the CONSORT statement designed for trial reporting and outlined the key elements of an economic evaluation to be reported and was aimed primarily at researchers and journal editors.

Table2 Methods Framework for Economic Evaluation (International scope)

TITLE	WHO GUIDE TO COST-EFFECTIVENESS ANALYSIS	REFERENCE CASE GUIDELINES FOR BENEFIT-COST ANALYSIS IN GLOBAL HEALTH AND DEVELOPMENT	THE INTERNATIONAL DECISION SUPPORT INITIATIVE REFERENCE CASE FOR ECONOMIC EVALUATION: AN AID TO THOUGHT	RECOMMENDATIONS FOR CONDUCT, METHODOLOGICAL PRACTICES, AND REPORTING OF COST-EFFECTIVENESS ANALYSES
Objectives	To provide policymakers and researchers with a clear understanding of the concepts and benefits of Generalized CEA (GCEA)	To clarify important concepts, aid in implementation, and provide default values for key parameters including options for standardized sensitivity analysis	To support decisions aimed at improving population health from within available funding while acknowledging the relevance and trade-offs associated with the incorporation of social values into those decisions	To improve quality and comparability of CEAs
Intended audience	Analysts with some background in CEA	Practitioners with some training and experience in conducting economic evaluations	Policymakers	Multiple stakeholders
Methodological approach	GCEA	BCA	Standardized principle-based methodology	CEA (with cost-consequences analysis (CCA) components)
Structure	A standard set of methodological choices on how to perform GCEA	Stepwise analytical approach, specific guidance in key areas with focus on approaches implementable in LMIC settings	11 principles with associated methodological specifications and reporting standards	Central topics in economic evaluation in book format
Key features (non-exhaustive)	Maximize generalizability of results across settings by comparing to starting point of doing none of the current interventions	Updated and coordinated guidance on key BCA concepts including valuation of morbidity and mortality risk reduction, standardized VSL benefits transfer approach	Each principle accompanied by corresponding methodological specification and reporting standard	Dual perspectives – health system and societal, introduction of the Impact Inventory concept (cross-sectoral CCA)

Source: Authors, World Bank.

4.3 Digital Health Interventions and the Concept of Value

The concept of value in health and health care is highly contested in the literature with an abundance of value conceptualizations that seek to accurately reflect the "things that matter" to payers, patients, and the population impacted by the investment or implementation of a health intervention. While approaches to value conceptualization necessarily change over time and continual work in this area is required, a consistent approach to decision making and economic evidence requires a consistent approach to the representation of outcomes. Figure 5 depicts major areas of value that are expected to be in common to a greater or lesser degree with the three main audiences: investors, product developers, and researchers. These value areas are also informed by the findings of the DHI literature as described in Section 2. An economic evaluation that can appropriately reflect value is a critical input to the decision-making process and therefore any analytical methodology should be tailored to the conceptualization of value related to DHIs.

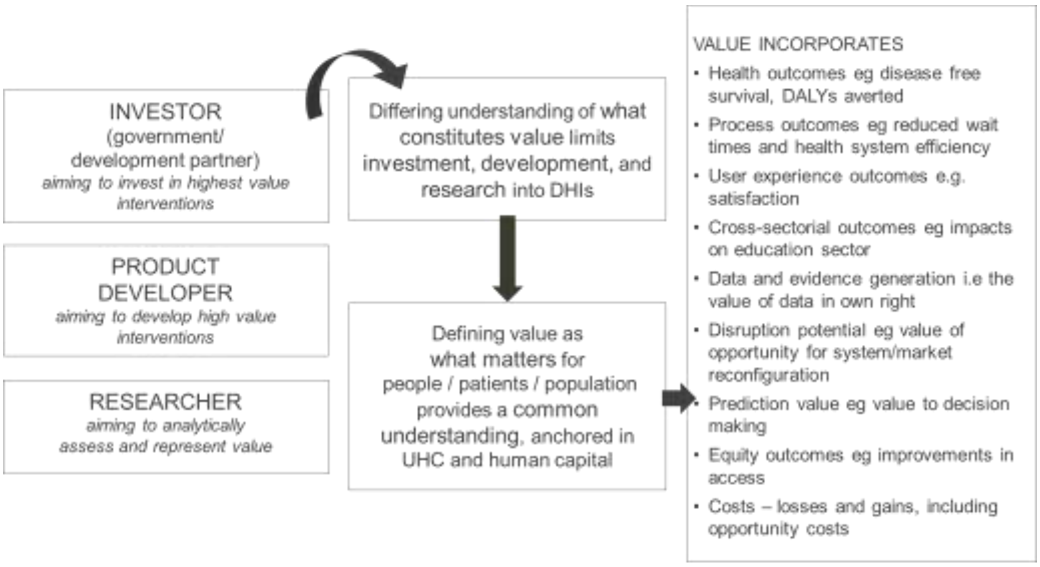

Figure 5 The Concept of Value in an Economic Evaluation

Source: Authors, World Bank.

Note: DALY = disability adjusted life year, a composite measure of burden of disease incorporating morbidity and mortality.

4.4 Existing Guidance on the Evaluation of Digital Health Interventions

The World Health Organization guide to Monitoring and Evaluating Digital Health Interventions (2016) provides a strong foundation for the assessment and generation of evidence on DHIs [21]. It provides an extensive guidance on approaches and tools for monitoring and evaluation, with a description of evidence types and evaluation methods. This framework expands on the concepts introduced in the guide, tailored towards the concept of economic evaluation, and broadened to incorporate dynamics related to predictive analytics and DHIs.

A systematic review conducted by Kolasa et al (2020) noted a central challenge in the value assessment of DHIs was due to the complexity of the simultaneous evaluation of clinical, organizational, and economic aspects [6]. The review identified Value Assessment Frameworks for DHIs in the literature and synthetized the criteria as well as the methods being used in the evaluation of DHIs.

The review summarized assessment frameworks across telemedicine and mobile health and generic conceptualizations for DHI assessment. Using the structure of the EUnetHTA core model (a methods framework developed by European national HTA agencies [22]), a list of categories was elicited to screen against and map the scope of included publications. The list included 12 categories: health problem and current use of technology; safety; clinical effectiveness; patient and social aspects; economic; legal; ethical; organizational; usability; data security; and technical aspects and stability. In total, 11 value assessment frameworks were identified. The most often discussed assessment domains were safety, clinical effectiveness, usability, economic aspects, and interoperability which were discussed in seven of the frameworks, while organizational, usability, data security, and technical considerations were discussed in six of the frameworks. The least attention was devoted to the understanding of ethical (two out of 11 frameworks) and legal aspects (one out of 11 frameworks).

The review proposed several recommendations regarding the criteria to be considered in the measurement of value of DHIs, as detailed in Table 3.

Table 3 Summary Findings from Review of Existing DHI Analytical Frameworks

FIVE RECOMMENDATIONS	
Choice of comparator	The value of a DHI should be determined by the incremental advantage compared to the current standard of care.
Multi-stakeholder perspective	The value of a DHI should quantify incremental differences it delivers to all beneficiaries.
Organizational impact	The value of DHIs should be conditional on the health care system preparedness to consume efficiency gains and assurance that data generated by DHI will be accessible by health care professionals.
Multidimensional outcomes	The value delivered by digital health solutions should consider multiple dimensions such as clinical, organizational, behavioral, and technical.
Interoperability	Connectivity to other data sources must be considered in the evaluation of a digital health solution.
ONE METHODOLOGICAL SUGGESTION	
Value aggregation function	Given the multiple criteria to be considered in the pricing and reimbursement process, each value attribute should be weighted in any calculations based on the preferences of chosen stakeholder groups, and the aggregate value score for each DHI should be estimated.

Source: Adapted from Kolasa et al (2020).[6]

4.5 Digital Health Interventions and the Role of Evidence

Digital Health Interventions are unique among other health interventions in the way that evidence is generated and synthesized. An economic evaluation typically views the range of evidence for a health intervention at a single point in time in a particular context and many methodological approaches are used to reflect how health intervention parameters may change over time, context and scale. However, there are unique attributes to the evidence base for DHIs, including the way that the implementation of a digital intervention will in itself commonly generate further evidence and that the specific attributes of a digital intervention can evolve over time to adapt to context.

The distinction between efficacy (the effect observed in a controlled and ideal environment) and effectiveness (the effect in routine conditions in a specific context) is a key component of any type of economic evaluation intended to inform decision making. Evaluation of non-complex health interventions, such as pharmaceuticals, will typically adopt a primary effect estimate from clinical trial data and then attempt to reflect any variation from the observed effect that is expected in the context of the economic evaluation. This information is also supplemented by observed data and real-world evidence (RWE) but a common objective is to utilize efficacy information to estimate effectiveness.

The regulatory environment for a particular type of health intervention is a major driver of the approach to evidence generation. The strong traditional focus of health intervention regulation on patient safety and efficacy means that evidence generation is commonly focused on whether potential benefits outweigh potential harms rather than cost effectiveness or value for money. When conducting economic evaluation for an intervention that already has regulatory approval (such as pharmaceutical), evidence required at the point of regulation will have demonstrated efficacy and safety for a particular clinical indication. In contrast, economic evaluations of non-regulated products (or where the regulatory environment does not routinely require demonstrated efficacy through clinical trial) may not have baseline evidence of impact to parameterize an analysis.

DHIs, in common with many other complex interventions, are intricately integrated into the context within which they are used. This diminishes the importance of the concept of efficacy within an economic evaluation of a DHIs and limits the applicability of the traditional evidence hierarchy. The randomized controlled trial (RCT) remains a gold standard for isolating the treatment effect of a digital or non-digital interventions, however RCTs are unlikely to be able to routinely provide the intervention effectiveness estimate required in an economic evaluation of a DHI.

However, published peer-reviewed reporting trials remains an essential element of evidence generation for DHIs and should be used where applicable. The SPIRIT-AI and CONSORT-AI statements provide clear specification for the design and reporting of protocols and trials of interventions involving Artificial Intelligence components [23] [24]. They expand on the established CONSORT and SPIRIT statements with specific components to enable the generation of evidence related to AI technologies. The expanded use of the statements will enable improved trial conduct and reporting and is likely to have a positive impact on the evidence available for use in economic evaluation.

A DHI, by its nature, generates data through its use. The data may be utilized in different ways depending on the information system in which the DHI operates, but the routine generation of data has implications for the management of evidence in an economic evaluation [7]. The data gathered will inform basic parameters such as utilization, but also enables linkages with other sources to impact on health and processes dynamically. As many DHIs include features of adaptive design, the generation of evidence can lead directly to changes in the effect size of the intervention. There is also substantive scope for DHIs to dynamically improve the effectiveness of other digital and non-digital interventions as a result of improved evidence. For example, a basic digital clinical decision support tool to aid prescribing may have direct process efficiency impacts on the time taken for clinicians to complete a patient consultation. Its efficiency impact may improve as it expands to more therapeutic areas or facilitates a greater granularity of decision support, and over time may also improve the effectiveness of the medical interventions being prescribed through improved targeting of indications and dosing and monitoring of side effects. Considerations for the assessment of digital predictive capabilities can build on existing economic evaluation of non-digital risk prediction models. A systematic review assessed the range and quality of economic evaluations of 39 risk prediction models in clinical practice, evaluating models such as the Framingham Risk Score (for risk of Coronary Health Disease) and Nottingham Prognostic Index (for prognosis following breast cancer surgery) [25]. The review found wide variation in methodological approaches particularly in areas such as comparator choice and inclusion of costs, with the majority adopting cost utility or cost effectiveness analysis (life/life year saved). Although limited, this existing evidence base can be used to inform approaches to evaluating digital prediction models that may incorporate predictive analytics functionality.

Existing initiatives have consolidated approaches to the consideration of the evidence base for DHIs. Murray et al developed a series of research questions that assist in designing research that is tailored to generating evidence for decision making related to DHIs [26]. The UK's National Institute for Health and Care Excellence (NICE) have developed an Evidence Standards Framework that details the approach to use of evidence when an evaluation is conducted to inform investments in the National Health Service in England and Wales [7]. While this initiative is targeted towards a particular high-income country context, it demonstrates an approach to the organization of evidence that can be applied more generally.

The dynamics of evidence generation and use for DHIs are clearly different from traditional health interventions, but any investment in DHIs, particularly by country governments, will immediately displace actual or potential spending on other interventions, be they digital or non-digital. Therefore, DHIs should not be seen as exceptional or warranting a lower standard of evidence to inform economic evaluation or decision making [7]. The framework concept attempts to ensure that the principles supporting evidence production and use are applied consistently across digital and non-digital interventions, while accommodating operational and pragmatic differences.

4.6 The Risks and Benefits of a DHI Economic Evaluation Framework

Any proposed methodological framework is intended to be widely adopted and incentivized to improve analysis and use of resulting information. The benefits of using a framework is linked to its capability to enhance both the quality of individual decisions through robust standards for the planning, conduct, and reporting of economic evaluations and the consistency in decision-making over time [27].

However, there are risks associated with the use of a prescriptive framework. Economic evaluation is a rapidly evolving field, and particularly in the case of DHIs, there are several areas where consensus does not yet exist or methods have not been developed to appropriately represent the full range of costs and benefits, their inherent uncertainty, and to accurately understand the nature of the opportunity cost of investment. Therefore, a framework carries the risk that certain practices and ways of representing value in economic evaluation of DHIs become "set in stone".

Importantly, the balance of advantages and disadvantages of the use of a framework will be strongly influenced by how prescriptive it is. The issue is not whether to use a framework but to assess (a) how prescriptive the framework should be, (b) how to interpret analyses that diverge from it, and (c) how to draw a balance between standardization of economic evaluation and methodological innovation [27].

5. The DHI Economic Evaluation Framework

5.1 The Digital Health Interventions Economic Evaluation Framework

This section describes the DHI Economic Evaluation Framework, outlining the main elements, principles and specifications as detailed in Figure 6.

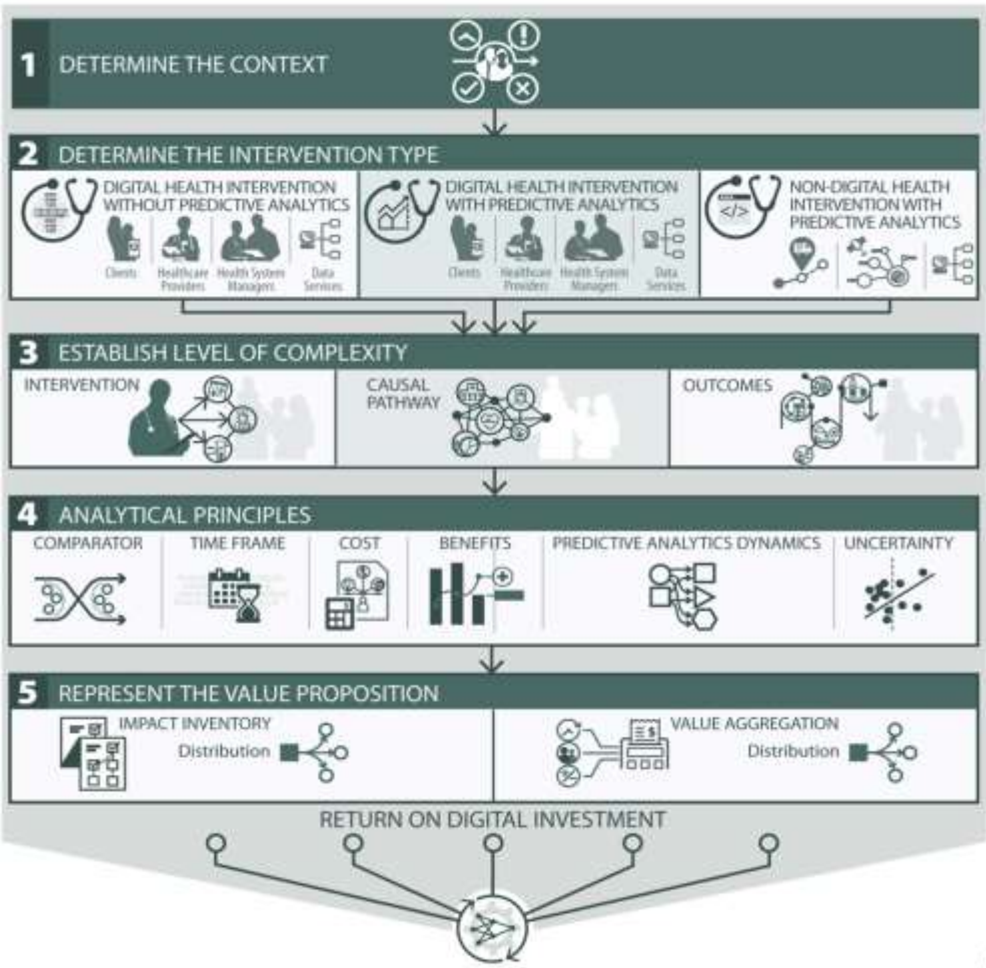

Figure 6 Digital Health Intervention Economic Evaluation Framework

Source: Authors, World Bank

STEP 1: Identify and represent the CONTEXT of the economic evaluation

The first step in an economic evaluation of a DHI is to determine the context in which the intervention is to be assessed. A DHI, by nature, is a context-specific intervention, with costs and effects highly contingent on the health system and digital architecture in which it is implemented. When determining context, it is important to specify general contextual elements including the health system characteristics, how the intervention would be used in the clinical or treatment pathway, and user type. The digital context of the intervention is also an essential element and includes digital architecture and equipment, information systems, and software where applicable. Analysts should use existing context analysis tools as appropriate to the intervention, such as the Digital Health Investment Review Tool developed under the Maternal and Child Survival Program [28].

The intended decision maker is the organization, body, or group of people who will make use of the information produced in the economic evaluation. It is an important component of the economic evaluation to specify as it informs interpretation and application of all other methodological concepts, including the perspective of the analysis. It is important to note that there is no generalized "correct" intended decision maker for an economic evaluation, but the production of an economic evaluation with an undefined decision maker can potentially lead to inappropriate or incorrect use of the evidence, and sub-optimal decisions. For example, an economic evaluation of a digital health technology intended to inform an investment decision by a secondary care provider (such as a private hospital group) is likely to approach the definition of the comparator, scoping of costs, time horizon and representation of clinical effects and summary impact measures very differently to a country government department of health considering rolling out the same intervention nationally.

Intrinsically linked to the intended decision maker is the context in which the DHI will be implemented. Given the major influence that the context will have on all aspects of a DHI functionality and impact, there is little utility in concepts of efficacy for the economic evaluation of DHIs. The context of the DHI should therefore be established from the outset, with clear explanation of key components of the context, including the country context, health system environment, and digital enabling environment.

Determining the intended decision maker and context does not preclude applicability of the outputs of an economic evaluation to other contexts. In practice, the clear definition of the decision maker, with explanation of how this perspective affects methodological decisions and parameterization will facilitate transferability of economic evaluation to other contexts as users will be able to assess which aspects of an analysis require modification to apply it to another context.

STEP 2: Determine the intervention type

Specification of the intervention, including its use-case and intended population is an essential element of the evaluation. This framework requires not only a general description of the intervention, but also alignment to established classification systems. In the first instance, the intervention should be assigned to a WHO DHI type, with specification as to whether there is a predictive analytics component. This creates three major categories of DHI, with further granularity for DHIs under the WHO classification system.

1. DHIs defined under the WHO DHI classification system without AI enabling functionality – e.g., simple text message to patients as a reminder for regular follow-up.
2. DHIs defined under the WHO DHI classification system with AI enabling functionality – e.g., machine learning radiology diagnostics.
3. Non-digital health interventions enabled by the used of AI technology – e.g., use of demographic and clinical information for identification of at-risk patients for health professional consultation.

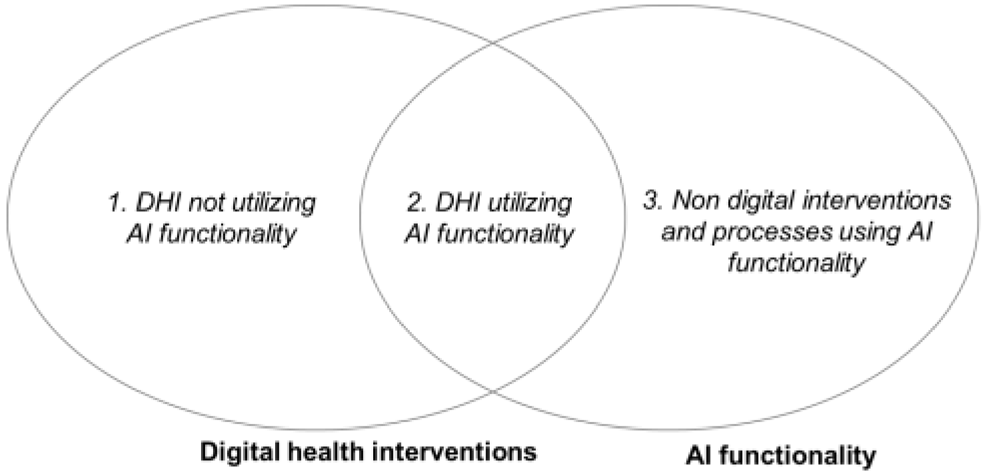

Figure 7 WHO Classification With and Without AI

STEP 3: Establish the level of complexity

The third step in the framework is to establish the level of complexity, clearly identifying key elements to inform the approach in steps 4 and 5. The concept of intervention complexity and implications for evaluation are not unique to DHIs and have been well established in the literature, particularly in the areas of program evaluations and public health interventions [29]. While the fundamental principles of evaluation apply whether an intervention is simple or complex, there are important considerations in the evaluation of complex interventions, and this has implications for the generalized approach to evaluation of DHI, which commonly fall under the definition of a complex intervention.

In a discussion paper on economic evaluation of DHIs, McNamee et al. highlighted the critical importance of reflecting DHIs as complex interventions within complex systems [5]. The paper noted the description of a complex intervention by the Medical Research Council and National Institute for Health Research [30] [31] as one that "contains several interacting components, and other characteristics, such as the number and difficulty of behaviors required by those delivering or receiving the intervention." The proposed approach to representing complexity in the framework is detailed in Table 4 and builds on the approach by Petticrew et al [32] and highlighted by McNamee et al [5].

The approach requires specifying three types of complexity. Intervention complexity relates to the intervention itself, as DHIs commonly have multiple parts which combine to produce the observed effect. Intervention complexity is compounded by the continual adaption and modification of DHIs over time. Causal pathway complexity refers to linking attribution of an impact to the initial input. Frequently the realization of the impact of a DHI involves multiple moderators and this complexity is also highly linked to implementation context. Outcome complexity refers to the complexity associated with determining impact, as DHIs may have externalities and spillovers. While the particular approach to representing the level of complexity of the DHI will likely depend on the requirements of the analysis, users should at least explicitly acknowledge the components of complexity and use it to inform the methodological choice in step 4.

Table 4 Digital Health Interventions and Complexity

TYPE	DESCRIPTION
Intervention complexity	Many DHIs are multi-attribute, with inter-dependent parts and delivery mechanisms, and often adapt over time.
Causal pathway complexity	The causal pathway linking observed outputs to outcome and impact can have substantial complexity, compounded by range of identified outcomes and complex attribution and feedback loops.
Outcome complexity	DHIs commonly have range of outcomes, including health and health system effects and involve spillovers and externalities, impact on patients, users, and the wider population.

Source: Petticrew et al 2013, McNamee 2016

STEP 4: Apply analytical principles

The fourth step of the framework is to apply economic evaluation principles as detailed in table 5. This builds on the extensive methods literature in economic evaluation [4][33]. It is beyond the scope of this framework document to replicate existing methodological guidance, but to identify the principles that should ideally be followed when making context and intervention-specific methods choices. An important consideration is that while there may be some variation to the particular approach of any economic evaluation, these principles are expected to apply for both digital and non-digital interventions (with the exception of predictive analytics that is expected to be unique to DHIs with a predictive analytics component).

Table 5 Analytical Principles

ITEM	PRINCIPLE
Comparator	The comparator(s) against which costs and benefits are measured should accurately reflect the decision problem
Timeframe	The time horizon of the analysis should be of sufficient length to capture all significant differences between the intervention and comparator
Costs	All differences between the intervention and the comparator in expected resource and costs of delivery to the target population(s) should be incorporated into the evaluation
Benefits	All differences between the intervention and the comparator in expected benefits to the target population(s), should be incorporated into the evaluation
Predictive analytics	The dynamics of predictive analytics on evidence generation, costs and benefits should be reflected in the analysis
Uncertainty	The uncertainty associated with the evaluation estimates should be appropriately represented

Establishing the comparator against which the intervention will be evaluated is fundamental to the approach and outputs of an analysis. Economic evaluation seeks to represent costs and consequences in context, which requires understanding of what the intervention would replace, including if the current practice is doing nothing or minimal action. However, it is often difficult to determine exactly what would be replaced by an intervention, and so this framework recommends the development of two comparator scenarios, drawing on the approach of the iDSI reference case. Analyst judgement is required to determine the validity of each scenario depending on the context and needs of the analysis. The first scenario is standard of care or existing process, and the second scenario is minimal supportive care. For example, an AI-enabled diagnostic tool that assisted interpretation of radiology scans to identify disease and referral would be compared to a scenario that consisted of a radiologist manually interpreting scans and a scenario where patients wouldn't routinely have access to radiology services. The costs and benefits of the new diagnostic tool could then be represented relative to these two scenarios to gain a full understanding of the impact that the new intervention represents.

The timeframe adopted in an analysis is important to ensure that any bias based on the timing of costs and benefits is minimized. The general principle is that the timeframe should be of sufficient length to capture all relevant costs and benefits between the intervention and comparator. Where there are observed mortality differences between the intervention and comparator, this will usually require that a lifetime time horizon is adopted even if mortality estimates are largely projected to a longer horizon. This is particularly relevant for DHIs where a significant proportion of costs are frequently incurred upfront at the stage of development or implementation, with benefits extending into the future. An important consideration for timeframe choice is the discount applied to future costs and effects. This framework does not prescribe a particular annual discount rate but requires that the discount rate chosen is clearly specified and incorporates local contextual evidence.

The framework requires a comprehensive approach to costing, with the principle specifying that "all differences between the intervention and the comparator in expected resource and costs of delivery to the target population(s) should be incorporated into the evaluation". The approach to collection, synthesis, and inclusion of costs in an economic evaluation is detailed extensively in the literature. The Global Health Costing Collaborative (GHCC) is an initiative dedicated to establishing best practice in costing and has made a significant contribution to methodological understanding and practice of costing health interventions and aligned to the costing principle of this framework. However, the nature of costing for DHIs is likely to require specific approaches and it is expected that dedicated research relating to DHI costing is required. In particular, assessment of a DHI is likely to require consideration of how fixed costs are to be attributed in relation to costs of development and future cost structures and at different levels of scale. A "legacy" intervention may pass costs or cost savings onto the health system over a long time period and the marginal cost per user may be substantially reduced with many concurrent DHIs.

In common with costing, estimation of the benefits associated with a DHI should take a comprehensive approach, with the principle that all differences between the intervention and the comparator in expected benefits to the target population(s) should be incorporated into the evaluation. The principle requires in the first instance that benefits are identified in natural units, with approaches to aggregation and synthesis considered in Step 5. The impact inventory details specific areas of value that will guide the organization of the benefits into meaningful categories, aiding interpretation by decision makers. A particular consideration for representing benefits of DHIs is the potential for intervention modification after initial implementation (see earlier section on 3.5 on role of evidence) [34].

The Predictive Analytics principle requires that the impact associated with the ability of a DHI to assist in prediction within the anatomy of a task is explicitly detailed and quantified where possible. This category is unique to DHIs with predictive functionality and enables the impact of the prediction component of an intervention to be isolated from the wider impact of intervention. For example, a radiology diagnostic intervention with predictive analytics may be used for assisting diagnosis of tuberculosis disease from radiology scans. The complete diagnostic intervention may have many components, such as digital information capture, automated messaging and referral functionalities, but the prediction value would be represented by the improved diagnostic accuracy in terms of sensitivity and specificity of the intervention compared to either standard of care or combinations of the intervention in conjunction with radiologist prediction and judgement. Representing the prediction value in isolation enables comparison of differences in prediction value between interventions and at different time points and contexts.

This principle interacts with other principles (particularly costs and benefits); however, the unique predictive impacts should be explicitly established in the analysis. The approach to representing predictive analytics dynamics will need to be tailored to the nature of the intervention, with specific consideration given to the identification of the impacts relating to the nature of evidence generation and the disruptive and process changes that may be enabled by the intervention.

STEP 5: Represent the value proposition

The fifth step in the framework is to represent the value proposition. The framework takes a dual approach of disaggregation using an impact inventory (see Table 6), and aggregation in a form required by the intended decision maker. Throughout the representation of the value proposition, it is important that the uncertainty and distribution of impacts are clearly represented. Methods for representation of uncertainty in economic evaluation are well established, and analysis should reflect major aspects of uncertainty including parameter, structural, and methodological uncertainty [33], incorporating the unique attributes of DHI in relation to evidence generation and use (Section 2). The distribution of the costs and consequences of any health intervention is an essential input to an investment or implementation decision and should be represented as clearly as possible in both the impact inventory and in any aggregation. Representing the distribution of impacts across populations enables consideration of factors beyond efficiency or health maximization and includes equity and priority to disadvantaged populations. While the methodological approaches to representing distribution of impacts is likely to be common across digital and non-digital interventions, [35] [36] DHIs have important and unique distributional considerations. Firstly, in many instances DHIs have the potential to mitigate existing access barriers to traditional health care provision particularly for marginalized or rural populations. However, there is also potential for DHIs to exacerbate existing inequalities where access to a DHI is limited to those that already have access to an enabling digital health infrastructure, such as a smartphone or a local clinic with an internet connection. While an economic evaluation may not be able to quantify the full distributional impacts of a DHI investment, the Value Proposition step should attempt to represent impacts using established economic evaluation methods.

5.2 Disaggregation: Impact Inventory

An impact inventory is a series of health and non-health outcomes and costs presented in a tabulated form and is a common approach in the cost consequence analysis form of economic evaluation. The US Panel on Cost-Effectiveness in Health and Medicine proposed the use of an impact inventory when conducting CEAs, noting that many outcomes and costs were important to a decision maker, but could not reasonably be incorporated within an Incremental Cost Effectiveness Ratio (ICER). In addition, the US Panel Impact Inventory enabled the representation of both a health services perspective and a wider societal perspective to the analysis, enabling a versatile approach that could address the economic evidence requirements of a range of potential decision makers. This framework builds on this approach with specification of particular impacts that should be represented in an economic evaluation of DHIs.

The Impact Inventory specifies five broad inventory areas, and it is expected that depending on the nature of the DHI being evaluated, additional sub-categories will be required. Application of the impact inventory will facilitate further specification and detail on the approach for comprehensive representation of appropriate sub-categories, building on existing methodological guidance on outcomes measurement of general health interventions. Identification of the “type” of DHI in Step 2 will guide a generalized approach to representing components of the impact inventory.

An essential element of the impact inventory is that outcomes are represented in units that are applicable to the particular category. It is likely that the different categories may not be completely mutually exclusive, and some categories will have downstream attribution effects on summary metrics and costs. However, the main aim of the impact inventory step is to facilitate transparent representation of results in disaggregated form and to provide a common framework for value aggregation. Considerations about the appropriate level of aggregation and development of summary metrics are addressed in the value aggregation step.

The Health impact component requires representation of the observed health effects relative to the identified comparator(s), with a requirement for both positive and negative health effects to be represented. The Health System impact requires representation of the changes in processes that are expected to be facilitated by the introduction of the DHI, and is likely to incorporate intermediate outcomes such as user experience and acceptability, where these effects lead to meaningful process change.

The User-experience impact will commonly be reported in a qualitative synthesis or through use of established tools for measurement of the user experience. It is beyond the scope of this framework to recommend specific user experience measurement tools; however, the analyst should ensure that any tools are validated and appropriate for use in the context in which the intervention is to be used. The User-experience impact may be a fundamental contributor to the observed health impact or health system impact. For example, a recent DHI that utilized a personalized digital counseling application to enable consumer choice in contraceptive use found a positive effect on increasing adoption [37]. As a joint patient-and-provider facing intervention, a substantive proportion of the impact was contingent upon the provider/user being able to effectively interact and use the application. Economic evaluation that specifically identifies user experience (whether that user is a patient, the general population, a health provider or a health services manager) will be essential to accurate representation of the overall intervention impact.

The Cross-sectoral impact enables a wider societal impact of a DHI to be incorporated, including impacts on other social sectors such as education. While the “Health Impact”, “Health System Impact” and “Cross-sectoral impact” components of value representation are methodologically similar whether the intervention is non digital or digital, the Value of Data and “User experience impact”   value categories are more specifically tailored towards the unique impact of   DHIs It is likely that there may not be quantitative information for some aspects of the Impact Inventory, in which case a qualitative description of the impacts should be provided.

Table 6 Components of Impact Inventory

COMPONENT	DESCRIPTION
Health impact	Health impacts: natural units, generalized units, and net health benefits that incorporates the opportunity cost of lost health
Health system impact	The impact on the health system as a result of intervention implementation, includes direct measurable impacts and longer term predicted process changes or disruption potential.
User experience impact	Impacts on patients, population and health workforce not represented within health or health system impact such as trust, privacy, choice, satisfaction and knowledge.
Value of data	The value represented by the availability of data gathered through the use of the DHI that is not represented in direct impacts on health, health system or user experience
Cross-sectoral impact (beyond health system)	Impacts on non-health sector e.g., education and wider societal benefits e.g., human capital

Value of data requires explicit identification of what evidence is being generated through the use of the DHI. This could range from simple utilization data to comprehensive datasets on clinical pathways, processes, and patient outcomes. This value is likely to be contingent on the context of the DHI and in particular the existing information systems and digital infrastructure, further highlighting the need for explicit specification of context in Step 1.

5.3 Value Aggregation

A central consideration in an economic evaluation is what, if any, aggregation of the costs and consequences is required and is appropriate. The ability for an economic evaluation to coherently summarize costs and benefits into easily understood metrics is often a fundamental component of communication, comparability and usefulness in decision making. However, the approach to aggregation often involves a series of assumptions, simplifications and underlying conceptual frameworks that if not represented explicitly can lead to misunderstanding or misrepresentation of results.

Representing an incremental cost effectiveness ratio (ICER) with a generalized measure of health such as the Quality-Adjusted Life Year (QALY) is commonly used for marginal health budget decisions under an extra-welfarist framework where improvement in population health is a major, but not only, consideration of decision makers. Use of the ICER with a generalized measure enables representation of the local opportunity cost of the decision with fixed or known budget where the marginal productivity of the health system can be estimated. This approach is commonly used by national priority setting institutions such as Health Technology Assessment agencies within a broader decision making framework for investment decisions about individual interventions or groups of interventions and in development of clinical guidelines and treatment protocols.

Estimating a return on investment (ROI), Benefit-Cost Ratio (BCR) or net benefits utilizing a benefit-cost analysis (BCA) analytical approach is commonly applied to major policy or regulatory changes or large-scale investments where there are multiple impacts across a range of constraints and sectors and the objective is to represent the impact and distribution on societal welfare. Health impacts of interventions can be monetized utilizing concepts such as the value of statistical life (VSL) informed by valuation of mortality or morbidity risk reduction. Converting all positive and negative impacts into monetary form enables aggregation using a common metric which enables clear and simple communication of results. Applying a BCA framework to DHI investments will require consideration of how the value categories introduced in the impact inventory can be monetized. Methodology for monetization of Health Value, Processes Value, and Cross-sector Value impacts are largely in common with established BCA methods. If a total monetized summary metric is required in the analysis, the bespoke approaches would need to be adopted for monetization of Evidence generation value, Prediction Value, and Disruption Potential, with careful consideration required to ensure correct attribution and avoidance of double counting. Some emerging approaches that may be applicable for monetization of DHI-specific impacts are detailed in table 7, and it is expected that ongoing methods research will generate more tailored methods.

Table 7 Approaches for monetization of DHI-specific impacts

COMPONENT	ESTABLISHED/EMERGING APPROACHES TO MONETIZATION
Health value	Value of Statistical Life (Robinson et al 2019) [38]
Process value	Efficiency analysis (Durrett et al) [39] Time preference (Whittington & Cook 2019) [40] Contingent Valuation (Bayoumi 2004) [41]
Cross-sectorial value	Economy-wide effects (Strzepek et al 2018) [42]
Evidence generation value	Data Shapley (Ghorbani & Zou 2019) [43] Value of Data Assets (First San Francisco) [44]
Prediction value	Scenario analysis – preliminary development (Agrawal et al 2019) [45]
Disruption potential value	Analysis of potential future scenarios with valuation

A social return on investment (SROI) analysis is a form of BCA. Building on the approach for estimating ROI it explicitly attempts to identify and isolate wider social benefits from immediate financial returns of an investment and is useful where there are particular benefits that are considered socially desirable by a decision maker and that should be explicitly represented.

The major types of value aggregation are detailed in table 8.

Table 8 key features of different types of value aggregation

TYPE OF VALUE AGGREGATION	REPRESENTATION	INCORPORATES	INTERPRETATION REQUIRES	FURTHER METHODOLOGICAL GUIDANCE
Cost-benefit analysis	Benefit-cost ratio: net benefits, return on investment	Full costs and benefits converted to monetized form to extent possible	Understanding of valuation technique, particularly health benefits	Benefit-Cost Analysis Reference Case [4]

Cost utility analysis (a subset of CEA)	Incremental cost effectiveness ratio (effect in generalized units) Net health/monetary benefits	Costs and effects from a defined perspective	Marginal productivity of health system (k); ICERs of competing investments	iDSI Reference Case; [2] Drummond 2015 [46]
Cost effectiveness analysis	Incremental cost effectiveness ratio	Costs and effects from a defined perspective Effects in natural units	ICERs of competing investments (in same effect units)	Drummond 2015 [46]
Social return on investment	SROI	All costs and benefits converted to monetized form, plus wider social impact	Understanding of relative weight or importance of additional social benefits	Olsen 2003 [47]

The complexity and variety of DHIs and wide range of potential decision makers necessitates that the concept of "Return on Digital Investment" utilizing the DHI economic evaluation framework does not prescribe a particular form of aggregation, however it does require that researchers and decision makers are cognizant of the options available, the additional information required for their interpretation, and the risk of creating false certainty through obfuscation of important gaps in evidence, attribution, and valuation.

The use of the Impact Inventory in addition to value aggregation will assist in transparent representation of results, and it is expected that a comprehensive assessment of the Return on Digital Investment will involve a series of quantitative and descriptive outputs.

6. Using the DHI Economic Evaluation Framework

6.1 Case Example ONE

The overarching concept when applying the framework is to "comply or justify" – that is, users should attempt to adhere to the steps and principles of the framework, but where this is not feasible or appropriate, should provide justification for the methodological choice. In this way, the framework will facilitate methodological transparency over conformity, improving the overall usefulness of economic evaluations of DHIs.

This section details the initial planning steps for application of the framework to the evaluation of an information system and predictive analytics intervention to serve as an example of how the framework can be used to plan the methodological approach. As more evaluations utilize the framework, further examples will enable refinement of methods specifications.

The example analytic plan utilizing the framework is an economic evaluation of the Orbit data capture system and the Zenysis data analytics platform compared to a pre-implementation scenario in a targeted population HIV prevention program in South Africa. Orbit is a web-based data collection and information management system for capturing programmatic performance information. Data entry is completed by health workers and enables the generation of multiple monitoring and evaluation indicators. The pre-implementation scenario against which Orbit is compared is an Excel-based information management system, where health workers would electronically send completed Excel templates monthly, entered from paper-based point of care data collection forms. Data collected through the Orbit system is utilized by Zenysis, an analytics interoperability platform that can integrate multiple sources of data, enabling data triangulation, processing, and visualization of fragmented data to improve planning and decision making. The economic evaluation would aim to assess Orbit and Zenysis as a combined intervention to inform potential scale-up or use within other country contexts.

Table 9 Framework Example 1- Data Analytics Intervention

FRAMEWORK COMPONENT	SPECIFICATION	CONSIDERATIONS
STEP 1: IDENTIFY THE DECISION MAKER AND CONTEXT		
	Global Fund and subsidiaries. South African HIV program targeting specific groups	Although the direct intention is to inform Global Fund programs, there is a secondary objective to provide general information to inform other potential funders or country governments
STEP 2. DETERMINE THE INTERVENTION TYPE		
	Data collection and organization and predictive analytics	Type: Data Services. (WHO Class 4.1: "data collection, management and use") with predictive analytics component.
STEP 3. ESTABLISH THE LEVEL OF COMPLEXITY		
Intervention	Low complexity	Relatively simple data capture and analytics intervention
Outcomes	High complexity	A range of process and health outcomes
Causal pathway	High complexity	The pathway establishing links from immediate outcome measures to summary metrics of health are highly complex and uncertain with limited evidence of attribution. This complexity likely to affect the approach to value aggregation
STEP 4: ANALYTICAL PRINCIPLES		
Standing	Patients; health workers; program implementers and sub-recipients; Global Fund	
Comparator	Routine data collection (paper based with manual entering into MS Excel)	
Timeframe	3-year time horizon; disaggregated (determined based on decision maker's requirements)	
Costs	Implementation and recurring costs; cost offsets for reduced health system utilization	
Benefits	Process efficiencies in data gathering, Health worker user experience, Speed and precision of decision making, Health impact of improved care	
Predictive analytics	Identification of "high need" recipients and targeted package of interventions based on characteristics	

Table continued…

Table 9 Framework Example 1- Data Analytics Intervention (continued)

FRAMEWORK COMPONENT	SPECIFICATION	CONSIDERATIONS
STEP 5: VALUE PROPOSITION		
Value aggregation	Limited to financial return on investment	With available evidence and complexity of causal pathway it is not possible to represent aggregated health impacts required for net benefits or an incremental cost effectiveness ratio. Decision maker can observe financial return on investment with clear identification of impacts beyond the value aggregation
Impact inventory	*Health value*	No directly identifiable health impact
	Process value	Valuation of process improvements and program efficiencies
	Value of data	Estimate the unique value that the generated information represents to identified process efficiencies, and potential value to other similar programs*
	Prediction value	Qualitative representation of predictive analytics functionality*
	Cross-sectoral value	No directly identifiable cross-sector value
Uncertainty	One-way parameter sensitivity analysis; scenario analysis	
Distribution		Limited sub-population analysis as this is targeted program; represent the costs and consequences annual over analytical period

S*ource:* Authors, World Bank.

*Note: *Current methods likely limited to descriptive statistics, further methodological development on quantification of value required*

6.2 Case Example TWO

The second case example is a potential approach for economic evaluation of an intervention to utilize deep learning neural networks to interpret chest radiography (CXR) to screen and triage patients with pulmonary tuberculosis (TB). The example draws on an existing assessment of the diagnostic accuracy of three systems "CAD4TB", "qXR" and "Lunit" for a hypothetical decision for scale up of the interventions in the Cameroon country context, funded by domestic health resources. The existing retrospective assessment investigates the diagnostic accuracy of the systems using CXR images collected from Nepal and Cameroon as part of different studies. Included populations were adults (aged 15 years or older) with symptoms suggestive of TB (cough more than 2 weeks, fever, night sweats, weight loss). While the existing evaluation provides useful insights into the accuracy of the different systems compared to existing methods of radiologist interpretation of CXR, an economic evaluation would increase the understanding and representation of the different components of value for the interventions. In this example, it is assumed the government seeks information to guide allocatively efficient use of the limited health budget.

Table10 Framework Example 2 – AI-enabled diagnostic:

FRAMEWORK COMPONENT	SPECIFICATION	CONSIDERATIONS
STEP 1: IDENTIFY THE DECISION MAKER AND CONTEXT		
	Country government, Cameroon	Existing evidence is from Nepal and Cameroon, analysis would isolate Cameroon data only and/or consider the transferability of data from Nepal to Cameroon setting
STEP 2. DETERMINE THE INTERVENTION TYPE		
	Deep Learning system to analyze digital chest radiographs for TB-related abnormalities	Type: Health Service User targeted. (WHO Class 2.10: "laboratory and diagnostics imaging management") with predictive analytics component.
STEP 3. ESTABLISH THE LEVEL OF COMPLEXITY		
Intervention	High complexity	Depending on operation and interface the intervention may represent low complexity to the user
Outcomes	High complexity	For immediate diagnostic accuracy there is low complexity. Improved TB management and increased health is high complexity
Causal pathway	High complexity	

Table continued…

Table 10 Framework Example 2 – AI-enabled diagnostic (continued)

FRAMEWORK COMPONENT	SPECIFICATION	CONSIDERATIONS
STEP 4: ANALYTICAL PRINCIPLES		
Comparator	Human radiologist with Xpert as confirmatory	
Time horizon	Lifetime	Additional, shorter time horizon depending requirements of decision maker
Standing	Patients, Radiologists, other hospital staff, government; general population	
Costs	Equipment and introduction costs of an DL system; cost offsets of improved targeting and reduction in TB care	
Benefits	Increase capacity to aid TB diagnosis; Reduce the number of follow-on tests	
Predictive analytics	The predictive analytics component is a key feature of the intervention	ROC curves from initial evaluation provides baseline accuracy estimates compared to the human radiologist + Xpert (the established comparator)

COMPONENT	APPROACH
STEP 5: VALUE PROPOSITION	
Health value	Intermediate impacts on improved TB treatment outcomes; identify any impact on reduction in mortality and morbidity to develop utility measure
Process value	Improved diagnostic accuracy, and patient flow through diagnostic and treatment pathway
Value of data	Generate data on localized TB prevalence
Prediction value	The unique prediction value can be assessed using ROC curve approach
Cross-sectorial value	Limited immediate cross sectorial impacts including labor supply and worker productivity, however any impact on social sector through improved TB outcomes should be considered
Uncertainty	One-way parameter sensitivity analysis; scenario analysis, probabilistic sensitivity analysis depending on information/analytical resources
Distribution	Represent impact distribution by income and socioeconomic status

7. Implications for Research and Policy

Any methodological guidance is only as good as the extent to which it improves the information available to decision makers. As a global public good, it is envisioned that investors, product developers and researchers will adopt and utilize this framework as a thought aid when planning, conducting and interpreting economic evaluations for DHIs, building on the principle-based concepts of the International Decision Support Initiative Reference Case for Economic Evaluation [2].

The framework is intended to integrate with existing guidance, and the ongoing application of the framework into specific evaluations of digital health interventions will provide methodological consistency and further opportunity to refine the framework specifications. The WHO Guide to Monitoring and Evaluating DHIs [21] notes the role of economic evaluation within the stages in the maturity life-cycle of an intervention. Within the life cycle, an intervention can be at pre-prototype/protype; pilot; demonstration; scale-up; or integration and sustainability. It is expected that this framework will be used in conducting economic evaluation predominantly at the demonstration and scale-up phases and following pilots. However, economic evaluation can be a useful tool for informing ongoing investment and modifications to established DHIs within the integrations/sustainability phase.

The framework is not developed as a "rule book" to constrain economic evaluation practice or create an unachievable standard to which economic evaluations of DHIs should be conducted. Nor is it a "textbook", seeking to replicate the extensive field of methodological practice in economic evaluation. It is intended that the framework can be applied to evaluations conducted across a range of resource, time, and evidence constraints. As the DHI landscape and evidence base is rapidly evolving, methodological innovation in the economic assessment of DHIs will be necessary and it is expected that the framework will enable the development of a prioritized methodological research agenda. The overarching concept when applying the framework is to "comply or justify" – that is, users should attempt to adhere to the steps and principles of the framework, but where this is not feasible or appropriate, should provide justification for the methodological choice. In this way, the framework will facilitate methodological transparency and improve the overall usefulness of economic evaluations of DHIs.

References

[1] M. F. Drummond, M. J. Sculpher, K. Claxton, G. L. Stoddart, and G. W. Torrance, *Methods for the Economic Evaluation of Health Care Programmmes*, 4th ed. Oxford: Oxford University Press, 2015.

[2] T. Wilkinson *et al.*, “The International Decision Support Initiative Reference Case for Economic Evaluation: An Aid to Thought,” *Value Heal.*, vol. 19, no. 8, pp. 921–928, 2016, doi: 10.1016/j.jval.2016.04.015.

[3] G. D. Sanders *et al.*, “Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: Second panel on cost-effectiveness in health and medicine,” *JAMA - J. Am. Med. Assoc.*, vol. 316, no. 10, pp. 1093–1103, 2016, doi: 10.1001/jama.2016.12195.

[4] L. A. Robinson *et al.*, “Reference Case Guidelines for Benefit-Cost Analysis in Global Health and Development,” *Boston, MA Cent. Heal. Decis. Sci. Harvard TH Chan Sch. Public Heal.*, 2019.

[5] P. McNamee *et al.*, “Designing and Undertaking a Health Economics Study of Digital Health Interventions.,” *Am. J. Prev. Med.*, vol. 51, no. 5, pp. 852–860, Nov. 2016, doi: 10.1016/j.amepre.2016.05.007.

[6] K. Kolasa and G. Kozinski, “How to Value Digital Health Interventions? A Systematic Literature Review,” *Int. J. Environ. Res. Public Health*, vol. 17, no. 6, p. 2119, 2020.

[7] NICE, “Evidence Standards Framework for Digital Health Technologies,” 2019. doi: 10.1007/978-1-349-95810-8_867.

[8] K. Ho, “Health in the Digital World: Transformational Trends,” in *Healthcare and the Effect of Technology: Developments, Challenges and Advancements*, IGI Global, 2010, pp. 1–12.

[9] World Health Organization, “Classification of digital health interventions v1. 0: a shared language to describe the uses of digital technology for health,” World Health Organization, 2018.

[10] B. Wahl, A. Cossy-Gantner, S. Germann, and N. R. Schwalbe, “Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings?,” *BMJ Glob. Heal.*, vol. 3, no. 4, p. e000798, 2018.

[11] C. Krittanawong, H. Zhang, Z. Wang, M. Aydar, and T. Kitai, “Artificial intelligence in precision cardiovascular medicine,” *J. Am. Coll. Cardiol.*, vol. 69, no. 21, pp. 2657–2664, 2017.

[12] D. L. Poole and A. K. Mackworth, *Artificial Intelligence: foundations of computational agents*. Cambridge University Press, 2010.

[13] A. Agrawal, J. Gans, and A. Goldfarb, *Prediction machines: the simple economics of artificial intelligence*. Harvard Business Press, 2018.

[14] USAID, “ARTIFICIAL INTELLIGENCE IN GLOBAL HEALTH Defining a Collective Path Forward,” 2019.

[15] M. Görgens *et al.*, “Study Protocol: An Evidence Gap Map for Digital Health Interventions,” *World Bank Gr.*, 2020.

[16] World Bank, “Cost Benefit Analysis in World Bank Projects,” Washington DC, 2010.

[17] M. Gold, J. Siegel, L. Russell, and M. Weinstein, *Cost-effectiveness in health and medicine*. Oxford University Press, 1996.

[18] A. Vassall, S. Sweeney, J. Kahn, and J. Gomez, “GHCC | Global Heath Cost Consortium,” 2017. .

[19] R. C. Hutubessy, R. M. Baltussen, T. T. Torres-Edejer, and D. B. Evans, “Generalised cost-effectiveness analysis: an aid to decision making in health,” *Appl. Health Econ. Health Policy*, vol. 1, no. 2, p. 89—95, 2002.

[20] D. Husereau *et al.*, “Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement,” *Eur. J. Heal. Econ.*, vol. 14, no. 3, pp. 367–372, 2013, doi: 10.1007/s10198-013-0471-6.

[21] WHO, *WHO | Monitoring and evaluating digital health interventions*. 2016.

[22] EUnetHTA Joint Action Work Package 8, “EUnetHTA HTA Core Model version 3.0,” 2016.

[23] X. Liu, S. C. Rivera, D. Moher, M. J. Calvert, and A. K. Denniston, “Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI Extension,” *Bmj*, vol. 370, no. October 2019, p. m3164, 2020, doi: 10.1136/bmj.m3164.

[24] S. C. Rivera, X. Liu, A.-W. Chan, A. K. Denniston, and M. J. Calvert, “Consensus statement Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension The SPIRIT-AI and CONSORT-AI Working Group*, SPIRIT-AI and CONSORT-AI Steering Group and SPIRIT-AI and CONSORT-AI Cons,” *Nat. Med.*, vol. 26, no. 9, pp. 1351–1363, 2020, doi: 10.1038/s41591-020-1037-7.

[25] A. van Giessen *et al.*, “Systematic Review of Health Economic Impact Evaluations of Risk Prediction Models: Stop Developing, Start Evaluating,” *Value Heal.*, vol. 20, no. 4, pp. 718–726, 2017, doi: 10.1016/j.jval.2017.01.001.

[26] E. Murray *et al.*, “Evaluating Digital Health Interventions: Key Questions and Approaches,” *American Journal of*

Preventive Medicine. 2016, doi: 10.1016/j.amepre.2016.06.008.

[27] T. Wilkinson *et al.*, "The Reference Case for Economic Evaluation. Methods for Economic Evaluation Project Final Report," 2014.

[28] S. Ollis, V. Dwivedi, M. Sitjar, A. Waugaman, B. Weiss, and M. Schaefer, "Digital Health Investment Review Tool," 2018.

[29] M. Petticrew, "When are complex interventions 'complex'? When are simple interventions 'simple'?," *Eur. J. Public Health*, vol. 21, no. 4, pp. 397–398, 2011.

[30] P. Craig, P. Dieppe, S. Macintyre, S. Mitchie, I. Nazareth, and M. Petticrew, "Developing and evaluating complex interventions: The new Medical Research Council guidance," *Bmj*, vol. 337, no. 7676, pp. 979–983, 2008, doi: 10.1136/bmj.a1655.

[31] Medical Research Council (MRC) and National Institute of Health Research (NIHR), "Updated guidance: developing and evaluating complex interventions [draft]," p. 99, 2019.

[32] M. Petticrew *et al.*, "Complex interventions and their implications for systematic reviews: a pragmatic approach," *J. Clin. Epidemiol.*, vol. 66, no. 11, pp. 1209–1214, 2013.

[33] M. F. Drummond, M. J. Sculpher, K. Claxton, G. L. Stoddart, and G. W. Torrance, *Methods for the economic evaluation of health care programmes*. Oxford university press, 2015.

[34] C. Gena, "Methods and techniques for the evaluation of user-adaptive systems," *Knowl. Eng. Rev.*, vol. 20, no. 1, pp. 1–37, Mar. 2005, doi: 10.1017/S0269888905000299.

[35] M. Asaria, S. Griffin, and R. Cookson, "Distributional cost-effectiveness analysis: a tutorial," *Med. Decis. Mak.*, vol. 36, no. 1, pp. 8–19, 2016.

[36] S. Verguet, J. J. Kim, and D. T. Jamison, "Extended Cost-Effectiveness Analysis for Health Policy Assessment: A Tutorial," *Pharmacoeconomics*, vol. 34, no. 9, pp. 913–923, Sep. 2016, doi: 10.1007/s40273-016-0414-z.

[37] S. Athey, Bergstrom K., Hadad V., Jamison J.C., Ozler B., Parisotto L., and Sama J.D. 2021. "Shared decision making: Can improved counseling increase willingness to pay for modern contraceptives?," World Bank Policy Research Working Paper Series #9777.

[38] L. A. Robinson, J. K. Hammitt, and L. O'Keeffe, "Valuing Mortality Risk Reductions in Global Benefit-Cost Analysis," *J. Benefit-Cost Anal.*, pp. 15–50, 2019, doi: 10.1017/bca.2018.26.

[39] R. Durrett *et al.*, "Measuring Efficiency in Health Care," 2006, Accessed: Dec. 19, 2022. [Online]. Available: https://econpapers.repec.org/RePEc:cup:cbooks:9780521851442.

[40] D. Whittington and J. Cook, "Valuing Changes in Time Use in Low- and Middle-Income Countries," *J. Benefit-Cost Anal.*, vol. 10, pp. 51–72, 2019, doi: 10.1017/bca.2018.21.

[41] A. M. Bayoumi, "The measurement of contingent valuation for health economics," *Pharmacoeconomics*, vol. 22, no. 11, pp. 691–700, 2004, doi: 10.2165/00019053-200422110-00001.

[42] K. M. Strzepek, C. Amanya, and J. E. Neumann, "Assessing Economy - wide Effects of Environmental and Health Interventions in Support of Benefit - Cost Analysis," 2018.

[43] A. Ghorbani and J. Zou, "Data shapley: Equitable valuation of data for machine learning," *36th Int. Conf. Mach. Learn. ICML 2019*, vol. 2019-June, pp. 4053–4065, 2019.

[44] Kelle O'Neal, "How to Measure Data as an Enterprise Asset | First San Francisco Partners." https://www.firstsanfranciscopartners.com/blog/how-to-measure-data-as-an-enterprise-asset/?cn-reloaded=1 (accessed Sep. 11, 2020).

[45] A. Agrawal, J. S. Gans, and A. Goldfarb, "Artificial intelligence: The ambiguous labor market impact of automating prediction," *J. Econ. Perspect.*, vol. 33, no. 2, pp. 31–50, 2019, doi: 10.1257/jep.33.2.31.

[46] S. G. L. and T. G. . Drummond M.F., Sculpher, M.J., Claxton K., *Methods for the economic evaluation of health care programmes*. Oxford University Press, 2015.

[47] S. Olsen, "Social return on investment: Standard guidelines," 2003.