# IMPACT OF SAFE MOTHERHOOD ACTION GROUPS ON USE OF MATERNAL HEALTH CARE IN ZAMBIA

October 2014

This publication was produced by Ben Johns, Nikita Ramchandani, Alison Comfort, and Slavea Chankova for the Zambia Integrated System Strengthening Project.

**Recommended Citation:** Johns, Ben, Nikita Ramchandani, Alison Comfort, and Slavea Chankova. October 2014. *Impact of Safe Motherhood Action Groups on Use of Maternal Health Care in Zambia.* Bethesda, MD: Zambia Integrated Systems Strengthening Project, Abt Associates Inc.

# TABLE OF CONTENTS

# ACRONYMS

| | |
|---|---|
| ACNM | American College of Nurse Midwives |
| ANC | Antenatal Care |
| CARMMA | Campaign for Accelerated Reduction of Maternal Mortality |
| CBMNLSS | Community-based Maternal and Newborn Life Saving Skills |
| CSO | Central Statistical Office |
| D-in-D | Difference in differences |
| D-in-S | Difference in slopes |
| DHMO | District Health Management Office |
| DHS | Demographic and Health Survey |
| EmONC | Emergency Obstetric and Newborn Care |
| FP | Family Planning |
| GEE | Generalized Estimating Equation model |
| HBLSS | Home-based Life Saving Skills |
| HIV/AIDS | Human immunodeficiency virus/Acquired immunodeficiency syndrome |
| HMIS | Health Management Information System |
| IRR | Incidence rate ratio |
| JSI | John Snow International |
| MCDMCH | Ministry of Community Development, Mother and Child Health |
| MDG | Millennium Development Goals |
| MMR | Maternal Mortality Ratio |
| MNCH | Maternal, Newborn, and Child Health |
| MOH | Ministry of Health |
| PMTCT | Prevention of mother to child transmission of HIV |
| PNC | Postnatal Care |
| SMAG | Save Motherhood Action Group |
| SMGL | Saving Mothers Giving Life |
| UN | United Nations |
| UNFPA | United Nations Population Fund |
| USAID | United States Agency for International Development |
| ZISSP | Zambia Integrated Systems Strengthening Program |

# ACKNOWLEDGMENTS

The authors of this report would like to acknowledge the contributions of the many individuals who provided advice on the design of the evaluation, assisted with the data collection process, and provided valuable comments to earlier drafts of the report.

We would like to thank Victoria Musonda, Douglas Johnson, Anna Maria Speciale, Suzanne Stalls, Karen Rowe, Kathryn Stillman, Kathleen Poer, Benson Bwalya, Bernard Kasawa, and Jeanette Kesselman for their thoughtful comments and input during the design of the evaluation and/or to earlier drafts of this report.

We are particularly grateful to Peter Funsani and Patrick Mwansa for carrying out the data collection from the HMIS and to Anna Maria Speciale for gathering information on the SMAG program background and implementation.

Special thanks to all SMAG members, health workers, and ZISSP field staff who were generous with their time during the field visits by the evaluation team.

# EXECUTIVE SUMMARY

## BACKGROUND

Zambia has been making progress in improving maternal health but the maternal mortality ratio remains high, at 398 deaths per 100,000 live births. One of the main reasons is that over a third of women are not delivering at a health facility and are not delivering with skilled attendants. One of the interventions implemented by the Government of Zambia to improve maternal and neonatal outcomes is the Safe Motherhood Action Group (SMAG). SMAGs are community-based volunteer groups that aim to reduce critical delays in decision-making at the household level about seeking life-saving maternal care at health facilities. SMAG members deliver essential information on safe motherhood to men and women; encourage pregnant women to go for antenatal care, delivery, and postnatal care in a health facility; identify maternal and newborn complications during pregnancy, delivery, and the postnatal period; offer first aid care, and refer cases with maternal and newborn problems for management at health facilities.

The Zambia Integrated Systems Strengthening Project (ZISSP) is a five-year USAID-funded project that works closely with the Ministry of Health (MOH) and the Ministry of Community Development, Mother and Child Health (MCDMCH) to strengthen skills and systems for planning, management, and the delivery of health services. Since February 2012, ZISSP has worked with the MOH and the MCDMCH to establish and train SMAGs in 17 districts, covering the catchment areas of 139 health facilities.

The SMAG training provided with ZISSP's support utilizes the Home Based Life Saving Skills (HBLSS) curriculum and training methodology, developed by the American College of Nurse Midwives. HBLSS utilizes a community education approach designed for low-literacy populations that emphasizes the respectful consideration of the existing knowledge and solutions before attempting to instill biomedical practices, and builds community consensus for acceptable, feasible actions to take in response to maternal and neonatal complications. These actions enable the community and family members to provide a "first responder" care while reaching health facilities for life-threatening problems. The skills are taught using role-play demonstration, problem identification, problem solving, negotiation, and practice. SMAG members teach these skills during community meetings that they organize; these meetings are the cornerstone of the SMAG program.

This report presents the results from an evaluation of the impact of SMAGs on utilization of maternal care at health facilities. The evaluation covers the SMAG program supported by ZISSP. Other organizations support SMAG programs in different parts of the country, many using training approaches different from the HBLSS approach; these other SMAG programs are outside the scope of this evaluation.

## METHODS

The evaluation measured the impact of SMAGs on five outcome indicators, measured at each health facility that has the SMAG program implemented in its catchment area:

1. Number of institutional deliveries
2. Number of first antenatal care visits before 20 weeks after conception

3. Number of first antenatal care visits at 20 weeks after conception or later
4. Number of postnatal care visits within six days of delivery
5. Number of postnatal care visits between six days and six weeks of delivery

To measure the effect of the SMAG program, we compare changes in the five outcome indicators in facilities after the SMAGs training was conducted in their catchment area to changes in a comparison group of facilities without SMAGs during the same time period. The comparison group was selected using a propensity score matching process that aimed to replicate the selection process of the facilities covered by the ZISSP-supported SMAG program.

The study utilized several existing datasets; no primary data collection occurred for this report. To match SMAG facilities with comparable non-SMAG facilities, we used data from three sources that had information on district and facility characteristics: the 2012 MOH Health Facility Listing, the ZISSP District Selection Plan, and the national MOH Health Management Information System (HMIS). We used the national HMIS database as the source of our five outcome indicators. The evaluation includes facilities for which data were available for two to 18 months *after* they received the SMAG program.

We used regression analyses and two difference-in-differences models to measure the effect of the SMAG program on the five outcomes: (i) a model that compared changes in outcomes between the SMAG and the comparison group; and (ii) a model that compared changes in the trend (slope) of outcomes before and after the SMAG program. The regression analyses used negative binomial models. In addition, a number of sensitivity analyses were performed to test the robustness of the main results.

## RESULTS

The main results show that there is:

- Evidence for a positive association between the SMAG program and an increased utilization of institutional deliveries from about 17.8 per facility per month without the SMAG program to around 20 with the SMAG program, representing an increase of about 12%;

- Limited evidence for a positive association between the SMAG program and an increased utilization of PNC;

- Inconclusive evidence of an association between the SMAG program and utilization of ANC.

These results are consistent with the focus of the SMAG program – the contents of the community meetings held by the SMAGs was heavily weighted on delivery at a health facility, and this is the indicator for which we found the most consistent evidence of an effect for the program. The lack of a larger effect size for institutional deliveries may be due to low intensity of the program across SMAGs: for example, SMAGs may be present in a small proportion of communities in a facility catchment area, or SMAGs may not hold monthly meetings, adhere to the HBLSS curriculum, or be effective in engaging their audience. In some areas, lack of skilled staff in facilities and transport barriers (such as unavailability of transport options at night) may discourage or limit women's ability to deliver in the facility. It is also possible that the increase in institutional deliveries measured in this initial time period of the SMAG program is primarily a result of care-seeking by women who experience complications; in that case,

there would still be scope in the long run for the SMAG program to effect an increase in facility-based deliveries for non-complicated cases.

## CONCLUSIONS

This evaluation aimed to contribute to the evidence on the effect that training SMAGs to educate communities on safe motherhood practices has on the utilization of facility-based maternal care. The positive impact of SMAGs on institutional deliveries is a notable and encouraging result, as it is during delivery that a large proportion of life-threatening complications occur.

An assessment of specific implementation practices and lessons of the HBLSS-based SMAGs training was not part of this evaluation but would be a valuable study. It can shed light on changes in SMAG program implementation that could lead to even greater impact on utilization of facility-based care for mothers and newborns.

# 1. BACKGROUND

## 1.1 MATERNAL HEALTH SITUATION IN ZAMBIA

With the introduction of the Millennium Development Goals (MDGs), maternal mortality was given new importance on the world stage. MDG 5 calls for a reduction by three-quarters in the maternal mortality ratio (MMR) between 1990 and 2015. As the 2015 deadline quickly approaches, it is clear that of all the MDGs, MDG5 is the least attainable, especially in Sub-Saharan Africa where there have been the smallest MMR declines (Hogan et al. 2010).

Zambia, in particular, faces significant challenges in achieving its MDG 5 target. The MMR has fallen from 729 maternal deaths per 100,000 live births in 2001/2002 to 398 by 2013/2014 (CSO, 2014; CSO 2007). Compared to 1990, the reduction in MMR represents a 45% decline, which is substantial but insufficient to reach MDG5 by 2015.

Some of the reasons for the lack of progress in reducing MMR may be due to barriers to accessing early antenatal care (ANC) and care during and after delivery. Almost all (96%) pregnant women in Zambia receive some ANC from a skilled provider (CSO, 2014). However, many seek ANC late in their pregnancy; only 19% have their first ANC visit during the first trimester of pregnancy (CSO, 2007).[1] Important barriers to appropriate and timely ANC continue to include cost of transport, poor road networks, late recognition/denial of pregnancy, reliance on traditional/tribal (herbal, spiritual) pregnancy beliefs, lack of perceived benefits from ANC services, influence of male partner and a dissatisfaction with the attitude of staff at antenatal clinics (Miaffo et al., 2004; Larsen et al., 2004; USAID, 2013).

In recent years, the Ministry of Health (MOH) and Ministry of Community Development, Mother and Child Health (MCDMCH) have advocated facility delivery to reduce maternal and neonatal morbidity and mortality. This was emphasized by the MOH during the 2010 launch of the Campaign for Accelerated Reduction of Maternal Mortality in Zambia (CARMMA) with a running slogan "Zambia cares: No woman should die while giving life" (UNICEF 2010). Trends in facility-based deliveries and use of skilled attendance at delivery have only recently been improving in Zambia. The proportion of women who delivered in a health facility increased substantially, from 44% in 2001/02 to 67% by 2013/2014 (CSO 2014). Similarly, the proportion who delivered with a skilled birth attendant also rose from 43% in 2001/02 to 64% in 2013/2014 (CSO 2014). However, these trends still imply that more than a third of women in Zambia are not delivering in a health facility and are not delivering with skilled attendants, two major contributors to maternal mortality and morbidity.

[1] These data for 2013/2014 are not yet available.

## 1.2 SAFE MOTHERHOOD ACTION GROUPS

One of the interventions implemented by the Government of Zambia to improve maternal and neonatal outcomes is the Safe Motherhood Action Group (SMAG). SMAGs are volunteer groups based in the community that aim to reduce critical delays in decision-making at the household level about seeking life-saving maternal care at health facilities. The SMAGs aim to improve the links between men and women of reproductive age and the health facility for maternal and newborn care seeking. SMAG members deliver essential information on safe motherhood to men and women; encourage pregnant women to go for antenatal care, delivery, and postnatal care in a health facility; identify maternal and newborn complications during pregnancy, delivery, and the postnatal period; offer first aid care, and refer cases with maternal and newborn problems for management at health facilities.

SMAG members typically include men and women who are not healthcare professionals but are already involved in some community level health-related activities and who are generally respected by their community. They may include community health workers, traditional leaders, traditional birth attendants (who are often no longer practicing, after the change in government policy on home delivery which promotes deliveries in a health facility for all women), and Neighborhood Health Committee members. SMAG members are volunteers who do not receive any monetary support, but may receive various materials that assist them in their work. They usually have other work and responsibilities in their households and communities, but are on-call for the community members 24 hours a day for emergencies related to maternal and newborn complications.

The SMAG program was established in Zambia in 2003 as part of a national safe motherhood program. SMAGs were initially supported by the United Nations Population Fund (UNFPA) in North-Western Province before the Government of Zambia adopted them as a national program in 2008. Since 2003, SMAGs have been established and trained across Zambia, with support from a range of development partners. By 2010, SMAGs had been established in 45 of Zambia's 72 districts[2] with provincial level coverage ranging from less than 20% of districts in Copperbelt Province to all districts in Luapula, Northern, and North-Western Provinces. However, because of a lack of a central coordination for the SMAGs, communities often did not know how to use these groups after they were established and many SMAGs have remained inactive (Ensor et al. 2013). The reason for this is that there was no specific person assigned for this coordination at the MOH, and the SMAG program has been donor-driven.

## 1.3 ZISSP SUPPORT FOR THE SAFE MOTHERHOOD ACTION GROUP PROGRAM

The Zambia Integrated Systems Strengthening Program (ZISSP) is a five-year USAID-funded project that works closely with the MOH and MCDMCH at the national, provincial, district and community levels to strengthen skills and systems for planning, management, and the delivery of health services through capacity building strategies and technical assistance. ZISSP also works with communities to foster increased use of public health services by Zambians and their families. ZISSP interventions are designed to strengthen and expand the delivery of high-impact health services for malaria, maternal, newborn and

[2] Until 2013, Zambia was subdivided into 72 districts. In 2013, 17 new districts were created.

child health, family planning, nutrition and HIV/AIDS. ZISSP has targeted most of its technical assistance work in 27 districts in Zambia.

The work of ZISSP supports the eight priority health areas identified by USAID for Zambia: 1) Increase access to reproductive health, sexual health, and family planning services, especially in rural areas; 2) Strengthen reproductive health and family planning policies and improve planning and resource allocation; 3) Increase access to and education about family planning; 4) Increase access to high quality antenatal care; 5) Increase access to skilled delivery care; 6) Provide prompt postpartum care, counseling, and access to family planning; 7) Improve post-abortion care; 8) Strengthen health promotion activities.

As part of its work targeted at improving maternal and neonatal health, ZISSP has worked with the MOH and MCDMCH to establish and train SMAGs in 17 districts, covering 139 health facilities. The SMAG training provided with ZISSP's support utilizes the Home Based Life Saving Skills (HBLSS) curriculum and training methodology, developed by the American College of Nurse Midwives.[3] This model (described in detail in the following section) is used for training community volunteers in community mobilization to promote safe motherhood outcomes. The purpose of the training is to: provide SMAG members with information about safe motherhood and newborn health; develop SMAG members' skills in group facilitation and education, community mobilization, and information gathering and reporting; and build networks of cooperation and support among SMAGs, health center staff, and other community-based organizations. This training is expected to translate into skills and knowledge among community members around maternal and newborn care, including: birth preparedness, complication readiness, ability to provide initial emergency care until reaching the facility, and accessing care at a health facility.

The SMAG program aims to:

1. Strengthen community participation in maternal, newborn, and child health
2. Improve community knowledge on safe motherhood issues through health education
3. Increase male involvement in safe motherhood activities
4. Strengthen partnerships between the community and health care system
5. Enhance utilization of reproductive health services

The primary objectives of the SMAG program are to:

- Create a group of skilled HBLSS SMAG trainers who conduct regular community meetings with leaders of the SMAGs that are within the health center catchment areas in their districts.
- Increase the understanding and use of life-saving measures to be utilized while en route to the health facility, in case of complications.
- Increase access to facilities through referrals and through improved community understanding/identification of problems requiring referral.

The ZISSP SMAG training began in February 2012 and utilized a cascading training approach beginning with the training of Master Trainers, who trained SMAG trainers, who in turn trained the SMAG

[3] This curriculum was adapted for the local context and was subsequently re-titled Community Based Maternal and Newborn Life Saving Skills.

members. By 2014, 40 Master Trainers, 92 SMAG trainers and over 3,000 SMAG members have been trained.

The MOH and MCDMCH have also been supported by several other development partners, such as UNFPA for example, to establish and train SMAGs throughout Zambia. Some of these organizations use training curriculums and approaches broadly similar to the HBLSS-based model used by ZISSP. Others, however, have used teaching approaches that are more didactic (i.e. focused on teaching the technical subject matter), without a prominent participatory skills-building component such as the one that is central to the HBLSS model. They have targeted health workers from selected health facilities and health workers were expected, in turn, to work with the community to identify and train SMAG members in their catchment areas using the same methodology. By contrast, the approach supported by ZISSP was to train the SMAG trainers who in turn directly trained SMAG members using the HBLSS curriculum; the approach to this training is more participatory: the SMAGs themselves are left to identify incidents that happened in their community which ended in a maternal death, and assess the gaps and work out solutions.

In addition, 10 of the 17 districts in which ZISSP supported SMAGs are part of the Saving Mothers Giving Life (SMGL) initiative, a five-year public-private partnership to rapidly reduce maternal mortality. Interventions through SMGL include equipping facilities to provide high quality emergency obstetric and newborn care, training skilled birth attendants in quality and respectful delivery care, and improving equipment and supplies needed at facilities for quality delivery care.

## 1.4 EVALUATION OF THE IMPACT OF SAFE MOTHERHOOD ACTION GROUP PROGRAM

This report presents the results from an evaluation of the effect of the SMAGs trained with ZISSP support. The evaluation aims to answer the question: *Does the training and support of SMAGs lead to an increase in utilization of maternal health services at health facilities?*

The evaluation aims to inform the MOH, MCDMCH, ZISSP, and other partners about the effect of the SMAG program on use of health facilities for antenatal care, delivery, and postnatal care. To our knowledge, evidence on this topic is scarce. One study that we are aware of found that SMAGs in rural areas who were trained to conduct community discussions on safe pregnancy and delivery plus the provision of emergency transport led to significant increases in knowledge about danger signs and maternal care seeking at health facilities, as measured by a population survey (Ensor et al. 2013). However, that study did not measure separately the effect of SMAG training alone.

This evaluation focuses on assessing the impact of the SMAGs trained with ZISSP support using the HBLSS curriculum. It is important to note that results may be different for SMAGs trained by other organizations or with other training approaches (which are outside of the scope of this study).

Although the evaluation was supported through the ZISSP project, the evaluation team members were all external to the project and did not have any involvement in the design or implementation of the intervention.

The remainder of this report is structured as follows. Section 2 describes the SMAG program in ZISSP-supported districts (i.e. the intervention being evaluated); Section 3 presents the evaluation methods; Section 4 presents the results; Section 5 includes discussion of the results; and Section 6 provides some overall conclusions.

# 2. PROGRAM DESCRIPTION

SMAGs throughout Zambia serve as a vehicle to spread health awareness to communities. SMAGs are supported by many government and development partners' organizations to achieve a wide variety of health objectives and they are trained in multiple health topics. Although SMAGs trained and supported by ZISSP undergo a training using a specific curriculum and have a set of prescribed activities and objectives, on the ground these trainings and activities have been influenced and modified by a variety of external factors which vary from one SMAG to the next (e.g. other development partners' trainings and support materials, on-the-ground program adaptation initiated by SMAG members, etc.)

This evaluation did not systematically collect implementation evidence on the extent to which the ZISSP-trained SMAGs followed, in reality, each component of the prescribed model on how they should operate after their training. Anecdotal evidence suggests that some SMAGs had modified or chosen not to use certain methods prescribed in their training (e.g. the picture books during the community meetings), while others have added components that they considered appropriate to deliver the messages (e.g. songs and dance).

Team members for this impact evaluation visited a small number of ZISSP-trained SMAGs to observe their activities and speak with SMAG members about their work. These learning visits aimed to generally familiarize the team with the SMAG program, but provided only limited insights about the extent to which program implementation followed the prescribed model. The program description below outlines the prescribed model and notes examples of some differences in its implementation that were employed by some SMAGs.

## 2.1 HOME BASED LIFE SAVING SKILLS TRAINING CURRICULUM

The central component of the ZISSP support for the SMAGs included in this evaluation is the training of SMAG members on the HBLSS package. HBLSS was developed by the American College of Nurse-Midwives in 1998 as a response to the high levels of maternal and neonatal mortality in areas of the world with elevated levels of home births and limited access to skilled care (Lori et al. 2012). HBLSS is a community-based skills-building program that aims to reduce maternal and neonatal mortality by increasing use of basic life-saving measures within the home and community, and by decreasing delays in reaching health facilities that can provide care for life-threatening problems (Sibley et al. 2001).

Objectives of the HBLSS curriculum include: (1) decreasing delays in the recognition and response to major complications during pregnancy and delivery, (2) increasing access to emergency maternal and neonatal care, and (3) encouraging timely, appropriate emergency referral. HBLSS utilizes a community education approach that emphasizes community consensus for acceptable, feasible actions to take in response to maternal and neonatal complications. These actions essentially provide a "first responder" approach for the community which enables them to act while seeking care in facilities for maternal and neonatal problems. The skills are taught using role-play demonstration, negotiation, and practice, offering a skill set that enables and empowers community members to act in life threatening situations.

HBLSS was designed for use in low-literacy or non-literate populations utilizing teaching techniques such as role-playing, story-telling, skills checklists, and picture cards (known as Take Action Cards). These

cards show drawings of particular problems, such as bleeding during pregnancy, with pictures of emergency actions to be taken in the process of referral to a health facility beginning with the call for help and transportation.

The HBLSS approach uses adult learning principles and a step-by-step process of problem identification, problem solving, negotiation, and practice. The approach emphasizes the respectful consideration of the existing knowledge and solutions before attempting to instill biomedical practices. The HBLSS manual incorporates this step-by-step process during each community meeting. The strategy is to build on existing practices and to negotiate acceptable and feasible safe practices with the family and community members. The end goal is to build a consensus on the appropriate action and a capacity to take that action when needed (Sibley et al.2004).

The skills-building targets all members of the community and the family, including mothers-in-law, tribal leaders, husbands, and traditional healers. Community members who attend the meeting are repeatedly asked (by the trainers) educators, "Is this acceptable?" until safe actions are agreed upon. It is understood that community members are more likely to use skills or take actions that they have agreed are reasonable and acceptable.

In the Zambian context, the HBLSS curriculum was adapted to serve as the curriculum for the SMAGs. It is also adapted to fit local context and priorities. For example, Zambia has prioritized and established a national goal of facility-based births for all women. This is promoted in the HBLSS/SMAG curriculum, where arranging for transport prior to a birth or an emergency is a reoccurring key lesson for each of the community meetings. In this context, SMAGs teach the life-saving interventions as initial actions to be taken prior to and while waiting for the transport to a facility, in the event that a woman is far from a facility or experiences a complication prior to reaching the facility.

## 2.2 FORMATION AND TRAINING OF SAFE MOTHERHOOD ACTION GROUPS

This section describes the details of the SMAG program implemented with ZISSP support.

The program aimed to train two people from each zone to become SMAG members. The number of villages that make up a zone varies significantly, but a zone typically covers about 100 to 500 households. Each village has a headman and each zone may have a senior headman who represents the chief. The senior headman, with the Neighborhood Health Committee, decided which villages the SMAG members would come from. The criteria used for this selection varied across communities with some villages chosen because of low levels of maternal care seeking at facilities, and others chosen because there were capable community members to serve as SMAG members. After the villages were chosen, the village headmen, together with the Neighborhood Health Committee members, chose SMAG members based on any prior community involvement, previous training in public health issues, as well as their position and standing in the community. The qualifications of SMAG members vary within each group and across regions. In regions with a strong donor presence, such as SMGL districts for example, SMAG members have participated in numerous trainings on a variety of health topics.

The selected SMAG members underwent a one-week training based on the HBLSS curriculum. The curriculum has 12 modules that cover the primary causes of maternal and neonatal death. The topics include: maternal and neonatal mortality and morbidity, antenatal care, labor, hemorrhage, newborn resuscitation, infections, and post-partum care. Across all topics, SMAG members are trained to identify danger signs and provide first response care. A large component of the training involves building skills on

how to engage audiences at community meetings and how to counsel men and women on safe motherhood issues, including provision of life-saving care.

SMAG members were also trained to track pregnancy-related activities for pregnant women (i.e. date of first ANC visit, immunizations, place of delivery, etc.) as well as newborn milestones; they track these in two separate registers: (i) the pregnancy and delivery register, and (ii) the postnatal register (for care for the mother and the baby). Since literacy levels vary amongst SMAG members, these notebooks include pictures of each activity/milestone instead of printed words.

Some SMAGs were also trained on the USAID/MCDMCH Pregnancy Care Planner, a pictorial guide used as a teaching aid that is not part of HBLSS. This planner uses pictures to illustrate the chronology and steps a family needs to take when a woman becomes pregnant.

SMAG participants were given a pre-test prior to the training and a post-test; those who did not receive a satisfactory score (at least 50% correct) were given another opportunity to re-take the test. Those who did poorly in the post-test also received special attention and mentorship in the field.

There were no refresher trainings after the initial one-week session. SMAGs received mentorship visits by the Master Trainers and government staff, such as district MCH coordinators. These visits were meant to serve as a platform to reinforce the HBLSS lessons. However, the frequency of such support and supervision visits has varied.

## 2.3 KEY ACTIVITIES OF SAFE MOTHERHOOD ACTION GROUPS

The main responsibilities of the SMAGs are to: conduct community sessions on maternal health care seeking; refer pregnant women to health facilities for antenatal care, delivery, and postnatal care; identify maternal and newborn complications and refer such cases to health facilities; and provide first-responder life-saving care when needed. SMAG members trained with ZISSP support were also instructed to escort women in labor to health facilities, providing any first-response life-saving care en route as necessary (this is not an element of the standard HBLSS curriculum).

### 2.3.1 COMMUNITY MEETINGS

The community meetings are the centerpiece of the SMAGs program. The SMAGs were instructed to hold community meetings once a month. However, some SMAG members have reported that community members are less willing to attend meetings in the rainy season, especially if no indoor space is available. Therefore, meetings may not always be held on a monthly basis.

When possible, the SMAGs plan their meetings to coincide with other community events (for example, Mobile Health Day). Community attendance at SMAG meetings varies but a typical meeting is attended by about 50 people, of whom usually about two-thirds or more are women. This size of meetings, however, is larger than what the HBLSS training recommends (smaller groups ensure a better facilitated process of exploring, adapting, and gaining consensus of the community first aid actions).

Meetings are usually held at the village level and in a central location such as a schoolyard, and people from across the zone attend. Some village headmen have encouraged attendance and some SMAG members go door-to-door to promote their role and invite people to the meetings (however, there is no information to what extent this was happening across all SMAGs). Although SMAGs were trained to

keep records of attendance at such meetings, record keeping has generally been incomplete (as discussed below).

During the training, SMAGs were instructed to have at least two SMAG members present when conducting each community meeting. SMAG members were also given uniforms, which consisted of a SMAG branded t-shirt and a matching *chitenge* (traditional wrap). During the community meetings, SMAG members conduct short plays (dramatizations). Although the HBLSS training had a prescribed format for these meetings, SMAGs often modify it to what they considered an effective way to deliver the messages around life-saving care.

Each HBLSS-prescribed meeting has a different format, depending on the topic and goal of the meeting. Each meeting's format prescribed by HBLSS starts with a discussion about what community members know about the topic of the meeting (e.g. how a pregnant woman should take care of herself during pregnancy), then adds in items not mentioned spontaneously by participants, and proceeds towards reaching an agreement – through a negotiating question-and-answer process with the audience – about the importance of all practices discussed during the meeting. Picture cards that illustrate each practice (e.g. the woman should drink plenty of liquids) should be used in the process and given to participants to take home as reminders about what to do. The prescribed format for a typical meeting includes role-play depicting what was learned. Opportunity is subsequently presented for all participants to practice the actions, with the meeting concluding after a discussion is held regarding the feasibility and acceptability of the suggested practices illustrated in the role play.

In reality, SMAGs often amend the format prescribed by their HBLSS training. One meeting observed by the evaluation team during a field visit, for example, had the following format (which is different from the HBLSS-prescribed agenda):

> *SMAG members conducted a short skit or play: A pregnant woman or recent mother shows signs of having a health problem: post-partum bleeding. The family and neighbors around her are concerned and do not know what to do. So they call a SMAG member. The SMAG member comes, recognizes the danger signs, and performs first respondent care. The SMAG member escorts the woman to the health facility to be taken care of by professionals.*
>
> *At the end of the play a SMAG member re-iterates the danger signs and first respondent lessons shown in the plays.*
>
> *Then SMAG members held a question and answer session and initiated a discussion with the community to generate their thoughts and build consensus on practices going forward.*

The SMAGs often modify the plays to fill them with humor, songs, and dances to engage the audience. The humor, singing and dancing elements are not part of the HBLSS curriculum but are something that SMAGs have added to make the meetings more engaging. Some SMAGs have adapted the role-play by having men play the part of a pregnant woman, which was meant to sensitize the community to men being involved in what are traditionally termed "women's issues".

Building consensus on what behaviors the community will practice moving forward is perhaps the most important and distinctive feature of the SMAG training approach through HBLSS. The SMAGs are trained to discuss how the community should respond to health emergencies, being careful not to dismiss 'incorrect' answers and to guide the community in the right direction. The SMAGs are also trained to use the picture books in these meetings. However, it is not known to what extent SMAGs adhere to all the elements prescribed by the training manuals for the community meetings. Some

SMAGs visited by the evaluation team, for example, neither used the picture books nor initiated the consensus building discussion prescribed by HBLSS.

### 2.3.2 TRACKING PREGNANT WOMEN'S STEPS TO DELIVERY

One objective of SMAGs is to identify pregnant women within the community, to enable effective targeting of health promotion and prevention interventions. At community meetings and through other interactions with community members, SMAG members aim to find out who in their community is pregnant. The trainings have provided the SMAGs with a timeline of activities that pregnant women should complete including seeking ANC services, saving money for the birth, having a plan regarding transport to the health facility etc. The SMAG members are instructed to track, in their logbooks, pregnant women's progress on each of these activities and remind women and their families of what is needed.

### 2.3.3 REFERRING PREGNANT WOMEN TO THE HEALTH FACILITY

SMAGs are trained to identify signs of complications during the pregnancy and to refer pregnant women to the health facility prior to delivery in such cases. SMAGs also encourage all pregnant women to deliver at health facilities.

Another key activity by the SMAGs, whenever possible, is escorting pregnant women to deliver at health facilities. SMAG members would either walk with these women to the facility or ride the *zambulance* (bicycles with attached stretchers) when one is available; in most villages, however, there are no *zambulances*. If the woman shows any danger signs, the SMAG members are trained to provide first-respondenr care and then provide referrals; in some cases, SMAG members provide first-respondent care en route to the health facility.[4]

### 2.3.4 REPORTING TO HEALTH FACILITY

SMAG members are expected to submit monthly reporting forms to the health facility, recording attendance at community meetings (including how many men and women attended), how many meetings were held, and how many pregnant women they had referred to the health facility for maternal care services. Field observations revealed that at times these records were incomplete or erroneous. One reason for this, which was brought up by health facility workers, is the low literacy levels of some SMAG members. Records kept in the pictorial-based registers, however, were generally more complete.

## 2.4 MENTORSHIP, MONITORING, AND OTHER SUPPORT

Master Trainers, who are health professionals such as district health officers, nurses, or midwives from the facility that covers the SMAG's area, are scheduled to visit SMAGs quarterly. Some SMAGs are also visited at times by MOH provincial level officials accompanied by ZISSP field staff or by SMGL

[4] Some male SMAG members report that although they do not provide first-response care to women, they instruct female relatives how to do it, by standing right outside.

coordinators (for those in the ten SMGL program districts). Master Trainers, in collaboration with ZISSP field staff, review aggregated monthly SMAG reports to identify villages with low rates of facility deliveries or ANC attendance. During monitoring and mentoring visits, ZISSP or SMGL staff try to investigate the causes and provide targeted guidance to the SMAGs to improve the situation (e.g., provide guidance on how to counsel women; or attending the community meetings in such areas to provide on-site coaching).

The package of materials that SMAGs receive to help them in their work has varied over time and across districts. Differing combinations of the following materials have been supplied to the SMAGs: t-shirts, bags, *chitenges,* flashlights, gum boots, reflector vests, megaphones, bicycles, and rain coats or umbrellas.

# 3. EVALUATION METHODOLOGY

## 3.1 RESEARCH QUESTIONS

The evaluation aims to answer the following question: *Does the training and support of SMAGs lead to an increase in utilization of maternal health services at health facilities?*

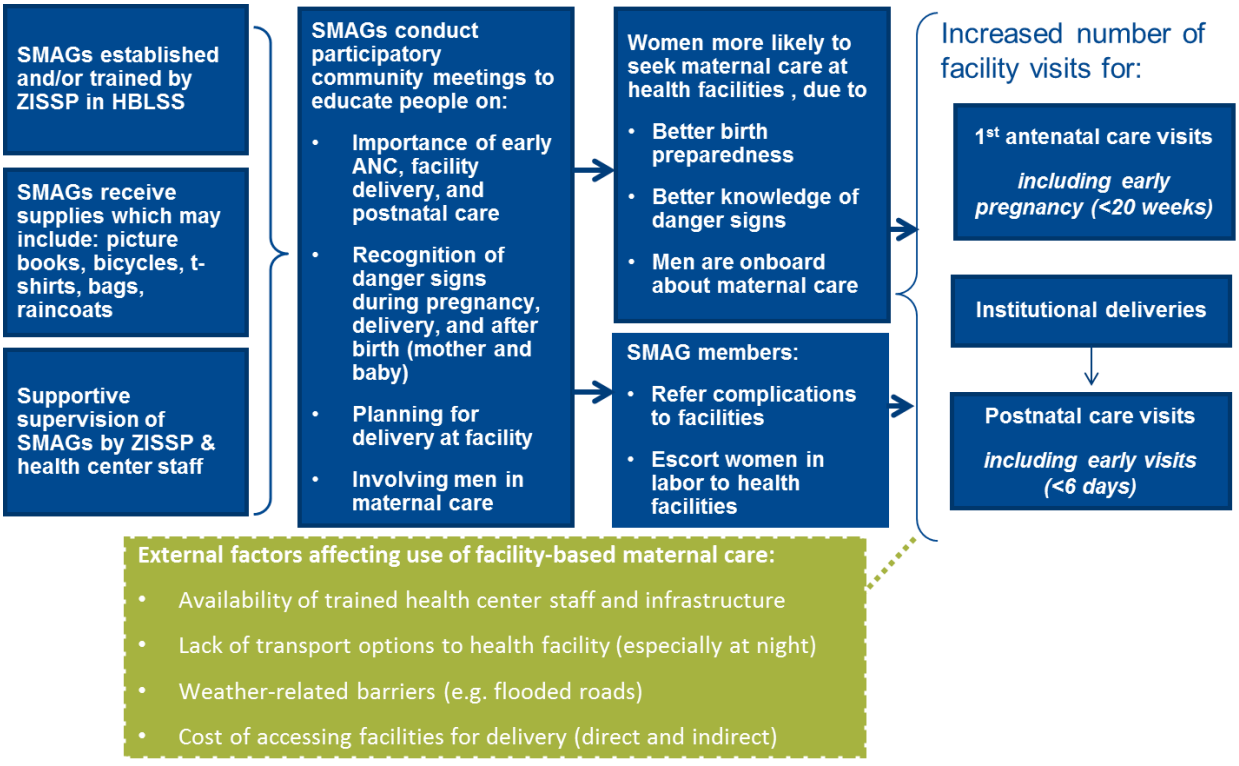
The maternal health services included in this evaluation are measured by the following five facility level outcome indicators:

1. Number of institutional deliveries
2. Number of first antenatal care visits before 20 weeks after conception
3. Number of first antenatal care visits at 20 weeks after conception or later
4. Number of postnatal care visits within six days of delivery
5. Number of postnatal care visits between six days and six weeks of delivery

The specific outcome indicators we use in this evaluation are defined by the indicators available in the national HMIS for the maternal health services promoted by the SMAG program.

Figure 1 summarizes the expected results pathway that is addressed by this evaluation (i.e., how the ZISSP-supported SMAG program is expected to lead to a change in the five outcome indicators).

**Figure 1. Expected Results Pathway**



While there are other elements of the SMAG training and other possible outcomes beyond the results depicted in this pathway, Figure 1 only focuses on the program aspects of interest that are covered by this specific evaluation.

In this report, "SMAG facilities" refers to health facilities that have in their catchment areas communities with SMAGs that were trained with ZISSP support using the HBLSS curriculum. This evaluation takes the 'intention-to-treat' perspective and counts any health facility which had any SMAGs trained with ZISSP support in its catchment area as a 'SMAG' facility. It does not track after that point what activities did or did not occur within the SMAGs associated with each facility. Rather, it assesses the average effect across all the varieties of activities that ensued after the SMAG training. Part of the rationale for community engagement activities is that they allow for local adaptation and innovation – thus, differences across communities are expected in the design of the program itself. Further study would be necessary to derive 'lessons learned' or 'best practices' about which activities have more or less success at improving knowledge or health seeking behavior in association with SMAGs. In addition, the evaluation does not capture variation across facilities in the coverage of the catchment area with the SMAGs program; the evaluation does not differentiate between facilities with a higher proportion of the catchment area supported by SMAGs relative to a lower proportion.

To measure the effect of the ZISSP-supported SMAG program, we compare before-after changes in the five outcome indicators in facilities with SMAGs supported by ZISSP in their catchment area to changes in a comparison group of facilities without such SMAGs. It should be noted that some of the comparison facilities may have had SMAGs established or trained by other organizations at some point in the past, using an approach that is different than the HBLSS model. However, there was no data on the extent to which such SMAGs may have existed or were functional during the study period in the districts not covered by ZISSP-trained SMAGs.

The SMAGs promote ANC seeking early during pregnancy, particularly in case of complications. However, since over 96% of women seek ANC at some point during pregnancy (CSO, 2014), any increase in first ANC visits earlier during pregnancy (i.e. ANC before 20 weeks after conception) means that there would be fewer women who have their first ANC visit later in pregnancy (after 20 weeks).

The rate of PNC seeking in Zambia is much lower (49%), with 39% of women accessing PNC within two days after delivery (CSO, 2007 – which is the latest data available). This means that there is scope for the SMAG program to effect an increase in PNC seeking both earlier (within six days) and later on (six days to six months) after delivery.

The following sections describe the data sources, the process used to select the comparison group (using propensity score matching), and the analytic methods used to estimate the effect of the program.

## 3.2 DATA

The study utilized several existing datasets; no primary data collection occurred for this report. To match SMAG facilities with comparable non-SMAG facilities, we used data from three sources that had information on district and facility characteristics: the 2012 MOH Health Facility Listing, the ZISSP District Selection Plan, and the national MOH Health Management Information System (HMIS).

We used data from the Health Facility Listing for the facility matching process (this process is described in detail in the following section). This listing is a census of health facilities that includes extensive facility level information including catchment population, number of beds, types of services offered, infrastructure such as operating theater, etc. We excluded hospitals, since SMAGs were associated only

with lower-level facilities. To select and match comparison group facilities to SMAG facilities, the following Health Facility Listing indicators were used:

- Number of beds
- Catchment population
- Distance to District Health Management Office (DHMO) (kilometers)
- Number of outreach sites
- Average distance to outreach sites (kilometers)
- Facility has radio capabilities
- Type of facility (health post, Rural Health Center, urban health center)
- Facility ownership (MOH or non-governmental organization [NGO]).

In the matching process, we also used data from the ZISSP District Selection Plan. This Plan includes data on each district's human resources for health availability, economic ranking, living conditions ranking, and the presence of any development partners supporting health programs as of 2011. Prior to the matching process, we eliminated any facilities that had SMAG programs that were supported by other, non-ZISSP, partners with trainings that were also based on the HBLSS curriculum, to avoid any confounding of our results. However, some facilities with SMAGs trained using non-HBLSS curricula were not excluded from the comparison group (as discussed later).

We used the national HMIS database, which is a repository of monthly data on the volume of services provided at health facilities, as the source of our five outcome indicators both prior to the SMAG intervention (for matching) and after the SMAG intervention (to assess the effects of the SMAG intervention). Table 1 provides a more detailed definition of these indicators. These variables are available on a monthly basis in the HMIS system.

**Table 1: Definitions of utilization (outcome) variables used in the analyses**

| Variable | Short form | Definition |
|---|---|---|
| Institutional deliveries | Delivery | Sum of normal, assisted and cesarean section deliveries in the facility |
| Antenatal care before 20 weeks | ANC < 20 weeks | A first visit by a pregnant woman to a health facility for the primary purpose of receiving antenatal care that occurs before 20 weeks after conception |
| Antenatal care after 20 weeks | ANC > 20 weeks | A first visit by a pregnant woman to a health facility for the primary purpose of receiving antenatal care that occurs after 20 weeks after conception |
| Postnatal care before six days | PNC < 6 days | Postnatal care within 6 days of delivery |
| Postnatal care between six days and 6 weeks | PNC > 6 days < 6 weeks | Postnatal care between 6 days and 6 weeks post-partum |

*Source:* Republic of Zambia Ministry of Health (2008)

Data from the HMIS were extracted for January 2010-December 2013. ZISSP-supported SMAG interventions were introduced on a phased basis starting from April 2012, and continued through

November 2013 (the cut-off date for this study was end of 2013). Thus, the evaluation includes facilities for which data were available for two to 18 months *after* they received the SMAG program.

## 3.3 ANALYTIC METHODS

### 3.3.1 MATCHING OF HEALTH FACILITIES

To establish a set of comparison facilities, we employed propensity score matching (Rubin 2001). Propensity score matching seeks to mimic the process by which facilities (and their associated communities) were selected to receive the SMAG intervention; that is, it seeks to use quantifiable variables to statistically explain the selection process for the facilities to receive the SMAG intervention. Since these variables are often also associated with the outcomes, finding facilities with similar measures on these variables (before the start of the intervention) as the facilities selected for SMAGs help control for confounding due to selection bias.

**Variables used to select facilities during the SMAG program**

The process of selecting communities (and therefore facilities) for the SMAG program followed a two-step process. First, ZISSP and government staff selected districts from within provinces based on four quantifiable criteria related to health services capacity and socio-economic profile:

1. District performance (combining data on ANC coverage, data reporting completeness, full immunization coverage of children under one year, and the percentage of deliveries attended by skilled personnel)

2. Human resources completeness score: The score takes data from the district plans of 2009 and measures the proportion of established staff positions that are filled, for four categories of personnel: doctors, clinical officers, nurses and midwives.

3. District Living Conditions Quartile: The MOH used the Living Conditions Monitoring Survey data to categorize districts into quartiles (A, B, C, or D) based on demographic, poverty, epidemiological profiles, and macroeconomic indicators.

4. The presence of collaborating partners which support Government of Zambia health programs.

The ZISSP program used these variables to select districts which were nominated for the SMAG program. Districts were selected to represent both districts with relatively weak and relatively strong health services and socio-economic profile in terms of the criteria listed above. The government then reviewed the nominated districts and made the final determination as to where SMAGs would be implemented. The ZISSP program nominated 26 districts; in the end SMAGs were implemented in 13 of these districts, and expanded to 4 districts not nominated originally by the ZISSP program.

Once districts were selected, community health coordinators selected communities within the selected districts for the SMAG intervention. This process was less structured than the selection of districts, and quantitative data were not purposively or uniformly used for the selection of communities for SMAGs; but the selection was intended to include consideration of the characteristics of the health facilities serving the communities. The selection of districts and communities for inclusion in the ZISSP SMAG training occurred in November 2011.

**Matching process for the evaluation**

We employ a two-step process to select comparison facilities for the evaluation: estimating the propensity score and then performing the matching based on the propensity score.

First, a logistic regression is run with 'selection into the SMAG program' as the dependent variable, and the variables used to select districts and / or facilities (described below) as the independent variables in this regression. At the district level, we use the four variables described above. For matching of individual facilities, we used the average monthly utilization of institutional deliveries, ANC, and PNC per capita from the 12-month period before the SMAG selection process took place (i.e., November 2010 to October 2011), as well as the extent of missing data on these indicators for the 12-months prior to selection (which reflects the completeness of the record-keeping[5]). In addition, we considered facility characteristics that may have influenced the selection, such as the number of beds, availability of certain types of infrastructure and services, distance to the District Health Management Office, facility type, and ownership (e.g., MOH or faith-based) (see Table 2 for a complete list). Based on interviews with program staff, we constructed this latter list of facility characteristics as variables that may have influenced community health coordinators when they selected communities for the SMAG intervention. The probability that a facility would be selected for inclusion in the SMAG program is predicted from the regression; this probability is referred to as the propensity score.

Second, facilities not in the SMAG program were matched to facilities that had the SMAG program. Matching was done by finding facilities with similar propensity scores to the SMAG facilities.

Given that, in some districts, a high proportion of facilities included at least one SMAG in their catchment area, we consider it possible that the relatively few facilities in these districts that were not SMAG facilities may be systematically different from those that were SMAG facilities, and there would be too few to form an adequate comparison. Thus, we include facilities in both non-SMAG and SMAG districts as potential comparison facilities. However, we also consider that district level indicators may also influence outcomes, and we need to balance the need to find comparable facilities with the need to find comparable district-level indicators. The two-step selection process (first districts were selected, then facilities) used for SMAG facilities implies that the logistic regression estimating the propensity scores should reflect a multi-level model. However, the methodology for propensity score matching using multilevel models is not well established (Li 2013). For this reason, we tried three different logistic regression models in order to assess which possible method works best for these analyses. First, we employed a two-step process where we ran one logistic regression at the district level and then a second one using only facility level variables, and then multiplying the two propensity scores together for use in the matching (Hong 2006). Second, we ran a multilevel logistic model with both fixed effects and random effects (Arpino 2011, Li 2013). And third, we ran a regular logistic model but included both district and facility level data, with facility level selection into SMAG as the independent variable (Arpino 2011).

[5] Note that all facilities had at least some data for the 12 month period, and thus an average could be calculated for the five indicators. Thus, no facilities were excluded do to missing data at this point, but we assess the extent of missing data as an input into the selection process. In fact, community health coordinators were instructed to consider the availability of HMIS data at facilities when they selected communities for the SMAG program; however, it is not known to what extent they considered this point when they did determine their selection.

In order to choose which of the different models works best, we assessed the average 'absolute standard bias' of the resulting matching across the variables used in the selection model. The absolute standard bias is a measure of the difference in the average value between the SMAG and comparison facilities. The absolute value of this difference is then divided by the standard deviation of the variable among the SMAG facilities in order to put all the variables into one common metric (i.e., the number of standard deviations separating the mean of the comparison and the SMAG facilities) to enable comparison across all variables. The two models that produced the lowest average absolute standard bias are included in the analyses, with the second best model used for sensitivity analysis.

The final consideration in this methodology is to determine how many comparison facilities to include in the analyses for each SMAG facility. We considered 1:1 and 2:1 matching, and assessed whether 2:1 matching inflates the absolute standard bias compared to 1:1 matching (or vice-versa). We excluded from the matching process non-SMAG facilities that were located in close proximity to SMAG facilities in order to minimize the potential for a spill-over effect from SMAG groups to neighboring facilities. Hospitals are also excluded from the analyses since SMAGs are associated with and coordinated through health centers (and in very rare cases health posts).

### 3.3.2 MEASURING THE EFFECT OF THE SMAG PROGRAM

Once we established which facilities to include in the analyses as the comparison group, we proceeded to the analyses that measure the effect of the SMAG program on the outcomes of interest. These analyses were run using separate regression models with each of the outcome indicators included as the dependent variable, respectively (see Table 1 for definitions of the indicators).

The regressions include the variables used in the propensity score selection model, as well as the outcome indicator. The five outcome indicators are available on a monthly basis from January 2010 through December 2013. We performed the regression analyses separately for each indicator.

The association between being in the SMAG program and service utilization is assessed as a difference-in-difference model (Box 1975). Figure 2 provides a general schematic that illustrates this approach.

**Figure 2: Illustrative difference-in-difference example**

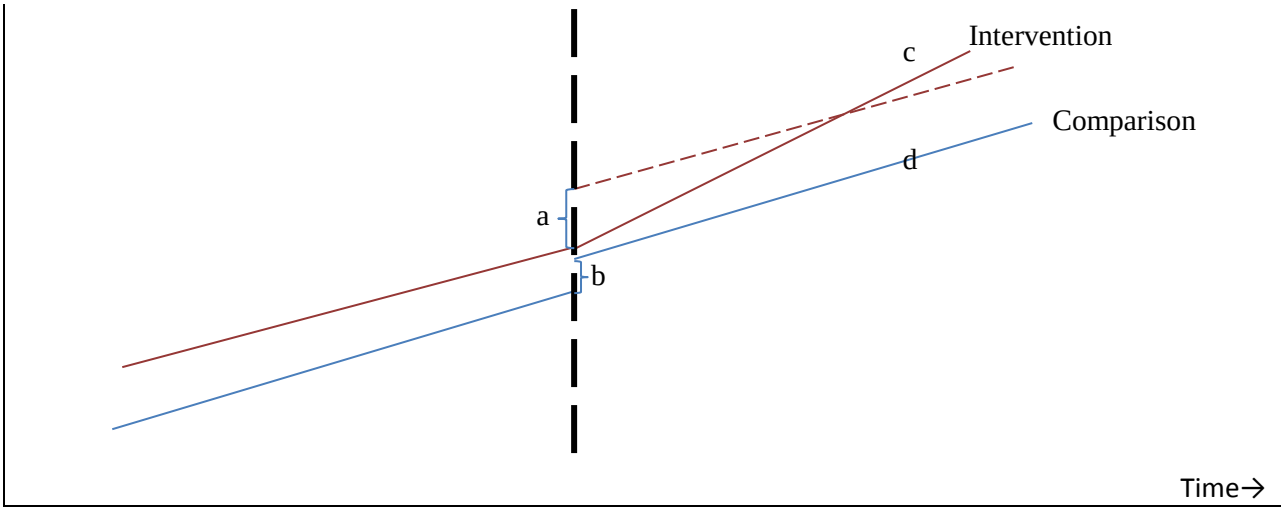


Figure 2 illustrates hypothetical changes in outcomes. The dashed vertical line represents the start of an intervention (e.g., SMAG), while the horizontal lines represent changes in utilization of health services over time (x axis). The red line represents average utilization at intervention facilities, while the blue line represents average utilization at comparison facilities. Note that simply comparing intervention and comparison facilities after the start of the program will lead to false conclusions because intervention facilities (in this example) had higher utilization before the start of the program, which needs to be accounted for in the analysis. We have matched comparison facilities to intervention facilities, such that the facilities should have similar utilization levels and trends before the start of the intervention. However, matching may not be perfect – that is, the utilization patterns before the intervention between the comparison and intervention facilities may be more similar than they would be if we used all facilities, but some difference may remain between the two groups. Thus, the difference-in-difference approach is still needed to control for this potential 'residual confounding'.

Two scenarios are presented to illustrate the effect of the program:

- **Difference-in-differences (D-in-D)**: First, we consider the dashed red line after the start of the intervention. In this example, the gap label with the letter 'a' represents the average increase in utilization after the start of the program in intervention facilities across the time measured after the start of the program, compared with before the program. The average increase measurement is based on the assumption that the trend (in this example, an upward trend) observed before the intervention would have continued after the intervention. However, in the illustrative example above, the comparison facilities also had a slight increase in use at the start of the program, represented by the letter 'b'. Thus, the association between the intervention and utilization is measured as 'a – b'; this is the 'difference in difference' indicator. This indicator can be interpreted as the difference in the utilization rate between intervention and comparison facilities, each month, on average, after the start of the program.

- **Difference-in-difference slope (D-in-S)**: The intervention may not have an immediate effect on utilization, but may have changed the trajectory of utilization over time. This is presented with the solid red line in Figure 2. In this case, the association between the intervention and utilization is assessed as the change in the slope of line 'c' after the start of the program (compared to before the program) minus the change in the slope of line 'd' after the start of the program (compared to before the start of the program); in Figure 2 there is no change in the slope of line 'd' before and after the program, which is what we would expect unless there were some other contemporaneous event occurring in the comparison area that influence the utilization trend over time. We refer to this as the 'difference in difference slope' estimate. This indicator can be interpreted as the difference in the change in the utilization rate between intervention and comparison facilities, per month, on average, after the start of the program.

The SMAG program was rolled out in a phased approach over 18 months, so not all communities formed and implemented SMAGs at the same time. Thus, we consider the 'start' of the SMAG program independently for each facility, depending on when the first community in the facility's catchment area received the SMAG training. We apply the 'start time' from a SMAG facility to the comparison facility or facilities matched to each SMAG facility.

## Regression models

The outcome variables of interest are all count variables (number of visits to health facilities). Thus, we run Poisson or negative binomial models using the catchment population of the health facility as the 'exposure'[6]. Pearson's goodness of fit test was used to determine whether to use Poisson or negative binomial models. For these models, we included a fixed effect at the facility level to control for time-invariant factors, which include facility characteristics, district characteristics, and help account for missing data. In addition, we control for the SMGL program using both a dummy variable equal to 1 when SMGL was implemented in a district (and 0 otherwise) and a time trend for the period when SMGL was implemented in a district. Thus, the results we report reflect the associations between SMAG and the utilization outcome variables independent of the SMGL program.

## Sensitivity analyses

We also ran three sets of sensitivity analyses:

- While a fixed-effect count variable regression should serve to control for the longitudinal nature of the data, there exists the possibility of residual autocorrelation, which may bias the standard errors of the regression (Bertrand 2004). Thus, we also re-ran each regression using a generalized estimation equation (GEE) of the same family [i.e., negative binomial or Poisson] as the base model (Liang 1986). Given the unknown relationship between utilization in one period and subsequent ones, we assess the correlation between variables over time to determine the best structure to use for these models (StataCorp, 2011).

- First, as described above in section 3.3.1, we re-ran the analyses using the second best matching method. HMIS data were known to contain data entry errors and other inaccuracies (Heywood et al., 2012; Ashraf et al. 2010). Further, there exists the possibility that some data that were not reported by facilities appear in the HMIS database as zeroes (rather than missing), resulting in an excess of zeroes or 'zero inflation'.

- To assess the potential influence of such data issues on the results, we re-ran the analyses excluding observations judged to be outliers or excess zeroes.

    - Outliers are initially identified as observations that are more than two standard deviations away from the mean utilization across all months included in the data set for a particular facility. Once these observations were identified, we inspected each to determine if the data point was plausible based on the other observations at the facility. For example, if the average number of visits at a facility was 50, and one month had 500 visits with no other months having over 100 visits, we would consider this to be a potential data entry error and consider it an outlier.

    - Excess zeroes were identified by summing across the five available utilization variables; if the sum was zero and none of them were recorded as missing, then we considered this a possible case of excess zeroes since it is unlikely that a particular facility had no maternal

[6] The 'exposure' variable is included in the regression analysis as an independent variable with the coefficient for the natural log of the exposure variable set to equal 1.

health utilization of any kind for an entire month. We again inspected each of these cases identified in this manner and determined if a month with no visits was plausible given the usual utilization of the facility.[7]

We use a cutoff of $p < 0.05$ to determine statistical significance. All analyses were done in Stata 12.1 MP (StataCorp, 2011).

# 4. RESULTS

## 4.1 COMPARISON BETWEEN INTERVENTION FACILITIES AND FACILITIES SELECTED FOR COMPARISON

### 4.1.1 DESCRIPTIVE STATISTICS OF FACILITY SAMPLE BEFORE THE START OF THE SAFE MOTHERHOOD ACTION GROUP INTERVENTION

Table 2 provides summary data for the facilities included in the matched sample. The data represent data at baseline, or, in the case of utilization variables, for the period before the start of the SMAG program for which we collected data. Baseline values are defined as the average of all data before the start of the SMAG program; the start of the SMAG program is defined on a facility by facility basis depending on when the SMAG training occurred for the groups in the facility catchment area.

The comparison facilities were selected using a single logistic regression on both district and facility level data, with one-to-one matching. This method performed better than other matching approaches; it reduced the average absolute standard bias from 0.42 standard deviations across variables between all possible comparison facilities and SMAG facilities to 0.09 standard deviations.[8] There were 138 facilities with SMAGs included in the time period of the analysis, and 1,208 possible matches (but 138 included in the analyses). See Annex A for more details on the matching.

Utilization in Table 2 is reported in rates; for example, SMAG facilities had an average of 2.64 visits per month for PNC in the first six days after birth per 1,000 people in their catchment population before the start of the SMAG program, while matched comparison facilities had 2.73.

[7] Note that fixed effect zero-inflated models are not currently available in Stata 12.1 and including a dummy variable for each facility caused the estimation equations to exceed computational memory.

[8] The method with the second lowest absolute standard bias was the 'two-step' method, where separate regressions were run at the facility and the district levels, with the results were multiplied together for one-to-one matching. This reduced the absolute standard bias to 0.11 after matching, and is used in sensitivity analysis.

At both the facility and the district level, there were no statistically significant differences between the SMAG facilities and the matched comparison facilities across the variables used in the matching. Implementation of SMGL could not be included in the matching analyses because, in the first round of implementation of SMGL, all district with SMGL also had SMAG facilities. With 54% of SMAG facilities in SMGL districts in 2012, and a further 21% in districts that started SMGL in 2013, we could not exclude SMGL districts from the analysis without reducing the sample size to about 34 SMAG facilities. Thus, there was a statistically significant difference between SMAG and comparison facilities in their likelihood of being located in SMGL districts.

**Table 2: Summary statistics**

| Variable | N | SMAG Facilities (138 Facilities) | | | Comparison Facilities (138 Facilities) | | | p-value |
|---|---|---|---|---|---|---|---|---|
| | | Mean | SE | | Mean | SE | | |
| *Utilization rates per month (per 1,000 people)* | | | | | | | | |
| Institutional delivery | 7,501 | 1.43 | (0.08) | | 1.61 | (0.1) | | 0.16 |
| First ANC visit before 20 weeks | 8,816 | 1.30 | (0.09) | | 1.33 | (0.07) | | 0.74 |
| First ANC visit after 20 weeks | 8,644 | 1.39 | (0.11) | | 1.32 | (0.08) | | 0.61 |
| Postnatal visit less than 6 days after birth | 8,998 | 1.39 | (0.08) | | 1.35 | (0.07) | | 0.74 |
| Postnatal visit 6 days to 6 weeks | 9,086 | 2.64 | (0.1) | | 2.73 | (0.11) | | 0.54 |
| | | | | | | | | |
| Missing data (percentage missing across 6 indicators above) | 275 | 0.09 | (0.02) | | 0.10 | (0.01) | | 0.70 |
| | | | | | | | | |
| *Facility characteristics* | | | | | | | | |
| Number of beds | 275 | 9.9 | (1.14) | | 10.5 | (0.89) | | 0.61 |
| Catchment population | 274 | 8,585 | (436.5) | | 9,167 | (543.) | | 0.38 |
| Distance to DHMO (kms) | 266 | 59 | (5.1) | | 59 | (4.9) | | 0.95 |
| Participation in SMAGs trained with HBLSS curricula by other partners (proportion) | 275 | 0.04 | (0.02) | | 0.02 | (0.01) | | 0.47 |
| Number of outreach sites | 275 | 6.67 | (0.39) | | 7.41 | (0.48) | | 0.18 |
| Average distance to outreach sites (kms) | 258 | 34.4 | (7.84) | | 21.9 | (2.07) | | 0.13 |
| Proportion of facilities operated by NGO | 275 | 0.04 | (0.02) | | 0.03 | (0.01) | | 0.54 |
| | | | | | | | | |
| *Type of facility* | 275 | | | | | | | 0.78 |
| Health post | | 7% | | | 8% | | | |
| Rural Health Center | | 87% | | | 84% | | | |
| Urban health center | | 7% | | | 8% | | | |
| | | | | | | | | |
| *Proportion of facilities with / offering:* | | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Skilled delivery | 275 | 0.89 | (0.04) | | 0.88 | (0.03) | | 0.76 |
| Reported capacity for EmONC at baseline | 275 | 0.65 | (0.09) | | 0.53 | (0.07) | | 0.24 |
| Laboratory | 275 | 0.36 | (0.12) | | 0.31 | (0.05) | | 0.64 |
| Operating theater | 275 | 0.01 | (0.01) | | 0.01 | (0.01) | | 0.99 |
| Radio | 275 | 0.44 | (0.13) | | 0.55 | (0.08) | | 0.40 |
| | | | | | | | | |
| *District characteristics* | | | | | | | | |
| Zambia ranking of district health performance | 275 | 38.57 | (3.41) | | 38.07 | (3.43) | | 0.91 |
| Number of donor-supported health projects | 275 | 4.62 | (0.49) | | 4.45 | (0.29) | | 0.71 |
| Participation in JSI drug pilot | 275 | 0.08 | (0.05) | | 0.12 | (0.04) | | 0.44 |
| | | | | | | | | |
| *Human resources for health ranking* | 275 | | | | | | | 0.49 |
| Lowest quartile | | 20% | | | 16% | | | |
| Second quartile | | 47% | | | 50% | | | |
| Third quartile | | 30% | | | 27% | | | |
| Highest quartile | | 4% | | | 7% | | | |
| | | | | | | | | |
| *District category based on living standards survey* | 275 | | | | | | | 0.14 |
| Lowest quartile | | 3% | | | 5% | | | |
| Second quartile | | 20% | | | 13% | | | |
| Third quartile | | 62% | | | 57% | | | |
| Highest quartile | | 16% | | | 25% | | | |
| | | | | | | | | |
| *World Bank Results-Based Finance project* | 275 | | | | | | | 0.94 |
| Comparison district | | 17% | | | 17% | | | |
| Received intervention | | 19% | | | 18% | | | |
| Not part of intervention | | 64% | | | 65% | | | |
| *SMGL participation* | | | | | | | | |
| District selected for SMGL (2011) | 275 | 0.54 | (0.14) | | 0.11 | (0.06) | | <0.001** |
| District selected for SMGL (2012) | 275 | 0.21 | (0.09) | | 0.06 | (0.03) | | 0.06 |

Results for two-tailed t-test unless percentages reported; these results for chi-squared test.
*Statistically significant at $p<0.05$
** Statistically significant at $p<0.01$
DHMO: District Health Management Office; EmONC: Emergency Obstetric and Neonatal Care; JSI: John Snow International; kms: Kilometers; SE: Standard Error. Figures may not add up to 100% due to rounding.

A small proportion (2%) of ZISSP SMAG facilities also had other implementing partners' SMAG programs that were based on the HBLSS curriculum in their catchment area. Thus, a small percentage of the comparison facilities also had non-ZISSP SMAGs in their catchment area in order to balance these possible effects. In addition, SMAGs trained with curricula other than HBLSS may also be present in study areas; however, we have no detailed information on which facilities may have had such SMAG programs at baseline.

### 4.1.2 TIME TRENDS IN UTILIZATION RATES

Figure 3 presents the average utilization rates for SMAG facilities and comparison facilities for every month included in the analyses. The gray vertical line in Figure 3 marks the start of the SMAG program. Data for SMAG facilities are shown for before and after the start of the program. Different SMAG facilities started the SMAG program at different times; implementation was rolled out over a 22-month period. Month 0 is relative to each SMAG facility, depending on the program's implementation in that facility's catchment area. In the 'after' portion of the graphs, the data presented only reflects facilities that had already started the SMAG program in the month depicted.

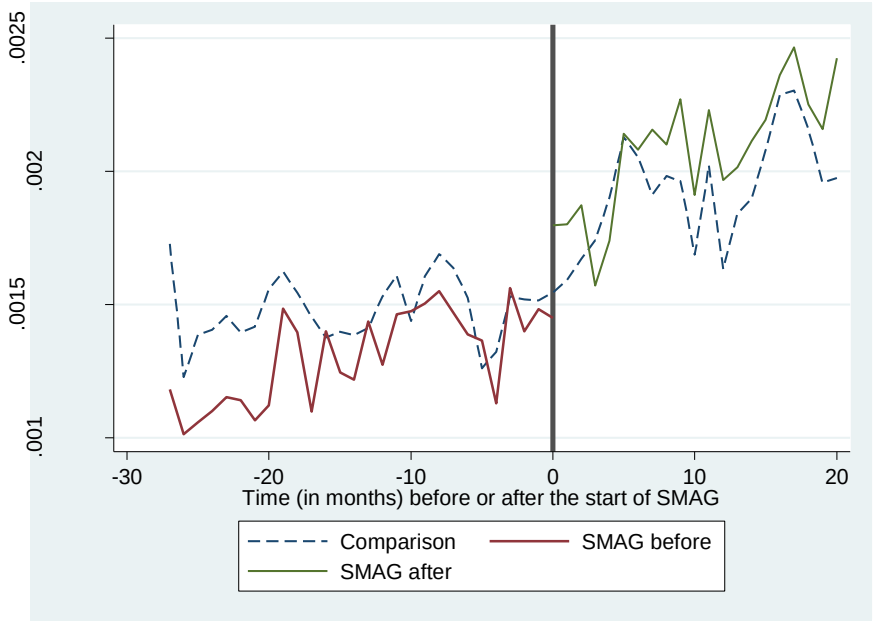
While the trends in Figure 3 show average utilization, they should be interpreted with a few caveats in mind. First, they do not account for missing data. Trends may change as relatively highly (or lowly) utilized facilities fail to report data for a given month; these 'entry and exit' of data may shift trends simply due to the availability of data and may not reflect true secular trends. Second, the graphs do not account for the fact that some facilities which did not record data may have been entered as a '0' instead of as missing data. This will artificially lower the average utilization rate when this happens. Based on interviews with staff in Zambia, efforts were made to eliminate this practice during the latter part of the period covered by the data. Thus, we view this as more likely in earlier time periods as compared with later time periods. This 'zero inflation' will bias results if it more common in SMAG facilities than in comparison facilities (or vice-versa). This may be possible since communities may have been, in part, selected into the SMAG program based on data reporting practices (i.e. that SMAG facilities had better reporting practices) and because the relative frequency of this practice was reported to be different in different provinces. Thirdly, the average could be influenced by data entry errors resulting in outliers. Lastly, the graphs represent the raw data which are not regression-adjusted and there may be differences remaining between the SMAG and comparison facilities, despite the matching exercise.

With these caveats in mind, the results show a marked upward trend over time for *institutional deliveries*, but especially for the time period after the introduction of SMAG. However, the rate appears roughly similar in SMAG facilities and comparison facilities. If the matching was successful, we would expect the SMAG facilities and comparison facilities to have similar rates before the start of the SMAG program in these graphs; if the intervention was successful, we would expect SMAG facilities to have either more deliveries or a sharper upward trend after the start of the SMAG program. The presence of SMGL in many SMAG districts, for example, means that other explanations for any changes observed are possible; these other explanations are controlled for in the regression analyses.

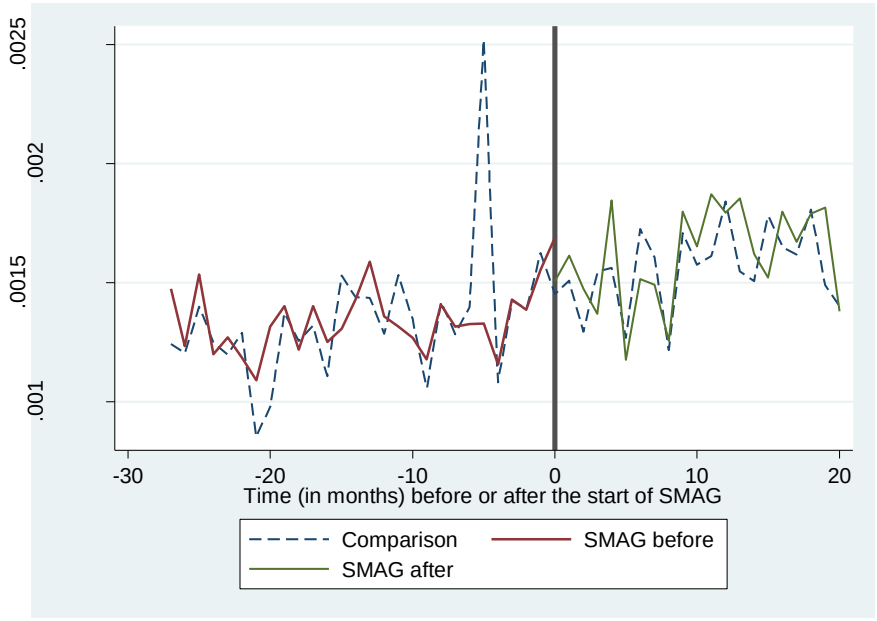
*First ANC visits before 20 weeks* show a slight upward trend over time, while the rate of *first ANC visits after 20 weeks* fell over the observation period. Given that over 96% of women seek ANC at some point during pregnancy (CSO, 2014) these results are consistent: as more women have their first ANC visit early into their pregnancy, there would be fewer women who have their first ANC visit later in pregnancy. There is little apparent difference between comparison facilities and SMAG facilities either before or after the start of the program for these two variables.

The utilization of PNC, both early and late, appeared to rise over the observation period, although perhaps more markedly so for PNC before 6 days. For both types of PNC visits, there appears to be little difference between comparison and SMAG facilities before the start of the intervention, but after the start of the intervention, SMAG facilities in general appear to have higher utilization rate.

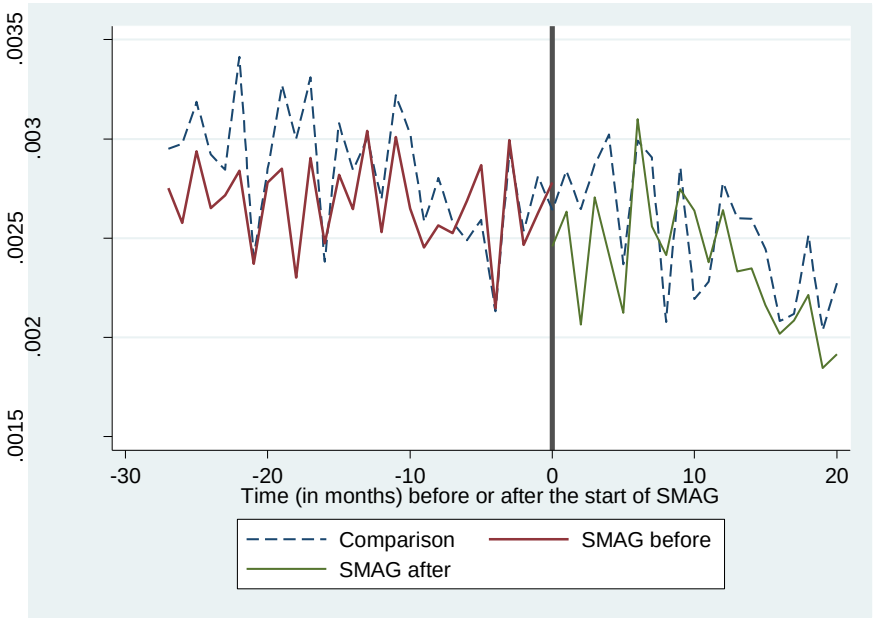
**Figure 3: Average utilization rates for the five outcome indicators**



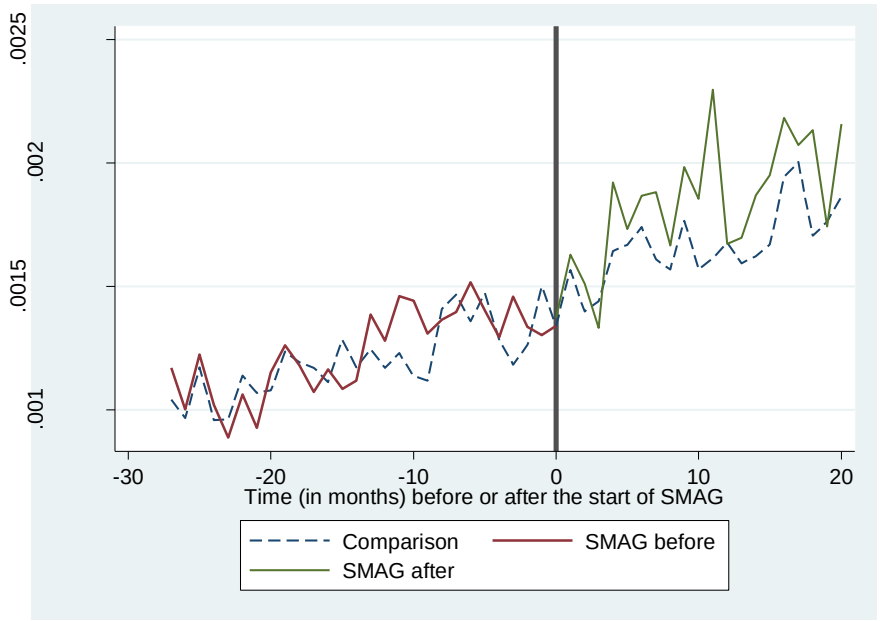

**Institutional deliveries**



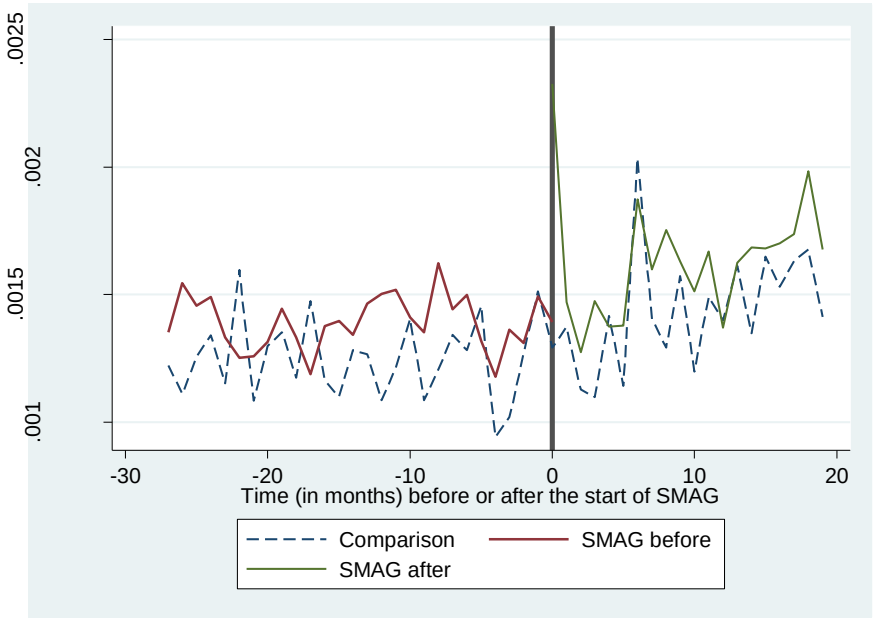

**1st ANC visit before 20 weeks**




**1st ANC visit after 20 weeks**




**PNC before 6 days**




**PNC 6 days to 6 weeks**

## 4.2 RESULTS OF THE REGRESSION MODELS

Statistical models assessing the association between the SMAG program and the five utilization indicators are presented in Table 3. The main coefficients of interest, showing the association between the SMAG program and the utilization variable, are presented using two different models for each utilization variable. The two models are the difference-in-difference model and the difference-in-slope model, which were described in the Methods section.

The coefficients reflect (1) D-in-D: the basic difference-in-difference estimation (assessing if there is an overall change in the utilization associated with the SMAG program) and (2) D-in-S: the difference-in-slope estimation showing whether the slope of the time trend in utilization changes after the implementation of the SMAG program when assessed against the time trend of comparison facilities, as well as whether there is an overall change in the utilization associated with the SMAG program.

**Table 3: Results of the regression models assessing the association between the SMAG program and utilization**

| Outcome variable / model | | n (facilities) | n (facility-months of data) | Difference in difference | | Difference in slope | | Number of visits per facility-month associated with SMAGs (at average facility) |
|---|---|---|---|---|---|---|---|---|
| | | | | IRR | p-value | IRR | p-value | |
| **Institutional Deliveries** | | | | | | | | |
| | D-in-D | 268 | 10,409 | 1.13 | <0.001** | | | 2.3 |
| | D-in-S | | | 1.07 | 0.02* | 1.005 | 0.03* | 2.0 |
| **ANC before 20 weeks** | | | | | | | | |
| | D-in-D | 274 | 12,387 | 1.03 | 0.41 | | | N/A |
| | D-in-S | | | 0.94 | 0.05 | 0.980 | <0.001** | -2.7 |
| **ANC after 20 weeks** | | | | | | | | |
| | D-in-D | 273 | 12,471 | 0.95 | 0.08 | | | N/A |
| | D-in-S | | | 0.95 | 0.11 | 0.998 | 0.57 | N/A |
| **PNC before 6 days** | | | | | | | | |
| | D-in-D | 273 | 12,140 | 1.14 | <0.001** | | | 2.0 |
| | D-in-S | | | 1.07 | 0.03* | 0.993 | 0.01* | 0.2 |
| **PNC 6 days to 6 months** | | | | | | | | |
| | D-in-D | 271 | 11,661 | 1.13 | 0.001** | | | 2.1 |
| | D-in-S | | | 1.05 | 0.21 | 0.995 | 0.19 | N/A |

Total possible observations are 274 facilities x 48 months = 13,152. All models include fixed effect at the facility level, quadratic secular trend, variables for presence of SMGL at the district level and time trend for SMGL, and for the catchment population of the facility.

*Statistically significant at p<0.05

** Statistically significant at p<0.01

Both estimations use a negative binomial model. In these models, the coefficient is reported as the incidence rate ratio (IRR); in this case, a coefficient greater than one (>1) reflects a positive association between the SMAG program and utilization of services (a higher incidence of people visiting facilities, measured relative to the facility's catchment population size), while a result between 0 and 1 reflects a negative association (a lower incidence of people visiting facilities). The last column in Table 3 translates the incidence rate ratios into the corresponding number of visits per facility per month (at the average facility) associated with SMAG for coefficients that are found to be statistically significant. D-in-S models are only considered to attain statistical significance if the linear combination of the two coefficients of interest in the model achieves statistical significance ($p<0.05$).

The results for the outcomes of interest are as follows:

- **Institutional deliveries**: both the D-in-D and the D-in-S models show a positive association between the SMAG program and utilization. The association is on the order of 2 or more deliveries per month at SMAG facilities compared with non-SMAG facilities, which is an increase from about 17.8 per facility per month without the SMAG program to around 20 after, representing an increase of about 12%.

- **ANC before 20 weeks**: the results of the D-in-D model do not show a statistically significant association between SMAGs and the number of first visits for ANC before 20 weeks. The D-in-S model suggests an association between SMAG and a decrease in utilization of ANC first visits before 20 weeks, on the order of 2.7 visits per month on average.

- **ANC after 20 weeks**: the results do not show a statistically significant association between SMAGs and the number of first visits for ANC after 20 weeks.

- **PNC before 6 days**: both the D-in-D and the D-in-S models show a positive association between the SMAG program and utilization. The association is between 0.2 and 2 visits per facility per month (although the D-in-S model results suggest a larger association immediately after SMAGs started, which wanes over time).

- **PNC between six days and six weeks**: The D-in-D model suggests an association between SMAG and an increase in utilization of PNC 6 days and 6 weeks of about 2.1 visits per month on average. The results of the D-in-S model do not show a statistically significant association between SMAGs and the number of PNC visits between 6 days and 6 weeks.

Table 4 summarizes these results. Overall, for institutional deliveries and both types of PNC, these models show that the SMAG program was associated with increased utilization of the measured utilization indicators; these results are statistically significant in all but one model (the PNC between 6 days and 6 weeks D-in-S model). The results for ANC are less clear, with either no statistical significance, or, in the case of ANC before 20 weeks, a negative association in the D-in-S model.

### Sensitivity analyses

Table 4 also presents a summary of the sensitivity analyses around these main results; the full analyses are provided in Annex B. Three types of sensitivity analyses were done. First, we re-ran the models presented in Table 3 but eliminating suspected outliers and observations which were recorded as zeroes but that we suspected were actually missing (Annex B also includes data on the extent of outliers and

missing data). Second, we re-ran the models using a different comparison group, derived from a different matching method (as described in the Methods section). Third, we re-ran the models using GEE.

**Table 4: Summary of findings**

| Utilization indicator | Strength of evidence from main analyses | Results of sensitivity analyses |
|---|---|---|
| Institutional deliveries | Association of the SMAG program was **positive** and statistically significant in both models. | Association of the SMAG program was **positive** in all models, statistically significant in all but one model. **Strong evidence** for an association between SMAGs and utilization. |
| ANC < 20 weeks | Association of the SMAG program was **positive** but not statistically significant in the D-in-D model, and **negative** and statistically significant in the D-in-S model. | D-in-D and D-in-S models consistently show opposite associations; D-in-S achieves statistical significance in two models. **Inconclusive evidence**. |
| ANC > 20 weeks | Association of the SMAG program was negative but not statistically significant in both models. | Results consistently negative but not statistically significant. **Inconclusive evidence**. |
| PNC < 6 days | Association of the SMAG program was **positive** and statistically significant in both models. | Results are sensitive to model specification. D-in-D results always positive and statistically significant; one D-in-S model had negative association. **Limited evidence** for a positive association between SMAGs and utilization. |
| PNC > 6 days and < 6 weeks | Association of the SMAG program was **positive** in both models and statistically significant in D-in-D model. | Positive association in all but the GEE models; 3 D-in-D models statistically significant and positive; D-in-S models not statistically significant. **Limited evidence** for a positive association between SMAGs and utilization. |

The results for institutional deliveries appear to be robust to the sensitivity analyses, consistently showing a positive association of the same general magnitude across all models, and achieving statistical significance in 7 out of the 8 models assessed

For the two PNC utilization variables, the results are less robust to model specification than for institutional deliveries, with some of the models showing negative, but not statistically significant, associations. However, for PNC before 6 days, 7 of the 8 models have a positive association and are statistically significant. However, since the GEE D-in-S model has a negative association, albeit not statistically significant, we conclude there is only limited evidence for a positive association between SMAG and this variable. PNC between 6 days and 6 weeks shows positive associations in all but the GEE models, but is only statistically significant in 3 models. Thus, there is limited for an association between SMAGs and this variable.

For ANC before 20 weeks, all the D-in-D models show a positive association (none statistically significant), while all the D-in-S models show a negative association (with 3 out of the 4 attaining

statistical significance). Thus, again, there is insufficient evidence to draw any conclusions about the association between SMAG and this variable.

For ANC after 20 weeks, all models indicate a negative association but none achieve statistical significance. Thus, the evidence is not sufficient to reach conclusions about the association between SMAG and this variable.

### 4.2.1 LIMITATIONS

There are multiple limitations to the analyses and the interpretation of the results. Below, we highlight some of the more prominent limitations.

While we have attempted to control for potential confounding due to selection bias by using propensity score matching and subsequent regression models that control for secular trends, our quantitative analysis may not have adequately captured the overall programmatic processes of selecting districts and communities for SMAG interventions. While the initial ZISSP district nomination processes followed quantitative selection processes, subsequent government approval and then selection of communities followed a less transparent process. Thus, our matching may have missed important variables used in the selection processes. For example, SMAGs were introduced into the SMGL districts specifically to support/reinforce the SMGL program, but we have no information on how SMGL districts were selected. Thus, if SMGL districts were selected for reasons that favor the success of maternal health programs, and these reasons influence SMAGs' success as well, we may not have adequately controlled for selection bias. This would bias results in favor of the SMAG program. Further, SMGL districts may have improved maternal health indicators due to SMGL and not SMAGs. While we controlled for SMGL in the regressions, the temporal overlap of the programs was such that we may still have captured some effects of SMGL in the SMAGs analysis. This again would bias the results in favor of the SMAG program if the SMGL program had positive impacts on maternal health utilization. On the other hand, SMAGs implemented in SMGL districts may have received less attention from program implementers (since there were multiple interventions occurring simultaneously), and the SMAGs in these districts may not have performed as well as other SMAGs. This would bias the results against the SMAG program.

Since we controlled for SMGL in the regression models, the results should be interpreted as the effect of SMAGs independent of SMGL. That is, the results represent the average effect of SMAGs in the areas surveyed, additional to whatever effects other SMGL interventions may have. This analysis was not intended to be an evaluation of the SMGL program, and we did not have adequate information to fully assess the plausible effects of the SMGL program. Lacking that data, we did not try to assess these interactive effects between SMGL and SMAGs.

Some of the comparison facilities may have had SMAGs established or trained by other organizations at some point in the past, using an approach that is different than the participatory HBLSS model used by ZISSP. If such SMAGs were functional and effective in increasing maternal care utilization during the study period, the results from our study about the impact of the ZISSP-supported SMAGs would be underestimated as compared to having no SMAG at all.

Finally, these analyses relied on HMIS data, which are suspected to be of low quality. While we have tried to assess the validity of the data and the resulting conclusions through a variety of sensitivity analyses, there were many missing data from the facilities. For example, 44% of facilities reported fewer than 40 months (83%) of utilization data for institutional deliveries. Thus, our models may not have been able to fully or successfully capture secular trends due to lack of full data availability.

# 5. DISCUSSION

From the results of this study, we broadly conclude that there is:

- Evidence for a positive association between the SMAG program and an increased utilization of institutional deliveries from about 17.8 per facility per month without the SMAG program to around 20 with the program, representing an increase of about 12%;

- Limited evidence for a positive association between the SMAG program and an increased utilization of PNC;

- Inconclusive evidence of an association between the SMAG program and utilization of ANC.

These results are consistent with the focus of the SMAG program – the contents of the community meetings held by the SMAGs was heavily weighted on delivery at a health facility, and this is the indicator for which we found the greatest evidence of effect from the program. The SMAG program comprises several direct and indirect interventions that encourage institutional deliveries. First, many SMAG members are former TBAs who no longer practice, thus decreasing or eliminating the option to deliver at home and providing credibility to the health facilities from a trusted community source. Second, SMAGs spread awareness about the life-saving benefits of institutional deliveries through community meetings, as well as instruct families on birth preparedness planning that involves men. Third, when SMAG members choose to escort women in labor to health facilities, the women may be more likely to agree to travel to the facility knowing that they are accompanied by someone who can provide first response care along the way if needed.

Several factors that are internal and external to the SMAG program may explain the lack of a larger effect size on uptake of institutional deliveries. First, we cannot determine the intensity of the program across SMAGs. That is, we do not know what proportion of SMAGs hold monthly meetings, how many adhere to the HBLSS curriculum, to what extent each SMAG is effective at engaging audiences, etc. Second, SMAGs and beneficiaries have both reported that in some areas health facilities do not have skilled staff on hand which discourages women from delivering in the facility. Cost of transport to health facilities, remoteness, and unavailability of transport options at night may also pose a barrier to accessing facilities in some areas. These barriers may be too difficult to overcome despite the SMAGs' educational activities around birth preparedness.

It is likely that the increase in institutional deliveries and, to some extent, PNC may be attributed mostly to a higher number of women who seek such care in case of serious complications. In this case, there would still be scope for the SMAG program to effect an increase in the number of women going to health facilities for normal deliveries as well.

# 6. CONCLUSIONS

This evaluation aimed to contribute to the evidence on the effect that training SMAGs to educate communities on safe motherhood practices and to promote deliveries in health facilities has on the utilization of facility-based maternal care. The positive impact of SMAGs on institutional deliveries is a notable and encouraging result, as it is during delivery that a large proportion of life-threatening complications occur.

An assessment of specific implementation practices and lessons of the HBLSS-based SMAGs training was not part of this evaluation but would be a valuable study. It can shed light on changes in SMAG program implementation that could lead to even greater impact on utilization of facility-based care for mothers and newborns.

# ANNEX A: RESULTS FROM MATCHING ANALYSES

Table A1 shows the average standard bias for selected variables for all available comparison facilities and SMAG facilities, as well as for the facilities included after matching (for both methods presented in the analyses). Both matching methods reduce bias by over 0.3 standard deviations when considering all facilities as comparators. The single logistic regression method reduces the bias on 23 out of the 26 variables considered for the matching. The largest increase at the facility level was for the number of outreach sites.

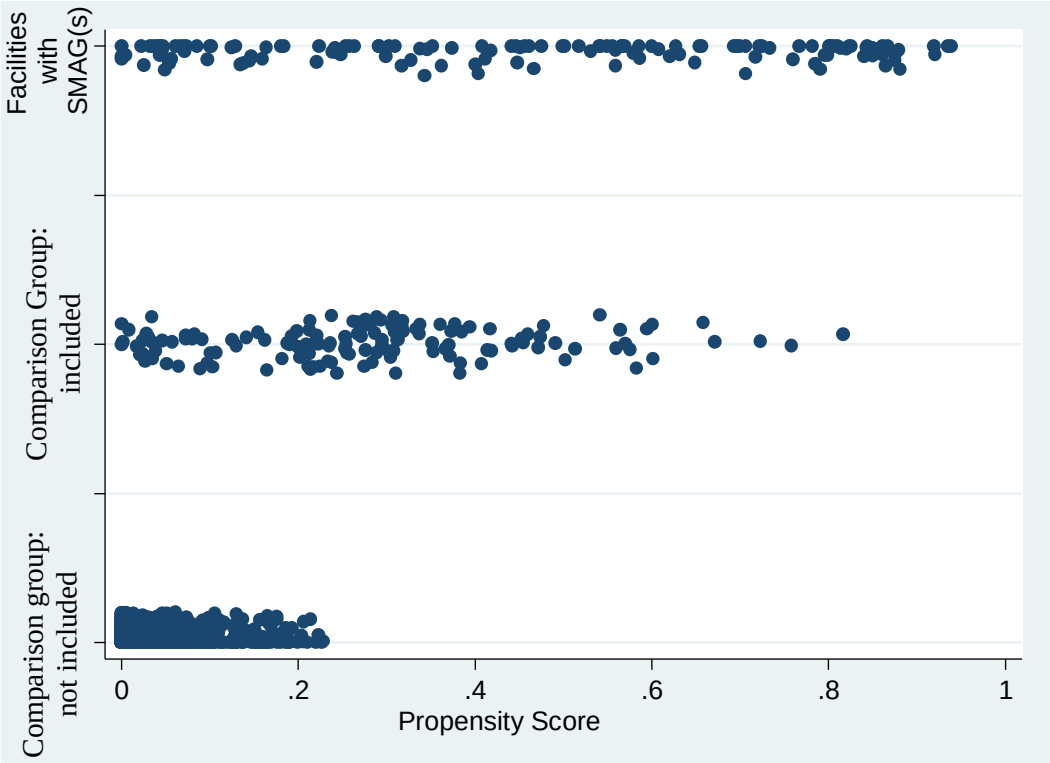
**Table A1: Absolute standards bias for selected indicators and matching methods**

| Variable | All available comparison facilities | Single logistic regression; 1-to-1 matching without replacement | 2-step logistic regression; 1-to-1 matching without replacement |
|---|---|---|---|
| *Utilization: visits per capita per month in the year before selection* | | | |
| Institutional deliveries | 1.66 | 0.12 | 0.10 |
| ANC < 20 weeks | 0.94 | 0.04 | 0.03 |
| ANC > 20 weeks | 0.34 | 0.08 | 0.16 |
| Post natal < 6 days | 0.78 | 0.03 | 0.07 |
| Post natal 6 days to 6 weeks | 1.39 | 0.04 | 0.05 |
| | | | |
| Percentage of utilization data missing | 0.36 | 0.07 | 0.14 |
| *Facility characteristics* | | | |
| Number of beds at facility | 0.20 | 0.04 | 0.07 |
| Number of outreach sites | 0.12 | **0.25** | **0.16** |
| Average distance to outreach sites | 0.19 | 0.13 | **0.22** |
| Catchment population | 0.19 | 0.13 | **0.22** |
| Distance to DHMO | 0.03 | 0.01 | **0.17** |
| Has radio | 0.19 | 0.13 | **0.22** |
| Has telephone | 0.01 | **0.02** | **0.08** |
| Urban | 0.35 | 0.04 | 0.14 |
| NGO ownership | 0.00 | **0.07** | **0.04** |

| | | | |
|---|---|---|---|
| Facility offers delivery (reported at baseline) | 0.33 | 0.05 | 0.05 |
| Facility offers EmONC (reported at baseline) | 0.96 | 0.27 | 0.30 |
| Facility offers PMTCT | 1.02 | 0.12 | 0.07 |
| Facility has a laboratory | 0.28 | 0.11 | 0.02 |
| Facility has a theater | 0.02 | 0.00 | 0.00 |
| | | | |
| Zambia government rank on health performance | 0.17 | 0.02 | 0.07 |
| Number of partners working on health in district | 0.41 | 0.10 | 0.03 |
| Participation in World Bank Results-Based Financing pilot | 0.09 | 0.02 | **0.13** |
| Human resources for health ranking | 0.14 | 0.10 | 0.10 |
| Participation in JSI drug pilot | 0.56 | 0.16 | 0.03 |
| Zambia government rank on living standards | 0.25 | 0.15 | 0.24 |
| Overall average | 0.42 | 0.09 | 0.11 |

Figure A1 shows the propensity scores for three classes of facilities. First, it shows the propensity scores estimated for the SMAG facilities along the top of the graph, which range from close to zero to over 0.90. The group below includes the facilities that were selected to be included in the analyses – the 'matched' comparison facilities. They have a similar range of propensity scores as the SMAG facilities, although their maximum propensity score is only about 0.85. The bottom group in the figure includes the potential comparison facilities that were not 'matched' – that is, they are not included in the analyses. These facilities all had low propensity scores (below 0.25). Intuitively, we expect SMAG facilities in general to have higher propensity scores since they were, in fact selected, and non-SMAG facilities to have lower propensity scores since they were not, in fact, selected. Figure A1 suggests that the matching process achieved good 'balance' between the SMAG and the matched comparison facilities, since almost all of the propensity scores for the SMAG facilities are represented in the matched comparison facilities – that is, selection bias based on the measured variables has been controlled for.

**Figure A1: Propensity scores for SMAG facilities, matched facilities, and facilities excluded from the analyses**



# ANNEX B: RESULTS FROM SENSITIVITY ANALYSES

**Table B1: Results of the regression models assessing the association between the SMAG program and utilization: Sensitivity analyses**

| Outcome / model | n (facilities) | n (facility-months of data) | Difference in difference | | Difference in difference: Slope | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Coefficient | p-value | Coefficient | p-value | Slope Coefficient | p-value |
| **Institutional Deliveries** | | | | | | | | |
| *Baseline model* | *268* | *10,409* | *1.13* | *<0.001*** | *1.05* | *0.04** | *1.01* | *0.01** |
| Generalized estimation | 273 | 10,541 | 1.12 | 0.04* | 1.03 | 0.63 | 1.003 | 0.52 |
| Outliers removed | 262 | 9,908 | 1.13 | <0.001** | 1.05 | 0.04* | 1.01 | 0.01* |
| Alternative matching | 253 | 10,070 | 1.16 | <0.001** | 1.10 | 0.001** | 0.998 | 0.41 |
| | | | | | | | | |
| **ANC after 20 weeks** | | | | | | | | |
| *Baseline model* | *273* | *12,471* | *0.95* | *0.08* | *0.95* | *0.09* | *1.00* | *0.54* |
| Generalized estimation | 273 | 12,471 | 0.96 | 0.37 | 0.95 | 0.33 | 0.99 | 0.05 |
| Outliers removed | 267 | 11,949 | 0.96 | 0.13 | 0.95 | 0.09 | 0.998 | 0.54 |
| Alternative matching | 254 | 11,877 | 0.98 | 0.39 | 0.99 | 0.85 | 0.996 | 0.12 |
| | | | | | | | | |
| **ANC before 20 weeks** | | | | | | | | |
| *Baseline model* | *274* | *12,387* | *1.03* | *0.41* | *0.93* | *0.03** | *0.98* | *<0.001*** |
| Generalized estimation | 273 | 12,387 | 1.01 | 0.82 | 0.96 | 0.41 | 0.98 | 0.001** |
| Outliers removed | 268 | 11,865 | 1.02 | 0.51 | 0.93 | 0.03* | 0.98 | 0** |
| Alternative matching | 254 | 11,782 | 1.06 | 0.05* | 0.96 | 0.29 | 1.00 | 0.54 |
| | | | | | | | | |
| **PNC before 6 days** | | | | | | | | |
| *Baseline model* | *273* | *12,140* | *1.14* | *<0.001*** | *1.08* | *0.02** | *0.99* | *0.02** |
| Generalized | 273 | 10,059 | 1.10 | 0.05* | 0.96 | 0.42 | 0.99 | 0.05* |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| estimation | | | | | | | | |
| Outliers removed | 267 | 11,618 | 1.15 | <0.001** | 1.08 | 0.02* | 0.99 | 0.02* |
| Alternative matching | 254 | 11,565 | 1.20 | <0.001** | 1.12 | <0.001** | 0.99 | 0.01* |
| | | | | | | | | |
| **PNC 6 days to 6 months** | | | | | | | | |
| *Baseline model* | *271* | *11,661* | *1.13* | *0.001*** | *1.04* | *0.30* | *1.00* | *0.27* |
| Generalized estimation | 273 | 11,663 | 0.99 | 0.83 | 0.80 | <0.001** | 0.99 | 0.19 |
| Outliers removed | 265 | 11,145 | 1.13 | 0.001** | 1.04 | 0.30 | 1.00 | 0.27 |
| Alternative matching | 254 | 11,098 | 1.12 | 0.002** | 1.01 | 0.81 | 0.99 | 0.09 |

All models control for variables fixed effect at the facility level, time and the square of time, SMGL and a time-trend for SMGL.

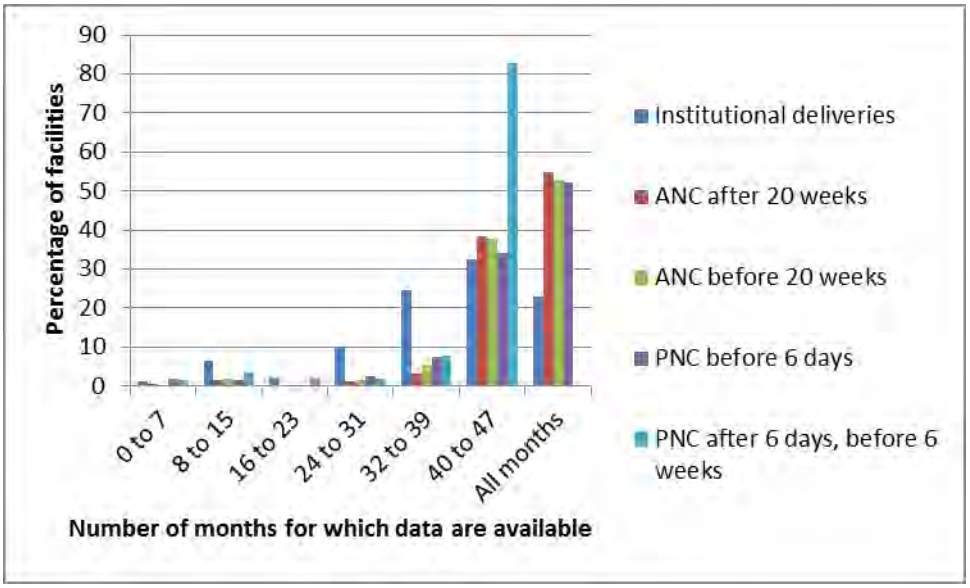
*Statistically significant at p<0.05

** Statistically significant at p<0.01

**Table B2: Percentage of facility-months of observations determined to be an outlier or and excess zero**

| Indicator | Percentage of observation identified as outliers or excess zeroes |
|---|---|
| Institutional deliveries | 4.8% |
| Family planning | 4.2% |
| ANC < 20 weeks | 4.2% |
| ANC > 20 weeks | 4.3% |
| PNC < 6 days | 4.4% |
| PNC > 6 days < 6 weeks | 4.8% |

**Figure B1: Missing data by outcome indicator**



# ANNEX C: REFERENCES


Arpino B, Mealli F. 2011. The specification of the propensity score in multilevel observational studies. *Computational Statistics and Data Analysis*. **55**: 1770-1780.

Ashraf N, Fink G, Weil DN. *Evaluating the effects of large scale health interventions in developing countries: The Zambian malarian initiative*. Brown University, Department of Economics 2010 Contract No.: 2010-9.

Bertrand M, Duflo E, Mullainathan S. 2004. How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics* February: 249-275.

Box GE, Tiao GC. 1975. Intervention analysis with application to economic and environmental problems. *Journal of American Statistical Association*. **70**: 553-64.

Central Statistics Office. *Preliminary Key Findings 2013-14 Zambia Demographic and Health Survey* Powerpoint presentation, 2014.

Dehejia RH, Wahba S. 1999. Causal effects in non-experimental studies: Re-evaluating the evaluation of training programs. *Journal of the American Statistical Association*. **94**: 1053-1062.

Drake C, 1993. Effects of misspecification of the propensity score on estimators of treatment effect. *Biometriks*. **49**: 1231-1236.

Ensor T, Green C, Quigley P, Badru AR, Kaluba D, Kureya T. 2014. Mobilizing communities to improve maternal health: results of an intervention in rural Zambia. Bulletin of the World Health Organization. **92**: 51-59.

Republic of Zambia Ministry of Health. December 2008. *HMIS Procedure Manual: Primary Health Care, Version 1.4*

Heywood A, Chishimba P, Lanjouw S. 2012. *Evaluation of the HMIS: Evaluation Report*. European Union's EDF Programme for Zambia

Hogan MC, Foreman KJ, Naghavi M, Ahn SY, Wang MR, Makela SM, Lopez AD, Lozano R, Murray CJL. Maternal mortality for 181 countries, 1980-2008: a systematic analysis of progress towards Millennium Development Goal 5. *Lancet* 2010; **375**; 1609-1623.

Hong G, Raudenbush SW. 2006. Evaluating kindergarten retention policy: A case study of causal inference for multilevel observational data. *Journal of the American Statistical Association*. **101** (475): 901-910.

Imbens G, Lemieux T. 2008, Regression discontinuity designs: A guide to practice. *Journal of Econometrics*. **142**:615-35.

Li F, Zaslavsky AM, Landrum MB. 2013. Propensity score weighting with multilevel data. *Statistics in Medicine*. **32** (19): 3373-3387.

Liang KY, Zeger S. 1986. Longitudinal data analysis using generalized linear models. *Biometrika* 73 (1): 13-22.

Lori JR, Amable EE, Mertz SG, Moriarty K, Behavior Change Following Implementation of Home-Based Life-Saving Skills in Liberia, West Africa. *J Midwifery Women's Health*. 2012; 57(5) Ñ495-501

Rubin DB, 2001. Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services & Outcomes Research Methodology*. **2**: 169-188.

StataCorp. 2011. Stata 12.1. College Station, Texas.

Sibley L, Buffington ST, Beck D, Armbruster D. Home-Based Life-Saving Skills: promoting safe motherhood through innovative community-based interventions. *J Midwifery Women's Health*. 2001; 46 (4):258-266.

UNICEF. 2010. Launch of campaign on accelerated reduction of maternal mortality in Zambia. http://www.unicef.org/media/media_53945.html Accessed October 14, 2014.