

# Fostering the use of quasi-experimental designs for evaluating public health interventions: insights from an mHealth project in Malawi

Jean Christophe Fotso<sup>1,a</sup>, Amanda Lea Robinson<sup>b</sup>, A. Camielle Noordam<sup>c</sup>, and Jessica Crawford<sup>d</sup>

<sup>a</sup> Concern Worldwide US, New York, USA

<sup>b</sup> Department of Political Science, The Ohio State University, USA

<sup>c</sup> Maastricht University, Maastricht, The Netherlands

<sup>d</sup> VillageReach, Balaka, Malawi

## Abstract

*The evidence base to support the growing field of mHealth is relatively nascent, with most studies lacking the level of rigor needed to inform scale up of interventions. This paper investigates the impact of a maternal, newborn and child health (MNCH) mHealth project in Malawi, comparing the intention-to-treat (ITT) and the treatment on the treated (TOT) estimates, and discussing the implications for future evaluations. Services offered included a toll-free case management hotline and mobile messaging service for women and children. The evaluation methods included a quasi-experimental pre-test post-test design, consisting of cross-sectional household surveys. A total of 4,230 women were interviewed in the intervention area and 2,463 in the control site. While the intervention did not have any ITT effects of the MNCH outcomes studied, there were large TOT effects. Rigorous evaluation designs can be successfully applied to mHealth pilot projects, helping to understand what works and what does not.*

**Keywords:** Maternal, newborn & child health; Evaluation; Quasi-experimental design; Malawi

## Résumé

*Les évidences pour soutenir le domaine de plus en plus croissant de la mSanté sont relativement nouvelles, la plupart d'études manquant la rigueur nécessaire pour informer les différentes interventions. Cet article analyse l'impact d'un projet mSanté sur la santé maternelle, néonatale et infantile (SMNI) au Malawi, en*

---

<sup>1</sup> Corresponding author

Concern Worldwide US, 355 Lexington Avenue, 19th floor; New York NY 10017, USA.

Tel: +1-212-557-8000; Fax: +1-212-557-8004

Email: [jean-christophe.fotso@concern.net](mailto:jean-christophe.fotso@concern.net)

*comparant l'estimateur de l'intention de traiter (IDT) et l'effet du traitement sur les traités (TST), et discutant les implications pour les interventions futures. Le projet offrait deux services : une ligne téléphonique d'appels sans frais, et un service mobile de messagerie pour les femmes et les enfants. L'évaluation a utilisé une méthode quasi-expérimentale avec comparaison avant-après à partir des enquêtes-ménages transversales. Un total de 4230 et 2463 femmes ont été interviewées dans le site d'intervention et de contrôle, respectivement. Bien que l'intervention n'ait eu aucun effet IDT sur les indicateurs de SMNI étudiés, il y a eu au contraire un fort impact TST. Des méthodes rigoureuses d'évaluation peuvent être appliquées dans les projets pilotes mSanté, afin de mieux comprendre ce qui pourrait ou ne pas fonctionner.*

**Mots clés:** mSanté ; Santé maternelle, néonatale et infantile ; Evaluation; Méthodes quasi-expérimentales; Malawi

## **Introduction**

As governments and international organizations strive to meet the health-related Millennium Development Goals (MDGs), the rapidly expanding field of mobile health (mHealth), broadly defined as the use of mobile phones to improve health and health systems, presents an unprecedented opportunity to increase access to health care and improve health outcomes (mHealth Alliance, 2012a; Al-Shorbaji & Geissbuhler, 2012; WHO, 2011). The declining cost of mobile phones, growth in subscriptions, and rapid advances in technology have triggered an explosion of mHealth pilot projects and programs in low- and middle-income countries (LMICs) since the mid-2000s (Mechael et al., 2012). A recent report shows that between 2005 and 2010, mobile telephone subscriptions grew almost three-fold in sub-Saharan Africa and more than six-fold in South Asia (World Bank & ITU, 2012). Not surprisingly, the second World Health Organization (WHO) global survey on mobile technologies reported that 83% of the 112 participating countries had at least

one mHealth initiative, and of these, three quarters reported having four or more mHealth initiatives (WHO, 2011).

There is a large body of literature reporting that mobile phones have been instrumental in achieving healthy behavior change such as that related to smoking cessation, weight loss, diet and physical activity, treatment adherence, and disease management (Free et al, 2013a; 2013b) and for improving health worker and health system performance (Gurman et al., 2012). In the areas of maternal, newborn and child health (MNCH), mHealth has been shown to have the potential to support and strengthen existing efforts to improve timely access to care (mHealth Alliance, 2012b; Tamrat & Kachnowski, 2011).

Existing MNCH mHealth interventions in LMICs tend to fall under the following three broad objectives: (1) stimulating demand and awareness for uptake of healthy behaviors among pregnant women and mothers; (2) strengthening human resource capacity for health care delivery; and (3) improving health system capacity (mHealth Alliance, 2012b; Lemaire,

2011). Demand generation activities in particular seek to provide information services geared towards increasing awareness and encouraging use of available health resources. They also include complementary patient-support services addressing the management of health issues, and communication services that connect women to peer networks or expert resources in the community such as skilled birth attendants (Noordam et al., 2011; Gurman et al., 2012; Tamrat & Kachnowski, 2011). Mobile phones are also used to improve point-of-care decision-making by community- and facility-based health workers (Mitchell et al., 2013; Tamrat & Kachnowski, 2012), and to strengthen referral systems (Crawford et al., 2014). A key milestone in the expansion of MNCH mHealth was reached with the development of a strategic framework conceptualizing the use of mobile technology to improve MNCH, whose major pillars include gender mainstreaming and women empowerment, and more generally, the gender and social dynamics that impact access to quality health services for women (mHealth Alliance, 2012b).

### **How robust is the mHealth evidence base?**

While mHealth interventions are gaining traction in almost all parts of the world, the evidence-base to support this trend is relatively nascent (Tomlinson et al., 2013; Leon et al., 2012; mHealth Alliance, 2012b). The body of literature shows that many mHealth interventions are pilot projects with limited measures of effectiveness (Fiordelli et al., 2013). When evidence of mHealth effectiveness exists, it is primarily on interventions in developed countries or for chronic

disease prevention and management and life-style changes, with most studies on MNCH mHealth in developing countries still lacking the level of rigor needed to inform replicability and scale up (mHealth Alliance, 2013, Gurman et al., 2012).

A recent review of the effectiveness of mHealth to improve service delivery noted that out of the 42 controlled trials identified none was implemented in LMICs. The authors concluded that robust evaluations are needed to ascertain the effectiveness of mHealth interventions and provide evidence for scale-up (Free et al., 2013a). A systematic review of the effectiveness of mHealth to improve health behaviors or disease management found that of the 75 trials identified, only three were conducted in LMICs (Free et al., 2013b). From a compilation of publications on mHealth behavior change communications, the key findings from most articles reviewed were from formative research or process evaluations (Gurman et al., 2012).

Another recent review of maternal health focused mHealth project evaluations revealed a scarcity of quantitative assessments, with only four out of 34 articles and reports covered including a quantitative design (Tamrat & Kachnowski, 2011). More generally, many of the studies published on mHealth suffer from methodological weaknesses such as the lack of intervention and control groups, absence of a baseline sample for comparison, small sample sizes, or being primarily descriptive rather than analytical in nature (mHealth Alliance, 2013; Tamrat & Kachnowski, 2011; Noordam et al., 2011). A more scientific approach, which frames and answers critical evaluation questions, is needed to

support and inform the expansion of the mHealth field, particularly in sub-Saharan Africa. This assessment aligns with views from mHealth experts who acknowledge that “generating quality evidence through methodologically rigorous research has emerged as a priority for the broader mHealth community” (mHealth 2013).

As a result of our limited evidence base on the effectiveness of the mHealth interventions, and our limited understanding of the challenges associated with the integration of these interventions in the existing structures or in national health policies, strategies and regulations, only a few mHealth projects have been brought to scale (Noordam et al., 2015; Lemaire, 2011; WHO 2010).

Overall, with a proliferation of mHealth pilot projects in developing countries, there is an urgent need to document the benefits and lessons learned. Evaluation research which seeks to attribute changes in outcomes to an intervention, and to fill gaps in understanding of what works and what does not, is essential to promote accountability in the allocation of resources and to maximize the potential of mHealth to improve MNCH outcomes (Victora et al., 2011).

Against this backdrop, the aim of this paper is to contribute to building the evidence base on the effectiveness of mHealth interventions. It investigates the impact of an MNCH mHealth project on uptake of home-based and facility-based practices for MNCH in a rural district of Malawi. Of special interest is the comparison of two types of impact estimates: a) the intention-to-treat (ITT) estimate, obtained by applying the evaluation analysis on individuals to

whom the intervention was offered, regardless of whether or not they actually enrolled; and b) the effect of treatment on the treated (TOT), which, by contrast, focuses on individuals to whom the intervention was offered and who actually enrolled. Given the very large discrepancies we find between the two estimates, the paper also discusses the implications of not accounting for self-selection bias under conditions of low uptake.

## Data and Methods

### Study Settings

The intervention was conducted in Balaka District in the Southern region of Malawi. At the time of the design of the intervention in 2010, over 90% of its total population of about 320,000 inhabitants lived in rural villages; fewer than 5% of the homes had electricity, and only about 15% of individuals aged 15 years or above reported having attended school beyond primary level. Balaka was among the districts with the poorest MNCH outcomes in Malawi. For example, it had highest infant mortality (104 per 1000 live births) and under-five mortality (160 per 1000 live births), and more than 40% of children under the age of five years were undernourished (IKI 2013). The selection of the control district was based on proximity to ensure the characteristics of the populations in the control and interventions areas are as similar as possible. Out of the five districts contiguous to Balaka, Ntcheu District was selected, as the other four districts were at the time, or had recently had been, the sites of numerous health projects (IKI 2013).

## **The MNCH mHealth Project in Malawi**

The MNCH mHealth project called *Chipatala cha pa foni* (CCPF) – meaning health center by phone – was conducted in Malawi with the aim of increasing knowledge and use of home- and facility-based MNCH services. To achieve these objectives, the intervention offered a toll-free case management hotline and an automated and personalized mobile messaging service. Community volunteers, trained and provided with phones, conducted community mobilization in the intervention sites and facilitated access to services to those without phones (IKI, 2011). The project was implemented between July 2011 and June 2013.

The toll-free case management hotline provided protocol-based health information, advice and referrals. It allowed mothers or caregivers of children under the age of five years, who may not be able or willing to access in-person consultations at a health center, to connect via mobile phone to trained personnel based at the district hospital. While women using a personal phone could dial the hotline number directly, those without access to a phone were instructed to go to the nearest community volunteer's house and use the project phone to connect privately to the hotline center (Innovations, 2012). The automated and personalized mobile messaging service, on the other hand, offered the opportunity for registered pregnant women, guardians of children under five years of age, and women of child-bearing age to receive weekly text and voice messages on appropriate health care-seeking practices. The messages were delivered to a user's phone or could be retrieved

in the form of voicemails by calling the toll-free number and responding to voice prompts from the system. Registration of women for this service was usually done during the first hotline call or during antenatal care (ANC) visits.

The project was implemented in the catchment areas of four health centers in Balaka District, namely, Phimbi, Chiyendausiku, Kalembo and Mbera, the only health centers that met the selection criteria set by the program and research teams (presence of electricity, at least two maternity nurses, and cell phone coverage). The district hospital and Christian Health Association of Malawi (CHAM) facilities were not considered for the intervention as their clients are likely to have different socioeconomic and health profiles (e.g. services are not free at CHAM) (IKI, 2013). The catchment areas of these four health centers covered a total population of about 155,000 residents, or about 48% of the district's total population.

### **Evaluation design and source of data**

A two-arm quasi-experimental, pre-post design was used to quantify the impact of the intervention on knowledge and use of home-based and facility-based care for mothers and children. The neighboring Ntcheu District in the Central region was selected to serve as control, as it was deemed most likely to exhibit similar MNCH outcomes as Balaka district. With the same eligibility criteria as for the intervention area, the Bwanje and Kasinje health centers were selected among eligible facilities (IKI, 2013). Given the nature of the project, randomization of subjects or groups between intervention and control was

neither practical nor feasible. In this context, a quasi-experimental design - use of a control group with nonrandom assignment to the groups - is shown to approximate the randomized experiment (Duflo et al. 2008) when observed differences between the intervention and the control sites are accounted for in the analyses (Heckman, 2005; Meyer, 1995). However, in quasi-experimental designs there may be unobserved differences between the intervention and control areas, making crucial the selection of an appropriate control area. The core of the assessment were baseline and endline cross-sectional population-based surveys of mothers aged 15-49 who had children under five years of age, as well as pregnant women and caregivers of children under five years of age, conducted in June-July 2011 and April-May 2013, respectively. The questionnaires covered household characteristics and women's knowledge and use of home-base and facility-based MNCH services. Qualitative methods were also used, but they are beyond the scope of this paper.

Villages were the primary sampling units. At baseline, Geographic Information Systems information and maps from the health centers were used to define the catchment areas of the

health facilities and to create a comprehensive list of villages with information on total population, estimated number of women of child-bearing age and estimated number of children less than five years of age. Using a systematic sampling approach, villages were randomly selected to be included in the study such that villages in each health center catchment area had the same probability of being selected into the sample regardless of their population size. Of the 325 villages in the intervention area, 39 villages were randomly selected into our sample, and of the 66 villages in the control area, 10 villages were randomly selected. All households within each village with eligible mothers or caregivers were selected and questionnaires were administered to all eligible respondents (IKI, 2011).

The number of survey respondents is shown in Table 1. A total of 6,453 households (2,810 at baseline and 3,643 at endline) were successfully visited: the response rate of households varied from 83-91% across rounds and catchment areas, with non-response due to refusals or the absence of eligible respondents across three household visits. The 6,453 households yielded a total of 6,693 women aged 15-49 and 6,846 children under the age of five years.

**Table I.** Number of survey respondents by population group at baseline and endline

	Baseline survey			End-line survey			Total
	Control	Intervention	Total	Control	Intervention	Total	
Health facility catchment areas	2	4	6	Same as at baseline			6
Villages <sup>1</sup>	10	8	18	39	38	77	NA
Households <sup>2</sup>	1,112	1,698	2,810	1,284	2,359	3,643	6,453
Women 15-49 <sup>2</sup>	1,119	1,721	2,840	1,344	2,509	3,853	6,693
Under-5 children <sup>2</sup>	1,365	2,220	3,585	1,075	2,186	3,261	6,846

<sup>1</sup>Random selection of villages (primary sampling units) in the catchment areas of all six qualified health centers in the intervention (4 health centers) and control (2 health centers) sites

<sup>2</sup>All households in the selected villages, and all women and children under-five in those households

### Outcome and control variables

This paper analyzes 14 outcome variables which are grouped into four aggregate outcome measures as follows:

- **Use of home-based practices for maternal health**, derived from: 1) bednet use during pregnancy, and 2) initiation of breastfeeding within one hour of birth. Both questions were asked of women who had a live birth in the last 18 months (n=2813).
- **Use of home-based practices for child health**, constructed from: 1) exclusive breastfeeding until six month of age, (2) bednet use by under-five child during the previous night, and 3) oral rehydration salt (ORS) to under-five child sick with diarrhea in the previous two weeks. The last question was asked for children who had experienced diarrhea in the past two weeks (n=1269), while the two others were asked for all children (n=6846).
- **Use of facility-based services for maternal health**, aggregated from: 1) received the correct dosage of the tetanus toxoid (TT) vaccine during pregnancy, 2) received a Vitamin A dose during pregnancy; 3) received the recommended four ANC consultations, 4) started ANC in the first trimester, 5) gave birth under the supervision of a skilled birth attendant, and 6) received one postnatal care (PNC) check-up within two days of birth. All questions were asked of women who had a live birth in the last 18 months (n=2813).
- **Use of facility-based services for child health**, computed from: 1) child was fully immunized by the

first birthday, 2) health facility care was sought for child with symptoms of acute respiratory infections (ARI) in the previous two weeks, and 3) health facility care was sought for child with fever in the previous two weeks. The three questions were asked for children between 12 and 24 months of age (n=1610), children with symptoms of ARI in the past two weeks (n=1895), and children with fever in the past two weeks (n=2194), respectively.

The indicators were aggregated using the following three-step protocol described by Kling et al. (2007). First, each indicator was recoded so that a higher value indicates a better outcome, as encouraged by the intervention. Second, missing values on individual variables were imputed at the group (intervention or control) mean (missingness ranged from 0%-2%). Third, the aggregate values were computed as averages across the set of indicators that apply to a given respondent. For example, if a child had not had diarrhea in the last two weeks, the ORS variable will be excluded from the computation of the aggregate home-based care.

Multivariate models control for variables at the community level (mean distance to the health center), household level (household wealth, number of under-five children, and ethnicity and religion of the household head), woman level (access to phone, education, marital status and age), and child level (age and sex). The household wealth variable is constructed on the overall sample from household characteristics (presence of electricity and type of drinking water, toilets, wall, roof and

floor) and household possessions of durable goods (e.g. bicycle, TV, fridge and watch), using principal component analysis (Jolliffe, 2002). The variable was further recoded as a dichotomous variable using the median value as the cut-off point (low 50%, high 50%). The specification of these variables is detailed in Tables 2A and 2B.

### Methods of analysis

We assess the impact of the intervention using the difference-in-difference (DID) method, the most widely used method for impact evaluation in the context of quasi-experimental designs (Heckman, 2005; Meyer, 1995). The analysis is carried out in three steps. **First**, we estimate the simple DID for a given outcome  $Y$  as follows:

$$DID(Y) = (\bar{Y}_{IE} - \bar{Y}_{CE}) - (\bar{Y}_{IB} - \bar{Y}_{CB}) \quad (1)$$

where  $\bar{Y}_{IE}$  and  $\bar{Y}_{CE}$  represent the average outcome at endline in the intervention site and control area, respectively, and  $\bar{Y}_{IB}$  and  $\bar{Y}_{CB}$  represent the average outcome at base line in the intervention site and control area, respectively. The DID estimator allows for unobserved heterogeneity between the intervention and the control sites, but assumes this unobserved heterogeneity is time invariant; so the potential bias cancels out through differencing (Bertrand et al., 2004; Rubin, 1974). Since CCPF was offered but not compulsory, this estimate is to be interpreted as intention-to-treat (ITT) effect – as it entails comparing the intervention and control areas without regard for the actual use of CCPF.

The DID estimate in (1) can also be calculated within a regression framework as follows:

$$Y_{ivt} = \beta_0 + \beta_1 T_v + \beta_2 P_t + \beta_3 (T * P)_{vt} + \varepsilon_{ivt} \quad (2)$$

where  $Y_{ivt}$  is the outcome measure for woman/child  $i$ , in village  $v$ , at time  $t$ .  $T_v$  is a dummy variable taking the value 1 for individuals in treatment areas and 0 for individuals in control areas,  $P_t$  is a dummy variable taking the value 0 for the baseline data and 1 for the endline data, and  $\varepsilon_{ivt}$  is the idiosyncratic error, clustered by health center catchment area. The DID estimator of interest is the coefficient  $\beta_3$  of the interaction between  $T_v$  and  $P_t$  and is the same as the estimate obtained in equation (1) when no controls are included.

**Next**, we estimate the adjusted effect using regression-based DID controlling for possible confounders according to the following formula:

$$Y_{ivt} = \beta_0 + \beta_1 T_v + \beta_2 P_t + \beta_3 (T * P)_{vt} + W'_{ivt} \Phi + \gamma X_v + \varepsilon_{ivt} \quad (3)$$

where  $W_{ivt}$  is a vector of the controls at the household, women and child levels (as presented in Tables 2 and 3), and  $X$  is the village-level control variable (mean distance to health facility as described in Table 2).

**Finally**, to assess the impact of the intervention on women who actually used the services, we estimate the treatment effect on the treated (TOT) which, in contrast to the ITT, compares the individuals who actually used the services to similar individuals in the control area. The method uses instrumental variable analyses to construct a proper counterfactual –

women who would have likely used the services in control communities had they been offered (Have et al. 2008, Angrist et al., 1996). We use the characteristics of women in the intervention areas who actually used the services to identify women with those same characteristics in the control areas, whom we expect, would have used the services had they been offered.

### Ethical clearance

Ethical approval for the study was granted by the Ministry of Health's National Health Sciences Research Committee in Malawi.

## Evaluation Results

### Sample characteristics

As shown in Table 1, a combined (baseline and endline) total of 4,230 women were interviewed in the intervention area, compared to 2,463 in the control site. The corresponding figures for children under-five were 4,406 and 2,440, respectively. Tables 2 and 3 compare the distribution of women and children in the intervention

and control areas at baseline. The distribution of women by education and age is similar in both areas, with about three women in four having a primary level education, and approximately 55% being between 20-29 years of age. Women's marital status, the mean distance to the health center, and the number of under-five children in the household, are similarly distributed across the intervention and the control sites. Table 2 also shows that the proportion of women from wealthier households is comparable across the two areas - household wealth was constructed on the pooled sample to allow the comparison of socioeconomic status across the two sites. By contrast, women's access to phones was higher in intervention area (32%) than in control communities (22%). The variables displaying the largest distributional differences were ethnicity, and to a lesser degree, religion. Finally, the distribution of children by sex and age in Table 3 is similar across the intervention and the control groups.

**Table 2.** Percentage distribution of mothers/caretakers of children under five and pregnant women **at baseline**

	Control	Intervention
	%	%
<b>Community level covariates</b>		
Mean distance to the health center (km)	5.7	4.4
<b>Household level covariates</b>		
Wealth		
Poor (Lowest 50%)	56.6	53.0
Rich (Highest 50%)	43.4	47.0

	Control	Intervention
	%	%
Number of under-five children		
0	5.7	4.5
1	66.6	60.8
2+	27.7	34.7
Ethnicity of the head of household		
Lomwe	6.0	21.0
Ngoni	77.3	20.6
Yao	6.8	38.9
Other	9.9	19.5
Religion of the head of household		
Catholic	16.1	17.5
Other Christian	70.1	41.0
Muslim	5.5	37.0
Other/No religion	8.4	4.5
<b>Woman-level covariates</b>		
Access to a phone		
No	77.6	68.0
Yes	22.4	32.0
Education		
None	16.5	15.4
Primary	73.7	74.0
Secondary+	9.7	10.6
Marital status		
Not in union	11.0	17.0
In union	89.0	83.0
Age in years		
<20	10.6	11.3
20-29	57.2	53.9
30+	32.2	34.7
<b>N</b>	<b>1,119</b>	<b>1,721</b>

**Table 3. Percentage distribution of children under five at baseline**

	Control	Intervention
	%	%
Age in months		
< 12	25.4	23.2
12-23	21.2	19.7
24+	53.4	57.0
Sex		
Male	48.4	51.3
Female	51.6	48.7
<b>N</b>	<b>1,365</b>	<b>2,220</b>

At endline, awareness of the toll-free hotline service was high, with nearly 77% of respondents in intervention communities reporting that they had heard about it, as indicated in Table 4. In contrast, only a third of respondents in the intervention communities were aware of the mobile messaging service, all of whom had also heard of the

hotline. Use of the services was relatively low among women who were aware of the intervention, at 24% for the hotline services and 23% for the messaging services. Thus, overall only 18% of the 2509 women in the intervention area at endline had used the hotline service, and only 8% had used the messaging services.

**Table 4.** Awareness and use of the services among women of child bearing age **at endline**

	Intervention area		Control area	
	%	N	%	N
Awareness and use of the hotline services				
Heard about the services	76.9	2,509	2.8	1,344
Used the services	23.8	1,929	3.5	38
Awareness and use of the mobile messaging services				
Heard about the services	33.3	2,509	0.4	1,344
Used the services	22.6	835	0.0	5

### **Effects of the intervention on home-based and facility-based MNCH care**

Table 5 presents two estimates of the impact of CCPF on the four aggregate MNCH outcomes of interest: the ITT and the TOT effects. The subsample serving as the denominator for each outcome is also described in the table.

#### **Unadjusted ITT effects**

The results of the unadjusted DID of the ITT model are presented in the first column of Table 5. They show a strong, negative average treatment effect of the project on facility-based care for child health ( $p < 0.01$ ), which results from a decrease of service use in the

intervention area, and an increase in the control site (not shown). Likewise, there is a negative and statistically significant effect on home-based care for children ( $p < 0.05$ ), though of lower magnitude, compared to the effect on facility-based care for child health. However, in this case, the negative effect is a result of a steeper increase in the control areas compared to the intervention site (not shown). The results in Table 5 also show that the intervention did not have any significant impact on maternal health, the small, positive effect on facility-based care failing to reach statistical significance, and the effect on home-based care appearing negligible.

**Table 5.** Effects (coefficients) of the intervention on aggregate home-based and facility-based MNCH outcomes

	Intention-to-treat (ITT) effect		Treatment on the treated effect	Sub-sample	
	Simple DID <sup>1</sup>	Adjusted <sup>2</sup> DID		N	Description
Use of home-based care for maternal health <sup>3</sup>	-0.006	-0.012	0.479***	2,813	Women 15-49 who had a live birth in last 18 months
Use of home-based care for child health <sup>4</sup>	-0.059**	-0.071**	0.603***	6,846	All children under five
Use of facility-based care for maternal health <sup>5</sup>	0.092	0.085	0.239**	2,813	Women 15-49 who had a live birth in last 18 months
Use of facility-based care for child health <sup>6</sup>	-0.172***	-0.171***	-0.499***	4,068	Children who are 12-23 months, or had ARI or fever in last 2 weeks

Statistical significance: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

<sup>1</sup>Difference in difference

<sup>2</sup>Controlling for variables at community, household, woman and child levels (in Tables 2 & 3)

<sup>3</sup>From the variables on bednet use during pregnancy, and initiation of breastfeeding.

<sup>4</sup>From the variables on exclusive breastfeeding, bednet use, and oral rehydration salt use.

<sup>5</sup>From the variables on tetanus toxoid vaccine, Vitamin A, ANC 4+, ANC initiation, skilled birth attendance, and PNC.

<sup>6</sup>From the variables on full immunization, care for acute respiratory infections, and care for fever

### **Adjusted ITT effects**

The second column of Table 5 presents the adjusted DID of the ITT estimates controlling for community, household, women and child level variables. As can be seen, the inclusion of potential confounders in the model did not result in major changes in the magnitude, direction or statistical significance of the effects. While the size of the effect on home-based care for child health increased slightly (from -0.059 to -0.071), the level of significance did not change. The impact of the three other outcomes remained largely unchanged.

### **Treatment effects on the treated**

Because of the low uptake of the services (only 18% of women in the intervention communities used CCPF) and the associated dilution of the treatment effect, it is critical to evaluate the impact of the intervention on women who actually used the services. The results in the third column of Table 5 reveal striking differences between the TOT and the ITT effects in terms of magnitude, direction and statistical significance. We note a large, positive and statistically significant TOT effect of the intervention on aggregate home-care for child health ( $p < 0.01$ ), as opposed to the negative ITT effect described above. There is a large, positive TOT effect on home-based care for maternal health ( $p < 0.01$ ). CCPF also recorded a positive, significant effect on facility-based care for maternal health among women who used the services ( $p < 0.05$ ). The negative ITT on facility-based child health care is amplified among women who used CCPF ( $p < 0.01$ ). In short, the treatment on the treated analyses suggest a much more

optimistic evaluation of the impact of CCPF.

### **Summary and Discussion**

As robust studies providing evidence on the impact of mHealth interventions on health outcomes are still lacking (mHealth Alliance. 2013, Tomlinson et al. 2013), this study used quasi-experimental evaluation data to assess the impact of an MNCH mHealth intervention on home-based and facility-based care for mothers and children. The analyses show dramatic differences between two approaches: the ITT model which uses a DID estimator to deal with unobserved differences between control and intervention communities, and the TOT model which employs instrumental variable analysis to address the self-selection bias inherent to projects with voluntary enrolment (Angrist et al., 1996; Bertrand et al., 2004).

Our analyses show a large, positive effect of the intervention on the utilization of home-based care for child health among those who used the services offered, in contrast to a modest, negative ITT effect observed on the same outcome. Similar differences are observed on home-based and facility-based care for maternal health, with large, positive TOT effects on both outcomes, and virtually no ITT effect on either of the outcomes. Only on facility-based care for child health do the two approaches yield similar results. There is a large, negative ITT effect of the intervention on facility-based care for child health, an effect which is amplified among women who used the services offered by CCPF. Separate analyses in another paper of this series suggest that this negative effect is driven solely by the

use of facility-care among children who had fever in the two weeks preceding the survey, and conclude that the intervention contributed to strengthening the home-to-facility continuum of care by reducing the unnecessary visits to health facilities for conditions that can be adequately managed through home-based care (Fotso et al. 2015).

The finding that less than 3% of individuals in the control district had heard about the intervention suggests that contamination across the intervention and control areas did not occur. To some extent, the variables on home-based and facility-based care for mothers and children were limited in number and scope. For example the inclusion of variables on facility-based treatment of diarrhea and home-based treatment of fever, would have been ideal. A major weakness of the intervention itself was the breakdown of community volunteer's phones, which limited, to some degree, the access to the services among women who did not have a personal phone, and may explain a part of the gap between awareness and use.

While contributing to the evidence base on the impact of mHealth interventions, our findings point to the need to carry out advanced evaluation techniques beyond the DID approach, especially when the proportion of the target individuals who enrolled in the services offered is low (Have et al. 2008; Bertrand et al. 2004). The merits and limitations of the TOT and ITT approaches are discussed below.

### **ITT model: Differences between the control and intervention sites**

The DID approach based on the ITT principle is designed to deal with differences at baseline between the control and the intervention communities. Indeed, although the two areas were adjacent to each other, notable disparities were observed for a few health outcomes prior to the intervention (IKI, 2012). The key assumption allowing DID to identify a causal effect is that in the absence of the intervention, the unobserved differences between intervention and control communities would remain the same over time, also referred to as the parallel trends assumption (Heckman 2005, Meyer 1995).

In this study, the parallel trend assumption was likely to be violated in two different ways. When an indicator was much lower in the control area than in the intervention site at baseline, as was the case for the proportion of children who slept under a bednet the previous night (71% versus 82%), the parallel trends assumption may not hold. Indeed at endline, 93% of children in control communities slept under a bednet the previous night, an improvement of 22 percentage points. According to the parallel trend assumption, in the absence of the intervention, the indicator would have also had to improve by 22 percentage points (to 104%), which is obviously not possible.

A second avenue through which the assumption may have been violated relates to programmatic changes between baseline and endline in either or both areas, like differential implementation of other MNCH interventions. Our data indicate that

new MNCH programs were introduced at a higher rate in control communities compared to intervention communities (IKI, 2013), which violates the parallel trends assumption since we would expect greater improvements in the control communities in the absence of the CCPF intervention. This ultimately suggests that the ITT effects reported in the analyses are likely to be underestimates.

### **TOT Model: Non-compliance and self-selection bias**

The uptake of different components of the CCPF intervention was low, at 8% for the messaging system and 18% for the hotline. As a result, the ITT estimates are not likely to provide a good indication of the impact of the intervention. The TOT method attempts to adjust for two critical self-selection biases. There are individuals in the control area who would not be able or willing to use the services even if they were offered. A mere comparison of the control and intervention areas as in the ITT principle, without accounting for this selection would lead to biased impact estimates (Angrist et al., 1996, Bertrand et al., 2004). The second self-selection bias accounted for in the TOT method is the fact that only a subgroup of the individuals assigned to the intervention area actually used the services, and this selection was non-random (Have et al 2008).

### **Conclusion**

This study has contributed to the evidence base on the effectiveness of mHealth to improve MNCH outcomes. The analyses of the datasets generated in this evaluation provided invaluable insight for the program in terms of

understanding the impact on the population and among those who used the services. Furthermore, this paper demonstrates that rigorous quasi-experimental evaluation designs can be successfully applied to mHealth pilot projects, helping to understand what works and what does not. The substantial differences between the ITT and TOT estimates observed in this study suggest that evaluations of programs like CCPF should go beyond control-intervention comparisons, and leverage the insights from TOT-type analyses.

These differences are driven by the fact that there was fairly low uptake in intervention areas and that self-selection into the use of services in intervention communities introduced substantial bias. Once non-compliance is accounted for, and the likelihood of service uptake modelled, the TOT estimate provides an unbiased estimate of the impact of CCPF services, furthering our understanding of the impact if there were full compliance.

The strong, positive TOT effects of the intervention on most of the MNCH outcomes covered suggest that for the replication or scale up of an intervention like CCPF, redoubled efforts should be made through community mobilization and improved access to phones, to ensure that the enrolment in the use of the services is high.

### **Acknowledgements**

The CCPF project is part of *Innovations for Maternal, Newborn & Child Health*, an initiative of Concern Worldwide U.S. funded through a multi-year grant from the Bill & Melinda Gates Foundation. The Government of Norway and the United Nations

Foundation also supported the Malawi mHealth project (CCPF) through the Innovation Working Group Catalytic mHealth Grants program as part of the UN Secretary General's Every Women Every Child strategy. CCPF was implemented by VillageReach, an international NGO headquartered in Seattle, USA. We would like to give special thanks to the Reproductive Health Unit and its Director, Mrs. Fannie Kachale, and the Balaka District Health Office for their support of CCPF. The evaluation was conducted by Invest in Knowledge Initiative (IKI), a Malawi-based research institution, with the leadership of Professor Susan Watkins of University of Pennsylvania and Dr. Amanda Robinson of Ohio State University. The authors are grateful to Dr. Linda Vesel of Concern Worldwide U.S. for her insightful comments on the manuscripts, and Ms. Lindsay Ross of Columbia University for her contribution to the literature search. They are also grateful to the anonymous reviewers for their comments.

### Authors' Contributions

JCF designed the paper; ALR and JCF conducted the analyses; JCF, ALR, ACN and JC all contributed the writing of the manuscript. All authors discussed the results and commented on the manuscript at all stages.

### References

- Al-Shorbaji, N. and Geissbuhler, A. 2012. "Establishing an evidence base for e-health: the proof is in the pudding." *Bulletin of the World Health Organization* 90:322-322A.
- Angrist, J.D., Imbens, G.W. and Rubin D.B. 1996. "Identification of causal effects using instrumental variables." *Journal of the American Statistical Association* 91(434):444-455.
- Bertrand, M., Duflo E., and Mullainathan, S. 2004. "How much should we trust differences-in-differences estimates?" *Quarterly Journal of Economics* 119(1):249-275.
- Crawford, J., Larsen-Cooper, E., Jezman, Z., Cunningham S.C., and Bancroft, E. 2014. "SMS versus voice messaging to deliver MNCH communication in rural Malawi: assessment of delivery success and user experience." *Global Health: Science & Practice* 2(1):35-46.
- Duflo E, Glennerster, R. and Kremer, M. 2008. "Using Randomization in Development Economics Research: A Toolkit." In *Handbook of Development Economics*, vol. 4, ed. T. Paul Schultz and John Strauss. Amsterdam: North-Holland.
- Fiordelli, M., Diviani, N. and Schulz, P.J. 2013. "Mapping mHealth research: a decade of evolution." *Journal of Medical Internet Research* 15(5): e95.
- Fotso, J.C., Bellhouse, L., Vesel, L., Jezman, Z. 2015. "Strengthening the home-to-facility continuum of newborn and child health care through mHealth: evidence from an intervention in rural Malawi." *African Population Studies* 29(1): 1663-1682.
- Free, C., Phillips, G., Galli, L., Watson, L., Galli, L., Felix, L., et al. 2013a. "The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review." *PLoS Med* 10(1):e1001362.
- Free, C., Phillips, G., Watson, L., Galli, L., Felix, L., et al. 2013b. "The effectiveness of mobile-health technologies to improve health care service delivery processes: a

- systematic review and meta-analysis.” *PLoS Med* 10(1):e1001363.
- Gurman, T.A., Rubin, S.E. and Roess, A.A. 2012. “Effectiveness of mHealth behavior change communication interventions in developing countries: a systematic review of the literature.” *Journal of Health Communication* 17:82-104.
- Have, T.R.T., Normand, S.L.T., Marcus, S.M., Brown, C.H., Lavori, P. and Duan, N. 2008. “Intent-to-treat vs. non-intent-to-treat analyses under treatment non-adherence in mental health randomized trials.” *Psychiatric Annals* 38(12):772-783.
- Heckman, J. 2005. “The scientific model of causality.” *Sociological Methodology* 35(1):1-97.
- Innovations for MNCH. 2012. “Bridging Women and Children to Better Health Care: Maternal, Newborn and Child Health (MNCH) Access through Mobile Technology in Balaka District, Malawi.” Concern Worldwide.
- Invest in Knowledge (IKI). 2011. “Baseline Report for an Evaluation of the Information and Communications Technology to Improve Health Services for Mothers and Children Project.” Malawi.
- Invest in Knowledge (IKI). 2013. “Evaluation of The Information and Communications Technology to Improve Health Services for Mothers and Children Project.” Malawi.
- Jolliffe, I.T. 2002. “Principal Component Analysis.” Series: Springer Series in Statistics, 2nd ed., Springer, NY.
- Kling, J.R., Leibman, J.B. and Lawrence, K.F. 2007. “Experimental analysis of neighborhood effects.” *Econometrica* 75(1):83-119.
- Lemaire, J. 2011. “Scaling up Mobile Health: Elements Necessary for the Successful Scale up of mHealth in Developing Countries.” Geneva: Advanced Development for Africa.
- Leon, N., Schneider, H. and Daviaud, E. 2012. “Applying a framework for assessing the health system challenges to scaling up mHealth in South Africa.” *BMC Medical Informatics and Decision Making* 12:123.
- Mechael, P., Nemser, B., Cosmaciuc, R., Cole-Lewis, H., et al. 2012. “Capitalizing on the characteristics of mHealth to evaluate its impact.” *Journal of Health Communication: International Perspectives* 17:sup1, 62-66.
- Meyer, B.D. 1995. “Natural and quasi-experiments in economics.” *Journal of Business & Economic Statistics* 13(2): 151-161.
- mHealth Alliance. 2013. “mHealth and MNCH: State of the Evidence - Trends, Gaps, Stakeholder Needs, and Opportunities for Future Research on the Use of Mobile Technology to Improve Maternal, Newborn, and Child Health.” Washington DC.
- mHealth Alliance. 2012a. “Leveraging Mobile Technologies to Promote Maternal & Newborn Health: The Current Landscape & Opportunities for Advancement in Low-Resource Settings.” Washington DC.
- mHealth Alliance. 2012b. “Leveraging Mobile Technologies for Maternal, Newborn & child Health: A Framework for Engagement.” Washington DC.
- Mitchell, M., Hedt-Gauthier, B.L., Msellemu, D., Nkaka, M. and Lesh, N. 2013. “Using electronic

- technology to improve clinical care – results from a before-after cluster trial to evaluate assessment and classification of sick children according to Integrated Management of Childhood Illnesses (IMCI) protocol in Tanzania.” *BMC Medical Informatics and Decision Making* 13:95.
- Noordam, A.C., Kuepper, B.M., Stekelenburg, J., Milen, A. 2011. “Improvement of maternal health services through the use of mobile phones.” *Tropical Medicine & International Health* 16(5):622–626.
- Noordam, A.C., George, A., Sharkey, A.B., Jafarli, A., Bakshi, S.S. and Kim, J.C. 2015. “Assessing scale up of mHealth innovations based on intervention complexity: Two case studies of child health programmes in Malawi and Zambia.” *Journal of Health Communication* 0:1–11.
- Rubin, D.B. 1974. “Estimating causal effect of treatments in randomized and nonrandomized studies.” *Journal of Educational Psychology* 66(5): 688-701.
- Tamrat, T., and Kachnowski, S. 2012. “Special delivery: An analysis of mHealth in maternal and newborn health programs and their outcomes around the world.” *Maternal and Child Health Journal* 16(5):1092-1101.
- Tomlinson, M., Rotheram-Borus, M.J., Swartz, L. and Tsai, A.C. “Scaling up mHealth: where is the evidence?” *PLoS Medicine* 10(2):e1001382.
- Victora, C.G., Black, R.E., Boerma, J., and Bryce, J. 2011. “Measuring impact in the Millennium Development Goal era and beyond: a new approach to large-scale effectiveness evaluations.” *The Lancet* 377(9759):85–95.
- WHO (World Health Organization). 2011. “mHealth: new horizons for health through mobile technologies: second global survey on eHealth.” Geneva.
- WHO (World Health Organization). 2010. “Atlas eHealth Country Profiles Global Observatory for eHealth series - Volume 1.” Geneva.
- World Bank & ITU (International Telecommunication Union). 2012. “The Little Data Book on Information and Communication Technology.” Washington DC.