

# Integrating Data to Evaluate a Global Health Grand Challenge

---

**Blen Biru, Andrea Taylor, Sowmya Rajan, Kate Crissman,  
Osondu Ogbuoji, Fernando Fernholz, Siddharth Dixit,  
Minahil Shahid, Pratik Doshi, and Krishna Udayakumar**  
*Duke University, North Carolina, United States*

**Amy Finnegan**  
*Duke University, North Carolina, United States;  
IntraHealth International, North Carolina, United States*

**Joy Noel Baumgartner**  
*University of North Carolina, North Carolina, United States*

**Abstract:** *This article describes the integrated, mixed methods (MM) design used to evaluate the Saving Lives at Birth (SL@B) program. SL@B is a multi-stakeholder, donor-supported global health initiative to tackle maternal and neonatal mortality via innovation. Since SL@B's launch in 2011, the program has supported 116 innovations through 147 awards around the globe. The evaluation for this large and complex program included a largely retrospective MM design aligned with principles of evaluating complexity. This paper highlights these MM evaluation strategies and integration dimensions employed to complete the SL@B evaluation that could inform future evaluations of portfolio-level global health programs.*

**Keywords:** *complex evaluation, data integration, data triangulation, global health, mixed methods evaluation*

**Résumé :** *Cet article décrit les méthodes mixtes (MM) intégrées employées lors de l'évaluation du programme Sauver des vies à la naissance (SL@B). SL@B est une initiative mondiale pour la santé, soutenue par des donateurs, visant à lutter contre la mortalité maternelle et néonatale grâce à l'innovation. Depuis le lancement de SL@B en 2011, le programme a appuyé 116 innovations par l'intermédiaire de 147 subventions à travers le monde. L'évaluation de ce programme vaste et complexe a compris un cadre de MM, largement rétrospectif, conforme aux principes d'évaluation de la complexité. Le présent article souligne ces stratégies d'évaluation par MM et les dimensions d'intégration utilisées pour compléter l'évaluation SL@B*

**Corresponding author:** Blen Biru, Research Associate, Duke Global Health Innovation Center, Duke Global Health Institute, 310 Trent Dr, Durham, North Carolina, United States, 27710; [blen.biru@duke.edu](mailto:blen.biru@duke.edu)

*qui pourraient éclairer les évaluations futures des programmes du portefeuille de santé mondiale.*

**Mots clés :** *santé mondiale, évaluation de méthodes mixtes, intégration de données, triangulation de données, évaluation complexe*

Mixed methods (MM) design has become increasingly critical for conducting high-quality evaluations, yet the literature indicates that a knowledge gap exists in understanding how to accomplish the requisite data integration (Farmer et al., 2006; O’Cathain et al., 2010). Integration has different definitions and is often used to describe integration at the data-combining stage (Fetters & Molina-Azorin, 2017). However, the data-focused definition may be limiting, as integration can take place at different stages or dimensions of an evaluation. In this article, we share illustrative examples of how our team designed a fully integrated MM study to evaluate a complex global health program. We use the term “fully integrated” as defined by Creamer (2017)—a MM approach where integration of qualitative and quantitative data happens at every stage of a research study.

The methods we employed align with published guidance and propositions for evaluating complexity which highlight that a system is always changing and could be dependent on other surrounding factors, thereby emphasizing the need for evaluators to understand the whole system, including components that make up the system and their corresponding interactions (Institute of Medicine, 2014; Preskill et al., 2014). Our study was unique in drawing upon program and cost data in addition to commonly used methods including qualitative data (interviews) and quantitative data (surveys) to produce multiple analyses that enhanced and complemented each other to answer the key evaluation questions.

The article is organized in the following sections: background context describing the program being evaluated; research design and deliverables, including the processes and integration techniques involved; challenges our evaluation team experienced navigating a complex evaluation; lessons learned that could be helpful to other evaluators working on similar evaluation projects; and a conclusion.

## EVALUATION CONTEXT

Our team at Duke University was tasked with evaluating a Global Health Grand Challenge—Saving Lives at Birth (SL@B) (<https://www.grandchallenges.ca/programs/saving-lives-at-birth/>). The SL@B program funds a community of innovators with diverse backgrounds working to reduce global maternal and newborn mortality through advancements in science and technology, improved service delivery models, and demand-driven approaches. SL@B is funded by a consortium of donors including the US Agency for International Development (USAID), Grand Challenges Canada (GCC), the Norwegian Agency for Development Cooperation (NORAD), the Bill and Melinda Gates Foundation (BMGF), the United Kingdom’s Department for International Development (DFID), and the Korea International Cooperation Agency (KOICA).

Since SL@B's launch in 2011, it has supported 116 innovations through 147 funding awards (with some innovations receiving multiple awards), addressing preventable causes of maternal and newborn mortality such as post-partum hemorrhage, pre-eclampsia, newborn sepsis, and hypothermia. The funded innovations included devices and diagnostics (61%), mHealth solutions (11%), drugs and vaccines (12%), and service delivery practices/approaches (16%). To source these innovations, SL@B used an open-call approach to solicit innovative ideas in the maternal and newborn health (MNH) field with no restrictions on type of organizations, geographic settings, or specific diseases. Most awards went to academic institutions (40%) and nonprofit organizations (38%), followed by for-profit (20%) and public international organizations (2%). The majority of these awards went to organizations with headquarters based in high-income countries (HICs) (83%) as opposed to low- and middle-income countries (LMICs) (17%). The SL@B program funded innovations at different stages of growth through three types of awards: seed—around \$250,000 US for up to two years to support development of early stage ideas, validation—around \$250,000 US for up to two years to test the effectiveness of an innovation, and transition-to-scale—around \$2 million US for up to four years to support scaling an innovation (Lalli et al., 2018).

From 2018 to 2020, two of the SL@B partners, USAID and GCC, sponsored an external evaluation to understand SL@B's impact on sourcing, supporting, and scaling innovations across the world to reduce newborn and maternal mortality. From this evaluation, they wanted to obtain data-driven insights on the program's achievements and unrealized potential to make informed decisions regarding a potential future iteration of the program. To that end, our team, comprising researchers with varied expertise including maternal and neonatal health, global health innovation, health economics, business, public policy, social work, demography, and evaluation methods, developed an evaluation plan. Our evaluation team also established an external advisory board composed of experts in the maternal and neonatal health and investment field, who provided guidance in designing and conducting the evaluation. Below were the primary evaluation questions of interest to USAID and GCC:

1. How does SL@B map onto the global landscape of maternal and neonatal (MNH) innovation?
2. Does SL@B fill a gap in MNH innovation funding?
3. What has been the impact of SL@B on sourcing and scaling MNH innovations between 2011 and 2018?
4. What is the potential impact of SL@B-funded innovations on MNH mortality and in the MNH ecosystem?

Having varied expertise not only allowed our team to design a MM study appropriate for the evaluation questions, but it also contributed to more robust analyses, interpretations, and dissemination. The team collaboration was made possible through organized weekly calls to check in on sub-group workstreams and also

whole-group brainstorming workshops when needed. [Fetters and Molina-Azorin \(2017\)](#) define this integration in the team dimension as “creating and orchestrating mixed methods teams, e.g. individuals of different disciplinary backgrounds, qualitative, quantitative, and mixed methods researchers and optimizing team contributions to produce best quality mixed methods research” (p. 294).

## EVALUATION RESEARCH DESIGN AND DELIVERABLES

### *Complex MM design*

To answer the multiple evaluation questions, our team designed a retrospective MM study, which aligns with other published literature on propositions for complexity evaluation ([Institute of Medicine, 2014](#); [Preskill et al., 2014](#)). These propositions emphasize the need to explore the intricacy and interdependencies within a system, given that a program’s success relies not only on its elements but also on its relationship with relevant stakeholders and the ecosystem in general. In evaluating SL@B, we recognized that there were numerous other players in the maternal and newborn health ecosystem contributing to the success, or lack thereof, of the program. To capture these nuances and different stakeholders’ perspectives, we employed a MM design. [Table 1](#) details the data sources and methods.

The complex MM design was intended both to lead to a deeper understanding of the program and its impact and to include a range of stakeholder perspectives, thereby increasing the rigor of the study. This mixing rationale for complementarity purposes is also the reason that most studies employ MM ([Archibald et al., 2015](#)). Our team also anticipated that different types of evaluation data and deliverables might resonate more or less with different types of stakeholders. This process of identifying the overarching purpose of the mixed methods and outlining what can be achieved by individual methods is described as integration in the research design dimension by [Fetters and Molina-Azorin \(2017\)](#). [Table 1](#) presents the purpose of each evaluation deliverable and corresponding evaluation methods used in this study. All the research methods and data collection strategies were approved by the Duke University Campus Institutional Review Board (IRB) under three separate protocol numbers (2018-0370, 2018-0617, and 2019-0546, for existing data, interviews, and quantitative survey, respectively).

Given the need to answer multiple evaluation questions with different methods, our team used different triangulation techniques. Triangulation is a method used for analysis when multiple sources, theories, and researchers are involved and is considered important to enhance the robustness and credibility of research findings ([Farmer et al., 2006](#); [Institute of Medicine, 2014](#); [O’Cathain et al., 2010](#)). According to [Denzin \(1978\)](#), triangulation techniques can further be classified as follows:

- methodological triangulation—when more than one research or data collection technique is used, e.g., interviews and program data;

**Table 1.** Data methods, integration and project deliverables

Evaluation methods	Data source, purpose, sample size	Project deliverables (each was a stand-alone document)					
		1) Portfolio review	2) Qualitative key findings	3) Partnership brief	4) Acceleration to impact brief	5) Four separate CEAs	6) Final report
<b>Qualitative</b>							
Key informant in-depth interviews	MNH funders and SL@B grantees, to understand funding landscape ( $n = 21$ ) MNH experts, to evaluate SL@B's success and areas for improvement ( $n = 40$ ) SL@B innovators, to understand their path to market/ acceleration ( $n = 11$ ) SL@B innovators and partners, to identify formed partnerships ( $n = 18$ )	X	X		X		X
<b>Quantitative</b>							
Online survey	Sent to all SL@B grantees, to benchmark innovators' scaling pathways and evaluate SL@B's support through their growth ( $n = 61$ , which represents a 54% response rate from all innovators)			X			X
<b>Other methods</b>							
Program data matrix ( $N = 300+$ documents)	SL@B grantees' reports, to describe SL@B's portfolio challenges SL@B grantees' reports, to describe SL@B partnerships	X	X	X	X		X
Desk review	Literature review, to understand global landscape of MNH innovations and scaling pathways	X			X		X
Cost data	SL@B grantees, to obtain historical costs and scale up projections ( $n = 4$ )					X	X
Impact models	GCC models, to validate number of lives saved/ improved estimates ( $n = 4$ )					X	X

- data triangulation—multiple data sources, e.g., different types of reports or multiple respondent groups like innovators and investors;
- investigator triangulation—involvement of two or more researchers in the analysis; and
- theoretical triangulation—alternative disciplinary or theoretical lenses to interpret research findings, e.g., greatest public health impact vs. the innovations’ business sustainability.

In our evaluation research, investigator triangulation was used for all of the deliverables, given that our team is composed of inter-disciplinary professionals. Theoretical triangulation overlapped with investigator triangulation in that our team members brought different lenses to the data interpretation and analysis depending upon their expertise. Methodological triangulation and data triangulation were also employed for different deliverables and are further discussed below. These triangulation methods served to produce a more comprehensive evaluation. Research completed by Archibald (2016) also supports the idea that investigator triangulation is beneficial in multiple ways: “expertise may be leveraged, lone-investigator deficits compensated for, and data and inference quality enhanced” (p. 246).

### ***Evaluation research deliverables***

#### *Portfolio review / landscape analysis*

The first evaluation product our team focused on creating was a Portfolio Review / Landscape Analysis (Taylor, Rajan, et al., 2020) that was designed to understand SL@B’s composition of innovations and how they were situated within the global landscape of maternal and newborn health innovations and funding opportunities. The methods used in the portfolio analysis relied heavily on reviews of program data to validate a matrix initially developed by the sponsors to categorize the SL@B portfolio into different types and stages of innovations. The matrix classified innovation stages as Ideation, Research and Development, Proof of Concept, Transition to Scale, Scaling, and Sustainable Scale (IDIA, 2017). Additionally, the matrix identified where the headquarters of the innovations are located, HIC versus LMIC, countries of innovation implementation, specific causes of morbidity/mortality targeted, types of innovation (devices, diagnostics, mHealth solutions, drugs and vaccines, service delivery practices/approaches), and other relevant variables. Our team validated the matrix and expanded upon the data included for further analyses. The program data matrix review was instrumental in facilitating purposeful sampling for the methods used in this study. We used important characteristics identified through the program review such as types and stages of innovation in order to have a representative sample for the different data-collection methods used. Fetters and Molina-Azorin (2017) label this process as integration in the data collection dimension, specifically “integration through connecting,” where findings from one type of data-collection method can be used to determine the data collection/sampling of another research method.

To further understand SL@B's innovation portfolio in a global context, desk research on publicly available data of 227 MNH innovations and 25 funders of MNH innovations was conducted. Desk research entailed looking for publicly available information on innovations to understand the MNH funding landscape, for example, number of awards, amounts invested, and prioritized innovation growth stages, as well as conducting subsequent interviews with 12 of the 25 identified funders. An additional nine interviews with SL@B grantees and intermediary organizations were also conducted. The descriptive landscape analysis of the SL@B program in the larger ecosystem was complemented by the qualitative findings from the key informant interviews (KIIs) to highlight where SL@B fit in the MNH space. To triangulate the data, we completed “core integration analytics”—an interactive back-and-forth analysis within these three data sets: program data matrix, desk research, and KIIs that led to complementary findings. Triangulating the data and synthesizing data-driven insights from the different methods lent credibility to the final product and increased our confidence in the findings and our subsequent recommendations.

### *Summary of qualitative key findings*

Our second evaluation product—Summary of Qualitative Key Findings—was used to answer the first three evaluation questions from the perspective of key stakeholders, which included MNH experts, USAID leadership, government representatives, SL@B implementation partners, MNH investors, and SL@B innovators. The purposeful sampling strategy used to identify different stakeholders illustrates integration through sampling dimension and is referred to as “multilevel strategy,” where different types of data can be gathered from various social levels or types of stakeholders, for example, survey data from innovators and interview data from funders (Fetters & Molina-Azorin, 2017).

A total of 40 KIIs were conducted in different ways: in person by our team, at the USAID-hosted DevelopmentXChange (DevX) event in Washington, D.C., in 2018 (annual networking event organized by SL@B), in Ethiopia and Kenya during site visits, and/or via phone calls. The semi-structured interview guide was developed based on SL@B's Theory of Change (TOC) (Lalli et al., 2018) and included questions such as these: What have been the major achievements of SL@B? Has there been any unrealized potential for the program? What are the markers of a successful innovation? What are some of the bottlenecks to scaling an innovation in MNH? Does SL@B's sourcing and procurement meet the needs of the MNH community?

All key informants provided written consent to participate in the study. Interviews were audio-recorded, transcribed, and analyzed using a pre-defined thematic framework. The team created a codebook based on the interview guide and initial reading of sample interviews from each key informant groups by four team members. Using the codebook, two members coded and analyzed the data using NVivo12 software. Inter-coder reliability was established at greater than 90% for the first 10% of interviews that were double-coded. Disagreements were addressed through discussions between the two coders until consensus was reached. During

the analysis of the coded interviews, there were two rounds of memoing. The first was completed by grouping themes based on different respondent groups: investors, innovators, MNH experts, and HIC-based versus LMIC-based key informants; and the second was by aggregating all respondent groups given that there were high levels of convergence in the themes across groups.

One of the techniques in the data analysis integration dimension stated by [Fetters and Molina-Azorin \(2017\)](#), “core integration analytics,” was used to integrate the matrix program data with the qualitative insights for complementarity purposes. As a result, the report provided rich qualitative insights including direct quotes from KIs on SL@B’s achievements and areas for improvement from various stakeholders. The sponsors were not accustomed to this level of analytic rigor for qualitative data in an evaluation, but we believed it was warranted, given the potential biases that could have surfaced regarding overemphasizing particular stakeholder perspectives. Furthermore, this rigorous analysis positioned our team to be able to publish these findings in peer-reviewed journals and contribute to the advancement of the MNH ecosystem literature ([Finnegan et al., 2021](#)).

#### *Acceleration to impact brief*

Our team completed a second qualitative product focused on innovators’ pathways to scale—the Acceleration to Impact Brief ([Taylor, Biru, et al., 2020](#)). This brief highlighted SL@B’s as well as other funders’ support in accelerating market entry for MNH innovations and identified areas for improvement related to supporting innovators in their pathway to scale ([USAID Center for Accelerating Innovation and Impact, 2016](#)). A sample of 11 SL@B grantees were purposively selected to represent a range of stages and types of innovations and organizations based in three of the SL@B target markets: Kenya, Ethiopia, and India. This sub-group sampling from a larger sample illustrates integration through sampling dimension, specifically referred to as “nested strategy” by [Fetters and Molina-Azorin \(2017\)](#). All interviews were conducted virtually and included questions such as the following: Do you feel that SL@B helped your organization get to market sooner than you might have otherwise? Do you have suggestions for how SL@B or other similar programs could better support innovations to get to market and scale?

Similar to the Summary of Qualitative Key Findings Brief, coding, inter-coder reliability, and thematic analysis were completed upon recording and transcribing interviews. This qualitative analysis was complemented by program data pulled from KIIs’ innovation reports to provide descriptive statistics of the portfolio, as well as desk research on scaling pathways of health-care innovations. Core integration analytics was employed for interactive and back-and-forth analysis from the different data sets, resulting in complementary insights.

The brief illustrated factors that SL@B innovators found helpful through their growth stages and scaling trajectories. It also provided recommendations for SL@B and other funders on addressing the gaps identified in supporting innovations on the path to market. This particular product was designed for deeper insights on selected innovations, not necessarily portfolio-level findings. The

triangulation of methods served to strengthen team members' responses to the sponsors' queries about innovation-specific insights.

### *Partnership brief*

The Partnership Brief ([Rajan et al., 2020](#)) explored the role that SL@B has played in establishing connections for its innovators and used data from interviews conducted as well as survey responses relevant to partnerships SL@B innovators created in their pathway to scale. Our team conducted 18 interviews (8 in person in India and 10 virtually) with SL@B grantees and other stakeholders. Illustrative interview questions included the following: Could you talk about the purpose of the partnerships formed? Have there been any challenges that you have overcome forming and sustaining the partnerships?

Insights from KIIs, quantitative survey data regarding partnerships formed by innovators, and program data/innovator reports were mixed to write the partnership brief, which included descriptions and facilitators of partnerships formed by SL@B innovations, gaps identified, as well as recommendations for SL@B and other funders to better promote partnerships for innovations in the MNH field. Intra-method analysis was employed where the different data sources were analyzed independently; interview data were analyzed qualitatively using Nvivo, and survey data quantitatively using R and Excel.

Insights from these data sources were presented independently in the first section of the report. In the last section of the report, the insights were consolidated and presented, prefaced by the following sentence: "We draw on interviews and survey data to highlight three data-driven recommendations applicable not only to the SL@B program, but also to the broader funding community and innovation ecosystem" ([Rajan et al., 2020](#), p.16). This integration in the interpretation dimension is defined as "confirmation" by [Fetters and Molina-Azorin \(2017\)](#)—"when the qualitative and quantitative findings give cause for drawing the same conclusion" (p. 302).

### *Cost effectiveness analysis (CEA)*

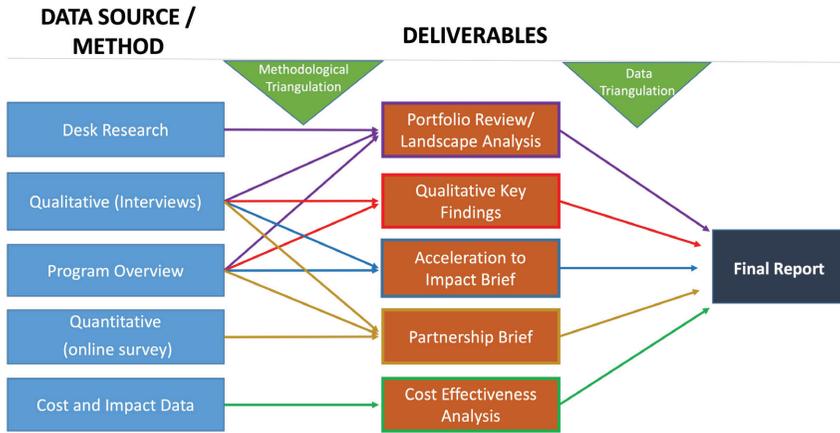
Cost Effectiveness Analyses were conducted for select SL@B innovations to evaluate SL@B's potential impact on MNH mortality. Most SL@B innovations were in the early stages of innovation: seed and validation, 63% and 19% of innovations respectively, which means that most of them did not have preliminary effectiveness data needed to conduct a CEA. SL@B funded a total of 26 transition-to-scale grants, four of which were included in the CEAs based on data availability (innovation's cost data and the estimated number of lives saved per innovation) and innovator's willingness to participate. Our team was not able to find an existing cost-data collection tool that was user-friendly for SL@B innovations or comprehensive enough to estimate the expected health impact of a project. We therefore created a new Excel-based costing tool that is applicable to different types of innovations and has minimal computational requirements, where innovators themselves could fill in the first draft ([Dixit et al., 2020](#)). The completion of the tool was followed by technical assistance from our team via phone and email as needed.

The estimated number of lives saved by a SL@B innovation was originally obtained from impact modeling work completed by GCC, and our team revised and validated the accuracy of these models before using them for further analysis in the CEAs using a newly created checklist. The checklist included standardized elements and assumptions (e.g., whether the model accounts for burden of disease, efficacy of innovation, quality of service delivery, etc.) that needed to be included or addressed in an impact model of an innovation. Once cost data were received from innovators and lives-saved estimates from GCC's impact models were validated, a CEA report was produced for each of the four innovations highlighting their health and social impact if scaled nationwide, projected until 2030, a year that was chosen to align with the Sustainable Development Goals (SDGs). The value of the CEAs was both at the individual level—for example, an innovator in India can convince the government or other funders that the innovation is cost-effective and worthy of investment—and at the portfolio level—for example, to determine if SL@B was able to source and support cost-effective MNH innovations. The CEAs were not able to contribute to the latter aim as much, given the small sample size and lack of generalizability.

### *Final report*

The deliverables discussed above drew mostly from at least two data sources and were released as intra-publications, that is, through multiple publications and crosslinking (Fetters & Molina-Azorin, 2017). Our team was tasked with a much more complex triangulation process when writing the final report (Biru et al., 2020). The team's strategy to write the final report followed a triangulation design, where all data sources (Figure 1) were initially analyzed independently and partially integrated for the stand-alone deliverables and later mixed and fully integrated into one final report (Denzin, 1978; Farmer et al., 2006). According to Erzberger and Prein (1997), triangulation helps researchers identify if convergence, complementarity, and dissonance exist between the multiple data sources considered. Most of the reports combined were found to be convergent and complementary. There were instances of silence—where one theme shows up in one of the reports and not in the other (Farmer et al., 2006; O'Cathain et al., 2010). The results in the final report were released as inter-publication—releasing findings through one publication (Fetters & Molina-Azorin, 2017). The report consolidated insights from the different reports and presented findings categorized by four overarching themes, following the structure of the four primary evaluation questions. To summarize the main take-away messages, an infographic was added in the executive summary of the final report.

The final report also had one major source of data that was not as prominent in the other deliverables: the online quantitative survey for SL@B innovators. The survey was designed to better understand innovators' pathway to scale. Some questions were adopted from existing funder surveys and some questions were added in order to fill a gap of information that the team observed during qualitative interviews to benchmark the time, amount of money, formed partnerships, and challenges of SL@B innovators in their pathway to scale. This process



**Figure 1.** Data integration process

falls under integration in the data collection dimension, specifically “integration through building,” where one type of data collection is informed by a previously collected data set (Fetters & Molina-Azorin, 2017). For example, while conducting the KIIs, we learned that little was known regarding the timeline for scaling global health innovations other than drugs. To further illuminate that qualitative finding, we included questions in our subsequent online survey to understand the average amount of time SL@B innovations take to progress through the different growth stages: seed-validation-scale. The survey was sent to all SL@B grantees, active and inactive ( $N = 116$ ), via Qualtrics and yielded a 54% response rate ( $n = 61$ ).

The main triangulation technique used in the final report was data triangulation, where multiple reports—Portfolio Review / Landscape Analysis, Summary of Qualitative Key Findings, Acceleration to Impact Brief, and Partnership Brief—were mixed. Similarly to all the other products, investigator and theoretical triangulation were also used where deliverable leaders took ownership of incorporating the analysis drawn from their deliverable. Core integration analytics was employed to pull information from the different data sets. The findings upon analysis illustrated integration at the interpretation dimension: confirmation—where a similar conclusion can be drawn from both the qualitative and quantitative results; complementarity—where distinct but non-contradictory stories emerge from the qualitative and quantitative data; and expansion—where combining the qualitative and quantitative data results in a more comprehensive data interpretation (Fetters & Molina-Azorin, 2017).

The integration of the four individual CEAs in the final report was difficult, given the small sample size and the lack of ability to generalize about cost-effectiveness and impact at a portfolio level. Therefore, the CEA insights that were included in the final report focused only on issues related to the process of conducting the CEAs and not actual findings from the CEA reports.

We had to reflect on the findings of each report and assess the added value that each method or report brought in order to make meta-inferences—conclusions drawn from integrated MM studies, as defined by [Teddlie and Tashakkori \(2009\)](#). This process resulted in a final report that the evaluation sponsors wanted and needed, which included valuable insights about the current SL@B portfolio, gaps identified, and achievements, in sourcing, supporting, and scaling MNH innovations, as well as recommendations for future directions, should the SL@B program continue. The strength of the evaluation was that it reflected the voices of different stakeholders involved and used rigorous methods and data integration to produce data-driven insights. However, there were limitations and challenges faced during the process that are discussed below.

[Table 2](#) summarizes the high-level triangulation techniques used in each deliverable.

**Table 2.** Summary of data and methodological triangulation techniques

Deliverable	Triangulation technique used
<b>Portfolio review / landscape analysis</b>	<p>Methodological triangulation</p> <ul style="list-style-type: none"> <li>multiple methods (program data matrix, interviews, desk research)</li> </ul> <p>Data triangulation</p> <ul style="list-style-type: none"> <li>multiple respondent groups (funders, intermediaries, and SL@B grantees)</li> </ul>
<b>Summary of qualitative key findings</b>	<p>Data triangulation</p> <ul style="list-style-type: none"> <li>multiple respondent groups (innovators, investors, and MNH experts)</li> </ul>
<b>Acceleration to impact brief</b>	<p>Methodological triangulation</p> <ul style="list-style-type: none"> <li>multiple methods (interviews, program data, and desk research)</li> </ul>
<b>Partnership brief</b>	<p>Methodological triangulation</p> <ul style="list-style-type: none"> <li>multiple methods (interviews, program data, and quantitative survey)</li> </ul> <p>Data triangulation</p> <ul style="list-style-type: none"> <li>multiple respondent groups (funders and SL@B grantees)</li> </ul>
<b>Cost effectiveness analysis (CEA)</b>	<p>Methodological triangulation</p> <ul style="list-style-type: none"> <li>multiple methods (cost data and impact modeling)</li> </ul>
<b>Final report</b>	<p>Data triangulation</p> <ul style="list-style-type: none"> <li>multiple reports (portfolio review/landscape analysis, summary of qualitative key findings, acceleration to impact brief, and partnership brief)</li> </ul>

## CHALLENGES FACED, LESSONS LEARNED, AND RECOMMENDATIONS

### *Challenge #1: Complexities inherent to retrospective portfolio-level evaluation*

The retrospective nature—starting an evaluation after the program had been ongoing and pivoting for eight years—added complexity to the evaluation. From 2011 to 2018, the SL@B program funded 116 unique innovations, creating a large and diverse portfolio. To that end, it was challenging to make findings generalizable given the nuances and outliers in the portfolio. To address this challenge, we had to acknowledge the diverse outcomes by each innovation within the portfolio, categorize the data into subgroups, and highlight findings accordingly, for example by types of innovations and locations of innovators.

On the innovators end, those who were involved in the evaluation study and the SL@B program more broadly, particularly the active SL@B grantees, may have experienced respondent fatigue due to multiple requests to contribute to different evaluation and program deliverables. Our team acknowledged that the innovators were receiving multiple asks, so we tried to reach out in advance and give them ample time to respond to the various requests. We also respected their decision if they were stretched and declined participation.

### *Challenge #2: (Un-)Incentivized primary data collection*

Our sponsors prioritized secondary data analyses when possible, and a significant portion of our data was gathered from existing resources (program data and desk research). However, some primary data collection was necessary. We had relative control over the key informant interviews, except for some travel limitations that made it difficult to gather data from the SL@B implementation partners in LMICs, where phone interviews were more challenging, for instance because of time zones and unreliable Internet connections. That said, the online survey and cost-data collection which were collected via email proved to be more difficult than the qualitative interviews. Trying to solicit participation in the quantitative survey from inactive grantees who did not have an incentive to participate, and synthesizing key stakeholder input from widely varying perspectives given their involvement with and understanding of SL@B at any given time over the program period, was an anticipated challenge.

For the online survey deployed to SL@B innovators, we anticipated that participation might be low, so we pre-emptively thought about a creative incentive—a stages of growth infographic where we aggregated a participant's responses on time spent, amount of money spent, and partnerships formed at each growth stage. At the beginning of the survey, we promised potential participants that this infographic would be sent to them within a few days to use as they wish. We do not know how much this offer may have influenced an innovator's willingness to complete the approximately one-hour survey, but it perhaps created some

goodwill in that we recognized that participation was a big ask for the inactive grantees in particular.

Besides the already noted challenges of the online survey, we needed cost data from innovators in order to conduct the CEAs. This required a significant time commitment from both our team and the innovators themselves. In all four CEAs pursued, the cost data took much longer to collect than either the research team or the innovators expected. The innovators were under no obligation to collaborate, so the timeline was dependent upon them. To address this challenge, our team gently and repeatedly reached out to offer technical assistance for those who recognized the value in completing the costing tool (Dixit et al., 2020), but it was still a laborious process. We do think that having our innovator-friendly costing tool (which we pre-tested with innovators) helped, but without financial incentives, we were beholden to innovators' interest in moving forward on the CEA evaluation deliverables.

We had one primary data-collection opportunity that did not pan out at all. Our team intended to support actively funded innovators to collect data from their beneficiaries using the Equity Tool (<https://www.equitytool.org/>), a country-specific, simplified version of the Wealth Index from the Demographic Health Surveys (DHS) (Metrics for Management, 2018). This tool would have helped our team assess whether SL@B innovations were reaching the most economically vulnerable populations and reducing health inequities as the program was intended to do. After holding focus group discussions with four innovators to introduce the tool and its ability to help innovators categorize the wealth quintiles of their beneficiaries, we decided not to pursue it due to lack of interest from the majority of our participants in employing it, except for one innovator. The innovators' reasons for lack of interest included thinking they already know the economic status of their beneficiaries based on the region or health center they serve or thinking it would be too time consuming and outside the scope of their current project implementation. The sponsors also seemed less interested in this equity-focused question over time.

For the CEAs and for the one innovator who did want to pursue using the Equity Tool, we think the key to obtaining the data was patience, relationship building by making ourselves continually available for assistance after gentle reminders, and making a strong case for the added value of the analyses for their innovations and articulating impact.

### ***Challenge #3: Shifting priorities of the sponsors funding the evaluation***

The priority questions for the evaluation shifted a number of times over the two-and-a-half-year evaluation period, due to leadership transitions within the funding sponsor agencies and their individual emerging priority areas and desired work products. These transitions disrupted our team's workflow and called for regular regrouping and revisiting of the workplan, to be responsive to the funders' requested changes. While organizational leadership changes are inevitable, the

impact on this project was perhaps felt more keenly due to the lack of an agreed-upon and up-to-date TOC that the evaluation could draw from. In the absence of a clear TOC that could help maintain the evaluation focus, evaluation pivots were frequent. For example, mid-way through the evaluation, we were asked to focus on a new workstream, which meant de-emphasizing other workstreams for a time period, but then with a subsequent leadership change, that new workstream was not prioritized, and we regrouped again. Priorities throughout the evaluation period were lowered or elevated depending on sponsor leadership. In addition to leadership change within the funding agencies, we incorporated suggestions from other stakeholders such as the SL@B evaluation advisory board, which were not always aligned with funders' interest, but we believed helped lend a broader perspective to our deliverables.

Not having an agreed-upon TOC allowed the sponsors to change the success indicators with every change in leadership and request a different angle of analyzing the data and presenting the deliverables. One of our key recommendations in the final report is to spend time aligning the SL@B donor stakeholders on a TOC that will inform future benchmarks of success and impact indicators. The importance of having an agreed upon TOC is also reiterated in designing complex global initiatives evaluations ([Institute of Medicine, 2014](#)).

An important factor that helped us while navigating these shifting priorities was our IRB approvals. Ethical review boards require protocols, approved questionnaires, and interview guides. Every time the sponsors wanted something added or changed mid-stream, we communicated with them the limitations based on our IRB approvals. While we did accommodate some amendments, having IRB approvals in place helped us with continuity and staying impartial within an evolving donor environment, which allowed us generate reliable evaluation evidence that would advance the MNH innovation field more broadly, regardless of leadership changes.

## **BROADER IMPLICATIONS FOR PRACTICE**

One of the discussed propositions for complex evaluations—exploring intricacy and interdependencies within a system—allows evaluators to avoid a simplistic approach in designing an evaluation. These principles are also what propelled our evaluation team to consider SL@B in the context of the ecosystem rather than just the programmatic elements. Most of our team was used to more straightforward evaluations with clear research questions and outcome measures, an unchanging environment, and sharing results with sponsors only at the end of the evaluation. However, given the complexity of the program and heavy involvement from the SL@B evaluation sponsors, the project required a learning curve for our team in order to adapt to uncertainty and to trust that we could produce a comprehensive external evaluation that was still responsive to individual sponsor needs. Instead of more independent workstreams as we would usually set up for other projects, our team grew more cohesive, more communicative, and more integrated as time

went on in order to adapt to this complex evaluation environment. Having multiple deliverables utilizing various triangulation techniques and integration dimensions aided our ability to be flexible and responsive to the four main evaluation questions when challenges arose. Overall, having a well-aligned evaluation team that can navigate and be comfortable with complexity in evaluation is paramount.

For evaluators embarking on a similar research project, we advise setting up a skilled interdisciplinary team with strong leadership that is able to navigate the different moving parts of a complex evaluation. We also recommend that the research team starts the evaluation with an up-to-date TOC to avoid unprecedented pivots and delays. Moreover, we suggest that the research team employs multiple research methods and triangulation of data techniques and consider the different dimensions of data integration to enhance the credibility of the research findings. It is important to note that MM warrants intensive data collection, which might be arduous, especially when not incentivized. Therefore, it is imperative to establish clear value, support, and relationships with research participants and other stakeholders involved in the study.

## CONCLUSION

To evaluate the SL@B Grand Challenge, our team designed a mixed methods study that used integrated data within multiple evaluation deliverables. Data triangulation techniques employed during analyses were critical for yielding robust, data-driven outcomes. The main takeaways from our evaluation research experience are the importance of systematic data integration, having diverse deliverables that reflect the diversity of the program and its stakeholders, highlighting nuances whenever possible, building an interdisciplinary research team to draw necessary expertise, ensuring the evaluation has ethical approvals, aligning on a final Theory of Change with the sponsor, and embracing creativity and flexibility in a changing evaluation environment. We hope that sharing the details of our work evaluating a large global health portfolio such as SL@B provides insights for other evaluators involved with similarly complex portfolio evaluation projects.

## REFERENCES

- Archibald, M. M. (2016). Investigator triangulation: A collaborative strategy with potential for mixed methods research. *Journal of Mixed Methods Research*, 10(3), 228–250. <https://doi.org/10.1177/1558689815570092>
- Archibald, M. M., Radil, A. I., Zhang, X., & Hanson, W. E. (2015). Current mixed methods practices in qualitative research: A content analysis of leading journals. *International Journal of Qualitative Methods*, 14(2), 5–33. <https://doi.org/10.1177/160940691501400205>
- Biru, B., Rajan, S., Taylor, A., Crissman, K., Dixit, S., Doshi, P., . . . Udayakumar, K. (2020). *Evaluating saving lives at birth: Evaluation report: Rounds one to eight (2011–2020)*. Duke Global Health Innovation Center and Duke Global Health Institute Evidence

- Lab at Duke University. <https://dukeghic.org/wp-content/uploads/sites/20/2020/09/Evaluating-SL@B-Final-Report-2020.pdf>
- Creamer, E. G. (2017). *An introduction to fully integrated mixed methods research*. Sage.
- Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods*. McGraw-Hill.
- Dixit, S., Fernholz, F., Ogbuoji, O., Finnegan, A., Biru, B., Baumgartner, J. N., & Udayakumar, K. (2020). *Evaluating saving lives at birth: Costing tool for healthcare innovators: An excel-based tool for capturing historical and scale-up costs of an innovation*. Duke Global Health Innovation Center and Duke Global Health Institute Evidence Lab at Duke University. <https://dukeghic.org/wp-content/uploads/sites/20/2020/12/Costing-Tool-for-Healthcare-Innovations-Brief.pdf>
- Erzberger, C., & Prein, G. (1997). Triangulation: Validity and empirically-based hypothesis construction. *Quality & Quantity: International Journal of Methodology*, 31(2), 141–154. <https://doi.org/10.1023/a:1004249313062>
- Farmer, T., Robinson, K., Elliott, S. J., & Eyles, J. (2006). Developing and implementing a triangulation protocol for qualitative health research. *Qualitative Health Research*, 16(3), 377–394. <https://doi.org/10.1177/1049732305285708>. Medline:16449687
- Fetters, M. D., & Molina-Azorin, J. F. (2017). The *Journal of Mixed Methods Research* starts a new decade: The mixed methods research integration trilogy and its dimensions. *Journal of Mixed Methods Research*, 11(3), 291–307. <https://doi.org/10.1177/15586898177114066>
- Finnegan, A., Biru, B., Taylor, A., Rajan, S., Udayakumar, K., & Baumgartner, J. N. (2021). Improving global maternal and newborn survival via innovation: Stakeholder perspectives on the Saving Lives at Birth Grand Challenge. *PLoS One*, 16(7), e0254589. <https://doi.org/10.1371/journal.pone.0254589>. Medline:34260640
- Institute of Medicine. (2014). *Evaluation design for complex global initiatives: Workshop summary*. The National Academies Press. <https://doi.org/10.17226/18739>
- International Development Innovation Alliance (IDIA). (2017). *Insights on scaling innovation*. <https://static1.squarespace.com/static/5b156e3bf2e6b10bb0788609/t/5b1717eb8a922da5042cd0bc/1528240110897/Insights+on+Scaling+Innovation.pdf>
- Lalli, M., Ruysen, H., Blencowe, H., Yee, K., Clune, K., DeSilva, M., . . . Lawn, J. E. (2018). Saving Lives at Birth: Development of a retrospective theory of change, impact framework and prioritised metrics. *Globalization and Health*, 14(1), 13. <https://doi.org/10.1186/s12992-018-0327-z>. Medline:29378667
- Metrics for Management. (2018). Introducing the equity tool: A simple way to assess if your program is reaching the poor. <https://www.equitytool.org>
- O’Cathain, A., Murphy, E., & Nicholl, J. (2010). Three techniques for integrating data in mixed methods studies. *BMJ*, 341, c4587. <https://doi.org/10.1136/bmj.c4587>. Medline:20851841
- Preskill, H., Gopal, S., Mack, K., & Cook, J. (2014). *Evaluating complexity: Propositions for improving practice*. <https://www.fsg.org/publications/evaluating-complexity>
- Rajan, S., Biru, B., Taylor, A., Crissman, K., Doshi, P., Baumgartner, J. N., & Udayakumar, K. (2020). *Partnerships for increased innovation impact: The role of the SL@B program in*

*facilitating partnerships in the maternal and newborn health ecosystem.* Duke Global Health Innovation Center and Duke Global Health Institute Evidence Lab at Duke University. [https://dukeghic.org/wp-content/uploads/sites/20/2020/07/ESLAB-Partnerships-brief\\_July-2020.pdf](https://dukeghic.org/wp-content/uploads/sites/20/2020/07/ESLAB-Partnerships-brief_July-2020.pdf)

Taylor, A., Biru, B., Rajan, S., Quivera, H., Dixit, S., Baumgartner, J. N., & Udayakumar, K. (2020). *Acceleration to impact: The role of the SL@B program in accelerating market entry for maternal and newborn health innovations.* Duke Global Health Innovation Center and Duke Global Health Institute Evidence Lab at Duke University. <https://dukeghic.org/wp-content/uploads/sites/20/2020/07/Acceleration-to-Impact.pdf>

Taylor, A., Rajan, S., Biru, B., Crissman, K., Quivera, H., Baumgartner, J. N., & Udayakumar, K. (2020). *Saving lives at birth: Program portfolio review and funding landscape analysis.* Duke Global Health Innovation Center and Duke Global Health Institute Evidence Lab at Duke University. Retrieved from <https://dukeghic.org/wp-content/uploads/sites/20/2020/07/Portfolio-Review-and-Funding-Landscape-Analysis.pdf>

Teddlie, C. B., & Tashakkori, A. M. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences.* Sage.

USAID Center for Accelerating Innovation and Impact. (2016). *Pathways to scale: A guide for early-stage global health innovators on business models and partnership approaches to scale-up.* [https://www.usaid.gov/sites/default/files/documents/1864/Pathways-to-Scale-Guide\\_20161013\\_online-508.pdf](https://www.usaid.gov/sites/default/files/documents/1864/Pathways-to-Scale-Guide_20161013_online-508.pdf)

## AUTHOR INFORMATION

**Blen Biru** is a research associate at the Duke Global Health Innovation Center. She is interested in using evaluation research methods to improve programs and ensure equitable access to healthcare.

**Andrea Taylor** is an assistant director at the Duke Global Health Innovation Center. She is interested in evaluating the scaling and adaptation of health-care innovations to address critical access and quality challenges.

**Sowmya Rajan** is a senior manager at the Duke Global Health Innovation Center. She is interested in evaluation research to build evidence that informs health-care programs and policies.

**Kate Crissman** is a research analyst at the Duke Global Health Innovation Center. She is interested in the evaluation and implementation of health-care innovations, with a concentration in Maternal and Child Health.

**Osondu Ogbuoji** is an assistant research professor at the Duke Global Health Institute. His primary research interest is making health systems work better in resource limited settings.

**Fernando Fernholz** is a Professor of the Practice at the Duke Sanford School of Public Policy. He's interested in evaluating policies for economic development.

**Siddharth Dixit** is a research associate at the Duke Global Health Institute. He is interested in evaluation and economic analysis of health-care innovations.

**Minahil Shahid** is a research associate at the Duke Global Health Institute. She is interested in using data analysis and evaluation research to inform evidence-based health policies.

**Pratik Doshi** is an internal medicine resident at the University of Southern California. He is interested in the intersection of health policy, business, and medicine to provide greater access to health care for all people.

**Krishna Udayakumar** is the founding director of the Global Health Innovation Center. He is interested in generating deeper evidence and support for scaling and adapting of health innovations and policy reforms globally.

**Amy Finnegan** is a senior data scientist at IntraHealth International. She is interested in designing data collection tools, analysis plans, machine learning, and other creative research design methods to address global health challenges.

**Joy Noel Baumgartner** is an associate professor at the UNC School of Social Work. She is interested in evaluating and strengthening the delivery of integrated health and social service interventions to address global health issues.