JANUARY 2021


# TRANSPARENCY FOR DEVELOPMENT

## PROJECT RESULTS & IMPLICATIONS


# Acknowledgements

The T4D project – from inception to design, data collection to results and outreach – is truly the product of the efforts, dedication, and expertise of a wonderful team and set of advisors. We want to begin by thanking the broader T4D team (past and present) based at Harvard and R4D, including Matthew Bombyk, Grant Bridgman, Akshay Dixit, Eric Englin, Hannah Hilligoss, Cara Myers, Rohit Naimpally, James Rasaiah, Niriksha Shetty, and J. Preston Whitt - and based in Indonesia and Tanzania, including Sarah Alphs, Jenna Juwono, Lindsey Roots, R. Reksa Samudra, and Astri Waluyo.

This work also integrates excellent research conducted by partners and colleagues in country, including ethnographic researchers Iqra Anugrah, Megan Cogburn, Mohmed Yunus Rafiq, Kankan Xie, and the late Chris Grorud and country researchers Chris Chapman, Fatu Conteh, the late Linda Gassama, Takondwa Iman, Agnes Lahai, Emmanuel Sapalado, and James Wumenu.

This work and our learning about social accountability would not have happened without the amazing contributions of and leadership by our design and implementation partners, the teams at the Clinton Health Access Initiative in Tanzania (Hilda Kigola, Esther Mtumbuka, Happy Ndomba), PATTIRO in Indonesia (Novita Anggraeni, Riska Amelia Hasan, Yulius Hendra Hasanuddin, Didik Purwandanu, Sad Dian Utomo), Center for Democratic Development in Ghana (Rexford Asiama, Mohammed Awal, Mavis Dome, Franklin Oduro), Malawi Economic Justice Network (Kelvin Chirwa, June Kambalametore, Kondwani Kaunda), and WASH-Network in Sierra Leone (Sandy Massaquoi, Patrick Momoh, Patrick Pokawa, Emmannuel Sandy, Musa Soko, Samuel Vandi).

# Acknowledgements (continued)

We are grateful to our institutional research and data collection partners: J-PAL Southeast Asia, SurveyMETER in Indonesia, EDI and Ideas in Action in Tanzania, and Innovations for Poverty Action (Sierra Leone and Tanzania).

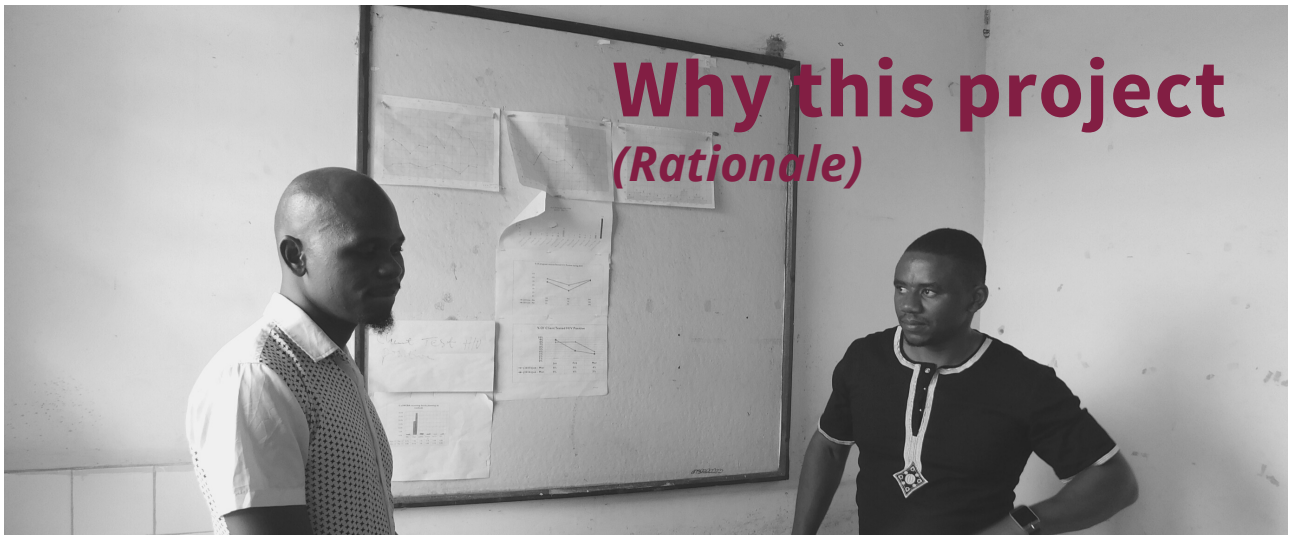
For generous funding for this project, we thank the William and Flora Hewlett Foundation (including Kevin Bohrer, Ruth Levine, Alfonsina Penaloza, and Pat Scheid), the UK Department for International Development (including Peter Evans, Tom McEnroe, the late Helen Richards, and Tom Wingfield), and the Bill and Melinda Gates Foundation (including Nushina Mir), as well as the Transparency and Accountability Initiative (including Florencia Guerzovich, Brendan Halloran, Vanessa Herringshaw, & Michael Jarvis), which conceptualized and brokered the project.

We thank our colleagues at our respective institutions (the Ash Center for Democratic Governance and Innovation at the Harvard Kennedy School, Results for Development, and the Evans School of Public Policy and Governance at the University of Washington) who provided endless support in the form of research and content feedback, outreach, communications, financial management, and operations.

We gratefully acknowledge the members of the T4D advisory committee: Yamini Aiyar, Jessica Cohen, Jonathan Fox, Anuradha Joshi, Dan Posner, and Bill Savedoff.

Finally, our most important debt is to the thousands of community members and government officials in Indonesia, Tanzania, Ghana, Malawi, and Sierra Leone who volunteered their time and energy to participate in the Transparency for Development activities. More on the project, including background documents and reports and other working papers, is available at t4d.ash.harvard.edu.

*Archon Fung, Jean Arkedis, Jessica Creighton, Steve Kosack, Dan Levy, and Courtney Tolmie (T4D Principal Investigators)*



# Why this project
## *(Rationale)*

In the past three decades, the concept of citizens playing a role in monitoring and holding government officials to account for poor services has gained prominence as a potential tool to fight mismanagement, poor prioritization of resources and corruption around the world.

The idea is extremely appealing – who better than citizens to observe breakdowns in government decisions and actions and to fight for solutions to these breakdowns when they occur? As members of communities where services are supposed to be delivered, citizens are closer to the frontline and thus better able to observe these breakdowns than a politician or technocrat living in the capital.

And as the intended beneficiaries of government funding and service delivery, the people in these communities who are being affected by these problems have clear incentives to get them solved.

**Social Accountability:** *Efforts by citizens and civil society to hold power holders to account for their actions.*

There is a definite logic to this argument. And as the field of social accountability has grown, more and more civil society organizations (CSOs) have taken on different types of social accountability to increase citizen empowerment and participation and ultimately improve development outcomes.

However, the research on social accountability has shown a decidedly more mixed picture. While cases such as Bjorkman and Svensson (2009) showed enormous impact of health-focused scorecards in Uganda, many other studies showed a bleaker picture of social accountability representing a big promise with little impact (Banerjee, Deaton, and Duflo (2004), Lieberman, Posner, and Tsai (2013)).

Why is there such a mixed picture? And are there things that can be learned – about who is involved and how information is shared and what actions are taken – that can better inform decisions about whether and how to undertake social accountability?

## Objectives

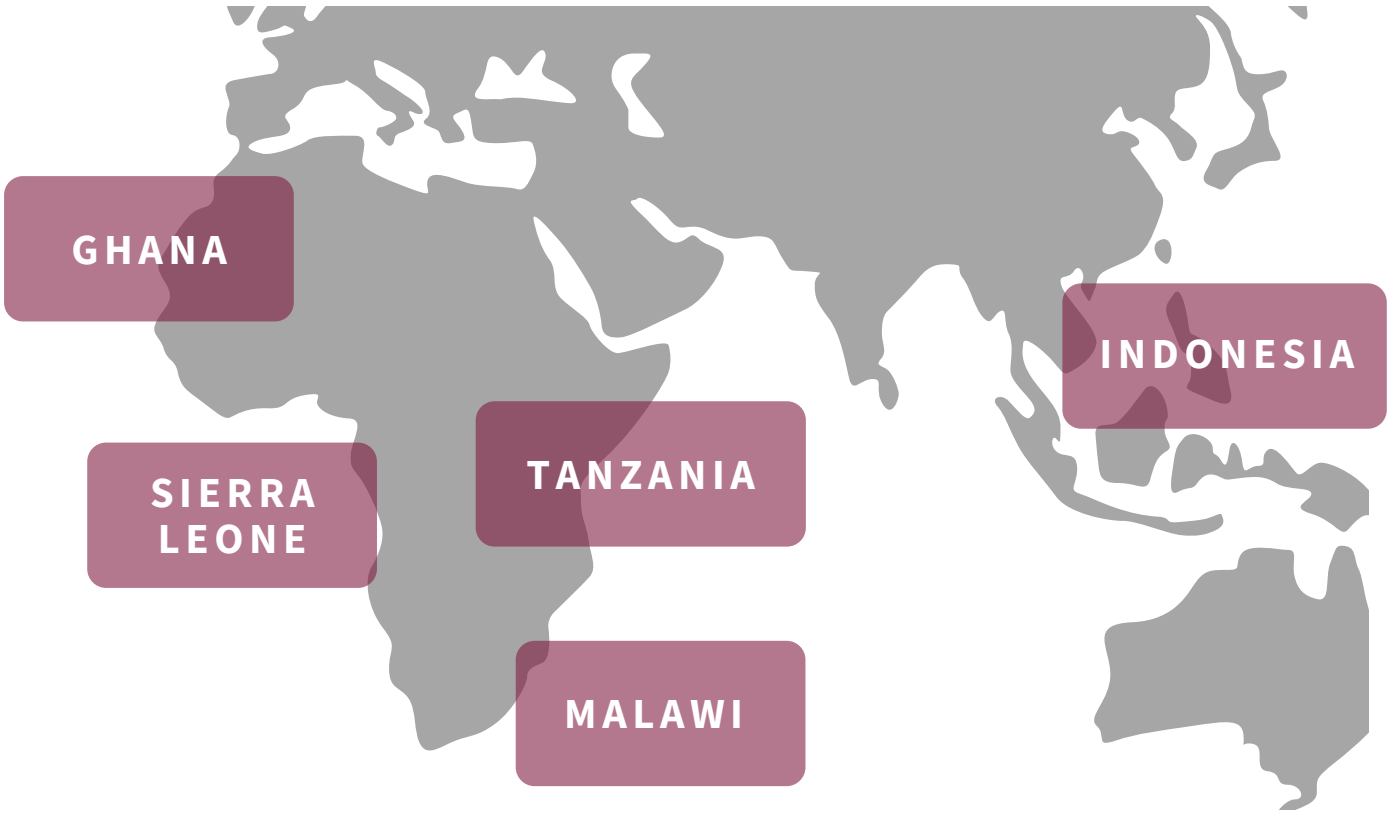
The Transparency for Development (T4D) project was launched in 2013 to try to answer these challenging questions and to make sense of a highly mixed evidence base. Ultimately, we sought to design research that could provide useful and actionable guidance to policymakers, donors, and practitioners alike seeking to improve the effectiveness of their work.

In the next section, we highlight the research questions that we sought to answer as part of the T4D project; however, in this section, we want to focus on the larger objectives of this project outside of the world of research. At its core, T4D was designed to offer evidence of whether community-led transparency and accountability programs can improve health outcomes – as well as how, why, through what mechanisms, and in what contexts. What is most valuable from this evidence varies for the diverse range of audiences who care or are curious about social accountability.

For donors, organizations, and policymakers seeking to improve **Health Outcomes**, the project hoped to provide information that can guide decisions about whether to undertake social accountability to improve health and how to better design programs that will achieve the goals of social accountability (including increased citizen voice, better information, and more responsive policymakers and service providers) while also improving health.

For **Social Accountability practitioners** (CSOs, community-based organizations, and other people involved in social accountability), we hoped to provide actionable and evidence-based recommendations for how programs can be designed better to be more responsive to complex settings, actors, and situations - and thus ultimately be more effective in achieving the diverse goals that practitioners design for this work.

For **Global Advocates of Social Accountability**, we hoped to share evidence that can help guide the discussion in this field moving forward, whether through a demonstration of positive results or as an "opening of the black box" to reveal how social accountability can be improved upon and made more effective as it continues to be a major tool that global development actors turn to.




Finally, for **Researchers,** while we hope to answer many questions related to the role of social accountability, we also know that this is a dynamic and complex field. As such, our objectives were to provide foundational evidence that encourages and points to new directions of research and to elevate the value of a deliberate and unique mixed-methods approach to understand what goes into and can come out of complicated accountability programs such as this one.

***Box 1. T4D Country Partner Organizations***

- *Ghana: Center for Democratic Development*
- *Indonesia: PATTIRO, J-PAL Southeast Asia, SurveyMETER*
- *Malawi: Malawi Economic Justice Network*
- *Sierra Leone: WASH-Network, Innovations for Poverty Action*
- *Tanzania: Clinton Health Access Initiative, EDI Group, Ideas in Action, Innovations for Poverty Action*

# What T4D sought to answer

## *(Research Questions)*

The T4D project took place over 6.5 years and was designed to answer six core research questions. The first four questions focus on issues of impact and whether community-led transparency and accountability can improve health and other important outcomes:

(1) What is the impact of the civic participation intervention on **utilization of health care services** related to maternal and child health?

(2) What is the impact of the civic participation intervention on **content of health care services** related to maternal and child health?

(3) What is the impact of the civic participation intervention on **child health outcomes**?

(4) What is the impact of the civic participation intervention on **citizen perception of empowerment and efficacy**?

In addition to understanding "whether," we also wanted to understand "why" and "how."

Recognizing that these types of programs are more complex and potentially more context-dependent than interventions that seek similar outcomes (vaccination for example), understanding what was happening behind the outcomes was a critical part of this work. To complement the questions of impact, we considered two further research questions:

(5) If there are significant effects, what are the **mechanisms** through which these effects occur? If there are not effects, why not?

(6) How did community activists **respond to and experience** the program?

To answer these questions, we worked with local partners to develop a new social accountability or "civic participation" intervention, which we describe in the next section. We then tested this program – in 100 communities each in Tanzania and Indonesia – and then piloted an adapted version of the program in three more countries (Ghana, Malawi, and Sierra Leone). What we uncovered when we sought to answer these research questions - and ultimately what this means - are described in the final two sections of this report.

# What was done to encourage participation *(The Intervention)*

## *How we developed the intervention*

Social accountability approaches can take many forms – contestational or collaborative, service quality-focused or corruption-focused, bottom-up or sandwiched. To answer the research questions outlined in the previous section, we had to begin with the intervention that we would study to understand how and why community-driven accountability works (or does not work) the way that it does. We also had several parameters that we needed to meet in the design of the social accountability program – and some design principles that we placed upon ourselves.

As parameters, we needed to develop a social accountability intervention that had the potential to improve health outcomes – and, specifically for this program, maternal and newborn health outcomes. Because we wanted to test something that had the potential to work in different countries, we needed to design an approach that was flexible to different contexts.

And because we wanted to test it in many locations even within a given country, it needed to be something that could be replicated widely.

Beyond these parameters, we wanted to learn from the experts in designing this work, which meant listening to past evidence and to CSOs who design and lead and learn from this work every day. All of this resulted in a set of intervention design principles that we used with our local civil society partners to design a social accountability program that we wanted to test at scale. You can read more about the intervention design principles here, but these boil down to:

- **Scalability** – something that, if successful, could be replicated to many locations and sectors.
- **Relevant to local problems** – something that did not only consider a specific problem that may be prominent in one country or one region or one village, but that would instead allow citizens to focus on the problems that they considered the biggest priorities in their lives.

- **Non-prescriptive** – an approach that would not dictate the type of actions that communities designed and undertook but instead would provide examples from other places and a space in which participants could figure out for themselves what they wanted to do to push for change.

- **Free of outside resources and additional authority** – ultimately, we wanted to understand whether encouraging participation in accountability, without payments or brokering relationships, could lead to actions and eventually impact.

And returning to the point about relying on the experts, the final social accountability intervention that we developed was done hand-in-hand and through many iterations and pilots with civil society organizations that have been leaders in transparency, accountability and health in their countries: PATTIRO (Indonesia) and the Clinton Health Access Initiative (Tanzania) for the original intervention and the Center for Democratic Development (Ghana), Malawi Economic Justice Network, and WASH-Net (Sierra Leone) for an adapted model in a second phase of the work.

## *Intervention Description*

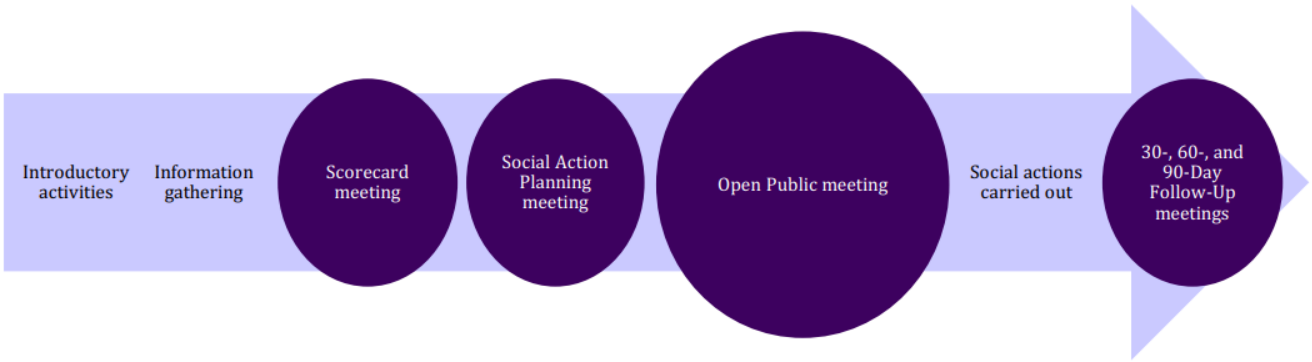
So what was the result of these parameters, design principles, co-creation, and iteration? An adapted community scorecard that looks like the visual in Figure 1.

And what does this stylized graphic look like in reality? A series of steps, led by a CSO facilitator in the hopes that those participating would design and take their own actions to improve health. Specifically:

- The facilitator was introduced to a community in a set of **Introductory Activities** to launch the program.

- The facilitator then conducted a set of surveys – one of the local health facility and another of women in the community who had recently given birth – to help **Gather Information** about the problems that might be underlying poor maternal and newborn health outcomes for this specific village. These issues could be related to corruption or poor governance, but they could also be related to lack of knowledge, cultural barriers, problems with access or problems with quality.

**FIGURE 1. ADAPTED SCORECARD INTERVENTION (PHASE 1)**



- At the same time, the facilitator also tried to find people who would be good **Community Representatives (CRs)**. These would be the core group leading the charge of designing and implementing actions as part of this project. CRs were deliberately selected to be a diverse group who cared about health and/or improving things in their villages – but who were not formal leaders or health providers (and thus people who already had power or clear roles in this work).

- After a few weeks of working on these prior steps, the facilitator led a first meeting with the CRs to **Present the Scorecard**. The scorecard focused on the information that the facilitator had gathered on the problems that could be underlying poor maternal and newborn health outcomes. The scorecard itself looked different in different places (Pictures 1 and 2) but in both countries they were designed to start a dialogue through which CRs could determine which problems they wanted to prioritize to solve.

- The day after completing the Scorecard Meeting, the facilitator met again with the CRs to **Develop the Social Action Plan**. In this meeting, the CRs would reflect on the problems that they prioritized the previous day and learn about actions that other communities have taken to work on accountability, such as protesting, collaborating with government officials, working with the media, naming and shaming, or community self-help. At the end of the meeting, the CRs developed a set of actions to target specific health problems.

- After developing the social action plan, several things would happen over the next several months and beyond. First, the CRs led an **Open Public Meeting** where they shared their social action plans with the larger community to gain buy-on. After this, it was up to the CRs whether they **attempted or completed the social actions** that they designed. The facilitator had no role in implementing the social actions, and so it was completely up to the CRs whether they worked on these.

- Finally, the facilitator returned to the community three more times for **Follow-Up Meetings** after 30, 60, and 90 days. During these meetings, the facilitator provided encouragement to CRs to keep working on the actions, celebrated their accomplishments, and helped them to brainstorm different solutions when an action did not work.

***Box 2. What is a "Social Action"***

*Some examples include:*

- *Leading an education campaign for pregnant women*
- *Working with the facility to post hours and cost information*
- *Advocating for a new facility or ambulance*
- *Building a placenta pit*
- *Creating a savings group for maternal health emergencies*
- *Helping women create birth preparedness plans*


Picture 1. Scorecard Poster for Indonesia


Picture 2. Scorecard Approach for Tanzania

## *How does this connect to ultimately improving health?*

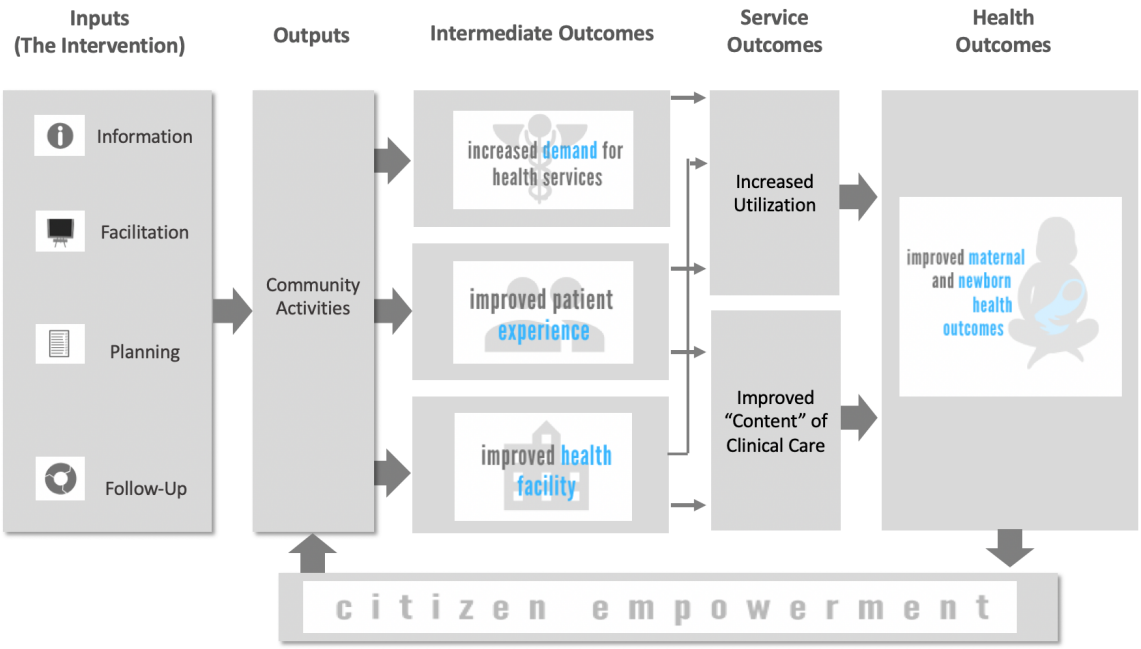
One of the challenges of linking social accountability programs to health outcomes is that the causal chain from start to finish is complicated and relies on a number of pathways actually activating. However, there are several lines that can be drawn between the activities undertaken by CSO facilitators described above – and an improvement in the outcomes that we describe in the research questions. For us, this pathway – or theory of change – looks like Figure 2.

For the T4D social accountability approach to ultimately change health outcomes, it needs to first lead to communities taking actions to fix problems they have identified (output), and then those actions have to mitigate problems associated with the demand for health services, patient experiences, and/or health facilities (intermediate outcomes).

Then the changes in these outcomes have to further make it more likely that women will utilize the formal health facilities and/or that the content of the care they receive will be better (service outcomes). There is a wide set of health literature that suggests that these changes in service outcomes are linked to improve health outcomes. And throughout, a key mechanism through which this takes place is citizen engagement and empowerment – we expect that citizen empowerment will increase if they see actions that lead to health improvements and that this increased empowerment will feed back into citizens choosing to take actions.

At its core, the theory of change is a visual of our research questions, and tracing the arrows in this theory of change helps us to understand whether social accountability can have the impact that people hope that it will.

**FIGURE 2. T4D THEORY OF CHANGE (PHASE 1)**



# What we found
## *(Results)*

The mixed-methods research that we undertook revealed many insights into how and where and through what mechanisms social accountability can and cannot work.
To help provide valuable information to the diverse audiences that are interested in social accountability, we bundle this into five key sets of results in this section:

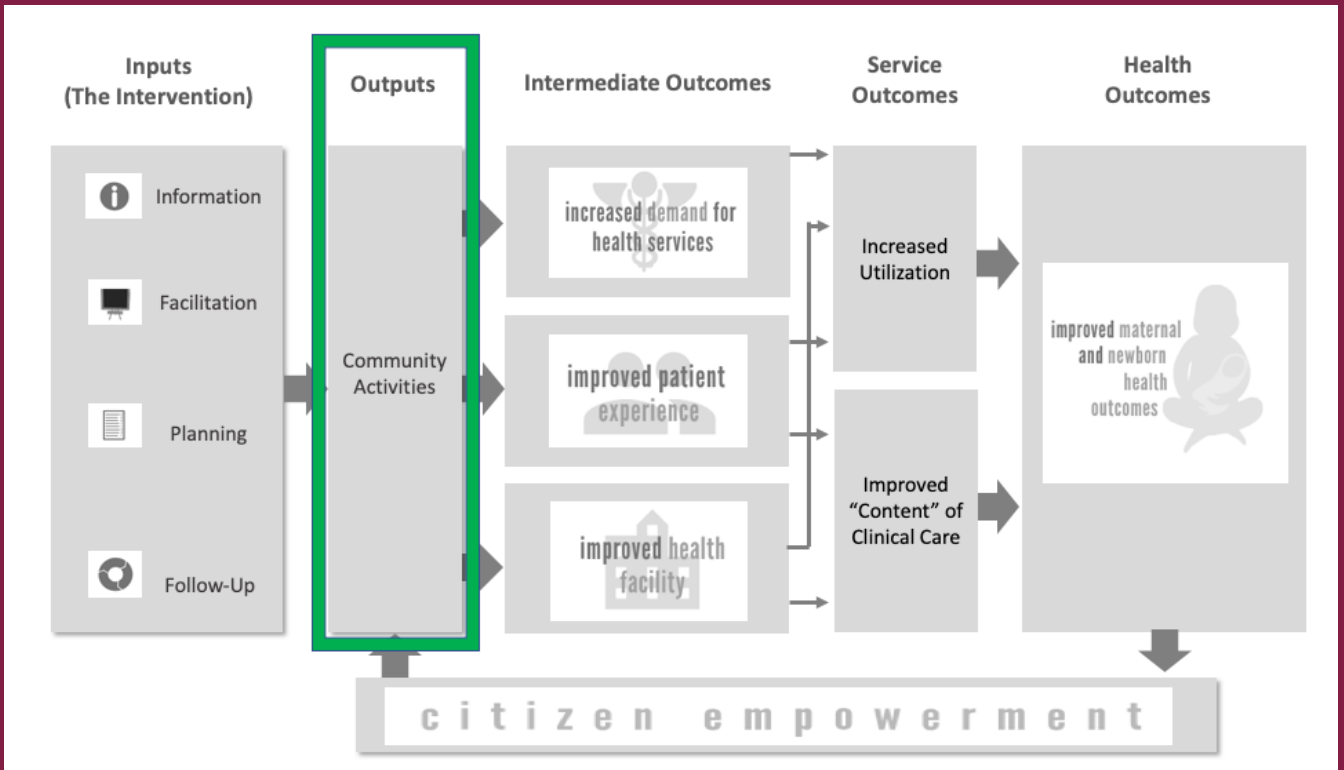
- Did communities take actions?
- Did actions lead to changes in health?
- Did empowerment and self-efficacy increase?
- How can we explain what happened and what did not?
- Are there ways to continue to improve social accountability design?

The rest of this section is organized around these five categories of research results and implications.

# Result 1 - People took action.

One of the first places in the theory of change where we could have observed a breakdown in the social accountability intervention is in communities designing and undertaking social actions. As described above, CRs were not paid or otherwise incentivized to complete social actions. And so after completing the design of the action plans in the Social Action Planning meeting, CRs could choose to undertake each action, or start but abandon actions, or not to start them at all.

***What did we see?*** The overwhelming majority of community representatives undertook actions, and these actions were incredibly diverse. However, the actions did highlight some important trends that may provide insight for those seeking to target social accountability designs to drive communities toward certain activities.



## *Key Findings*

- **There was high and sustained participation of CRs in the meetings to identify problems, design actions, and follow up with the facilitator after actions had been tried**. Communities in Indonesia and Tanzania started the intervention with an average of 14 and 15 CRs respectively (this was the number the facilitators attempted to recruit), and by the final meeting, an average of 11 and 10 CRs respectively attended. And the vast majority of those participants that attended these meetings participated by speaking up distinctively in the meetings (in Indonesia and Tanzania, 89 and 91 percent respectively spoke up in the first meeting and 94 and 68 percent respectively in the final meeting).

- **Almost every community attempted at least one action – and completed at least one action**. Communities planned anywhere from 2 to 17 actions during this short program, and we saw that the rate of completion of at least one of these actions was extremely high (see Table 1).

- **Actions were extremely diverse across and within communities**. While the Scorecard Meeting highlighted a range of different problems related to health outcomes, it was not guaranteed that different communities would choose to focus on different problems and develop diverse actions to address these. However, we observed a lot of variation across these outcomes, with 100 percent communities designing actions to address issues with demand for services, 60 percent designing actions to address patient experience, and 55 percent with actions focusing on improving the health facility. Within these major pathways, we saw even more variation, as highlighted in Table 2.

- **Despite the diversity across communities, there were few major differences between Tanzania and Indonesia on average**. There are exceptions to this, including a stronger focus on developing and proposing by-laws in Tanzania; however, for two very different countries, the general pattern of actions was quite similar.

**TABLE 1. ACTIONS DESIGNED AND COMPLETED IN TANZANIA AND INDONESIA**

| Percentage of villages in which participants... | | | | |
|---|---|---|---|---|
| | designed at least one action | attempted at least one action | completed at least one action (self-reported) | completed at least one action (verified in subset) |
| Indonesia | 100% | 98% | 91% | 78% |
| Tanzania | 100% | 100% | 98% | 100% |

- **Actions were collaborative rather than contestational.** While the intervention took a deliberately agnostic approach to whether actions should seek to work with or against actors, we observed that 91 percent of actions could be classified as collaborative in nature, with very few seeking a more confrontational approach.

- **Actions were locally-focused rather than longer route.** In designing actions, CRs focused the majority of their actions on local actors, including others in the community (60 percent of actions) and those at the health facility (35 percent of actions).

Only 8 percent of actions included a government target above the community level, even while some of the health problems on which communities were within the purview of higher level actors.

**TABLE 2. ACTIONS BY INTERMEDIATE OUTCOME TYPE AND PATHWAY**

| Intermediate Outcome | Pathway | % Villages |
|---|---|---|
| **Increased demand for health services (100%)** | Awareness, knowledge & attitudes | 94% |
| | Facility access | 71% |
| | Ability to pay | 45% |
| | By-laws, partnerships, or interventions aimed at health service uptake | 35% |
| **Improved patient experience (60%)** | Info & communication (cost, opening hours, etc.) | 39% |
| | Attitude, effort, or trust of provider | 36% |
| | Facility cleanliness | 6% |
| **Improved health facility (55%)** | Availability of drugs, supplies, other inputs | 28% |
| | Facility infrastructure | 28% |
| | Facility staffing | 18% |
| | Provider knowledge | 1% |
| **Actions Outside of the Theory of Change** | Non-health system directed | 9% |

- **Only one quarter (26 percent) of actions could be defined as "social accountability" actions.** While we define this program as a social accountability program, the design of the actions was left to the CRs to develop. As such, there were many cases in which the actions that were designed did not work through social accountability (by our definition, a social accountability action was one that sought to make a government official and/or service provider fix a problem that was within the responsibility of the government official/service provider to fix). This is an important finding for programs that may seek to encourage a wider set of civic engagement and participation (beyond traditional accountability) – and for those that explicitly want to bolster accountability actions.

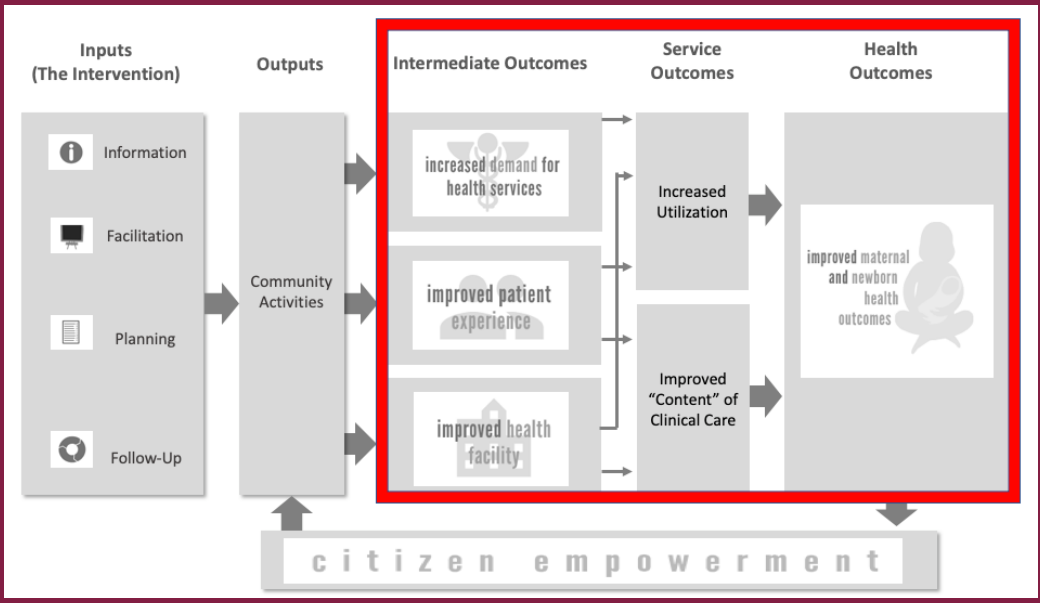
***How do we know this?*** We reviewed, coded and analyzed every social action developed in 200 communities, totaling more than 1,139 actions. In addition, we analyzed data from social action plans from all communities including from follow up meetings, key informant interviews from the 40 percent of communities where we collected this data, meeting observation data in the 18 percent of communities where we collected this data, and ethnographic observations and analysis from 8 communities total. You can read more about the methods and results in "Insights from the Transparency and Accountability Action Plans in Indonesia and Tanzania" (2020).

# Result 2 - The things that people did DID NOT improve health.

So, if communities designed, implemented, and completed many diverse actions, was this enough to lead to changes in health outcomes (including intermediate and service outcomes)? This overarching question is the focus of our first three research questions: What is the **impact** of the civic participation intervention on (1) **utilization of health care services** related to maternal and child health, (2) **content of health care** services related to maternal and child health, and (3) **child health outcomes**?

***What did we see?*** Across the board, we found that, while the intervention did catalyze actions, it did not result in an impact on health outcomes.

This doesn't mean that some of the health outcomes we were interested in did not improve between the time we collected baseline data and when we collected endline data; but it does mean that we saw the same improvements in outcomes in the places where there were no CSO facilitators doing the social accountability work as where we held the T4D social accountability intervention. Because of this, it is clear that this program was not the thing that drove the improvements in health. But this is an interesting point that we discuss later on in this report.



## *Key Findings*

- **The social accountability program did not have a statistically significant impact on health or service outcomes.** In total, we collected data on eight health-focused outcomes of interest – four related to utilization of health services (birth with a skilled provider, birth in a facility, and – in Tanzania only – two outcomes for proper antenatal care), two indices related to content of care (content of prenatal care and delivery and content of postnatal care), and two ultimate outcomes (stunting and underweight). Figures 3 and 4 demonstrate that the standard confidence interval for each of these outcomes crosses the middle line that represents an effect of zero. This means that we cannot say that any of these outcomes are different from zero. *The final two outcomes in each of these figures will be highlighted when we discuss empowerment results.*

**FIGURE 3. IMPACT OF THE T4D PROGRAM ON PRIMARY OUTCOMES IN INDONESIA**

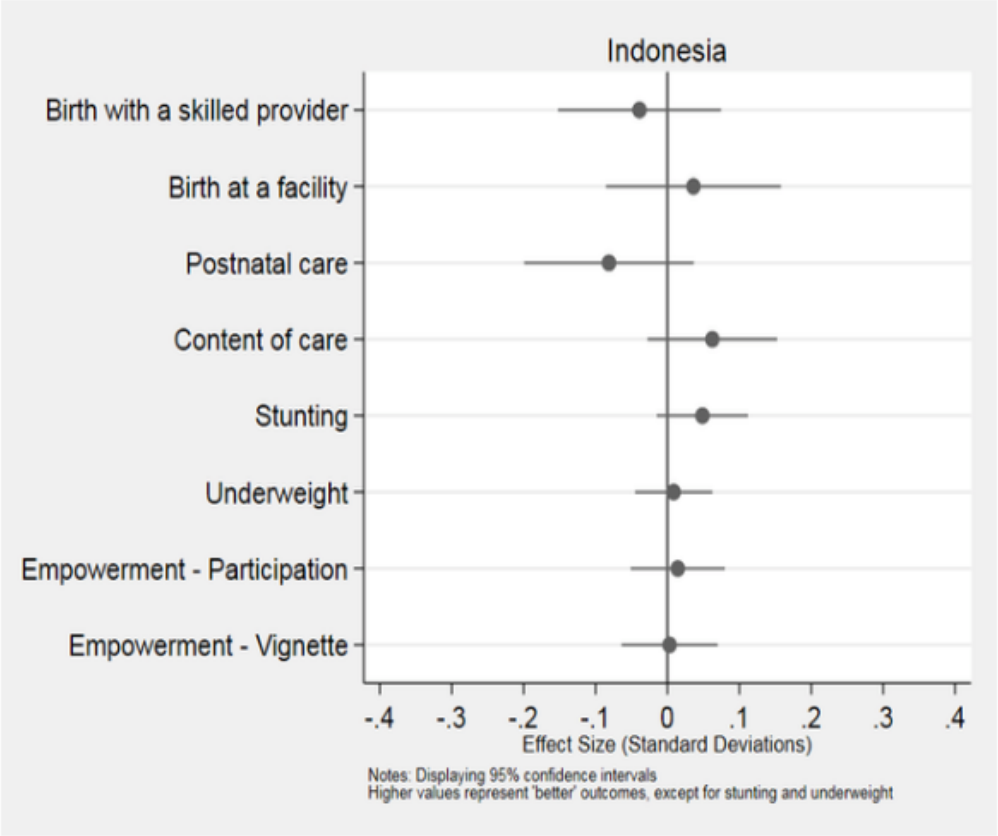
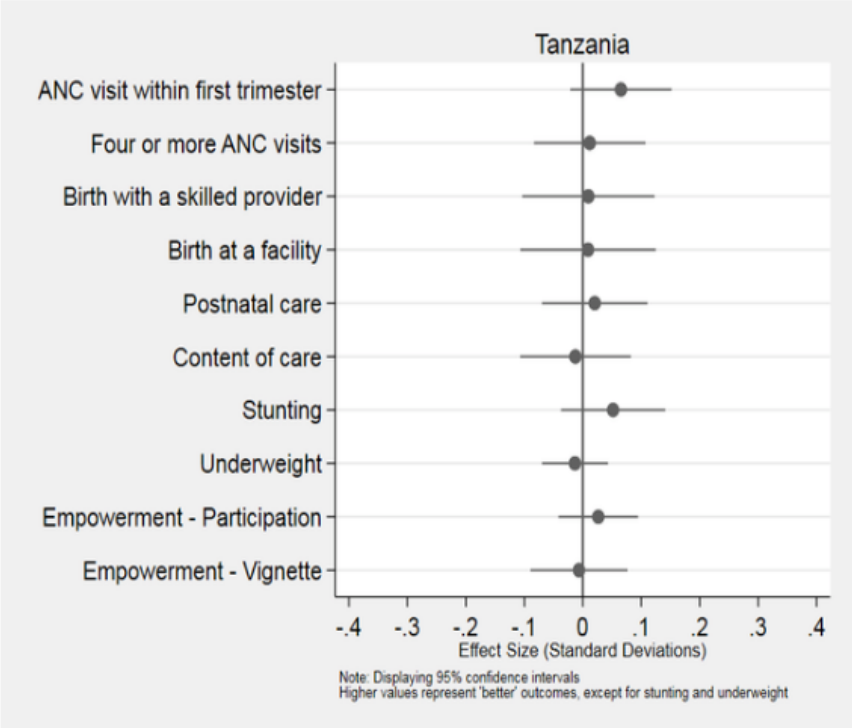


**FIGURE 4. IMPACT OF THE T4D PROGRAM ON PRIMARY OUTCOMES IN TANZANIA**



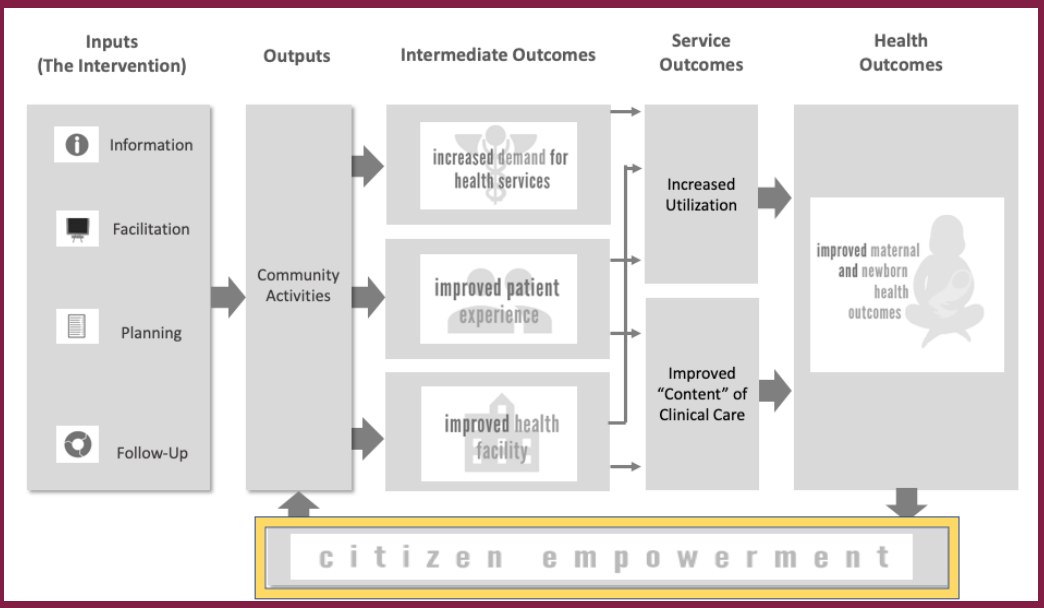

- **There were few statistically significant intermediate outcomes, and those that were significant were likely due to chance.** In addition to the health and service outcomes, we tested a set of over 100 intermediate outcomes in each country, ranging from ones related to increasing awareness to lowering cost to improving availability of drugs and supplies. In each country, the vast majority (over 90 percent) of these outcomes were not statistically significant. While it may be tempting to interpret those that did come back as significant as telling us something meaningful, we would expect to see around this number of outcomes as significant by random chance because we were testing so many. As such, we cannot say that the intervention moved the needle on any intermediate outcomes either.

***How do we know this?*** We undertook a baseline and an endline survey in each country with three major groups – women who recently gave birth, health facilities, and communities. These surveys were implemented in 200 total communities in each country, 100 that received the intervention (treatment) and 100 that did not (control), as part of our randomized controlled trial for this work. You can read more about the methods and results, including our randomization strategy and model, in "Can transparency and accountability programs improve health?" (2019).

# Result 3 - The things that people did increased empowerment - but only in some cases.

Based on the previous two sets of results, it looks like the breakdown in the theory of chain happened between actions (which were undertaken - and in many cases completed) and intermediate health outcomes related to demand, patient experience, and health facility. But the other set of outcomes that may play a role is those related to empowerment. Did communities in general experience improvements in their empowerment as a result of the social accountability program? How about the Community Representatives? And what was the experience like for those who participate?

***What did we see?*** It turns out the answer is: "It's complicated." The degree of optimism about the program and the actions that the CRs designed, feelings of self-efficacy, and empowerment itself varies across time (before and after the intervention), across types of community members (including those who were CRs and those who were part of the broader community), and across countries. While the picture is not simple, it does reveal some trends that may help explain the reason why we observe the outcomes (and lack thereof) that we do – something we highlight below and then dig into further in the next set of results.



## *Key Findings*

- **The broader community did not experience an increase in empowerment**. If you refer back to Figures 3 and 4, there are a set of outcomes at the bottom of each graphic that show two measures of empowerment that were collected from control and from treatment community members (not specifically CRs). One possibility that we could have observed in this program was that community members – even those not directly participating in the social actions – saw the work of the CRs and potential results of their actions and felt that they were more empowered. However, this is not what we observed – general community members did not experience an increase in empowerment as a result of the social accountability program.

- **Community Representatives in Indonesia largely increased their optimism about the program from start to finish, whereas Community Representatives in Tanzania became more skeptical**. At the start of the program, most CRs in both countries were optimistic regarding the program and the potential results of their actions, and we wanted to see if this changed as they implemented and observed the results of their actions. In Indonesia, we saw a general increase in optimism, with approximately 80 percent of those who start as skeptical becoming optimistic and only 10 percent of those who started as optimistic becoming skeptical.

  On the other hand, 60 percent of those who started as optimistic in Tanzania became skeptical by the end whereas approximately 50 percent of those that started as skeptical became optimistic.

- **A similar pattern emerged with regard to perceptions of civic efficacy among CRs in Indonesia and Tanzania**. We asked CRs in both countries the question of how able they felt to make improvements in their community. In Indonesia, those that felt somewhat or very able increased from 40 percent in the first meeting to almost 70 percent in the final meeting. Tanzania, on the other hand, showed little movement, with almost 100 percent reporting that they were somewhat or very able in the first meeting and then dropping to closer to 90 percent by the final meeting.

***How do we know this?*** For the general community, these results come from the baseline and endline data collection and regression analysis described in the previous set of results. You can read more about these results and our methods here. For the CRs, we conducted a brief survey, including several vignettes that demonstrated empowerment, at the very beginning of the program and immediately after the third following up meeting. The CR efficacy and optimism results come from our analysis of this data in approximately 40 percent of communities. You can read more about the methods and results in "Encouraging Participation" (2019).

***How did women experience the program differently than men***

The issue of gender disparities and differences in the results of the T4D project is complex. On the one hand, women participated at much higher rates than men, both speaking more and making up a greater percentage of CRs in both countries. While on the surface, this appears to be a positive finding, the potential underlying causes of these changes are decidedly mixed. CSO facilitators actively sought to recruit a diverse set of CRs, which meant that women likely were engaged more actively than they have been in similar programs in the past. There is some evidence from observations and focus group discussions that many female participants gained a sense of efficacy and empowerment from the experience that could be carried through to future activities. However, the greater participation of women may also have been due to the fact that maternal and newborn health is seen as a women's issue by many men who thus decided not to participate as actively and/or that men were discouraged by the lack of payment for their participation. More exploration is needed to understand the degree to which these positive and negative influences played a role in great female participation; however, it is likely that all of these potential root causes occurred to some degree.

# Result 4 - Unpacking why health did not improve, but empowerment may have.

So what can we make of the fact that a social accountability program like this one can catalyze diverse and numerous actions by community members, can in many cases increase optimism and a sense of self-efficacy, but that this ultimately does not result in any observable improvements in health above and beyond what we saw in control communities?

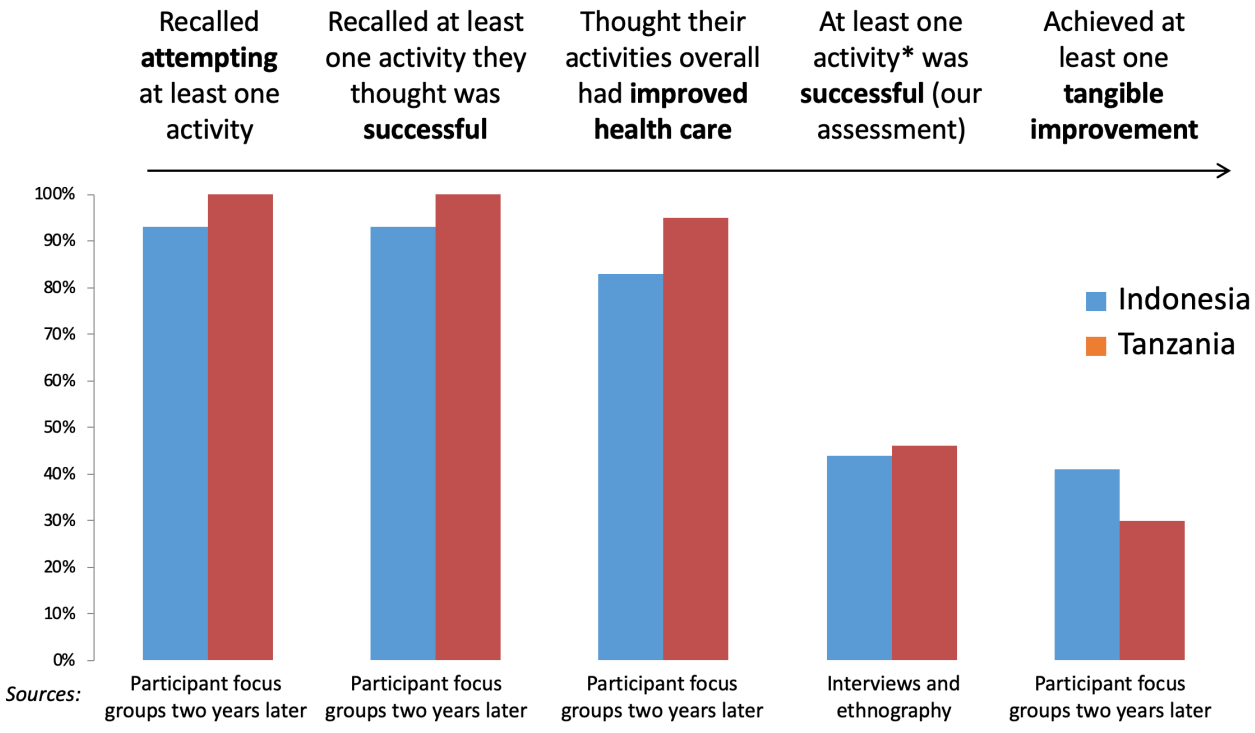
This is one area where this project can make a contribution to the social accountability field – by starting to open the black box to help us explain why programs can make progress, but not enough, and to help us understand what might work better (something we explore in the next results sub-section).

***What did we see?*** Most importantly, there is not one concrete and absolute reason why this social accountability program did not work. While this means that there is not a single thing that could be tweaked to make the difference in impacting health outcomes, it does mean that there are a lot of design choices that deserve further exploration that could make this work more effective.

# *Key Findings*

- **Many actions that were completed still did not lead to tangible outcomes.** The first thing we did to learn about "why" was to do a deep dive into the arrow between actions and intermediate outcomes in our theory of change; in other words, what can we say about the relationship between actions that happened and the intermediate outcomes that did not. What we see is a slow drop off in potential effectiveness of actions that is shown in Figure 5. When we spoke with CRs two years after the program, they recalled specific actions and recalled them being successful at rates of about 90 percent; however, when asked about tangible improvements from their actions, only 41 and 30 percent in Indonesia and Tanzania respectively could name a result of their actions. This closely mirrored our attempts to verify action outcomes; based on key informant interviews and ethnography, less than 50 percent of villages had actions that we could consider successful based on these external assessments. So the actions, even when completed, often could not make it to the point of having a tangible outcome.

**FIGURE 5. ACTIONS ATTEMPTED AND SUCCESSFUL IN INDONESIA AND TANZANIA**




* Excludes informing or educating others in the community about the importance of using biomedical maternal and newborn healthcare services

- **History with development plays an important role**. One of the major themes that we explored using the ethnographic component of this work was how CRs experienced the social accountability program. A finding that clearly emerged was that this was not the first experience that communities had with this type of community-driven development program in most (if not all) cases. The actions that were implemented closely mirrored things we observed also happening in control communities. And so while these actions still may have been valuable, the program did not catalyze actions that were unique.

- **Expectations regarding resources likely played a role as well**. In an earlier section, we shared that one of our design principles was to ensure that the program was largely free of outside resources, including per diems for participation. This became a hurdle that the facilitators spent significant time working to overcome, and it did result in some attrition of CRs. We would not suggest that this points to a need to include payment future models, but it is an important issue to be aware of in future design.

- **Context matters.** While this statement on its own is not revolutionary, there are specific hypotheses that emerged from our work that highlights ways in which context may matter in terms of places and settings where social accountability may be more effective. One contextual factor that emerges from both reviewing other studies and from diving into the results from T4D is the quality of healthcare. In comparing a set of evaluations of health scorecards, some with positive impact and some with no impact, one trend that stands out is that the quality of healthcare provision for those that saw significant improvements was lower than in those settings where no improvements were found. In doing a review of communities in the T4D sample that saw multiple tangible improvements in outcomes (ones we are considering positive deviants), we also see that people in those communities experience their healthcare as being poorer quality. This may suggest that these programs are better at catalyzing effective community action when people believe that their healthcare could be better.

***How do we know this?*** These findings came from the full range of data collected for this project, including an endline community survey, key informant interviews and ethnography. You can read more about the results shared in Figure 5 in "Can transparency and accountability programs improve health?" (2019); and the findings from the ethnography and the positive deviants are forthcoming and will be shared on our website when available.

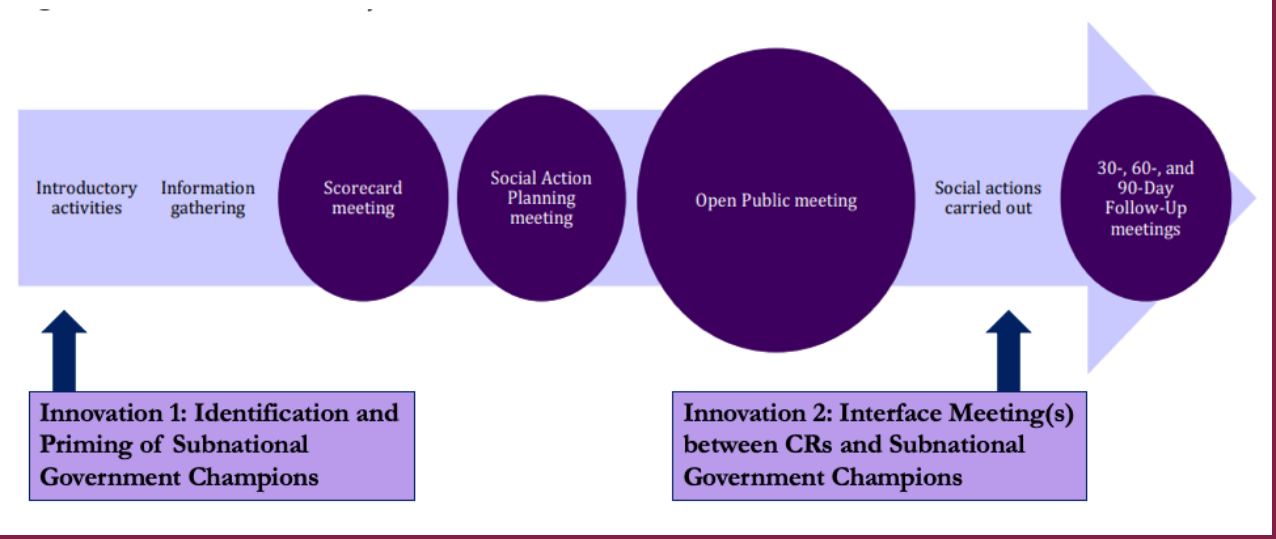
# Result 5 - What might work better in terms of social accountability design.

There is still a lot to explore on this topic, but we did take some of our early observations from this work to test one adapted approach. Everything we have discussed so far has focused on two countries – Indonesia and Tanzania. As discussed in an earlier results sub-section, one finding from these countries was that the actions that communities chose to undertake focused overwhelmingly on the local level. Which raises the question – could actions be more effective if we overcame obstacles preventing the design of actions targeting higher levels?

Working with three new CSO partners in Ghana, Malawi, and Sierra Leone, we co-designed an adapted version of the intervention. The resulting model (Figure 6) introduced two innovations:

First, in the new approach, the CSO facilitators and leaders began by **identifying and recruiting government champions** above the community level to be involved in the project. These individuals expressed a willingness and interest in working with communities to help them with health problems and actions that they designed. In Ghana and Sierra Leone, these were elected and appointed district officers, while the champions in Malawi were traditional authorities. You can read more from one of our partners about lessons for priming government champions in Box 3.

**FIGURE 6. ADAPTED SCORECARD INTERVENTION (PHASE 2)**

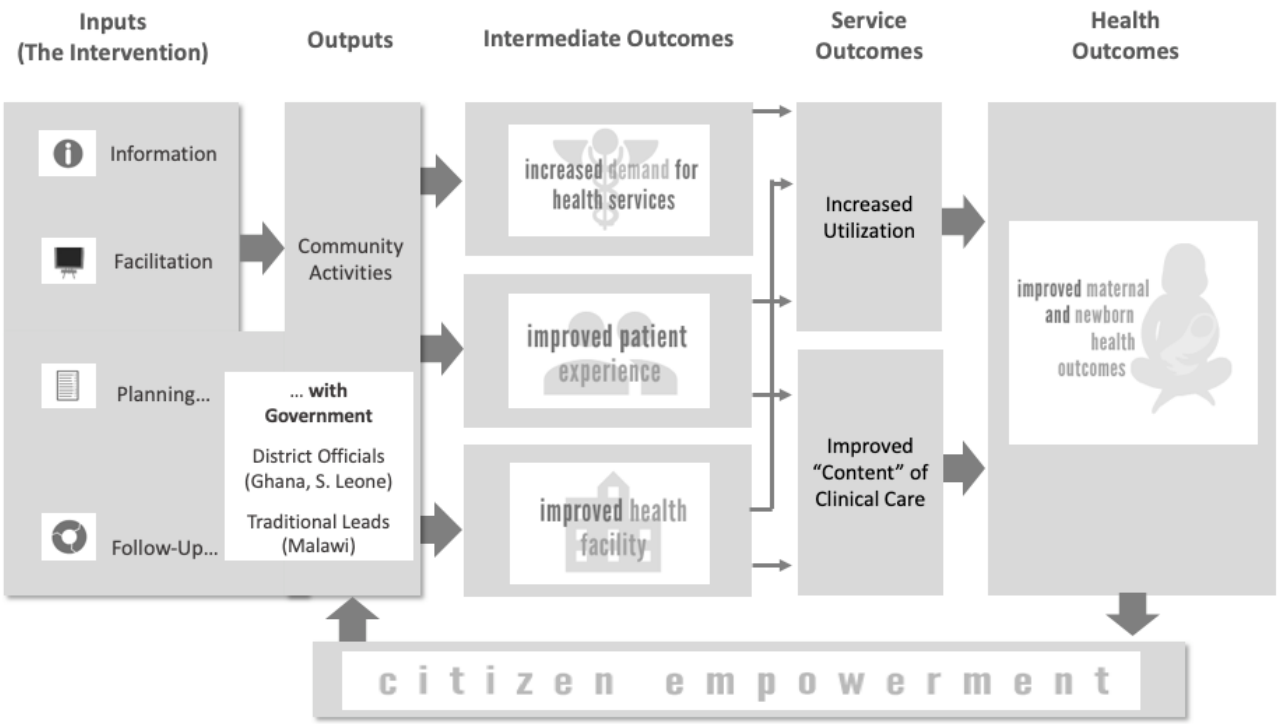


Second, during the action implementation period, CRs had 1-2 **meetings with these government champions**, to share their action plans, to get feedback from the government officials, and where appropriate to ask for their support.

***Why these innovations?*** Our sense was that directly connecting CRs with government champions could in some cases shorten the causal chain that the actions had to traverse to get to outcomes. CRs would be able to undertake actions that involved requests to higher level officials that they might not otherwise get access to, removing one hurdle in the chain to health outcomes. The resulting theory of change is represented in Figure 7.

***What did we see?*** The inclusion of government champions did show evidence of improving the program in most places, but not all. Ghana and Sierra Leone both experienced changes that we would expect to see (more actions geared toward higher level officials) as well as some changes that were not guaranteed. In Malawi, the resulting program looked more like the work in Indonesia and Tanzania.

**FIGURE 7. PHASE 2 (GOVERNMENT CHAMPIONS) THEORY OF CHANGE**



## *Key Findings*

- **Meetings with Government Champions were associated with more actions targeting higher level officials.** In Ghana and Sierra Leone an average of 57 percent of actions (and 70 percent of communities) targeted government officials above the village level, as compared to 8 percent of actions (and 44 percent of communities) in Tanzania and Indonesia. Malawi had more actions targeting government officials than in the Phase 1 countries (20 percent) but these came from a smaller number of communities.

- **Perceived tangible improvements were also higher in Ghana and Sierra Leone.** We collected data both from the CRs themselves and through independent researchers to assess the degree to which actions in these countries could be associated with tangible improvements. In Ghana and Sierra Leone, CRs in 60 percent of communities could recall a tangible improvement related to one or more of their actions, and that proportion is even higher according to independent researchers (70 percent). This is an increase from what we observed in Tanzania and Indonesia, (CRs in 36 percent of communities reported improvements, and 40 percent verified), and all are higher than what we saw in Malawi (0 and 20 percent reported by CRs and verified independently, respectively).

***Box 3. Insights on working with Government Champions***

*(Mohammed Awal, Center for Democratic Development in Ghana)*

Four characteristics that make good government champions are: (1) openness to and value of the idea and principles of social accountability, (2) belief in the mandate of public service in social service provision, (3) capacity for statesmanship, and (4) possessing social capital in the community.

Key steps to building these relationships include: (1) involving Government Champions early, including in the preparation phase, (2) building clear links of communications between the Government Champions and Communities, and (3) providing regular updates to the Government Champions, even when there is a break from meeting with CRs.

*You can read more in Awal's blog here.*

- **Ghana and Sierra Leone diverge somewhat in the CR perceptions of government responsiveness.** In Ghana, we asked CRs how responsive they experienced government officials to be before and after the program, and 26 percent increased their view of government responsiveness while 13 percent decreased from beginning to end of the program. While this can be seen generally as a positive change, this change pales in comparison to the change observed in Sierra Leone. Asked the same question, almost six times as many Sierra Leonean CRs experienced an increase in perception of government responsiveness as compared to a decrease (59 percent and 10 percent respectively). It is worth noting that a relatively small set of interactions and responses could be associated with such a significant improvement.

- **Communities were not the only ones who experienced the program as positive.** In interviews with government champions after the program, many highlighted ways in which the CRs and their interaction with this group of community members helped the government officials better perform in their jobs.

***How do we know this?*** We worked with CSO partners to test this program in five communities each in Ghana, Malawi, and Sierra Leone, and we worked with local researchers to collect data including meeting observations, action tracing, empowerment surveys, and government interviews. It is important to interpret these results with some caution because the program was only conducted in a small sample of communities. However, the evidence still provides helpful insights and a potential foundation for future work. The findings from Phase 2 are forthcoming and will be shared on our website when available.

## *What else can we learn from comparing more and less successful cases*

While the limited cases studied (especially in Phase 2) have to be interpreted with caution, the trends observed across the five different settings and adaptations reveal some patterns that may be valuable for future work and practice:

- As noted above, the **engagement of Government Champions** is incredibly promising in cases in which these individuals can be identified as well as primed to collaborate with citizens and communities. There were also more positive results when the officials acting as Champions had actual authority and/or decision-making power over at least some aspects of the part of the system than communities were seeking to improve. As such, we recommend a very deliberate set of selection criteria and priming process for Government Champions for social accountability.
- Further, observations that many actions were completed that did not result in tangible outcomes suggests that more could be done to **work with community representatives on the root causes of health system challenges and the types of actions that could overcome these**. The dedication and efforts of community members is promising, but more may need to be done to focus on "the right" actions.
- Analyzing trends in Indonesian and Tanzanian communities in which there was evidence of some concrete change related to social actions reveals other promising characteristics of "more successful" communities. First, in Tanzania specifically, we observed that communities that achieved positive changes had **more open and engaging experiences with the health system**. This was according to both subjective and objective measures, including citizen perceptions of trustworthiness of their facility and complaint management systems and outreach activities led by health facilities.

## *What else can we learn (continued)*

- In Indonesia, we observed that those communities with concrete positive changes were more likely than the larger sample to have greater awareness and value of the importance of modern maternal and newborn healthcare and thus were potentially more informed about the need to advocate for improvements in modern care.
- Finally, we observed statistically significant differences between how those communities who experience positive changes perceived their health system performance as compared to those who did not experience changes; however, the direction of the difference was not the same in each country. In Indonesia, positive change communities had **higher perceptions of health system performance** than the larger sample, whereas Tanzanian positive change communities had **lower perceptions of quality**.
- For the final three differences, these trends could suggest one of two directions for future work. One possible recommendation is to take these contextual characteristics as given, and so practitioners should be selective in which communities to work in based on the characteristics associated with positive changes. Alternatively, practitioners could build these factors into the design of their work, including design components such as trust building for citizens and capacity building for health providers to increase the likelihood of positive changes.


In the previous sections, we try to present what ends up being a very mixed picture of whether and how social accountability can improve health. Broken down into simple answers to our research questions, the story may feel deceptively clear: we don't have evidence that a program like this can improve health outcomes or wider community empowerment, but we do see that it can catalyze community action and (in some cases) the empowerment and self-efficacy of those directly involved.

But what does all of this mean for the audiences we started this project hoping to inform – CSOs and donors, health people and governance people, policymakers and researchers? Taking all of this evidence together, we see three possible ways to understand this project and corresponding paths forward on the surface. And, of these potential interpretations, we recommend one that should help guide how people use these results.

The first – and most dire – interpretation of these findings is that people should not continue to implement and support social accountability because we did not reveal an impact on health outcomes.

However, this is an oversimplification of these results for many reasons. First, we tested a single carefully-designed social accountability program in Phase 1 of this work and piloted an adapted version of the program in Phase 2; while our findings were designed to draw out more generalizable understandings for the field, the evaluation itself, and thus the results, still focused on a specific set of designs. Second, assessing the full theory of change for the program highlighted that some communities traversed several steps of the causal chain in getting to independently verified tangible outcomes – even more so in the cases of Ghana and Sierra Leone.

In this vein, a second interpretation may be that we should not stop doing social accountability – but should temper our expectations about what it can achieve. The evidence from this project highlighted the outcomes that even a relatively "light touch" program can achieve, including catalyzing the design and completion of social actions by community volunteers and improving civic efficacy for some CRs. When we began the program in 2013, we met with one CSO that noted that their goal in their programming was to encourage citizen empowerment, full stop.

And so for some organizations, it may be enough to support and undertake social accountability that strengthens engagement and empowerment of those whose voices are often ignored.

But there is a third interpretation, one that we would encourage CSOs and donors, policymakers and researchers alike to consider – that there is potential in using experience and evidence (including what was developed in the T4D project) to design better social accountability.

We believe that one of the biggest contributions of this project was developing a better understanding of what did not happen the way we expected:

- Social actions in Phase 1 were diverse – *but they did not target higher level government actors who may be necessary to enact change.*
- These same actions were overwhelmingly collaborative, which is undoubtedly appropriate in some settings – *but are there other problems that really necessitate more confrontational approaches?*
- Without the constraint of actions needing to focus on social accountability approaches, the majority of actions were community self-help – *but does this mean that we should be encouraging more traditional community-led "transparency and accountability actions" or instead focusing on integrating transparency and accountability for fully with community-driven development?*

These observations – as well as signs of what did work better – suggest some potential ways forward that do not equate to abandoning social accountability. For one example, adding meetings with government officials who were noted champions of community engagement shows promise, even when done in a relatively light touch way. And these benefits extended beyond the community members themselves. More examples are included in the Box on pages 31-32 of this report.

These results, among others, point to promising approaches to iterate on. We did not test any one model of social accountability that got it exactly right. But the T4D results do highlight that taking design elements that are promising while experimenting with overcoming gaps in the design may still help us to uncover the real potential of these types of approaches, even in moving the needle on health.

## *About the Ash Center*

*The Roy and Lila Ash Center for Democratic Governance and Innovation advances excellence and innovation in governance and public policy through research, education, and public discussion. By training the very best leaders, developing powerful new ideas, and disseminating innovative solutions and institutional reforms, the Center's goal is to meet the profound challenges facing the world's citizens. The Ford Foundation is a founding donor of the Center. Additional information about the Ash Center is available at ash.harvard.edu.*

## *About Results for Development*

*Results for Development Institute (R4D) is a non-profit organization whose mission is to support change agents around the globe to create self-sustaining systems that support healthy, educated people. Using multiple approaches in multiple sectors, including Global Education, Health, Nutrition, Accountability and Citizen Engagement, R4D supports the discovery and implementation of new ideas for reducing poverty and improving lives around the world.*

**********



HARVARD Kennedy School
ASH CENTER
for Democratic Governance and Innovation

RESULTS FOR DEVELOPMENT