Ernest Guevarra
Emmanuel Mandalazi
Safari Balegamire
Kristine Albrektsen
Kate Sadler
Khalid Abdelsalam
Gloria Urrea
Salma Alawad

Impact evaluation of the World Food Programme's moderate acute malnutrition treatment and prevention programmes in Sudan

July 2018

Impact Evaluation Report 79

Humanitarian

International Initiative for Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) promotes evidence-informed equitable, inclusive and sustainable development. We support the generation and effective use of high-quality evidence to inform decision-making and improve the lives of people living in poverty in low- and middle-income countries. We provide guidance and support to produce, synthesise and quality assure evidence of what works, for whom, how, why and at what cost.

3ie impact evaluations

3ie-supported impact evaluations assess the difference a development intervention has made to social and economic outcomes. 3ie is committed to funding rigorous evaluations that include a theory-based design, and use the most appropriate mix of methods to capture outcomes and are useful in complex development contexts.

About this report

3ie accepted the final version of the report, *Impact evaluation of the World Food Programme's moderate acute malnutrition treatment and prevention programmes in Sudan*, as partial fulfilment of requirements under grant TW6.1026 awarded under the Humanitarian Assistance Evidence Programme. The content has been copy-edited and formatted for publication by 3ie.

The 3ie technical quality assurance team for this report comprises Tara Kaul, Kanika Jha, an anonymous external impact evaluation design expert reviewer, and an anonymous external sector expert reviewer, with overall technical supervision by Marie Gaarder. The 3ie editorial production team for this report comprises Angel Kharya and Akarsh Gupta, with Beryl Leach providing overall editorial supervision.

All of the content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors and omissions are also the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Please direct any comments or queries to the corresponding author, Ernest Guevarra, at ernest@validinternational.org.

Research discussed in the publication is funded by the World Food Programme (Office of Evaluation) and UK aid through the Department for International Development (DFID). The views expressed in the report are not necessarily those of WFP or DFID.

Suggested citation: Guevarra, E, Mandalazi, E, Balegamire, S, Albrektsen, K, Sadler, K, Abdelsalam, K, Urrea, G and Alawad, S, 2018. *Impact evaluation of the World Food Programme's moderate acute malnutrition treatment and prevention programmes in Sudan.* 3ie Impact Evaluation Report 79. New Delhi: International Initiative for Impact Evaluation (3ie). Available at: doi: https://doi.org/10.23846/tw6IE79

Cover photo: UN Photo/Albert Gonzalez Farran

Impact evaluation of the World Food Programme's moderate acute malnutrition treatment and prevention programmes in Sudan

Ernest Guevarra
Valid International (VI)

Emmanuel Mandalazi
VI

Safari Balegamire
VI

Kristine Albrektsen
VI

Kate Sadler
VI

Khalid Abdelsalam
Independent consultant

Gloria Urrea
Independent consultant

Salma Alawad
Independent consultant

Impact Evaluation Report 79

July 2018

International Initiative for Impact Evaluation

Acknowledgements

We would like to acknowledge 3ie for funding this study and providing technical review and support throughout its development. We also wish to acknowledge the World Food Programme for commissioning this study, as well as its Sudan country office and the Kassala state office for their support in ensuring that the treatment and prevention programmes were running as per study requirements. We thank the Sudan Federal Ministry of Health and the Kassala State Ministry of Health for providing us with the operational and logistical support that enabled us to run such a complex study – particularly Ms. Duria Mohammed Osman, who was an invaluable liaison between Valid International and the Federal Ministry of Health. We also thank the mothers and children who agreed to participate in the study, as well as the study supervisors, enumerators and drivers.

Summary

Background

Acute malnutrition is one of the key drivers of child mortality in the developing world. A child suffering from Moderate Acute Malnutrition (MAM) has a three- to four-fold increased risk of dying compared to a well-nourished child. In Sudan, acute undernutrition is considered one of the most serious but least addressed health problems. Of 184 localities assessed in the 2013 Sudan national nutrition survey, 151 had a prevalence of Global Acute Malnutrition (GAM) above 10%, and 72 had a prevalence exceeding the international ‘emergency’ threshold of 15%. It is estimated that 500,000 children in Sudan (5.3%) suffer from Severe Acute Malnutrition (SAM).

In recent years there has been greater political commitment to integrated nutrition programming in Sudan. The World Food Programme (WFP), whose primary focus has been the treatment of MAM, has also tested various approaches to prevention. Knowledge gaps and questions remain, however, particularly regarding the impact of prevention interventions and the most effective programme design.

Our study aimed to assess the effectiveness of a food-based programme to prevent acute malnutrition within the framework of WFP Sudan’s Community-based Nutrition Integrated Programme. Specifically, we sought to evaluate the impact of implementing a targeted, food-based prevention of acute malnutrition programme in addition to a targeted supplementary feeding programme. We evaluated the impact by comparing the incidence and prevalence of GAM, MAM and SAM in study areas with a combined targeted supplementary feeding programme and a targeted food-based prevention programme (intervention) with study areas with only a targeted supplementary feeding programme in place (control). GAM is an internationally recognised indicator for the overall nutritional situation in a population and is the sum of all children aged 6 to 59 months with MAM and SAM. Prevalence evaluates all cases at a given time. Incidence measures new cases which occur.

Methods

We used different methods to address specific research questions.

We undertook a stepped wedge cluster controlled trial to assess the primary outcome of GAM prevalence. We staged a non-randomised rollout of a Food-Based prevention for MAM (FBMAM) alongside an existing Targeted Supplementary Feeding Programme (TSFP) in six localities in Kassala state. Specific localities switched from control (exposure to TSFP only) to intervention (exposure to TSFP and FBMAM) at two-month intervals between May and December 2016. At every ‘step’, we sampled villages in each locality (150 in all), which were selected using centric systematic area sampling. Study participants were eligible women and children.

We nested a two-arm parallel design cluster controlled study in the stepped wedge study to assess GAM incidence. We selected localities that were first exposed to FBMAM as the intervention group and then chose localities that were last to be exposed to prevention as the control. In each locality, we non-randomly selected 10 villages for

sampling. Study participants were eligible women and children. This cohort was then followed up four times over five months.

We conducted a qualitative sub-study to investigate coverage and the effects of Social and Behaviour Change Communication (SBCC) in sites across four localities selected according to programme status, available data and accessibility.

A cost-effectiveness analysis was planned but could not be completed due to lack of data. Comprehensive costing data should be collected as part of improved programme monitoring.

Similarly, it was not possible to assess the wider impact of different packages, partly as the SBCC was not fully implemented; we also acknowledge that change in behaviour takes time.

Key findings

The overarching research question of the impact evaluation focussed on the impact of different MAM treatment and prevention interventions on the incidence and prevalence of acute malnutrition in children and pregnant and lactating women. We observed no significant change in GAM, MAM or SAM prevalence or incidence between control and intervention groups, meaning that no programme impact was detected on these outcomes. However, a significant decrease (as much as 12%) in the prevalence of children (6–23 months old) at risk of malnutrition[1] was observed at rounds 2 and 3. We observed a similar pattern in pregnant and lactating women, with a non-significant decrease in PLW GAM but a significant decrease in PLW at-risk. We found no gender differences in programme performance outcomes.

Sub-research questions of this evaluation also aimed to examine the impact of performance (including coverage) of different intervention packages and the inclusion of an SBCC component on the effectiveness of MAM treatment and prevention. Performance for the TSFP met Sphere Project standards across all localities. It was not possible to fully assess Food-based Prevention of Moderate Acute Malnutrition (FBMAM) performance as data was patchy, but the limited data pointed to performance being generally good.

Coverage was fair for the TSFP with some localities reaching as high as 50% and overall coverage was up to 28%; however it was very low for FBMAM at no more than 10%. The qualitative investigation clearly identified the need for more effective case finding of MAM and at-risk children. It also highlighted issues with record keeping at the clinic level.

The SBCC actions encountered delays and contributed further to the already high workload of community mobilisers and programme staff, resulting in messages only reaching 12% of the target audience. Many of these did not then attend activities due to opportunity cost (i.e. loss of time, economic and other resources).

[1] Children 6-23 months old with a MUAC between 125mm to 135mm. This is the target group for the prevention programme.

Interpretation

Our results demonstrate that, in this context, the addition of a FBMAM programme onto a TSFP as a package intervention for the treatment and prevention of MAM, has decreased the prevalence of at-risk children but not of MAM and GAM incidence or prevalence directly.

The possible reasons for this are: (1) there is a time lag between at-risk reduction and prevalence reduction (as mediated by incidence reduction); (2) incident MAM cases may actually be previous SAM cases discharged from treatment; and (3) the level of coverage of the FBMAM programme is too low to support any change in prevalence or incidence at a population level.

Our results indicate the potential contribution of FBMAM in the reduction of moderate acute malnutrition prevalence via a decrease in the prevalence of children ‘at risk’ of acute malnutrition. We suggest further evaluation and research to provide a robust evidence base.

Although the findings did not enable us to identify the most effective intervention modality, they have highlighted the importance of quality and delivery of services in improving performance, coverage and nutritional status. We recommend improving the method and frequency of case findings; improving record keeping at clinics and the monitoring of MAM at the community level; and re-examining the relevance of messages and appropriateness of delivery platforms for SBCC actions in light of low participation levels in this programme component.

Contents

List of figures and tables

Abbreviations and acronyms

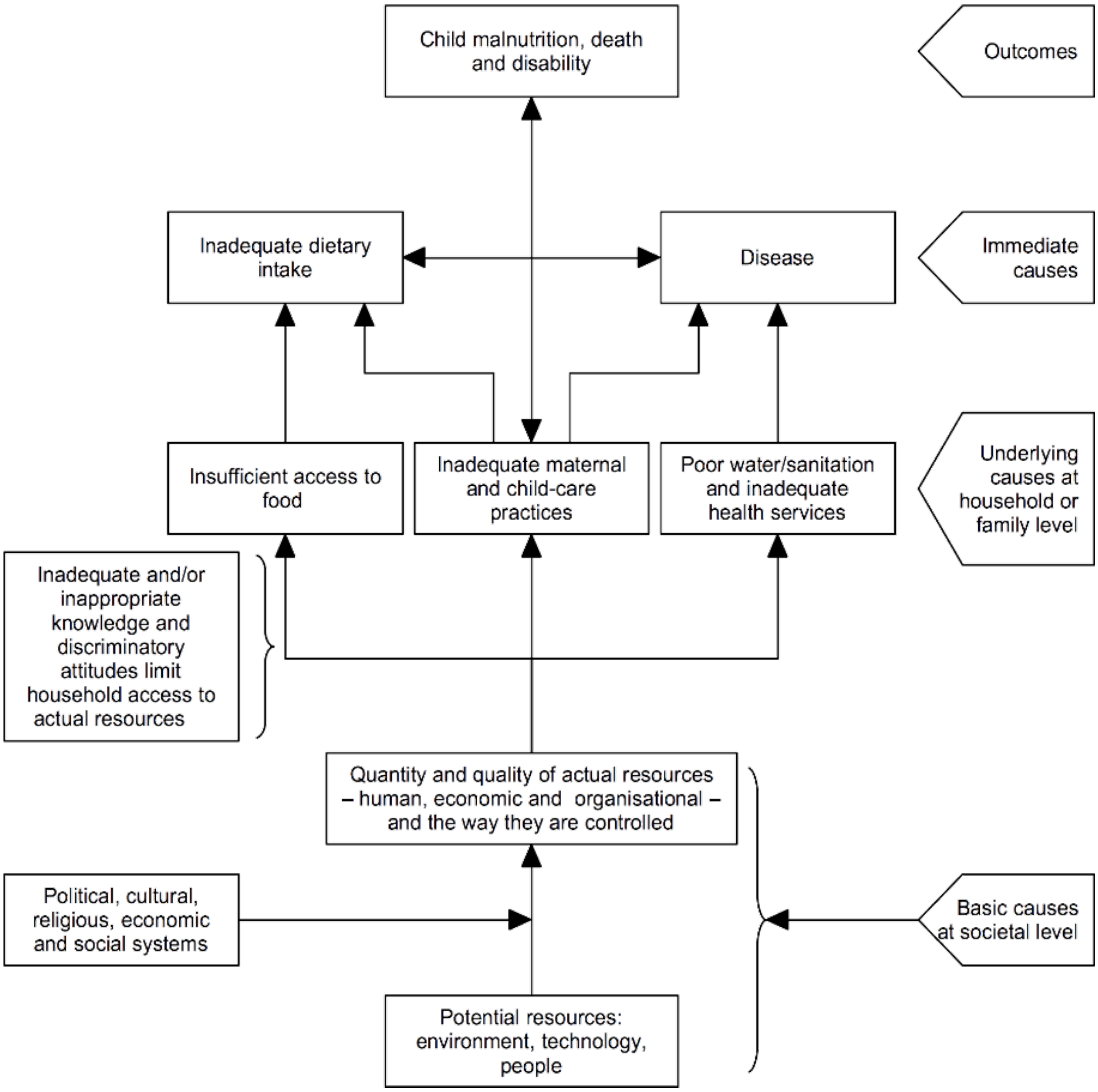
BCC	Behaviour change communication
CEA	Cost-effectiveness analysis
CMAM	Community-based management of severe acute malnutrition
CNIP	Community-based nutrition integrated programme
DALY	Disability-adjusted life year
eBSFP	Emergency blanket supplementary feeding programmes
FBMAM	Food-based prevention of moderate acute malnutrition
FMoH	Federal Ministry of Health
GAM	Global acute malnutrition
HF	Home fortification
ICFI	Infant and child feeding index
IYCF	Infant and young child feeding
LNS	Lipid-based nutrient supplement
MAM	Moderate acute malnutrition
MNP	Micronutrient powders
MUAC	Mid-upper arm circumference
NGO	Non-governmental organisation
OTP	Outpatient therapeutic programme
PLW	Pregnant and lactating women
PSU	Primary sampling unit
RUSF	Ready-to-use supplementary food
SAM	Severe acute malnutrition
SBCC	Social and behaviour change communication
SC	Super cereal
SC+	Super cereal plus
SNF	Specialised nutritious foods
SMoH	State Ministry of Health
S3M	Simple spatial survey method
TSFP	Targeted supplementary feeding programme
WASH	Water, sanitation and hygiene
WFP	United Nations World Food Programme
WDDS	Women's dietary diversity score

1. Introduction

1.1 Moderate acute malnutrition and maternal undernutrition at the global level and in Sudan

The 2008 Lancet series on maternal and child nutrition identifies acute malnutrition as one of the key drivers of child mortality in the developing world (Black et al. 2008; Bhutta et al. 2008; Victora et al. 2008; Bryce et al. 2008; Morris et al. 2008). Moderate Acute Malnutrition (MAM) is characterised by recent rapid weight loss or a failure to gain weight manifesting physically as wasting. In its more severe form (termed Severe Acute Malnutrition or SAM), wasting is more pronounced and/or bilateral oedema may be present. Acute malnutrition is the result of the interplay between myriad immediate, underlying and basic causes as described in the UNICEF causal framework of malnutrition shown in Figure 1 below.

Figure 1: UNICEF causal framework of malnutrition



Source: UNICEF (1997)

Children with acute malnutrition are at higher risk of dying than their well-nourished peers. Those with MAM are estimated to have a three- to four-fold increased mortality risk, while those with SAM have a nine-fold increased mortality risk (Black et al. 2008; James et al. 2016). Given that a greater number of children are affected globally by MAM, absolute mortality is higher for MAM than for SAM. When left untreated, children with MAM experience high rates of deterioration (MAM leading to SAM and possible death) and no improvement (James et al. 2016). MAM children are also at increased risk of disease and impaired physical and mental development (Black et al. 2013; Mucha 2014).

Together, MAM and SAM combine to measure Global Acute Malnutrition (GAM) in children. GAM prevalence is recognised as an important measure of the general nutritional status of a population and is used as an indicator of severity and a guide to the need for intervention in emergency situations.

The health and nutritional status of a mother is also an important, and related, global issue. The links between maternal undernutrition and poor prenatal and newborn outcomes such as intrauterine growth retardation, low birth weight or small for gestational age births, and pre-term birth have been well established (Zerfu and Ayele 2013; Abu-Saad and Fraser 2010; Katz et al. 2013). These adverse birth outcomes have, in turn, led to poor child health and nutrition outcomes, including an increased risk of being stunted at 24 months (Black et al. 2013), and an average two-fold increase in the risk of wasting in children 12–59 months (Jimenez and Stone-Jimenez 2014; World Health Organization 2017). Treating and preventing MAM and maternal undernutrition are therefore important in both emergency and non-emergency settings to reduce mortality and protect and improve livelihoods.

In Sudan, undernutrition is considered one of the most serious but least addressed socio-economic and health problems, and Sudan remains one of 34 countries contributing to 90% of the global burden of malnutrition. In particular, progress in reducing acute malnutrition is identified as lagging behind other African countries with trend data indicating little change since the 1980s (World Food Programme Sudan 2015; Sudan Federal Ministry of Health 2013). Out of 184 localities assessed in the 2013 Sudan national nutrition survey, 151 had a prevalence of GAM above 10% and 72 localities had a prevalence exceeding the international 'emergency' threshold of 15% (Sudan National S3M 2013). MAM prevalence was also well above 10% across large areas of the country. In addition, it is estimated that 500,000 children (5.3%) suffer from SAM in Sudan and that 2 million children are stunted. Maternal undernutrition is also very high, reaching 62% in some localities (World Food Programme Sudan 2015; Sudan National S3M 2013).

The specific multiple factors which lead to undernutrition in Sudan are known. These are related to low health status and health service coverage, water availability, food insecurity (including low production and fluctuating food prices), periodic humanitarian crises, limited safety nets, food taboos, and poor educational and economic performance at all levels. The impact of sector interventions has been limited due to a lack of reliable data to inform decision making, fragmented policies, lack of coordination between nutrition-relevant sectors and government and non-government initiatives, poor programme management, and a lack of capacity and resources. Moreover, static trends have been further reinforced by the fact that most donor funding for nutrition in Sudan is for short-term humanitarian action with limited effect on long-term outcomes (Sudan Federal Ministry of Health 2013; Sudan National S3M 2013).

In the last decade, however, there has been a growing awareness within Sudan of the central role that nutrition plays in public health and economic prosperity, which has led to greater political commitment to integrated nutrition programming. A range of guidelines and protocols have been developed, including Community-based Management of SAM (CMAM), hospital management of SAM, the Essential Nutrition Package,[2] and Infant and Young Child Feeding (IYCF). Although management of MAM is not included in the minimum Primary Healthcare Package due to its high cost, there has been an increasing focus on moderate malnutrition. This is evident in the National Nutrition Strategic Plan 2014–2018, which includes objectives to promote the management and prevention of the condition (Jimenez and Stone-Jimenez 2014; Sudan Federal Ministry of Health 2013; Sudan Federal Ministry of Health 2014).

Table 1: Nutrition programmes currently endorsed in Sudan and their known impacts and costs

Programme	Country context	Cost
Community-based Management of Acute Malnutrition – SAM treatment	Zambia (Bachmann 2009)	\$1760 per life saved \$53 per Disability-Adjusted Life Year (DALY) averted
Community-based Management of Acute Malnutrition – SAM treatment	Malawi (Wilford et al. 2012)	\$42 per DALY averted
Community-based Management of Acute Malnutrition – SAM treatment	Bangladesh (Puett et al. 2013)	\$26 per DALY averted
Community-based Management of Acute Malnutrition – SAM treatment	Nigeria (Frankel et al. 2015)	\$30 per DALY averted
Hospital management of SAM	Ethiopia (Tekeste et al. 2012)	\$284.56 mean cost per child treated \$21.01 opportunity cost per caretaker
Infant and young child feeding interventions[3]	Global estimates (Bhutta et al. 2013)	\$175 per life-year saved

1.2 Current MAM programming and knowledge gaps

The World Food Programme (WFP) has been operational in Sudan since 1963 and is currently the largest humanitarian actor in the country with over 4 million beneficiaries. Over the last two decades, WFP's response to continued, periodic nutrition crises in

[2] The Essential Nutrition Package includes promotion of maternal nutrition and child spacing; promotion of IYCF including optimal breastfeeding and complementary feeding practices; growth monitoring and health education with referral of SAM and MAM cases; control of micronutrient deficiencies and promotion of immunisation, family nutrition, dietary diversification and optimal hygiene and sanitation (see Sudan Federal Ministry of Health 2014).

[3] Includes interventions that promote early and exclusive breastfeeding for 6 months and continued breastfeeding for up to 24 months, as well as appropriate complementary feeding education in food secure populations and additional complementary food supplements in food insecure populations (see Bhutta et al. 2013).

Sudan has focused on the treatment of MAM (World Food Programme Sudan 2015) through the delivery of Specialised Nutritious Foods (SNF). In recent years, however, an increasing number of different approaches have been tested for the prevention of MAM.

In 2008, a MAM prevention intervention provided blanket SNF to all children from 6–59 months during the lean season in Darfur. While targeted treatment programmes met Sphere standards for all performance indicators across Sudan, blanket prevention underperformed and called into question the efficiency of this approach in reducing rates of acute malnutrition (Acharya and Kenefick 2012). In contrast, a 2010–2011 pilot project in Kassala that employed blanket SNF with increased community mobilisation and Behaviour Change Communication (BCC) resulted in a reduction in acute malnutrition prevalence (Acharya and Kenefick 2012), and led to the Food-Based prevention of MAM (FBMAM) programme, which focused on the first 1000 days of life. Nevertheless, questions remain regarding the impact and the most effective design of MAM programming.

Globally, while there is considerable evidence for the effectiveness of MAM treatment interventions in optimal conditions, there is insufficient and equivocal understanding of the variations in effectiveness in crisis and post-crisis conditions, especially with regard to impact differences of MAM treatment programmes when prevention interventions are also present. The nature of the two interventions differs (thus potentially complicating evaluability), but their interaction seems crucial. Children recovering after MAM treatment are at risk of reverting when there is no MAM prevention. Thus, there are a number of issues to consider such as prevention versus treatment, when to shift from one to the other, the cost-effectiveness of different approaches, the heterogeneity of impact based on sex, and how context affects the most appropriate way to deal with the issue.

A number of knowledge gaps in relation to MAM programming have been detailed latterly in reviews and international fora. Most recently an international review of MAM management found that, despite the associated mortality risk and the large global caseload for MAM, the low profile of MAM treatment was linked to the high unit cost of products currently used, low coverage and frequent defaulting experienced by programmes, and a focus on tackling generalised prevalence rates rather than season-specific incidence rates (Annan et al. 2014). The review also identified a lack of consensus around definitions, criteria and treatment protocols, an overfocus on products rather than implementation modalities, and a number of key evidence gaps, specifically: (1) the effectiveness of nutrition counselling either as a separate intervention or combined with provision of food products and medical intervention; (2) the cost-effectiveness of interventions targeting MAM; (3) appropriate comparisons of the different products used for addressing MAM; (4) studies on MAM in contexts other than Africa; and (5) how products are actually used by the beneficiaries within a community-based intervention, and how this usage affects outcomes (Annan et al. 2014). Overall the review echoed the conclusions of the 2014 International Symposium in Vienna on the subject of MAM: ‘There is a clear need for more and better-quality scientific evidence on effective interventions for the management of MAM in children’ (Becic et al. 2014, p.92).

Equally, as it becomes a feature of WFP’s international MAM strategy, MAM prevention is attracting more attention globally and raising questions about the specific nature of prevention programming. These include the effectiveness of food-based versus non-food-based approaches and how these differ from stunting reduction programmes, given that, in many contexts, high levels of stunting and wasting go hand in hand. The Vienna symposium concluded that 'the prevention of MAM should encompass both nutrition-specific and nutrition-sensitive interventions' (Becic et al. 2014, p.91). This was followed by two corresponding review papers, which highlighted limited rigorous evidence for the efficacy of different nutrition-sensitive and nutrition-specific approaches for MAM prevention and for the plausibility of replicability and scale-up (Mucha 2014; Jimenez and Stone-Jimenez 2014). In a workshop held in October 2015 during the development phase of the current study, these concerns were echoed by WFP Sudan and other key stakeholders, who identified the following knowledge gaps for existing MAM prevention programmes in Sudan:

- Appropriate admission criteria;
- The most impactful aspects of WFP interventions, including the effect of Social and Behaviour Change Communication (SBCC) and components for Pregnant and Lactating Women (PLW), and the integration of nutrition-sensitive programming such as Water, Sanitation and Hygiene (WASH);
- The most effective delivery of MAM prevention (community vs. clinic);
- Existing barriers to programme uptake at the community level and strategies to overcome these;
- The cost-effectiveness of different programme modalities and food products; and
- The impact of prevention programming on MAM incidence and prevalence.

1.3 Evaluation aims

This impact evaluation sought to address some of the key knowledge gaps listed above. The evaluation aimed to assess the effectiveness of WFP's MAM treatment programmes and also provide insight into the effect of additional prevention elements on MAM incidence and prevalence. Ultimately, effectiveness must be measured in relation to outcomes; therefore the study looked at the impact of interventions on rates of MAM, SAM, GAM and children at risk of malnutrition (see Section 2.2. on Theory of Change for more detail on these outcome measures).

The evaluation also aimed to answer questions on how to target and prioritise MAM treatment and prevention interventions and their specific components and how to identify the implementation modality that performed best in terms of intended results. In addition, the evaluation sought to gain insight into the impact that WFP MAM interventions had on the uptake of linked services (e.g. SBCC).

WFP Sudan is implementing a range of nutrition activities through the Community-based Nutrition Integrated Programme (CNIP) approach, which provides services to address key causes of acute malnutrition via single-service delivery platforms. CNIP places MAM treatment, MAM prevention, Home-based Fortification (HF) and SBCC components within a broader range of nutrition-specific and nutrition-sensitive interventions (World Food Programme 2015). Experiences from CNIP implementation will feed into the WFP Interim Country Strategic Plan, which is currently under development for action from July 2017. The evaluation therefore aimed to provide WFP Sudan with a better understanding of the quality, impact and cost effectiveness of their programmes and to inform decisions about continuation, scale-up and design.

WFP Sudan also works closely with the Sudan Federal Ministry of Health (FMoH) in the development and review of national health and nutrition policies. As mentioned previously, although political commitment to MAM management is increasing, perceived high costs have hampered the inclusion of interventions in the Primary Healthcare Package. The National Nutrition Strategic Plan will be revised in 2018 with key input and guidance from WFP. Findings and concrete data on the impact of MAM treatment and prevention programmes could therefore feed into future country-level nutrition policies and strategies. Additionally, with WFP assigning greater priority to MAM programming at the global level, the evaluation findings would come at the right time to influence internal strategic decision making.

1.4 Evaluation questions

The overarching research question of the impact evaluation was as follows:

> *What is the impact of different MAM treatment and prevention interventions on the incidence and prevalence of MAM and SAM in children under 5 years of age and PLW in Sudan (i.e. Targeted Supplementary Feeding Programmes [TSFP] for the treatment of MAM; targeted FBMAM; Emergency Blanket Supplementary Feeding Programmes [eBSFP] as rapid response to crises for the prevention of MAM; HF for the prevention of MAM; and SBCC for the prevention of MAM)?*

To answer this, the following sub-research questions were addressed:

- How are these impacts affected by intervention modality in terms of product used, delivery of service, duration of intervention and coverage?
- What is the impact of FBMAM on the effectiveness of MAM treatment (performance and coverage)?
- How timely and effective is are eBSFPs?
- How does the inclusion of an SBCC component impact the effectiveness of MAM treatment and prevention and what factors influence this?
- How appropriate are geographical and individual targeting criteria for each intervention?

A further two sub-questions were formulated at the planning stage of the study (however, due to a mixture of missing data and constraints on time and resources, comprehensive analyses were not feasible):

- What is the cost effectiveness of the different packages from a WFP perspective?
- What are the wider impacts, positive or negative, of the packages at household, community or institutional level (opportunity, social, economic, environmental)?

1.5 Evaluation strategy

We used a mixed methods approach to answer the research questions. Specifically, we designed a quantitative, quasi-experimental study that would provide factual and counterfactual data to assess programme impact. Additional data produced from routine monitoring were collected and analysed, and standard participatory methods were employed as part of a qualitative investigation. The investigation sought to provide contextualised and nuanced information regarding each programme, its implementation, and how it was perceived by the intended beneficiaries, thus adding another layer of corroborating evidence to explain the impact observed.

To answer the research questions related to outcomes on prevalence and coverage, we applied a stepped-wedge cluster controlled trial design. The design allowed for intra-cluster and inter-cluster controlled comparison both at specific time points of data collection and at each cluster's point of crossover from exposure to TSFP only (control), to exposure to both TSFP and FBMAM (intervention).

We nested a two-arm parallel design cluster controlled trial into the stepped-wedge study to assess and compare the incidence of acute malnutrition in areas with both MAM treatment and prevention (intervention), and with MAM treatment only (control).

We chose this survey design as the stepped-wedge approach allowed us to collect factual and counterfactual information in a study setting where the intervention programme to be evaluated was to be implemented widely and scaled up over time. The withholding of an intervention designed for wide implementation has ethical and political ramifications; hence we deemed a typical parallel study design to be contentious in a setting such as Sudan. A stepped-wedge design addresses these issues.

We applied an activity-based costing methodology and ingredient approach, accounting for all programme (provider) costs and key household/community (beneficiary) inputs/costs, with the intention of measuring the cost effectiveness of different intervention components and programme modalities.

Appendix D presents the pre-analysis plan we developed, including questions, in line with the above-mentioned evaluation strategy. This report provides information and results that address and discuss the majority of the pre-analysis points of inquiry that we set out to resolve. However, as indicated above, we have been unable to supply adequate responses to the cost-effectiveness analysis questions raised in the pre-analysis plan. Whilst provider cost data and budgets were provided to the evaluation team, it was not possible for the WFP country office team to provide data disaggregated by intervention package. Specifically, we have been unable to report on any of the four principal cost-effectiveness analysis metrics, namely: (1) cost per case of MAM averted in a child under five; (2) cost per case of PLW undernutrition averted; (3) cost per DALY averted; and (4) incremental cost effectiveness ratio. Therefore, rather than providing a comprehensive cost effectiveness analysis of WFP's MAM treatment and prevention programmes, the report outlines recommendations for securing necessary cost data through improved programme management and monitoring. It is clear that future evaluations that include a Cost-effectiveness analysis (CEA) component need to adopt mechanisms to better ensure that the format of data collected throughout the programme cycle is well-adapted to the data needs for this type of analysis.

1.6 Report outline

In the succeeding sections of this report, we first describe the intervention evaluated, its theory of change and our research hypotheses (Section 2). We then situate the intervention within the context of Sudan, specifically Kassala state, where we conducted the study (Section 3). This is followed by the timeline of events which occurred alongside the implementation of the programme and the evaluation (Section 4). We go on to describe the study design and the associated methods and their execution (Section 5), followed by our analytical approach and the results we obtained from this analysis (Section 6). We then discuss these results with regard to their internal and external validity and other factors which may explain what we observed (Section 7). Finally, we highlight and present specific findings with repercussions for policy and practice relevant to the treatment and prevention of MAM (Section 8).

2. Intervention, theory of change and research hypotheses

2.1 WFP MAM treatment and prevention programming

WFP MAM treatment and prevention programmes are implemented within the broader framework of CNIP and linked to other nutrition-specific and nutrition-sensitive interventions, including SAM treatment, IYCF, resilience, livelihoods and WASH.[4] MAM treatment and prevention programmes are delivered through a mixture of partners including government and international and national Non-Governmental Organisations (NGOs). The identification of beneficiaries is carried out by community workers via community-level screening and referrals. Key MAM components of the CNIP include:

2.1.1 MAM treatment

TSFP

This component includes distribution of SNF (1000kcal/d), instructions/key messages on use, provision of routine medicines, monitoring, practical preparation and cooking demonstrations, and referral to the SBCC component.

Targeted at:

- Children between 6–59 months with a Mid-Upper Arm Circumference (MUAC) greater than or equal to 115mm, but less than 125mm, and with no oedema;
- Children discharged from an outpatient therapeutic programme (OTP) for SAM treatment; and
- Pregnant (second or third semester) and lactating women (with infant less than 6 months old) with a MUAC greater than or equal to 185mm and less than 210mm.

Objectives:

- Prevent morbidity and mortality associated with MAM;
- Prevent targeted individuals with MAM from developing SAM;
- Prevent relapse of individuals who have been treated for SAM; and
- Improve maternal nutritional status.

Performance standards are based on children under five years and defined as:

- a cure rate greater than 75%
- a death rate less than 3%
- a default rate less than 15%
- a non-cured rate less than 15%

[4] Community ownership and empowerment is central to CNIP. Programme components are designed according to a community-based participatory planning approach, entailing a consultative process during which communities, WFP implementing partners and local government staff discuss and agree priority activities required to build food and nutrition security, address vulnerabilities and enhance community resilience.

2.1.2 MAM prevention

FBMAM

a) Blanket FBMAM: This component entails blanket distribution of SNF (500 Kcal/day) in localities with a GAM rate above 20%, instructions/key messages on use, provision of routine medicines, monitoring, practical preparation and cooking demonstrations, and referral to the SBCC component.

   Targeted at:
   - Children 6–23 months (height 65–87cm if birth date not known); and
   - PLW.

b) FBMAM: This component has a targeted distribution of SNF (500 Kcal/day) in localities with GAM rate of less than 20%, instructions/key messages on how to use, provision of routine medicines, monitoring, practical preparation and cooking demonstrations, and referral to the SBCC component.

   Targeted at:
   - Children 6–23 months with a MUAC greater than or equal to 125mm and less than 135mm[5]; and
   - PLW with a MUAC greater than or equal to 210mm and less than 230mm.

   Objectives:
   - Prevent acute malnutrition in children 6–23 months and in PLW;
   - Sensitise care-takers on inappropriate feeding, health, and hygiene practices and assist them to adopt optimal feeding practices;
   - Screen, identify and refer those requiring treatment for acute malnutrition;
   - Increase access and participation in child survival interventions such as de-worming, vitamin A supplementation, immunisation and/or measles vaccination campaigns.

HF

This component entails the distribution of micronutrient powders (MNP) consisting of 180 sachets per year (15 sachets per month) of product aligned to international standards, using a one-gram single dose MNP for point-of-use (post cooking) fortification, as well as monitoring and referral to the SBCC component.

Targeted at:
- All children 6–59 months with a MUAC greater than 13.5mm who are not eligible for TSFP or FBMAM, and who do not have oedema.

Objectives:
- Increase the intake of micronutrients in children aged 6–59 months;
- Bring about positive change in feeding practices and food hygiene, especially for infants and young children.

[5] This category of children has been defined as 'at risk' as they are within the age group most susceptible to acute malnutrition (under two years or within their first 1000 days of life) and have a MUAC close to the cut-off for MAM.

SBCC for improved health, IYCF and WASH
This component entails the delivery of information, key messages and consultation on health, IYCF and WASH through the establishment or use of existing communication platforms. These include health centres, nutrition programmes, mother and care groups, home visits, key informants, and folk and traditional media (e.g. local theatre, events, music, dance, skits, parades, storytelling, festivals, health fairs, puppet shows, loud speakers, fliers, etc.).

Implementing partners were trained on SBCC message content and delivering and communicating these messages effectively at the community level. Minimum-level messages were based on standard WFP SBCC materials (developed prior to this programme) that cover the following topics:

SBCC Messages:

a) Health
- Malaria prevention (impregnated mosquito nets)
- Immunisation, deworming and vitamin A supplementation
- Antenatal and postnatal care

b) IYCF
- Early initiation of breastfeeding within one hour of birth
- Exclusive breastfeeding up to six months
- Continued breastfeeding to at least two years
- Complementary feeding with nutrient- and energy-dense foods from six months, 3–5 times a day, in addition to breastfeeding
- Appropriate nutritional care for the sick and acutely malnourished children
- Adequate vitamin A intake for mothers and young children
- Adequate iron intake for mothers and young children
- Purchase and use of iodised salt
- Adequate use of micronutrient powders

c) WASH
- Wash hands with clean water and soap after defecation and prior to feeding children under three years;
- Safely dispose of children's faeces;
- Use an improved source for drinking water; if the source is not safe, treat or boil drinking water.

Targeted at:
- Mothers with children under five years (primary target group);
- PLW (primary target group);
- Husbands (identified as key influencers of primary target group);
- Community leaders (identified as key influencers of primary target group);
- Traditional birth attendants (identified as key influencers of primary target group);
- Religious leaders (identified as key influencers of primary target group);
- Traditional healers (identified as key influencers of primary target group).

Objectives:

- Sensitise care-takers on inappropriate feeding, health and hygiene practices and assist them to adopt optimal practices;
- Increase access and participation in child survival interventions such as de-worming, vitamin A supplementation, immunisation and/or measles vaccination campaigns.

Performance indicators are defined as:

- Proportion of mothers who initiate breastfeeding within one hour of birth;
- Proportion of mothers who exclusively breastfeed their infants up to six months;
- Proportion of mothers who continue breastfeeding up to at least two years;
- Proportion of children fed 3–5 times a day in addition to breastfeeding;
- Proportion of mothers who wash hands with soap at critical times;
- Proportion of households using an improved drinking water source;
- Proportion of households using an improved excreta disposal facility (improved toilet facility, not a shared latrine).

2.2 Theory of change

The CNIP's theory of change[6] is presented graphically in Figure 2. We based this on UNICEF's conceptual framework of the causes of undernutrition, as shown in Figure 1 (United Nations Children's Fund 1997). The FBMAM aims to impact insufficient access to food (cause 1: highlighted in green) by providing nutrient-dense food supplements to children who are considered greatly at risk, i.e. younger children (less than 24 months old) and with MUAC between 125mm and 135mm. Other complementary interventions, such as the SBCC, aim to impact on inadequate and inappropriate knowledge and discriminatory attitudes, which limit household access to actual resources (cause 2: highlighted in green). Changes brought about by the programme in relation to causes 1 and 2 are thought to act on a specific pathway through this framework (as outlined in green in Figure 2B), leading to the primary outcome of a decrease in child undernutrition and, in the case of the programme, decreased acute undernutrition.

The underlying assumptions of the MAM components of the CNIP are that targeted MAM prevention, aligned with the 'window of opportunity',[7] 'can mitigate the increase in MAM and associated risks related to mortality, morbidity and overall child development' (WFP 2015); and a participatory approach involving community consultation, sensitisation and mobilisation throughout the programme cycle will improve coverage and the resulting impact of MAM prevention and treatment programming (World Food Programme 2015). The primary research hypothesis of this impact evaluation was thus as follows:

> *The implementation of FBMAM, targeted at children 6–23 months and PLW, including SBCC components to improve IYCF and WASH behaviour in parallel with MAM and SAM treatment programmes, significantly lowers the incidence and prevalence of MAM in children under five years and in PLW over the course of the programme.*

[6] A theory of change document is still being developed by WFP; therefore, these figures have been established based on the CNIP field guide (WFP, 2015. *Community Nutrition Integrated Programme, A field guide for WFP supported Nutrition Projects in Sudan*. Version: 02.06.2015) and discussions with the WFP team.
[7] Defined as the first 1000 days from conception through a child's second year of life.

MAM is the focus of WFP's prevention and treatment interventions, but it cannot be evaluated in isolation. MAM and SAM programmes run alongside and feed into each other. Although MAM is the major component, SAM also contributes to GAM. GAM in turn measures the severity and overall nutritional status of children and is a key indicator in national nutrition surveys identifying problem localities. The primary outcomes and impacts of interest for this impact evaluation were therefore:

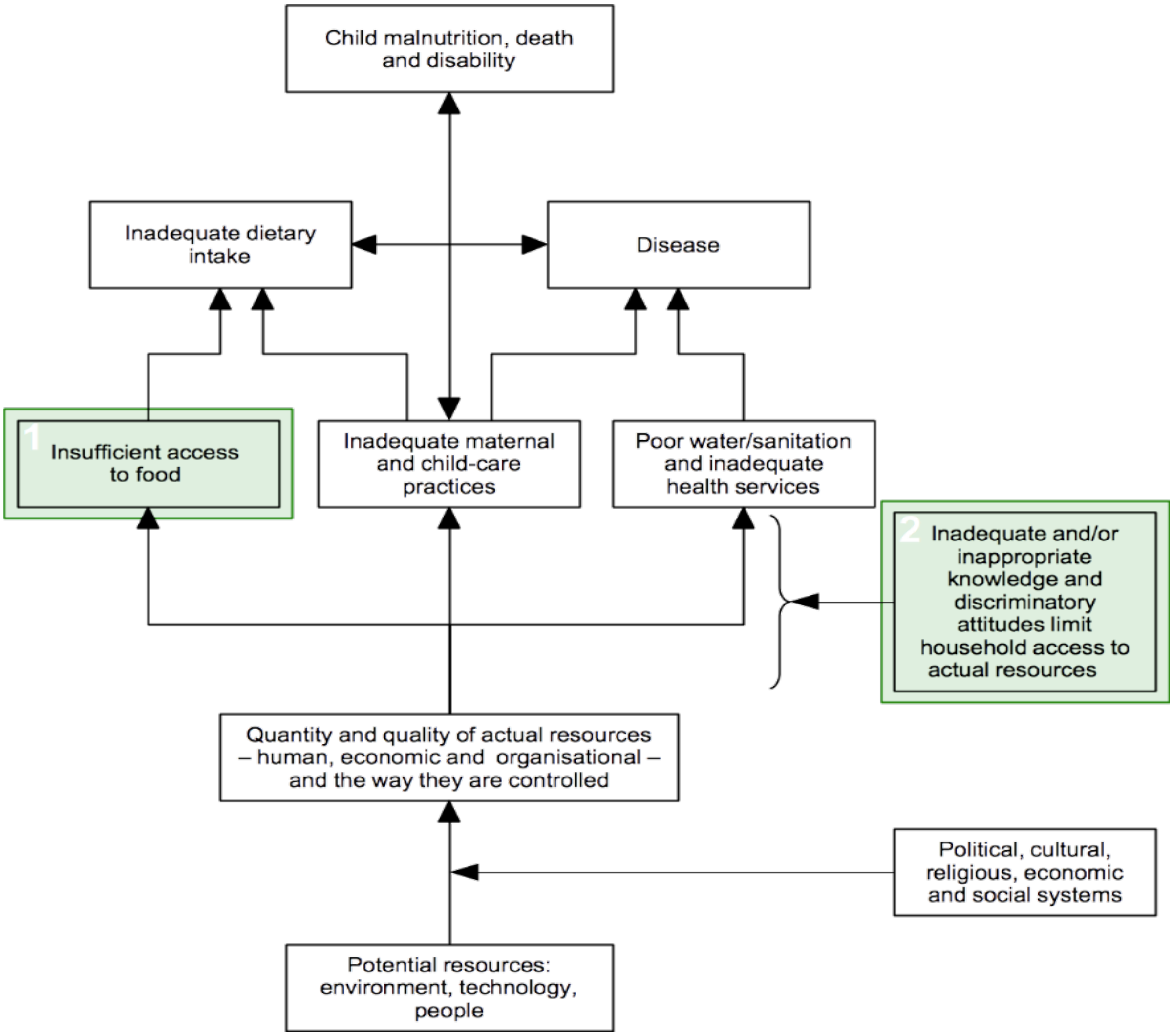
- GAM prevalence; and
- GAM incidence.

Also:

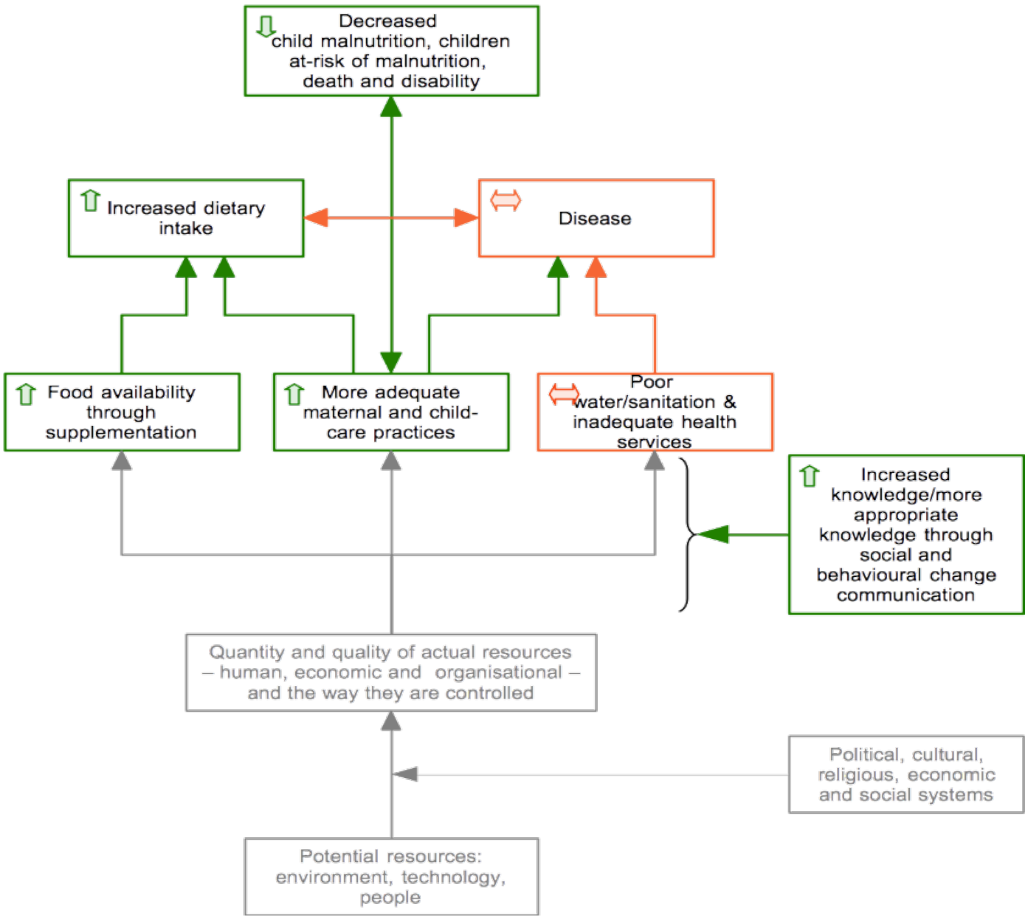
- MAM prevalence and incidence;
- SAM prevalence and incidence; and
- At-risk prevalence.

Figure 2: FBMAM programme's theory of change

A. Conceptual framework of the causes of malnutrition with causes acted on by FBMAM programme highlighted in green.



B. Theoretical pathway of change effected by FBMAM



Outcome indictors were defined as:

- GAM (encompasses cases of both MAM and SAM): MUAC less than 125mm or bilateral pitting oedema in children between 6 and 59 months old, and MUAC greater than or equal to 185mm in pregnant (2nd or 3rd semester) and lactating women (with an infant under 6 months);
- MAM: MUAC between 115mm and 125mm without bilateral pitting oedema in children between 6 and 59 months old, and MUAC greater than or equal to185mm and less than 210mm in PLW (with infant under 6 months);
- SAM: MUAC less than 115mm and/or bilateral pitting oedema in children between 6 and 59 months old, and MUAC greater than or equal to 210mm in PLW (with infant under 6 months);
- GAM prevalence was defined as the actual proportion of children or PLW with malnutrition (MAM and SAM combined) in the surveyed clusters at the time of each data collection. Similarly, MAM prevalence was the proportion of children or PLW with MAM at each round. Measuring prevalence would indicate how widespread malnutrition was, as it includes both new and existing cases;
- GAM incidence was defined as the number of new cases of malnutrition (MAM and SAM) that occurred over the course of the study. Again, MAM incidence would assess the number of new cases that manifested during the survey period. Incidence therefore would reveal the risk of children or PLW in relation to developing malnutrition;
- ‘At risk’ itself was important to quantify, and the FBMAM aims to target those directly ‘at risk’ of developing MAM but who are not yet cases. This indicator was

an outcome that needed to be measured alongside MAM, SAM and GAM in order to understand the malnutrition continuum and better assess the effect of combining prevention with treatment interventions. The CNIP Field Guide defines ‘at-risk’ as children aged 6–23 months with MUAC greater than or equal to 125mm and less than 135mm and PLW with MUAC greater than or equal to 210mm and less than 230mm.

Secondary outcomes that we considered as relevant covariates of wasting prevalence and incidence were included in the impact evaluation. These were:

- Programme coverage, defined as the proportion of the eligible target group (i.e. children and PLW meeting CNIP criteria as described above) receiving intervention (i.e. CNIP components) for the following indicators:
  - MAM case-finding effectiveness
  - MAM treatment (TSFP)
  - Targeted MAM Prevention (FBMAM)
  - HF
  - SBCC messages
  - Mother groups
  - Care groups
- Programme performance indicators:
  Defined for TSFP as:
  - Cured
  - Defaulted
  - Non-responder
  - Death
  - Referred to OTP

  Defined for eBSFP and FBMAM as:
  - Graduated
  - Defaulted
  - Death
  - Transferred to OTP/SC/TSFP
- Morbidity defined as period prevalence of most common childhood illnesses
- Cost effectiveness defined by the following metrics:
  - Cost per child treated
  - Cost per child cured
  - Cost per under-five case of MAM averted
  - Cost per PLW case of undernutrition averted
  - Cost per DALY averted
  - Incremental cost effectiveness ratio
- Knowledge, attitudes and practices, including
  - Healthy pregnancy
  - Child health and healthcare
  - IYCF
  - Women’s dietary diversity
  - Food supplementation
  - MNP
  - WASH

3. Context

3.1 Selection of study site and local context

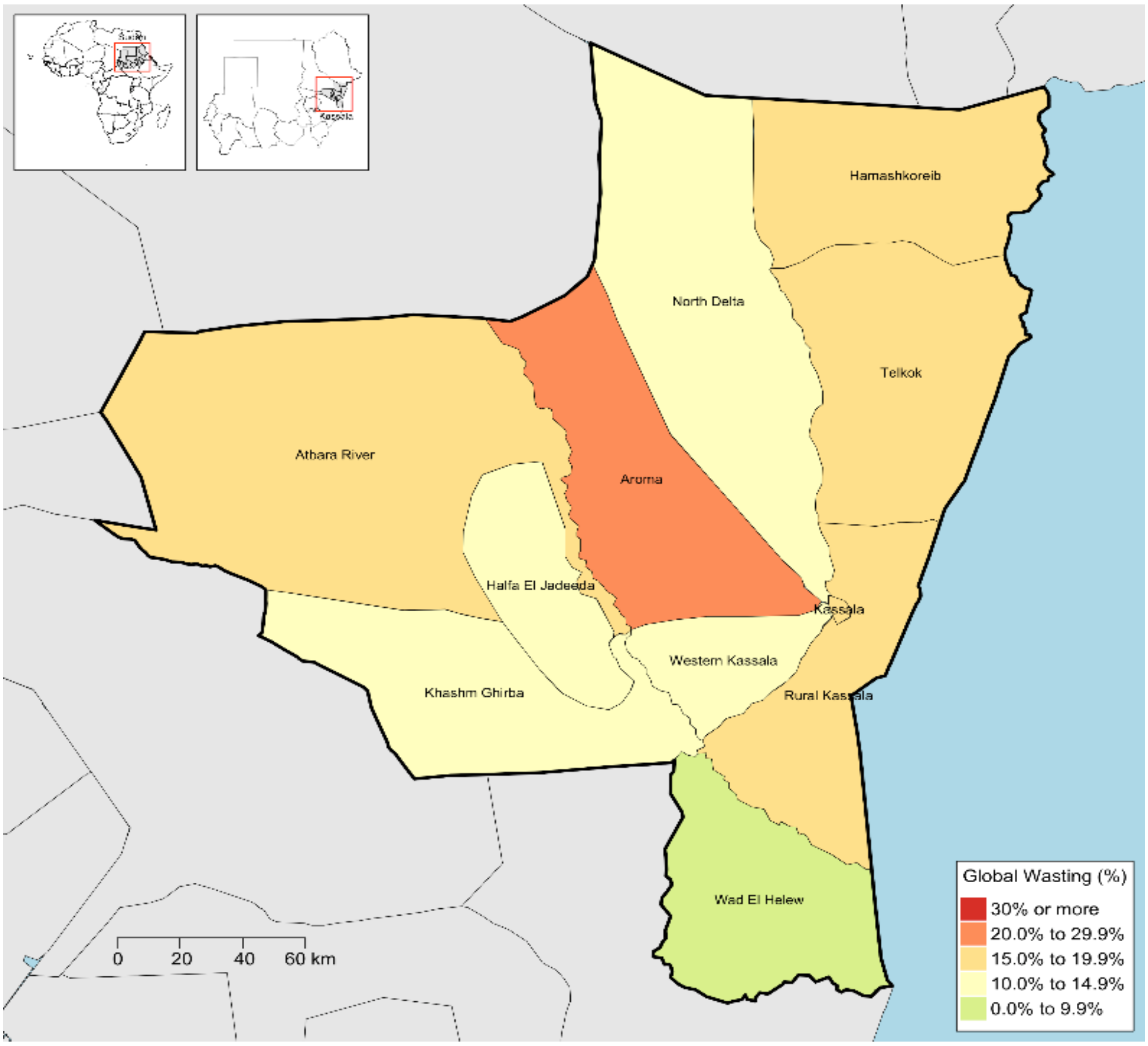
We selected Kassala state as the study site based on the WFP CNIP roll out plan, which initiated programmes in Kassala and Red Sea in early 2015. WFP were already operational in Kassala with well-established MAM treatment programmes and implementing partner agreements with the State Ministry of Health (SMoH) and national and international NGOs. This provided a good starting point for assessing the impact of adding a prevention package to MAM treatment with a range of different implementation modalities. Moreover, Kassala houses all of WFP's key target groups, the only exception being camps for internally displaced persons.

We also considered the fairly stable security situation in Kassala and the logistical needs of the impact evaluation, as reliable and frequent access to study localities was key to completing intensive data collection (see Section 5). States in which internally displaced person camps would be represented were considered too insecure to guarantee sufficient access for data collection by WFP Sudan. Kassala's overrepresentation in evaluations in Sudan was also considered a potential risk to the recruitment of study subjects. However, weighed against the drawbacks of other potential study locations, this concern was deemed minimal.

Kassala state is located in the eastern part of Sudan. It covers an area of 55,374 km,[2] and is split into 11 primarily rural localities. Kassala state is multi-ethnic and has a complex pattern of settlement as a result of various ethnic tribes settling in the area at different periods, bringing with them different histories and ways of life (Miller 2005). The predominantly sedentary population is estimated to total 1.8 million, with an annual growth rate of 2.5% (Central Bureau of Statistics 2008). Kassala has long suffered chronic poverty and has had acute undernutrition rates that are among the highest in Sudan, similar to those seen in Darfur, ranging from 15-19% over the past 10 years (Acharya and Kenefick 2012). In 2013, five of Kassala's localities had acute undernutrition rates of 15% or greater (Sudan National S3M 2013), as shown in Figure 3. These levels of acute undernutrition are associated with high food insecurity due to inadequate crop harvests frequently caused by droughts and floods; inadequate infrastructure; poor distribution of qualified human resources, which complicates interventions at scale; and cultural practices that undermine the nutritional status of children and women (World Food Programme 2012a).

A large influx of internally displaced persons (62,000) and conflict- and famine-induced refugees primarily from Eritrea and Ethiopia (75,000 in established camps) have put additional pressure on a fragile resource base (World Food Programme 2012a). Over the past decade Kassala has seen an increase in attention from development actors, and both national and international NGOs and civil society organisations operate in the state (Japan International Cooperation Agency et al. 2015).

Figure 3: Map of global acute malnutrition in Kassala, 2013



Source: Sudan National S3M (2013)

3.2 External validity

Given our study design, our results represent what can be expected from a combined targeted treatment and targeted prevention of acute malnutrition programme such as the one in place in Kassala state. Therefore, the findings and subsequent discussions and recommendations will likely only be relevant to a programme implemented in Kassala. Wider external application of this study's findings should be done carefully and should take into consideration the various contextual factors that make this programme difficult to compare with others, namely:

- Chronically high rates of acute and chronic malnutrition;
- Poor socio-economic situation within Kassala and throughout Sudan;
- Periodic crises linked to natural disasters or insecurity;
- Operational challenges in programme implementation brought about by state-level and country-level socio-political structures affecting supply chains, logistics and finance systems; and
- A still-evolving CNIP which is continually being adjusted and organised to be relevant to state- and country-level contexts.

As such, the study and its findings will likely not be transferrable to contexts with transitory food insecurity linked to seasonal or other fluctuating factors, or to settings with a highly mature and evolved community-based targeted feeding programme.

4. Timeline

The study implementation ran from the beginning of May to the end of December 2016.

The following table presents a timeline of events during the study period. These include programmatic, study implementation and other external events relevant to the evaluation.

Table 2: Timeline of events during the study period (May to December 2016)

Event	Location	May 2016				June 2016				July 2016				August 2016				September 2016				October 2016				November 2016				December 2016			
		W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4
Initiate FBMAM	Aroma																																
	Telkuk																																
	El Girba																																
	River Atbara																																
	Kassala City																																
	Rural Kassala																																
Blanket FBMAM	Aroma																																
	Telkuk																																
	El Girba																																
	River Atbara																																
	Kassala City																																
	Rural Kassala																																
Shift of product from Super Cereal+ to Plumpy Doz	Aroma																																
	Telkuk																																
	El Girba																																
	River Atbara																																
	Kassala City																																
	Rural Kassala																																
TSFP	Aroma																																
	Telkuk																																
	El Girba																																
	River Atbara																																
	Kassala City																																
	Rural Kassala																																
Home fortification	Aroma																																
	Telkuk																																
	El Girba																																

Event	Location	May 2016				June 2016				July 2016				August 2016				September 2016				October 2016				November 2016				December 2016			
		W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4
	River Atbara																																
	Kassala City																																
	Rural Kassala																																
Blanket home fortification	Aroma																																
	Telkuk																																
	El Girba																																
	River Atbara																																
	Kassala City																																
	Rural Kassala																																
Stockout (TSFP)	Aroma																																
	Telkuk																																
	El Girba																																
	River Atbara																																
	Kassala City																																
	Rural Kassala																																
Heavy rains and flooding	Aroma																																
	Telkuk																																
	El Girba																																
	River Atbara																																
	Kassala City																																
	Rural Kassala																																
Acute watery diarrhoea	Aroma (mostly around delta areas)																																
	Telkuk (mostly around delta areas)																																
	El Girba																																
	River Atbara																																
	Kassala City																																
	Rural Kassala																																

5. Evaluation: design, methods and implementation

The evaluation entailed the implementation of a stepped wedge cluster controlled trial with a nested incidence study, a qualitative impact sub-study and a cost effectiveness analysis. The following section outlines the design and data collection and analysis strategies of these studies.

5.1 Ethical review and approval

As per the 3ie principles of impact evaluation that require 'investigators to obtain ethical clearance with the appropriate institutions in conducting their studies' (International Initiative for Impact Evaluation 2016, p.2), we applied for ethical approval from the Sudan Ethical Review Board, who granted us clearance on the 21st of February 2016.[8] In order to ensure the ethical nature of the research and adhere to 3ie standards specifying the application of 'do no harm' principles in research conduct (International Initiative for Impact Evaluation 2016), the team undertook a number of measures, including:

- Research participation was based on informed, un-coerced and documented consent. An explanation was given to caregivers regarding the measurements and questionnaires and we asked permission to include each child in the study. We obtained consent from individual PLW and documented this in each questionnaire. We included training in humanitarian principles and in safeguarding children and vulnerable people for enumerators and supervisors prior to study implementation.
- Due to the controlled trial design, there was a risk that intervention could purposefully or knowingly be withheld from a specific segment of the population, subjecting it to the risk of adverse health and nutrition outcomes. However, since the study evaluated the impact of malnutrition prevention in a setting of existing malnutrition treatment (the study control status), this concern was not considered to be strictly relevant, as no treatment for MAM was withheld during the implementation of the study. Moreover, application of the stepped-wedge design ensured that intervention was not withheld during the period of study.
- The study involved thorough case-finding methods, particularly among the target groups relevant to the outcome measures being evaluated. When we identified cases that were eligible to receive intervention but not enroled, they were referred to appropriate programmes. If enumerators encountered children who were grossly and critically ill due to reasons other than malnutrition, they were referred to appropriate treatment facilities or available programmes.
- A key concern of the study was to ensure the utmost confidentiality of data/information provided by respondents. We have therefore anonymised data, and no identifying data or information for respondents has been kept unnecessarily or in a format or structure that can be traced back to its source. Some identifying data (i.e. village names, respondent names, etc.) were initially recorded for purposes of data verification, checking and cleaning, but we stripped these out once the processes had been finalised. As geospatial data was collected, this was summarised to a level where cluster location can be identified without revealing the specific location of an individual respondent. Moreover, to

[8] Endorsement number /.FMOH/DGP/RD/TC/2016

protect data, we used a local server mechanism in which a locally hosted computer received all data collected via mobile devices. Remote server-type data storage mechanisms were not used. We backed up data residing in the local server daily and stored it in two other physical locations (two portable hard drives). The local server and the portable hard drives were secured and password-protected with access only provided to the research and study team.

5.2 Evaluation design

5.2.1 Stepped-wedge cluster controlled trial

To answer the research questions related to outcomes on prevalence, coverage and morbidity, and to accommodate the different elements (FBMAM, HF, SBCC) and incremental roll-out of MAM prevention components, we applied a stepped-wedge cluster controlled trial design. The design allows for intra-cluster controlled comparison (horizontal comparison), in which each cluster is compared to itself at the start of the study and at each successive step at a two-month interval. This enables assessment of the impact of MAM prevention components when added to MAM treatment.

The design also allows inter-cluster controlled comparison (vertical comparison) between a number of clusters, with each cluster serving as a control at varying stages of the evaluation study. This enables the assessment of how the addition of varying combinations of MAM prevention components impact MAM treatment. It also allows for secular trend[9] to be accounted for in the horizontal comparison. Both horizontal and vertical analyses provide the necessary information to model the effects of time on the effectiveness of the MAM prevention packages, in terms of both when the intervention started and how long it has been ongoing. A graphical representation of a stepped-wedge design is found in Figure 4, showing the practical and analytic features that make it well-suited for this evaluation.

The time period available for data collection was nine months.[10] We therefore staged roll-out of the MAM prevention programme at two month intervals, implying four rounds of data collection with one measurement at each round. Due to time constraints, no baseline measurement was undertaken; however, given the above assumptions about the programme cycle and effects, intra-cluster control is still possible. However, any before-and-after difference observed in the study will not be able to account for the differences that existed between control and intervention groups prior to programme initiation. To account for delays in the initiation of different MAM prevention programme components (i.e. FBMAM, HF, SBCC), 'intervention' was defined as 'at least FBMAM with the possible addition of HF and/or SBCC' for two months or longer. Control was defined as 'MAM treatment/TSFP only'.

[9] Secular trend refers to the expected variation in a certain variable over a specific period or time scale. Secular trend is relevant to indicators that are seasonal, such as wasting prevalence and incidence, given that they are expected to change (i.e. increase or decrease) depending on the season with or without external input.

[10] The original proposal for the study period was 18 months, with the idea of allowing maximum exposure to a well-functioning and high-coverage programme after every step of the study design. We initially reduced this study period to 12 months given the time limitations set by the terms of the funding. We further cut down the study period to 8−9 months due to administrative delays in project setup, but felt that this duration would be sufficient (in conjunction with a well-functioning programme) to detect change in the primary outcome of acute malnutrition because of the relatively short known duration of this condition. We discuss the implications of these delays in detail in Appendix A.

The stepped-wedge approach allowed us to collect factual and counterfactual information in a study setting where the withholding of an intervention intended for wide implementation has ethical and political ramifications. We believed that a typical parallel study design where one group does not receive the intervention for the entirety of the study would be contentious in a setting such as Sudan. A stepped-wedge design addresses this issue. This design is now used more widely in the evaluation of service delivery interventions and therefore is suited for the FBMAM programme we are evaluating. In addition, stepped-wedge designs fare better than typical parallel designs in settings where cluster level effects are expected to be high. This was the case in our study, particularly as GAM was one of our outcome measures. By using a stepped-wedge design, we were able to make the important assumption that correlation within clusters remains roughly constant over time. We believe that this is a reasonable assumption given the outcome measures that we are assessing. There is no evidence to suggest that the correlation between GAM prevalence and incidence in each of the clusters we have selected would vary significantly within the study period.

5.2.2 Theoretical framework

Our study's theoretical framework is based on the Theory of Change presented in section 2.2. We theorise that FBMAM addresses a specific cause of acute malnutrition depicted in the theory of change, namely insufficient access to food. FBMAM targets those who are most at risk of acute malnutrition when there is insufficient access to food: children under 24 months with a nutritional status that is already borderline as manifested by a MUAC of 125mm to 135mm, and PLW with a nutritional status that is already borderline as manifested by a MUAC of 210mm to 230mm – with the rationale that these specific groups of individuals are most likely to become acutely undernourished without any supplemental food support.

However, as with the Theory of Change, the multi-causal nature of acute malnutrition means that other causes must be addressed; therefore, complementary interventions such as HF and SBCC are put in place.

With each component (food- and non-food based) of the intervention in place and working effectively, we expect that rates of acute undernutrition will decline. This is the basis of our survey design and subsequent analysis.

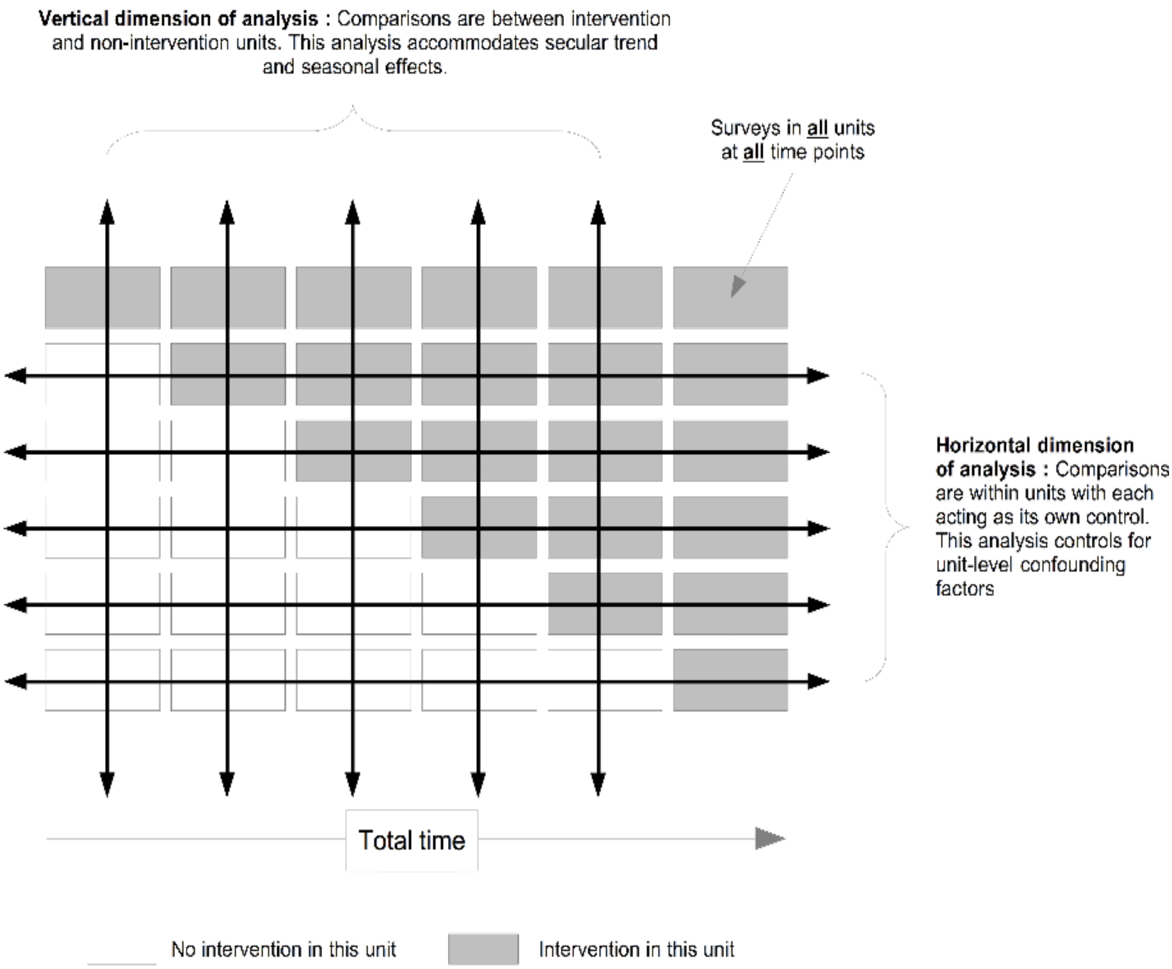
5.2.3 Sample size

Using the design specifications described above (a stepped-wedge cluster controlled trial with four rounds and one measurement after each round, but with no baseline measurement), we applied the sample size calculations proposed by Woertman and others (2013). We calculated the sample size based on an expected 5% decrease in MAM prevalence (from 20% to 15%), with an 80% power to detect a difference and a 5% level of significance. We took into account cluster-design effect based on an intra-cluster correlation of 0.034 as estimated by Kaiser and others (2006) and a cluster size of 192.[11]

[11] The cluster size of 192 is based on the smallest sample size that would allow for the estimation of GAM at the cluster level using a probit function estimator (see our pre-analysis plan in Appendix D for a detailed explanation of the probit function).

Our calculations indicated a total sample size requirement of 1,346. We planned to obtain this from 6 clusters with a sample size of 224 to be taken from each. For detailed sample size calculations, refer to Appendix E.

Figure 4: Illustration of the stepped-wedge design, where different clusters switch from control to intervention at different time points



5.2.4 Sampling structure

For the purposes of the study, we defined a cluster as a locality within Kassala state where FBMAM was to be implemented. We selected clusters that met control status (i.e. the existence of an already-operational TSFP). We also took into account status regarding other nutrition or nutrition-related interventions, such as the integrated management of childhood illness, in order to ensure that clusters had similar profiles. To minimise contamination, non-adjacent clusters were selected as much as possible. Finally, to accommodate the study's intensive data collection, which required frequent and reliable access, cluster selection was also based on accessibility.

To ensure sample sizes were achieved, clusters had to have a minimum of 24 villages or settlements. All localities in Kassala fit these criteria; therefore we considered them all. For cluster selection, we acquired a spreadsheet from WFP listing all programme sites, which included information on: (1) locality of the programme site; (2) implementing partner responsible for the programme site; and (3) types of services available from the site.

Based on this information and various considerations, we identified the following localities as the most feasible clusters for the impact evaluation:

1. ***Kassala (urban)*** – 20 sites with the highest catchment population of all the localities. One partner delivers services in this site (SMoH).
2. ***Kassala (rural)*** –17 sites. One partner delivers the package in this site (Talawiet Organisation for Development).
3. ***River Atbara*** – 9 sites with one partner implementing the programme (Plan International).
4. ***El Girba (rural)*** – 8 sites with one partner implementing the programme.
5. ***Aroma (rural)*** – 16 sites with two partners working in this locality (SMoH and Talawiet Organisation for Development).
6. ***Telkuk*** – 29 sites with three partners (SMoH, Sudanese Red Crescent, Sudan Vision).

Once we selected the localities, the next step was to determine the schedule for a locality to switch from control to intervention. Ideally, this should be randomised. However, WFP's operational requirements dictated that they implement the MAM prevention programme in rural Aroma and Telkuk first. Therefore, by default, we chose these two localities as the first to receive the intervention. Given that we were unable to randomise all clusters, we decided to purposefully select when the next clusters would switch to the intervention group based on discussions with WFP about the most convenient rollout schedule. Figure 5 presents our planned survey design and implementation. The randomisation of clusters into intervention and control was an important factor in establishing a truly experimental study design. Randomisation mitigates selection bias created by purposeful selection of study clusters based on pre-set criteria or ad hoc rules. We therefore took into account the possible impact of selection bias in the effects that we observed (see Section 7).

However, as indicated in Section 4 on the timeline of events during programme and study implementation, programme rollout in the different localities changed during the study period. Specifically, River Atbara and rural El Girba (the localities meant to institute interventions in rounds 2 and 3, respectively) did not receive interventions until round 4, together with Kassala and rural Kassala. Figure 6 illustrates the actual study design that was implemented.

Figure 5: Planned study design

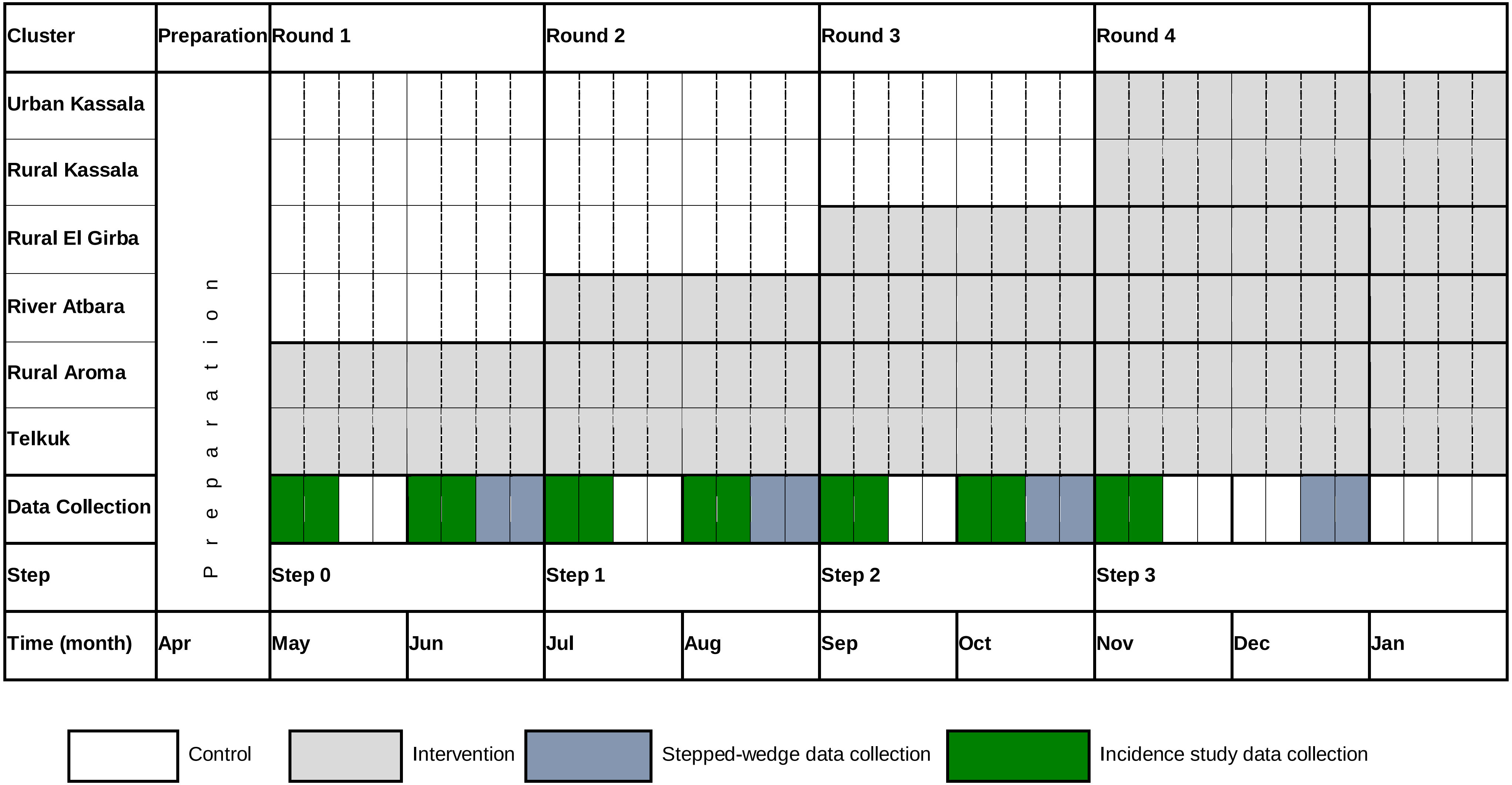
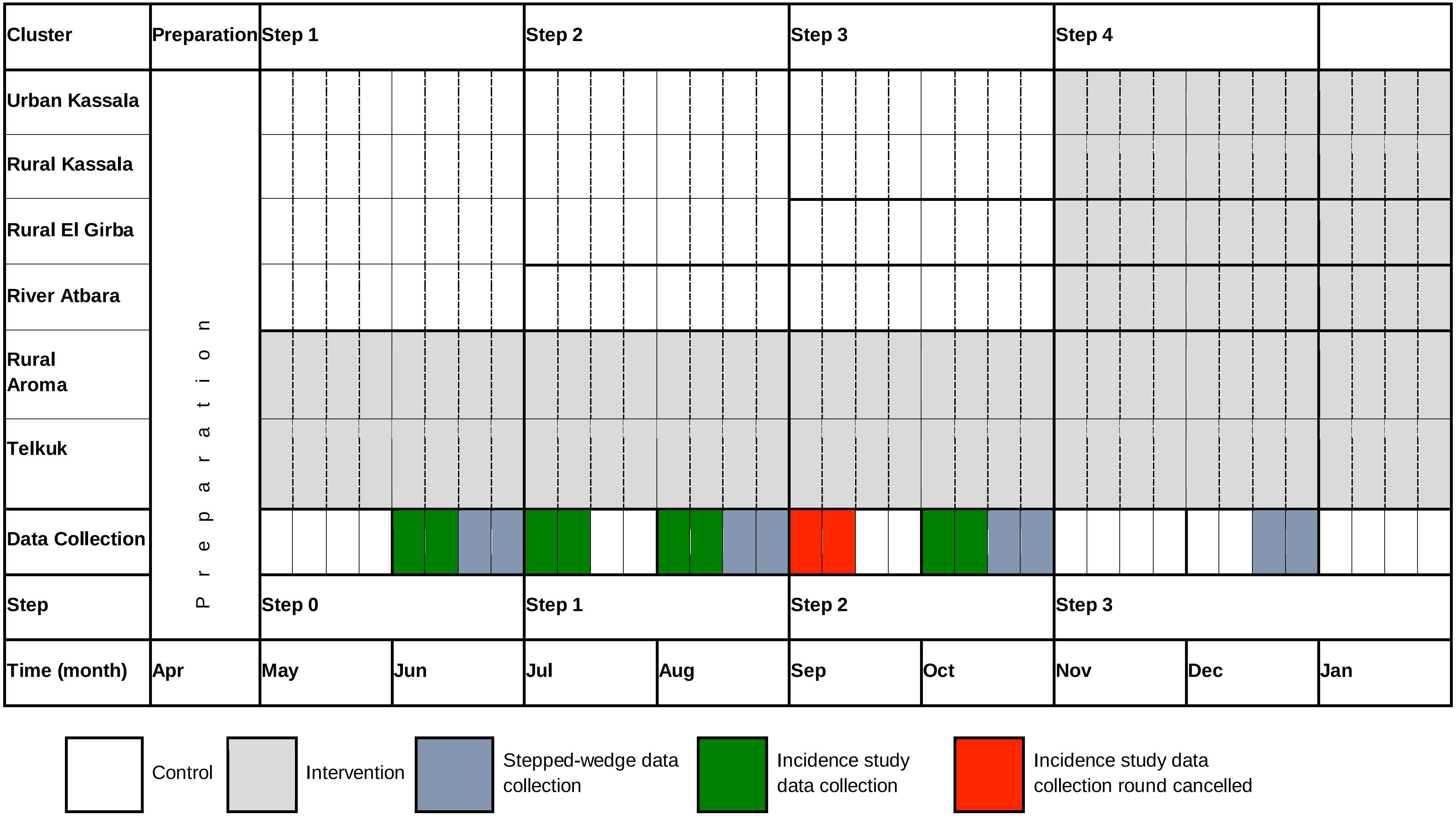


Figure 6: Actual study design



5.2.5 Sampling design

We used a two-stage spatial sampling design for the stepped-wedge cluster controlled trial to measure the primary outcome of prevalence for GAM, MAM and SAM, as well as the secondary outcomes presented in Section 2.

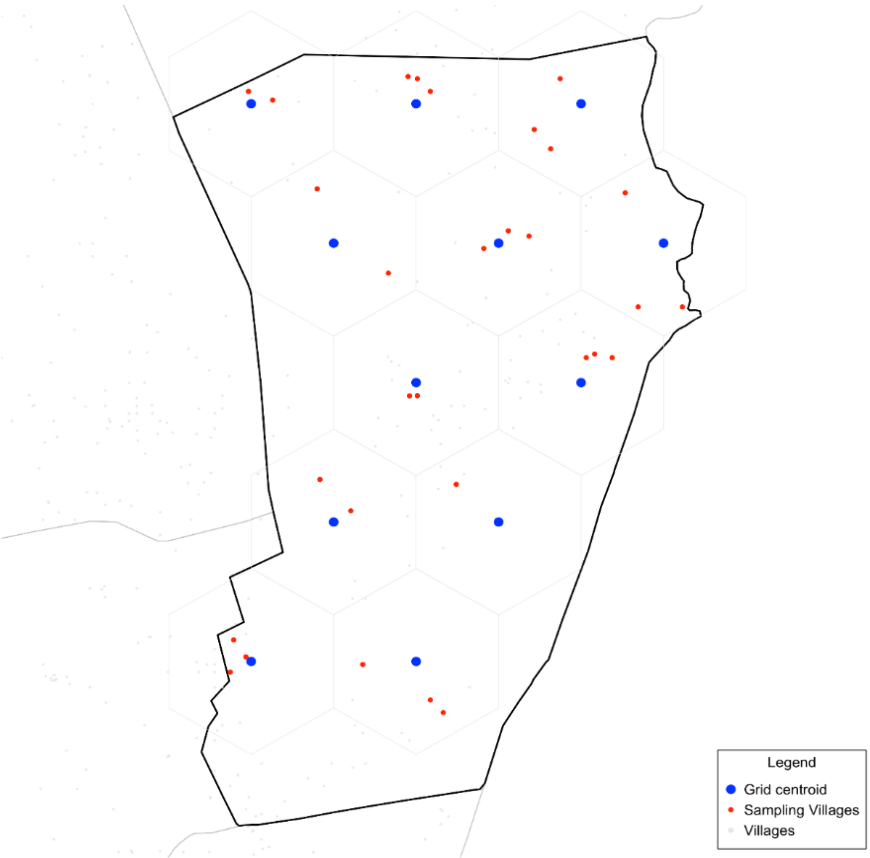
Stage 1 sampling

We applied a centric systematic area sampling approach (Milne 1959; Myatt et al. 2005), using a non-overlapping hexagonal grid to divide each of the selected clusters into 20 distinct sampling areas. In each of the hexagonal areas, we selected sampling locations from the nearest village / community / settlement to the centre of the hexagon for a total of 20 sampling villages per locality. We chose this number in order to provide an adequate sample size for the PLW target group based on the estimated number of PLWs per village, using an estimation approach recommended by Myatt (2012). This number of villages would also provide an adequate number of MAM cases for the programme coverage indicator. The stage 1 sampling plan is presented in Figure 7.

Stage 2 sampling

We carried out a full enumeration of children under five years and PLW in each of the selected sampling locations during stage 1 to guarantee efficiency. This ensured that we reached the sample sizes required for the PLW target group and programme coverage. For large sampling locations (i.e. urban towns, large villages), a systematic sampling approach was used to collect data on children under five for the prevalence, coverage and morbidity surveys. An active and adaptive case-finding approach was used to find all PLW in the sampling location.

Figure 7: Example of a stage 1 spatial sampling plan for a locality in Sudan



5.3 Incidence study

5.3.1 Design

We nested a two-arm parallel design cluster controlled trial into the main study to assess and compare the incidence of acute malnutrition in areas with both MAM treatment and prevention (intervention) and with MAM treatment only (control). Originally five rounds of follow-up were planned; however, one round had to be dropped as the timing of data collection was significantly delayed. The schedule of incidence study data collection that we were able to complete is shown in Figure 6 above.

5.3.2 Sample size

The sample size calculations in $y_{person\text{-}years}$, as proposed by Hayes and Bennet (Hayes and Bennett 1999) for an individually-randomised cluster controlled trial, was applied as follows:

$$y_{person-year} = \left(Z_{\alpha/2} + Z_{\beta}\right)^2 \times \frac{\lambda_0 + \lambda_1}{(\lambda_0 - \lambda_1)^2}$$

Where:
λ_0 = incidence rate in control group
λ_1 = incidence rate in intervention group.

We used a value of λ_0 = 0.32 (assuming a prevalence rate of 20% in the control group) and a value of λ_1 = 0.24 (assuming a prevalence rate of 15% in the intervention group). This gives a sample size for one arm of the incidence study of:

$$y_{person-year} = (1.96 + 0.84)^2 \times \frac{0.32 + 0.24}{(0.32 - 0.24)^2} \approx 686$$

For both arms, an overall sample size of 1,372 was therefore needed. To calculate the number of clusters required to meet this sample size, we used the following formula:

$$n_{clusters} = 1 + \left(Z_{\alpha/2} + Z_{\beta}\right)^2 \times \frac{\dfrac{\lambda_0 + \lambda_1}{y + k^2\left({\lambda_0}^2 + {\lambda_1}^2\right)}}{(\lambda_0 - \lambda_1)^2}$$

Where:
k = intra-cluster correlation coefficient, which we set at 0.034.

This meant that:

$$n_{clusters} = 1 + (1.96 + 0.84)^2 \times \frac{\dfrac{0.32 + 0.24}{1372 + 0.034^2(0.32^2 + 0.24^2)}}{(0.32 - 0.24)^2} \approx 2$$

Therefore, two clusters (one for each study arm) were required to achieve a sample of 1,372 (686 per arm). To allow for the nesting of the incidence study within the main stepped-wedge trial, we selected the two localities chosen to receive intervention in step 1 (rural Aroma and Telkuk) as the intervention cluster, and the two localities selected to receive intervention at step 4 (rural Kassala and Kassala) as the control cluster.

5.4 Data collection

5.4.1 Identification of study subjects

Stepped-wedge controlled trial

We used the MUAC tape as a measurement tool to assess children and PLW. Once eligibility for programme participation had been established, we administered a bespoke questionnaire to mothers/carers of children and PLW. Mothers/carers of children covered by the programme answered standardised questions on knowledge, attitudes and practices. Mothers/carers of children eligible for, but not covered by, the programme answered questions covering reasons for not being in the programme, programme history of the child and programme awareness.

Incidence

We identified a cohort of children under five years at the health-centre level and followed them over a five-month period, with measurement at monthly intervals. The number of children who developed acute malnutrition (MAM or SAM) were recorded in order to estimate the number of child-years required for acute malnutrition to develop in children under five in the control and intervention groups.

Programme monitoring data

We also collected routine programme-monitoring data using pre-defined programme databases created by WFP for each of the different components of the MAM treatment and MAM prevention programmes.

5.4.2 Data collection tools

We collected data for GAM prevalence, at-risk prevalence, coverage and GAM incidence using an electronic data entry system. This was based on the open data kit standard, which runs on the Android operating software platform for mobile devices. We encoded the questionnaire into the electronic data entry system platform and used a local computer server. We provided each data collection team with mobile devices running on Android that had been configured with an application to receive the electronic data form. The team recorded all measurements and answers from respondents on the mobile devices and transmitted these to the local server when a mobile phone and/or WiFi signal was available. Paper versions of the study instruments can be found in Appendix C.

5.4.3 Enumerators

Five teams were involved in data collection for the stepped-wedge and cohort studies. Each team was composed of three enumerators, including a team leader and two measurers. We trained teams in a 10-day classroom-based workshop in March 2016, covering:

- Study design
- Nutrition situation in Sudan and Kassala
- CNIP
- Open data kit data collection system
- Questionnaires and indicators
- Anthropometric measurement and standardisation, specifically regarding MUAC and nutritional oedema

A refresher training and a field exercise were conducted over five days in May 2016, prior to the start of data collection in June 2016. During this time, we provided practical training on sampling and cross-checking data collection forms, and questionnaires were refined based on weaknesses identified. We implemented an actual dry run of data collection in an urban block of Kassala not covered by the study. Field training focused on incidence data collection.

Economic remuneration, which corresponded to national salary scales, was provided for enumerators using research funding. Enumerators were recruited and contracted through Sudan FMoH to facilitate administration. The study's national co-primary investigator supervised data collection. Back and spot checks of respondents and questionnaires ensured consistency and accuracy of questions and anthropometric measurements.

5.5 Qualitative study

We performed a qualitative investigation between round 3 and round 4 of the stepped-wedge study data collection period. The qualitative investigation aimed to provide more nuanced and contextualised information related to the preliminary cross-sectional results that we obtained from the first 3 rounds of the study. The results of the qualitative information, in turn, helped us to understand and contextualise the quantitative results. The qualitative investigation focused primarily on two key streams of enquiry. First, it carried out a more in-depth examination of coverage of the various components of the prevention programme. Second, it collected additional information on the effects of the SBCC interventions, specifically the mechanisms that change/do not change current practices relevant to children's and women's nutrition.

Based on programme status, available data and accessibility, the study identified Kassala, rural Kassala, Aroma and Telkuk as localities of interest. The selection of study sites within these localities was based on a purposive and convenient sampling approach, aiming for equal representation of WFP implementing partners in each locality.

The investigation was executed using the following steps:

1. Information was collated and synthesised from WFP routine programme monitoring data, the on-going stepped-wedge cluster control trial and the MAM incidence study, specifically:
    - Admission and defaulter data collected through the WFP database;
    - MUAC-at-admission collected from programme beneficiary cards;
    - Discharge outcomes collected through the WFP database;
    - Spatial mapping of programme sites and catchment areas, location of volunteer workers, and home location for admissions and defaulters, using data from the stepped-wedge cluster control trial;
    - Distance to programme sites collected through programme registers and the stepped-wedge cluster control trial;
    - Timeline of key events developed (as presented in Section 4).

    Table 3 lists the study sites for which routine data was reviewed and further information collected. The original plan was to collect routine data from 20 sites, but, due to distance and time constraints, only 12 sites were covered. This did not compromise proportional distribution.

The analysis of routine monitoring data provided us with insight into three important factors that have been shown to impact coverage:

- The programme's effectiveness in case-finding, as well as beneficiaries' treatment-seeking behaviour, for which MUAC-at-admission acts as a proxy;
- The programme's responsiveness to people's needs, as indicated by a time-series analysis of admissions and defaulters;
- Beneficiaries' geographical access to the programme.

These factors were also related to preliminary findings observed for programme coverage in the quantitative component of the study.

Table 3: Sites for which routine data was reviewed and further information collected

Locality	Implementing partner	Planned number of sites	Actual number of sites visited
Telkuk	MoH	0	0
	Sudan Vision Organisation	Elatiot, Darasta, Haladet East & Hashaneit Masejit (4)	Elatiot, Darasta & Hashaneit Masejit (3)
	Sudanese Red Crescent	Tahdia Osisi, Gademeiet Almasgid & Edoret (3)	Tahdia Osisi & Gademeiet Almasgid (2)
Aroma	Waad Organisation for Development	Degain, Makali, Tindlahi & Jama (4)	Degain, Makali & Tindlahi (3)
Kassala	MoH	Mukram Alderwa, Althawra, Gareb Algash & Tarawa (4)	Hamid Wkiul & Salam 16 (2)
Rural Kassala	TOD	Hafarat, Demen, Amara, Wedisherfay & Gulssa (5)	Wedisherfay & Gulssa (2)
	SRC	0	0

2. We discussed these factors, and the mechanisms by which they impacted on coverage, with key informants. These included WFP Kassala, FMoH and WFP implementing partners.
   We formulated the following key hypotheses for further investigation:
   - Potential issues with how routine monitoring data are collected may have an impact on the veracity of routine data once reported at state and national level;
   - Access to the programme is not necessarily influenced by geographical distance, but rather by specific cultural norms related to gender relations;
   - Locations with good coverage are those with regular screening, low MUAC-at-admission and a community which is very familiar with and aware of the programme.
3. We then proceeded to investigate using the following methods: key informant interviews, focused group discussions, semi-structured interviews and documented case studies. Key informants at this level included parents of programme beneficiaries, PLWs, volunteers and health workers. Men and women were interviewed separately to allow both groups to speak freely.

We selected the four sites for investigation based on the following specific criteria:

Table 4: Selected study sites and selection criteria

Locality	Study site	Reason for selection
Telkuk	Darasta Haladet East	Low MUAC at admission, indicating poor screening performance
Aroma	Degain	Discharge data not reflecting defaulting and cured cases
Aroma	Jamam	Convenient
Kassala	Tarawa	Anecdotal reporting of good programme performance and model site for CNIP

5.6 Cost effectiveness

We implemented an activity-based costing methodology using an ingredient approach, accounting for all programme (provider) costs and household/community (beneficiary) inputs/costs (Puett et al. 2013; World Health Organization 2003), with the intention of measuring the cost-effectiveness of different intervention components and programme modalities.

We collected programme (provider) cost data through reviews of programme budgets and financial reports. Unfortunately, it was challenging for WFP and their implementing partners to provide the data requested, particularly separate costs for treatment and prevention programmes. Moreover, not all field activities related to MAM prevention were funded by WFP, as implementing partners undertook complementary SBCC activities at their own costs.

We performed a basic cost analysis to gain an understanding of actual programme expenses by administering questionnaires to WFP and their implementing partners. We asked finance and programme staff to retrospectively categorise costing information into comprehensive and mutually exclusive cost centres based on operational activities (e.g. start-up investments, personnel, materials and supplies, and food and supply chain costs, among others). We assumed the sum of estimates of all component activities as equal to the total programme (provider) costs. We only accounted for costs covered by WFP in order to simplify the data collection process. A detailed breakdown of the information comprising the CEA can be found in Appendix H.

We collected household/community (beneficiary) input/cost data prospectively during the last round of the stepped-wedge study.

In addition to the costing data, we used programme-monitoring data to account for programme effectiveness. This data reported the number of beneficiaries (children under five and PLW) that were admitted and cured, recovered or graduated when participating in the different programmes.

Regrettably, our efforts to collate retrospective estimates of provider costs failed. As a result, we are unable to report on important cost-effectiveness metrics, including cost per case of MAM in children under-five averted; cost per PLW case of undernutrition averted; cost per DALY averted; and incremental cost effectiveness ratio. Our inability to secure adequate and reliable cost data may be due to the considerable time input requested by WFP and implementing partners to fill in the provider cost questionnaire. Rather than collecting this data remotely, a better approach would have been for our cost effectiveness analyst to conduct interviews and focus group discussions in the field in person. However, considering travel restrictions within Sudan along with the study's limited funding and staff time, this approach was not feasible. Funding limitations are likely to be a recurring issue in future impact evaluations; therefore resources may be better spent securing the necessary cost data for a comprehensive cost-effectiveness analysis through improved programme monitoring. We provide recommendations for how this can be operationalised in Section 8.

6. Impact analysis and results of the key evaluation questions

This section presents the outputs of the impact analysis. These outputs are presented and organised based on the study's key evaluation questions.

6.1 Analytical framework

6.1.1 GAM prevalence

GAM prevalence was measured at each study area or cluster. GAM prevalence was reported for children under five and PLW. GAM prevalence was estimated with a probit estimator, using the observed mean and standard deviation of the collected MUAC data at every data collection round of the stepped-wedge study.[12] This indicator was measured to answer our main research question on the impact of FBMAM on GAM prevalence.

6.1.2 GAM incidence

Given the changes in the study design,[13] GAM incidence was not estimated classically. Instead, we proxied GAM incidence using a time-to-event analysis, which in this case is a time-to-undernutrition metric. A Kaplan-Meier survival curve analysis will be applied on the data collected for the incidence sub-study to report on indicators for the proportion of children becoming acutely undernourished at monthly intervals, and the average number of months before a child becomes acutely undernourished. This indicator was measured to answer our main research question on impact of FBMAM on GAM incidence.

6.1.3 Programme coverage

Various programme coverage indicators were measured as a nested survey within the GAM prevalence surveys in order to answer questions on programme effectiveness, timeliness and targeting. Eligibility for each of the various MAM treatment and prevention packages was determined and various coverage estimators were assessed. Given the multiple intervention components within the MAM treatment and prevention packages, various coverage estimators were used. Specifically, we assessed the following coverage indicators:

[12] One collection round per step with four steps in total.

[13] Changes such as a follow-up period of less than one year, and loss of data collection rounds due to delays.

a. MAM case-finding effectiveness for children – this was defined as children 6–59 months who are current MAM cases[14] in TSFP, out of the total number of children 6–59 months who are current MAM cases.
b. MAM treatment coverage for children – this was defined as children 6–59 months who are current or recovering MAM cases[15] in the TSFP, out of the total children 6–59 months who are current and recovering MAM cases.
c. MAM case-finding effectiveness for PLW – this was defined as PLW who are current MAM cases[16] in the TSFP, out of the total PLW who are current MAM cases.
d. MAM treatment coverage for PLW – this was defined as PLW who are current and recovering MAM cases in the TSFP, out of the total PLW who are current and recovering MAM cases.
e. Targeted MAM prevention coverage for children – this was defined as children 6–23 months old who are at risk[17] in the targeted FBMAM programme, out of all children 6–23 months old who are at risk.
f. Targeted MAM prevention coverage for PLW – this was defined as PLW who are at risk[18] in the targeted FBMAM programme, out of all PLW who are at risk.
g. Blanket MAM prevention coverage for children – this was defined as children 6–23 months old in the blanket FBMAM, out of all children 6–23 months old.
h. Blanket MAM prevention coverage for PLW – this was defined as PLW in the blanket FBMAM, out of all PLW.
i. Home fortification coverage – this was defined as children 6–59 months old not eligible for the TSFP or FBMAM[19] who are receiving home fortification, out of all children 6–59 months old not eligible for TSFP or FBMAM.
j. SBCC coverage – this was defined as mothers and/or caregivers of children 6–59 months old and PLW who have received or participated in at least one appropriate education session and/or individual counselling session in the past month, out of the total of mothers and/or caregivers of children 6–59 months old and PLW.
k. Mothers' groups coverage – this was defined as mothers and/or caregivers of children 6–59 months old and PLW enrolled in mothers' clubs, out of the total of mothers and/or caregivers of children 6–59 months old and PLW.
l. Care groups coverage – this was defined as mothers and/or caregivers of children 6–59 months old and PLW enrolled in care groups, out of the total of mothers and/or caregivers of children 6–59 months old and PLW.

---

[14] WFP Sudan's Community-based Nutrition Integrated Programme (CNIP) Field Guide defines MAM cases as children 6–59 months with MUAC greater than or equal to 115mm and less than 125mm and no oedema, or children 6–59 months discharged from OTP.
[15] Recovering cases are children whose MUAC is greater than 125mm, but they have not met the minimum of two consecutive visits for discharge verification criteria.
[16] WFP Sudan's CNIP Field Guide defines acute malnourished PLW as women in their second or third trimester or with a child under 6 months old, who have a MUAC under 210mm.
[17] WFP Sudan's CNIP Field Guide defines at-risk children as children 6-23 months with a MUAC greater than or equal to 125mm and less than 135mm.
[18] WFP Sudan's CNIP Field Guide defines at-risk PLW as those with a MUAC greater than or equal to 210mm and under 230mm.
[19] This includes children discharged from FBMAM and children discharged from TSFP where no FBMAM exists.

All coverage indicators mentioned above were calculated in each of the study areas or clusters at each of the data collection rounds.

6.1.4 Cost effectiveness

We used activity-based costing[20] with relevant costs for both provider and participant, grouped by activity and organised by cost centres for analysis and calculation of the total incurred for implementing the MAM treatment programme in Kassala and the total cost of implementing both the MAM treatment programme and the MAM prevention programme in the state.

a. *Provider costs*

   Costs incurred by the service provider (WFP and the implementing partner) were collected using semi-structured key informant interviews with relevant programme and administrative staff at WFP and implementing partners. Cost data gathered from the provider included: (1) personnel costs; (2) programme supplies; and (3) programme delivery.

   For personnel, we collected data on salary information and staff time spent by WFP and implementing partners in the implementation of the MAM treatment and prevention programmes. Costs for non-salaried personnel (whether or not incentivised) such as community health workers or community mobilisers were collected. For non-incentivised personnel, a shadow wage rate[21] was estimated based on current labour markets in Kassala and from previous studies that have estimated this rate.[22]

   For supplies, costs of all supplies and materials including the feeding product used for both treatment and prevention of MAM were collected from programme budgets and programme staff.

   For programme delivery, data on transport costs, training, rent and utilities were collected through programme budgets, other related documentation and programme staff.

b. *Participant costs*

   The direct cost of participation for beneficiaries, including transport costs to access treatment and prevention programmes, travel time and indirect costs such as opportunity costs incurred by family and/or caregivers were collected through the cross-sectional surveys for the stepped-wedge study. We performed spatial

[20] See Fiedler, JL, Villalobos, CA and De Mattos, AC, 2008. An activity-based cost analysis of the Honduras Community-Based, Integrated Child Care (AIN-C) programme. *Health Policy and Planning*, 23(6), pp.408–427; and Waters, H, Abdallah, H and Santillán, D, 2001. Application of activity-based costing (ABC) for a Peruvian NGO healthcare provider. *The International Journal of Health Planning and Management*, 16(1), pp.3–18.

[21] Shadow wage rate is described as the foregone output or wage of labour. It is also called the opportunity cost of labour. See Heckman, JJ, 1974. Shadow prices, market wages, and labor supply. *Econometrica*, 42(4), pp.679–694; and Jacoby, HG, 1993. Shadow wages and peasant family labour supply: an econometric application to the Peruvian Sierra. *The Review of Economic Studies*, 60(4), pp.903–921.

[22] See Babikir, OM, Babiker, I and Bauer, S, 2007. A test of agricultural labour market efficiency in the Gezira scheme, Sudan: a production function approach. *World Review of Science, Technology and Sustainable Development*, 4(4), p.376.

interpolation using travel-time data and geo-location data collected by the study, together with publicly available geographic data on elevation, roads, land use and water bodies in Kassala, to create a raster-based cost surface at a resolution of at least 10 sq. km. Cost was measured in terms of the time it takes to travel from a specific location on the raster map to the nearest health facility or distribution site. Then, using the collected data on average daily wages in Kassala, the time-to-travel metric was converted into opportunity cost, thereby creating an opportunity-cost surface.

c. *Allocation to cost centres*
Cost centres were developed and finalised in collaboration with WFP and relevant implementing partners. The estimated costs described above were then categorised under the relevant cost centres.

6.1.5 Knowledge, attitude and practices

This was measured using a specifically designed set of questions for mothers/caregivers of children 6–59 months and PLW on topics covered by the SBCC component of the programme (i.e. healthy pregnancy, child health and healthcare, breastfeeding, complementary feeding, dietary diversity, food supplementation, use of MNP and WASH. Wherever possible, standard question sets that have been developed and tested for knowledge, attitudes and practices assessment were used. The following are some standard question sets used for this purpose:

a. IYCF – A question set adapted from standard guidelines[23] was used to assess breastfeeding, complementary feeding and diet diversity. Indicators were adapted for simplicity and rapidity, as well as the small sample size (as compared with MICS, DHS, etc.), which will allow analysis at a local level.

   The approach used was intended to produce a single indicator defining good feeding practices for infants and young children as either:
   - Exclusive breastfeeding in children aged under six months; or
   - Age-appropriate feeding practices (defined in terms of continued breastfeeding, dietary diversity and meal frequency) in older children.

   Age-appropriate feeding practices were measured using an Infant and Child Deeding Index (ICFI) similar to that used in the 2000 DHS survey of Ethiopia and further developed by IFPRI and FANTA as a KPC2000+ indicator:

[23] See World Health Organization, 2008. Indicators for assessing infant and young child feeding practices: conclusions of a consensus meeting held 6–8 November 2007, Washington, DC; World Health Organization 2010. Indicators for assessing infant and young child feeding practices part 2: measurement, Geneva.; World Health Organization, 2010. Indicators for assessing infant and young child feeding practices part 3: country profiles, Geneva.; Arimond M, and Ruel M, 2003. Generating indicators of appropriate feeding of children 6 through 23 months from the KPC 2000+, Washington DC, FANTA / AED; Arimond M, Ruel MT, 2002. Progress in developing an infant and child feeding index: an example using the Ethiopia demographic and health survey 2000. Food consumption and nutrition division discussion paper #143, Washington DC, IFPRI; and KPC Module 2: Breastfeeding and infant and young child feeding. 2006 ed., June 29, 2006.

	Age group (months)							
	6 – 8		9 – 11		12 – 36		36 – 60	
	Value	Score	Value	Score	Value	Score	Value	Score
Breastfed	Yes	+ 2	Yes	+ 2	Yes	+ 1	Yes	+ 0
Food groups	1 ≥ 2	+ 1 + 2	1 or 2 ≥ 3	+ 1 + 2	2 or 3 ≥ 4	+ 1 + 2	3 or 4 ≥ 5	+ 2 + 3
Meal frequency	1 ≥ 2	+ 1 + 2	1 or 2 ≥ 3	+ 1 + 2	2 3 ≥ 4	+ 1 + 2 + 3	2 3 ≥ 4	+ 1 + 2 + 3

The ICFI score is a measure of appropriate child feeding practices:

ICFI = Breastfeeding + Dietary Diversity + Meal frequency

Age-specific weighting is used for each item. Children receive a score between zero and six. Children receiving a score of six are classified as receiving good infant and young child feeding. The ICFI can be extended to include older children if required. The shaded areas in Table 14 represent this extension of the standard ICFI score to include children aged between 36 and 59 months.

b. Women’s dietary diversity – A standard diet diversity questionnaire[24] was used to assess the diets of mothers with children 6–59 months old and PLW. Women's dietary diversity scores (WDDS) were calculated for mothers of the children sampled for IYCF. The WDDS indicator assesses the quality of women’s diets and indicates their micronutrient adequacy. The data collected on women’s dietary diversity resulted in six indicators.

- Women’s dietary diversity scores (WDDS) were calculated based on the 10 food groups (see Table 15) determined to be relevant and important for women. The potential WDDS score ranges from zero to 10, based on the number of food groups consumed by women out of the 10 food groups.

FG1	Starchy staples
FG2	Dark green leafy vegetables
FG3	Other vitamin A-rich fruits and vegetables
FG4	Other fruits and vegetables
FG5	Organ meat
FG6	Meat and fish
FG7	Eggs
FG8	Legumes
FG9	Nuts and seeds
FG10	Milk and milk products

[24] Women’s dietary diversity is assessed using a similar questionnaire as the one used for household dietary diversity. The main difference is that there are questions about certain foods and food groups of particular importance to women, particularly those of reproductive age. For more information, see Arimond, M et al. *Dietary diversity as a measure of the micronutrient adequacy of women's diets: results from rural Bangladesh site*. Washington, DC: Food and Nutrition Technical Assistance II Project, FHI 360, 2009; Kennedy, G, Ballard, T, and Dop, MC. *Guidelines for measuring household and individual dietary diversity*. Rome: Food and Agricultural Organization, 2011.

- The mean WDDS was calculated as:

$$\text{Mean } WDDS = \frac{\sum WDDS}{\text{Total number of women assessed}} \qquad (5)$$

- Consumption of vitamin A-rich foods was calculated based on which women consumed them in the past 24 hours. This indicator identifies women at risk of vitamin A deficiency.
- Consumption of iron-rich foods was calculated based on which women consumed them in the past 24 hours. This indicator identifies women at risk of iron deficiency.

c. Food supplementation and MNP – We built upon a set of questions we developed and used to assess the coverage of knowledge and practices regarding the use of complementary food supplements in eastern Ghana.[25]

6.1.6 WASH

We used some components of the standard WASH indicator set[26] that focus on WASH-related behaviours such as safe disposal of child's faeces, water treatment practices, hand washing practices and other variations related to WASH behaviours.[27]

6.2 Analytical approach

We approached the analysis for the study at two levels: (1) a cross-sectional data analysis of multiple indicators for each round of the main stepped-wedge study, using a blocked weighted bootstrapping approach; and (2) a comparative analysis used to detect changes between comparison groups and test their significance, which used a similar blocked weighted bootstrapping approach to perform a two-sample z-test between the groups being compared. For the nested incidence study, we applied survival analysis to the data to estimate the mean time to acute malnutrition in both the control and intervention groups. We then compared the difference in mean time between the two groups using Cox's proportional hazards model. A more detailed description of the various analytical approaches is outlined below.

6.2.1 Blocked weighted bootstrap

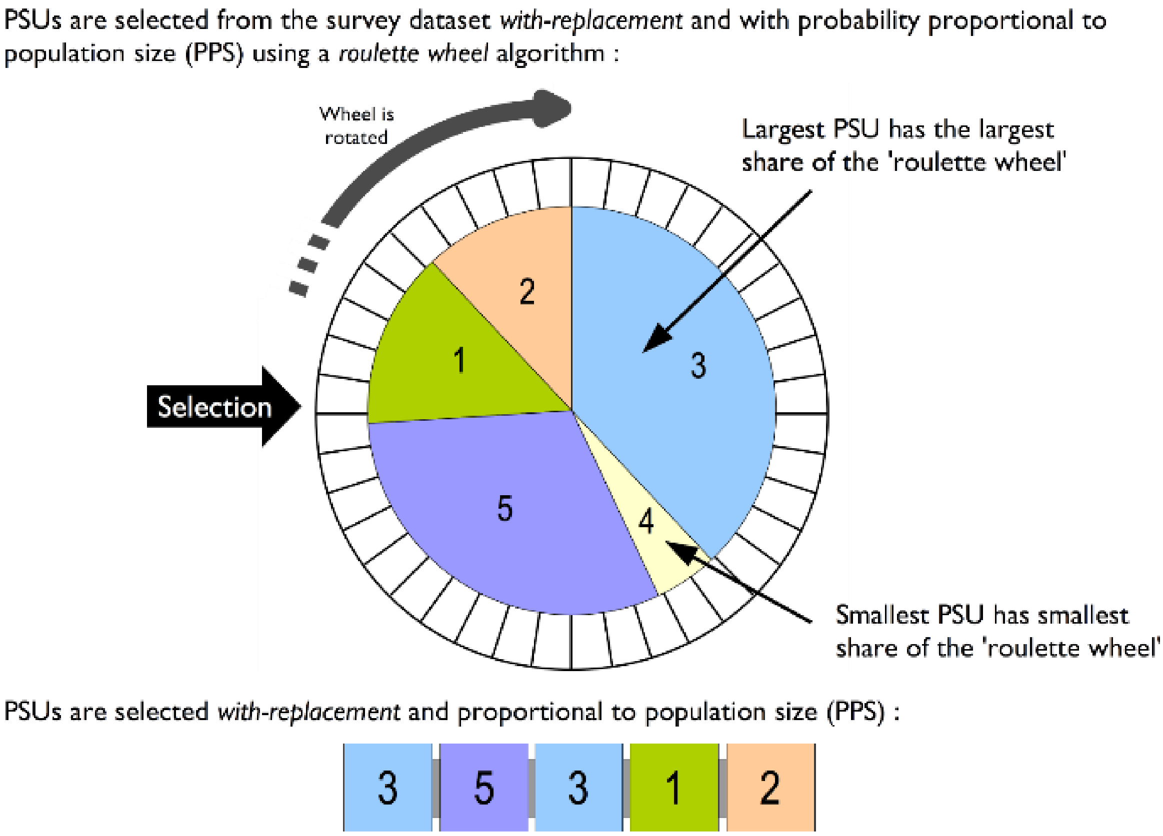
Bootstrap is a resampling technique (Diaconis and Efron 1983) used to calculate summary estimates and confidence intervals for all indicators previously discussed. In order to account for the two-stage sampling design of the study, we developed a bespoke analysis script. This took into consideration both the use of cluster sampling with villages as the primary sampling unit (PSU), and the appropriate posterior weighting of the cluster level data, given that we did not select PSUs proportional to population size. The *blocked* component of the bootstrap addresses the cluster sampling design, while the *weighted* component deals with the posterior weighting.

---

[25] See Aaron, G et al., 2014. Coverage of a market-based approach to deliver a complementary food supplement to infants and children in three districts in eastern Ghana: Use of the Simple Spatial Survey Method (S3M). *Faseb Journal* 28(1), pp.255–5.
[26] See World Health Organization. Core questions on drinking-water and sanitation for household surveys, Geneva: World Health Organization, 2006.
[27] See Hernandez, O. Access and behavioral outcome indicators for water, sanitation, and hygiene, Washington, DC: USAID, February 2010.

The blocked weighted bootstrap analysis uses a ‘roulette wheel’ algorithm to weight, by population, the selection probability of PSUs in bootstrap replicates. A total of *m* PSUs are sampled with replacement from the survey dataset, where *m* is the number of PSUs in the survey sample. Individual records within each PSU are then sampled with replacement. A total of *n* records are sampled with replacement from each of the selected PSUs where *n* is the number of individual records in a selected PSU. The resulting collection of records replicates the original survey in terms of both sample design and sample size. A large number of replicate surveys are taken (*r* = 1999 replicate surveys). The required statistic (e.g. the mean of an indicator value) is applied to each replicate survey. The reported estimate consists of the 50th (point estimate), the 2.5th (lower 95% confidence limit), and the 97.5th (upper 95% confidence limit) percentiles of the distribution of the statistic observed across all replicate surveys. The blocked weighted bootstrap procedure is graphically described in Figure 8.

Figure 8: Illustration of the blocked weighted bootstrap algorithm



6.2.2 Comparison of groups – vertical analysis

We compared groups using a blocked weighted bootstrapped *two-sample z-test*. Individual standard errors were calculated as:

$$SE_{point\ estimates} = \frac{UCL - LCL}{2 \times 1.96}$$

where *UCL* and *LCL* are the upper and lower 95% confidence limits on the indicator estimates. The resulting standard errors were pooled:

$$SE_{pooled} = \sqrt{SE_t^2 + SE_{t-1}^2}$$

and the test-statistic calculated as:

$$z = \frac{|Estimate_t - Estimate_{t-1}|}{SE_{pooled}}$$

A two-sided *p-value* was calculated. A *p-value* of < 0.05 was considered statistically significant.

6.2.3 Comparison of groups – horizontal analysis

This is a before-and-after comparison using a difference in differences analysis. We pooled data from round 1 for intervention clusters, then separately for control clusters, and estimated outcome measures from each set of data. These results served as the ‘before’ outcome measures for the control and intervention clusters. We then applied the same method of pooling intervention and control cluster data and estimated outcome measures for each at rounds 2, 3 and 4. Finally, we performed a series of before-and-after comparisons between round 1 estimates of the outcome measure and the outcome measure for each subsequent round.

As the horizontal comparison does not take secular trend into account, we estimated this by subtracting the before-and-after comparison for the control groups from the before-and-after comparison for the corresponding intervention group at each round. This provides an adjusted before-and-after difference, which accounts for any before-and-after difference due to secular trend.

We performed before-and-after comparisons using the same bootstrapped two-sample z-test employed in the vertical comparison.

6.3 Main study question: What is the impact on the incidence and prevalence of MAM and SAM in children under five and pregnant and lactating women of different MAM treatment and prevention interventions in Sudan?

6.3.1 Impact on prevalence - children

We first report the results of the vertical analysis:

Figure 9 below presents the vertical comparison of the prevalence of MAM, SAM, GAM and the ‘at risk’ category between intervention and control clusters at each of the four rounds of data collection.

Table 5 below presents the same information, but with confidence intervals for each outcome measure for intervention and control, along with the magnitude of change between intervention and control and the corresponding p-value.

The vertical comparison analysis controls for season, as each cluster is compared with each other at the same point in time. Hence, any effects due to time or season were controlled for in the analysis.

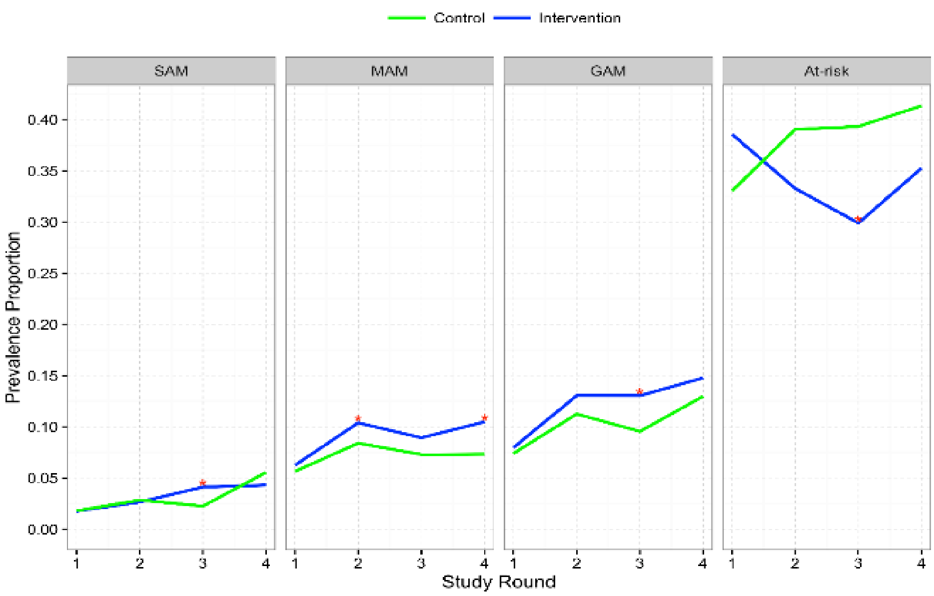
Figure 9 demonstrates that at round 1 of the study, both intervention and control clusters start off with roughly the same levels of prevalence for MAM, SAM and GAM. However, the intervention group in the ‘at risk’ category (the target group for FBMAM) has a higher prevalence for those at risk than the control group (though it is not statistically significant).

Over time, there is a general increasing trend in prevalence in each of the acute malnutrition categories for both intervention and control. Figure 5 shows that prevalence of acute malnutrition in the intervention group is not only increasing over time, but it is also consistently higher than the control group starting at round 2. Prevalence of SAM is similar in intervention and control groups in round 2 but increases in the intervention group in round 3, and slightly dips and crosses over with control in round 4. Prevalence of MAM and GAM are higher in intervention than in control groups from round 2 to round 4. This seems to indicate that when controlling for time and seasonal effects, the intervention group for the SAM, MAM and GAM categories has a higher prevalence of acute malnutrition starting at round 2, though this increase is not always statistically significant.

The trend is very different for the 'at risk' category. The intervention group starts off with higher prevalence of being 'at risk' than the control group at round 1. By rounds 2 and 3, there is a decrease in prevalence, with a slight increase again at round 4. The prevalence of being 'at risk' in the control group continues to increase over time. At round 3, there is a statistically significant decrease in the at risk prevalence for the intervention group compared to the control group.

The vertical analysis results from data collection round 1 to round 3 compares the same controls against the same interventions, given that El Girba and River Atbara did not switch from control to intervention in round 2 and round 3 respectively, as shown in Figure 6. By round 4, all other control areas have switched to intervention. In this round, we performed the vertical analysis by considering Aroma and Telkuk as the main intervention group and then taking rural Kassala, Kassala, El Girba and River Atbara as the comparison group. The round 4 vertical comparison therefore compares the areas with longest exposure to the treatment with the areas that have more recently been exposed.

Figure 9: Comparison of control and intervention groups per study round for each category of acute malnutrition



Note: Statistically significant differences between outcome measures are always indicated by a red asterisk on the plot.

Table 5: Intervention and control prevalence estimates and difference between estimates by child acute malnutrition category and study round

Acute malnutrition category	Study Round	Intervention			Control			Difference			p
		Estimate	95% LCL	95% UCL	Estimate	95% LCL	95% UCL	Estimate	95% LCL	95% UCL	
SAM	1	1.75%	1.03%	2.62%	1.79%	1.35%	2.34%	-0.0005	-0.0096	0.0094	0.9244
	2	2.65%	1.80%	3.64%	2.84%	2.19%	3.48%	-0.0017	-0.0125	0.0101	0.7631
	3	4.11%	3.28%	4.99%	2.25%	1.62%	2.94%	0.0185	0.0077	0.0296	0.0009
	4	4.30%	3.38%	5.24%	5.55%	4.03%	7.24%	-0.0122	-0.0319	0.0051	0.1951
MAM	1	6.24%	4.89%	7.91%	5.64%	4.77%	6.55%	0.0061	-0.0112	0.0254	0.513
	2	10.39%	9.05%	11.71%	8.40%	7.28%	9.43%	0.0201	0.0037	0.0367	0.0167
	3	8.93%	7.55%	10.41%	7.29%	6.27%	8.48%	0.0161	-0.0028	0.0343	0.0896
	4	10.49%	9.31%	11.70%	7.32%	5.55%	9.42%	0.0316	0.0084	0.0533	0.0058
GAM	1	7.97%	6.58%	9.65%	7.40%	6.48%	8.59%	0.0056	-0.0115	0.0251	0.5454
	2	13.05%	11.49%	14.65%	11.25%	9.78%	12.60%	0.0184	-0.0032	0.0393	0.0903
	3	13.04%	11.37%	15.07%	9.55%	8.22%	11.09%	0.0351	0.0113	0.0587	0.0037
	4	14.78%	13.20%	16.26%	12.98%	10.53%	15.46%	0.0177	-0.0128	0.0464	0.2415
At-risk	1	38.58%	32.71%	44.91%	33.03%	29.82%	37.02%	0.0549	-0.016	0.127	0.1321
	2	33.28%	28.94%	38.51%	39.05%	35.84%	42.45%	-0.0582	-0.1128	0.004	0.0507
	3	29.90%	26.25%	33.42%	39.34%	35.61%	42.81%	-0.094	-0.1443	-0.041	0.0004
	4	35.27%	32.06%	38.65%	41.36%	35.62%	46.93%	-0.0605	-0.1274	0.0064	0.0761

Figure 10 presents the results produced from the horizontal comparison analysis performed on the data for every study step.[28] Table 6 shows the same results along with the value for the correction factor to adjust the difference, all with 95% confidence intervals.

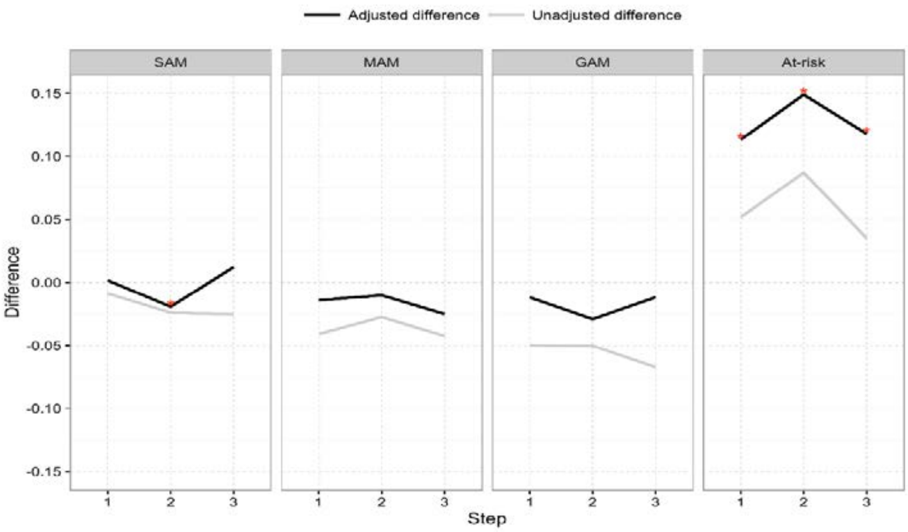
Both the unadjusted difference (before-and-after difference for intervention groups) and the adjusted difference (before-and-after difference for the intervention group adjusted by the before-and-after difference for the corresponding control group) are shown in Figure 6 and Table 5. For the outcome measure of prevalence, a negative difference indicates an increase in prevalence (‘after’ estimates are higher than ‘before’ estimates), while a positive difference indicates a decrease in prevalence (‘before’ estimates are higher than ‘after’ estimates).

For MAM, SAM and GAM, the difference is below 0 (negative) at almost every round (except for SAM, which in round 1 is almost at 0 (no difference) and in round 3 is just above 0 (slight decrease). This indicates that prevalence has increased over time at each round though it is not statistically significant.

For the ‘at risk’ category however, the difference is positive, which indicates a decrease in prevalence at every round, even after adjustment for secular trend. The decrease at each round is statistically significant.

Figure 10 also illustrates that the trend for unadjusted difference is consistent with the increasing trend for prevalence in the control group shown in the vertical comparison analysis. This indicates that the underlying prevalence of acute malnutrition was increasing over the time period of the study. This means that for any positive change (i.e. decrease in acute malnutrition) to manifest due to intervention, the effect would need to be quite large to counteract secular trend.

Figure 10: **Before-and-after difference (adjusted and unadjusted) in prevalence estimates for each outcome measure by study step**



[28] For the purposes of the analysis, we defined a study step as the study period between each round and round 1: step 0 is the starting round (i.e. round 1); step 1 is the time period between round 1 and round 2; step 2 is the time period between round 1 and round 3; and step 3 is the time period between round 1 and round 4.

Table 6: Before-and-after difference (unadjusted and adjusted) for each acute malnutrition category within each study step

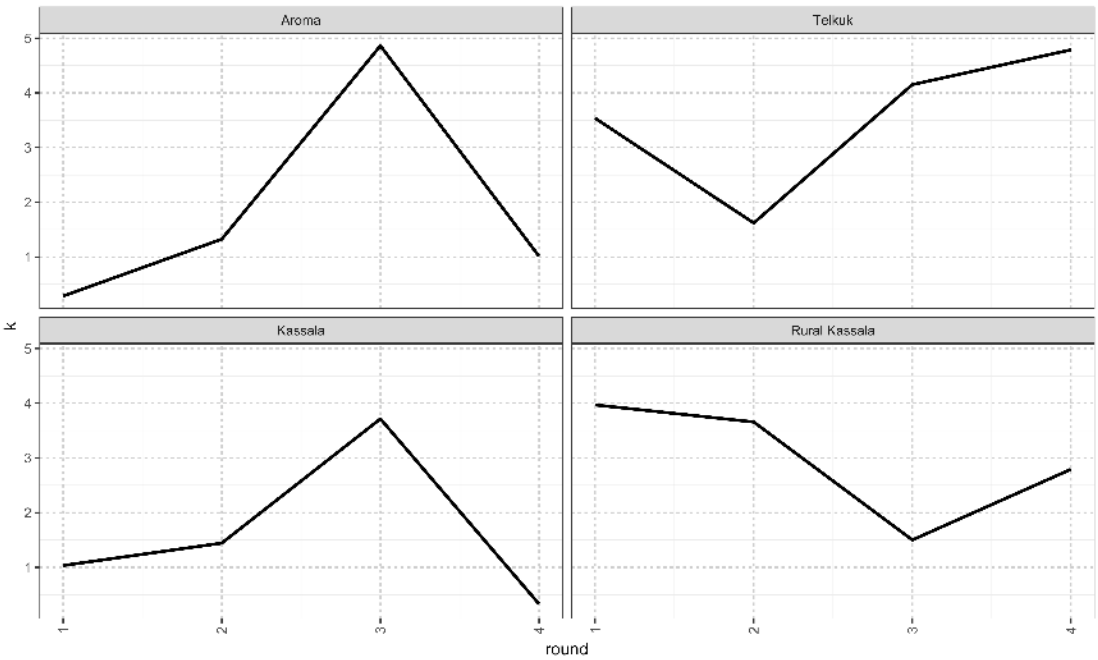
Acute malnutrition category	Study round comparison	Unadjusted difference			Difference correction (secular trend)			Adjusted difference			p
		Estimate	95% LCL	95% UCL	Estimate	95% LCL	95% UCL	Estimate	95% LCL	95% UCL	
SAM	Step 1	-0.0089	-0.0214	0.0038	-0.0106	-0.0184	-0.0017	0.0013	-0.0138	0.0167	0.8625
	Step 2	-0.0234	-0.0349	-0.0114	-0.0048	-0.0131	0.0032	-0.0186	-0.0328	-0.0041	0.0112
	Step 3	-0.0252	-0.0383	-0.0126	-0.0375	-0.0556	-0.0219	0.0126	-0.0074	0.0336	0.2287
MAM	Step 1	-0.0418	-0.0602	-0.0210	-0.0272	-0.0409	-0.0138	-0.0141	-0.0382	0.0101	0.2531
	Step 2	-0.0269	-0.0467	-0.0065	-0.0163	-0.0299	-0.0024	-0.0104	-0.0352	0.0134	0.3989
	Step 3	-0.0427	-0.0616	-0.0237	-0.0177	-0.0405	0.0032	-0.0252	-0.0534	0.0047	0.0884
GAM	Step 1	-0.0501	-0.0715	-0.0280	-0.0381	-0.0548	-0.0201	-0.0121	-0.0409	0.0164	0.4059
	Step 2	-0.0509	-0.0738	-0.0268	-0.0216	-0.0399	-0.0038	-0.0293	-0.0579	0.0027	0.0578
	Step 3	-0.0681	-0.0892	-0.0461	-0.0551	-0.0826	-0.0293	-0.0129	-0.0463	0.0240	0.4716
At-risk	Step 1	0.0519	-0.0272	0.1270	-0.0605	-0.1103	-0.0096	0.1117	0.0207	0.2052	0.0176
	Step 2	0.0889	0.0136	0.1632	-0.0620	-0.1100	-0.0090	0.1501	0.0624	0.2387	0.0008
	Step 3	0.0329	-0.0376	0.1011	-0.0825	-0.1500	-0.0110	0.1164	0.0167	0.2118	0.0193

6.3.2 Impact on incidence

Figure 11 presents the estimated incidence rate for each of the intervention (rural Aroma and Telkuk) and control (Kassala and rural Kassala) clusters studied and followed for approximately five months. Incidence is reported as an incidence correction factor *k* used to estimate new acute malnutrition cases based on prevalence estimates (Isanaka et al. 2016; Dale et al. 2017).

There is no discernible difference between control and intervention clusters with regard to incidence, and no recognisable pattern or trend in incidence. There is no statistically significant difference between the incidence rates in each cluster and between control and intervention.

Figure 11: Incidence rate of global acute malnutrition



6.3.3 Impact on prevalence – pregnant and lactating women

We now present study results for PLW. Figure 12 illustrates the prevalence of PLW GAM and PLW at risk per study round. We observed a general upward trend over time in the prevalence of PLW GAM and PLW at risk in both intervention and control groups. The prevalence in both acute malnutrition categories was higher in the intervention group compared to the control group in data collection rounds 1, 2 and 4 with a statistically significant difference between the two groups in round 2 and round 4. In round 3, we observed a drop in prevalence of PLW GAM and PLW at risk in the intervention group, with prevalence rates going slightly lower, though not statistically significant, than those in the control group. However, when taking into account secular trend, we noted that the results of the before-and-after comparison (see Figure 13) showed a general decrease in PLW GAM (up to almost 5% though not statistically significant) and PLW at risk (up to almost 15% with a significant decrease in step 2).

Figure 12: Comparison of control and intervention groups per study round for each category of PLW acute malnutrition

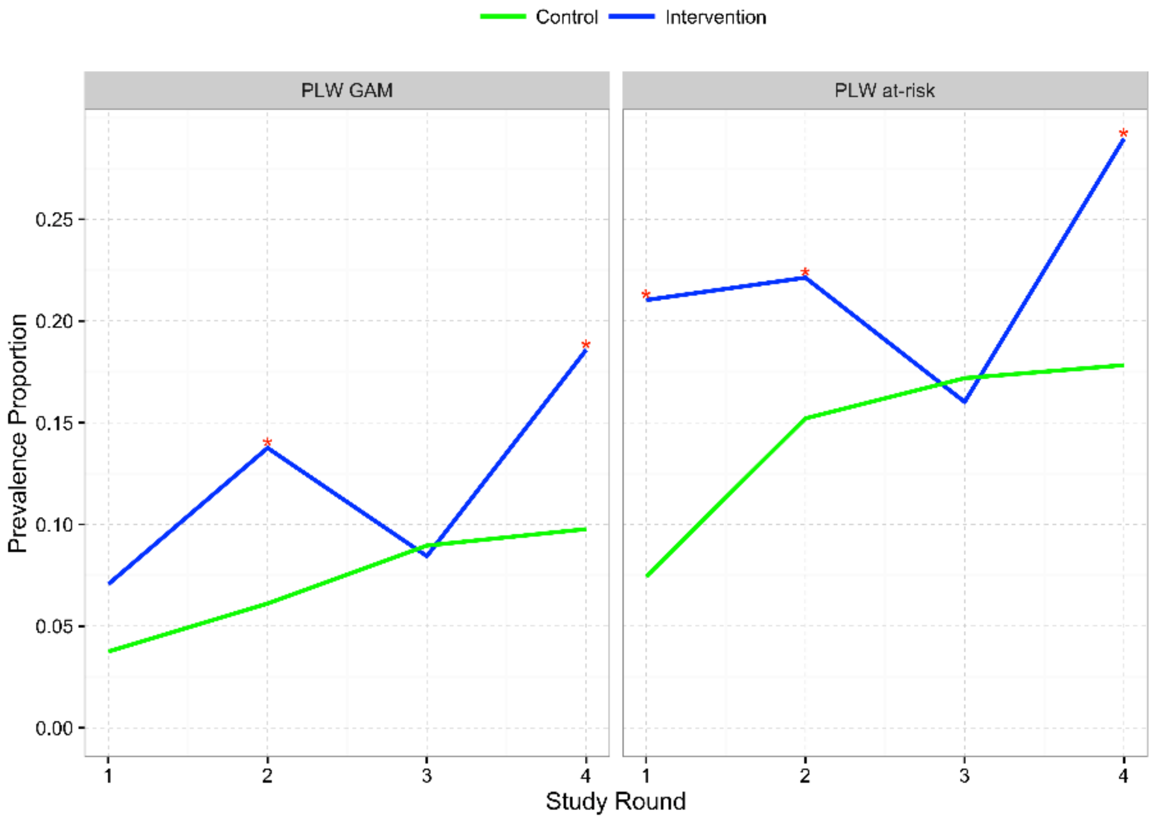
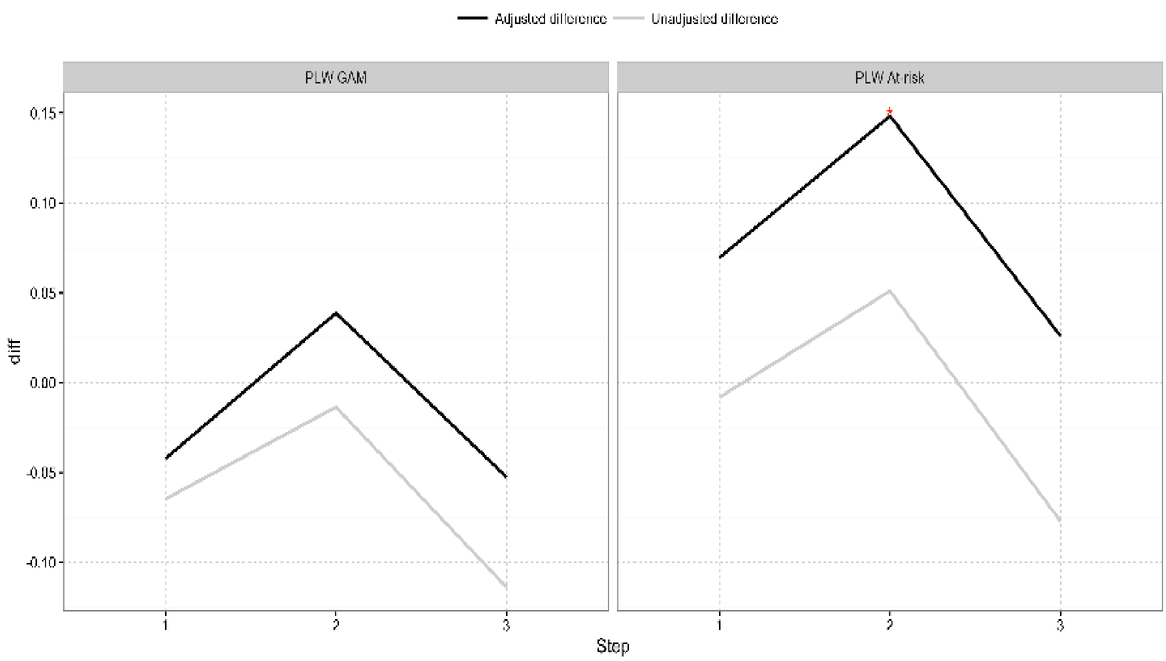


Figure 13: Before-and-after difference (adjusted and unadjusted) in prevalence estimates for each PLW outcome measure by study step



This pattern of decreasing prevalence in PLW GAM and PLW at risk in the intervention group is similar to what we found in the child sample.

Table 7: Intervention and control prevalence estimates and difference between estimates by PLW acute malnutrition category and study round

PLW malnutrition category	Study Round	Intervention			Control			Difference			p
		Estimate	95% LCL	95% UCL	Estimate	95% LCL	95% UCL	Estimate	95% LCL	95% UCL	
GAM	1	7.06%	3.95%	11.50%	3.75%	2.17%	6.08%	3.34%	-0.63%	7.91%	0.1255
	2	13.76%	11.14%	16.56%	6.11%	4.24%	7.84%	7.64%	4.67%	10.93%	0
	3	8.44%	6.89%	10.23%	8.96%	6.53%	11.50%	-0.53%	-3.54%	2.47%	0.7314
	4	18.58%	16.32%	21.27%	9.77%	7.04%	13.09%	8.82%	4.79%	12.65%	0
At-risk	1	21.03%	16.00%	25.88%	7.43%	5.46%	10.28%	13.56%	7.97%	18.83%	0
	2	22.13%	19.34%	24.91%	15.21%	12.04%	19.21%	6.95%	2.03%	11.17%	0.0029
	3	16.02%	14.00%	18.40%	17.19%	13.73%	20.86%	-1.26%	-5.24%	3.13%	0.556
	4	28.96%	25.50%	32.25%	17.83%	14.62%	21.61%	11.11%	5.89%	15.47%	0

Table 8: Before-and-after difference (unadjusted and adjusted) for each PLW acute malnutrition category for each study step

Acute malnutrition category	Study round comparison	Unadjusted difference			Difference correction (secular trend)			Adjusted difference			p
		Estimate	95% LCL	95% UCL	Estimate	95% LCL	95% UCL	Estimate	95% LCL	95% UCL	
SAM	Step 1	-0.0648	-0.1101	-0.0137	-0.0229	-0.0467	0.0069	-0.0423	-0.0964	0.0123	0.1267
	Step 2	-0.0136	-0.0509	0.0376	-0.0516	-0.0818	-0.0189	0.0387	-0.0114	0.0951	0.1539
	Step 3	-0.1139	-0.1571	-0.0624	-0.0609	-0.0963	-0.0275	-0.0526	-0.1098	0.0092	0.0833
At-risk	Step 1	-0.0072	-0.066	0.045	-0.079	-0.1238	-0.0341	0.0703	0.0012	0.1408	0.0483
	Step 2	0.0505	-0.0056	0.1033	-0.0964	-0.1392	-0.0524	0.1473	0.078	0.2161	0
	Step 3	-0.0753	-0.137	-0.0176	-0.1034	-0.1466	-0.0609	0.0292	-0.0504	0.0939	0.4281

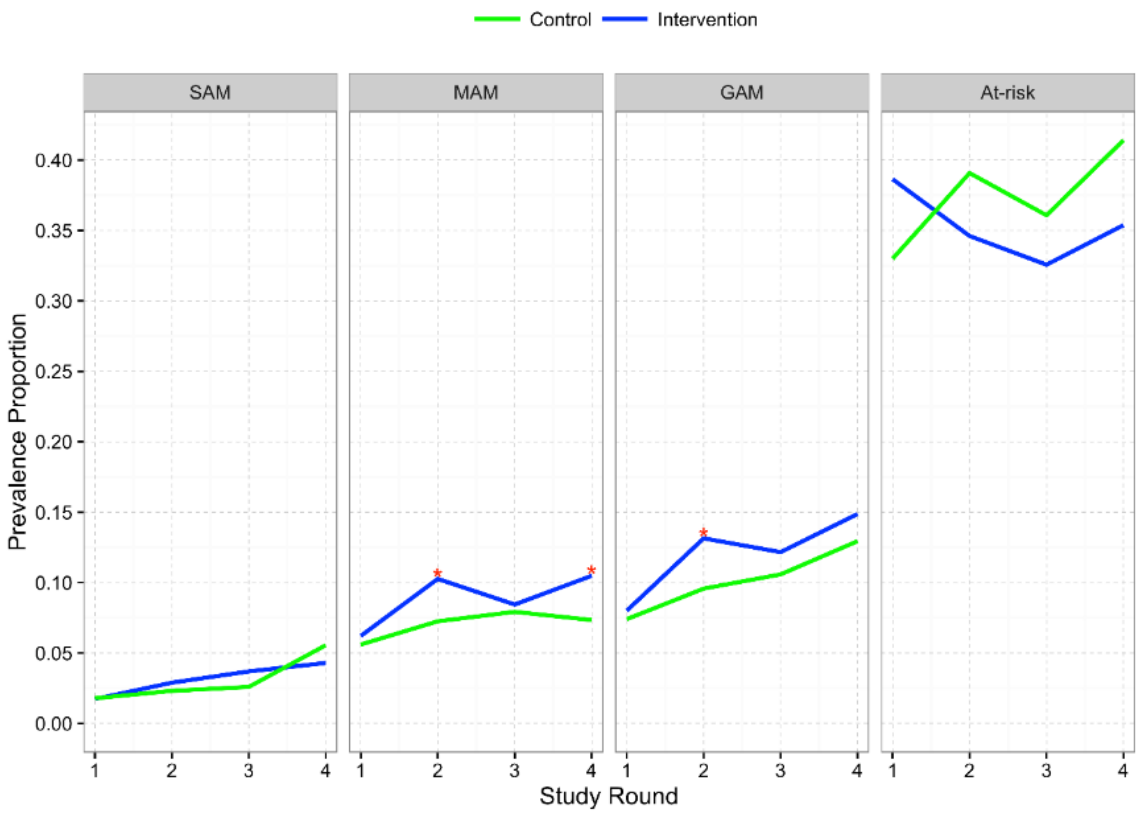
6.4 Sub-question 1: How are these impacts affected by different intervention modalities in terms of product used, delivery of service, duration of intervention and coverage?

To answer this sub-question, we performed additional analyses of data collected through the study and from secondary data (primarily routine programme monitoring data). In addition, a timeline of key programmatic events (see Section 4) was developed to contextualise the impact results, and to understand how any change in intervention modality may have affected the observed impact. This section therefore examines how the MAM treatment and prevention programme components were implemented, the performance of the delivery of the service, and the duration and coverage of the programmes.

6.4.1 Treatment and prevention intervention modality

Modality in terms of product used changed only once during the study period, as shown in the timeline. All clusters were exposed to TSFP using SC+ up to September 2016, after which Ready to Use Supplementary Food (RUSF) was used. The FBMAM programme was initially planned to roll out as per the stepped-wedge design; however this only occurred in rural Aroma and Telkuk (the first two clusters to receive intervention), and in rural Kassala and Kassala (the last to switch to intervention). El Girba and River Atbara, which were meant to switch to intervention in rounds 2 and round 3, respectively, only received intervention in round 4.

Figure 14: Comparison between children's acute malnutrition prevalence in intervention and control using an intention to treat analysis



These changes in the implementation rollout were considered in the previously-mentioned results; hence the impact reported is based on actual implementation. We also performed an intention-to-treat analysis. Figures 14 and 15 present results for vertical and horizontal analysis, respectively.

In the intention to treat analysis for the vertical comparison, the change in the number at risk that was previously noted is still present but not statistically significant. The increasing trend for SAM, MAM and GAM prevalence continued both for control and intervention, with intervention showing a higher prevalence than the controls. This is only significant in some data rounds.

We observed a similar trend in acute malnutrition prevalence and difference in prevalence over time, as shown in Figures 16 and 17.

Figure 15: Difference (adjusted and unadjusted) in children's acute malnutrition prevalence using an intention-to-treat analysis

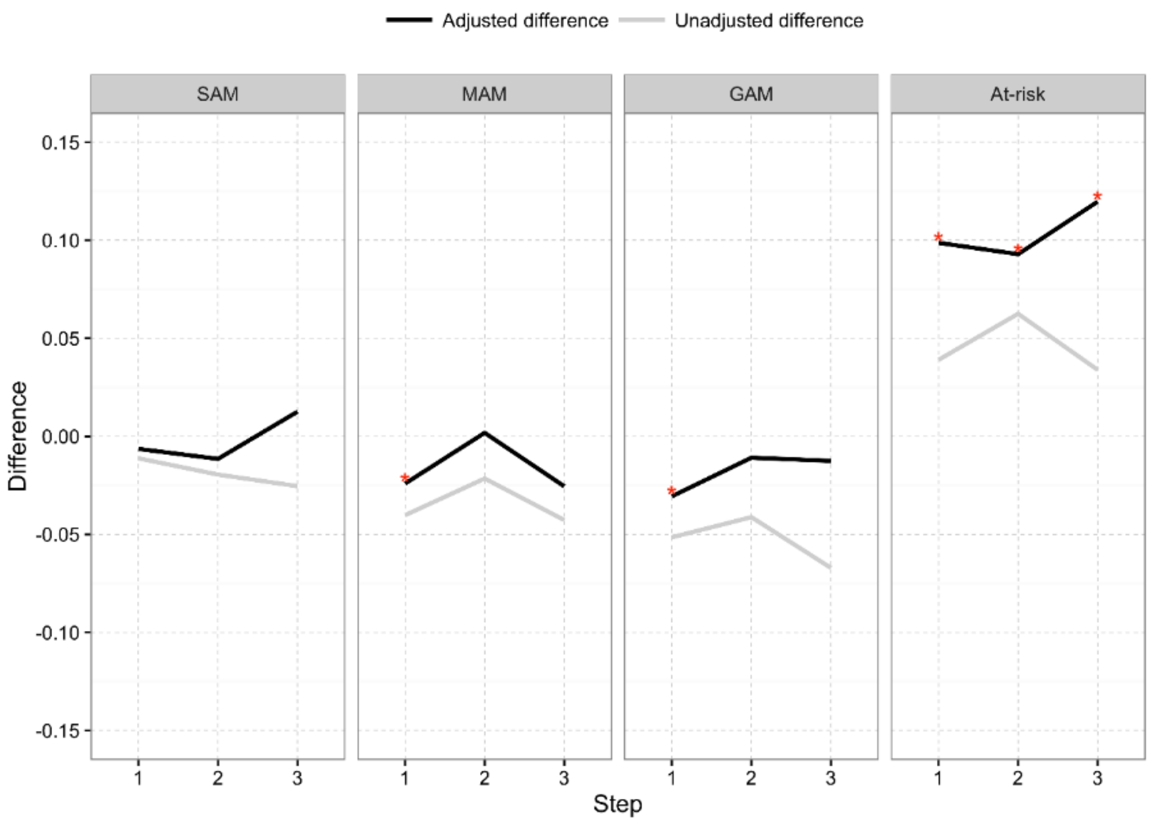


Figure 16: Comparison between PLW acute malnutrition prevalence in intervention and control using an intention-to-treat analysis

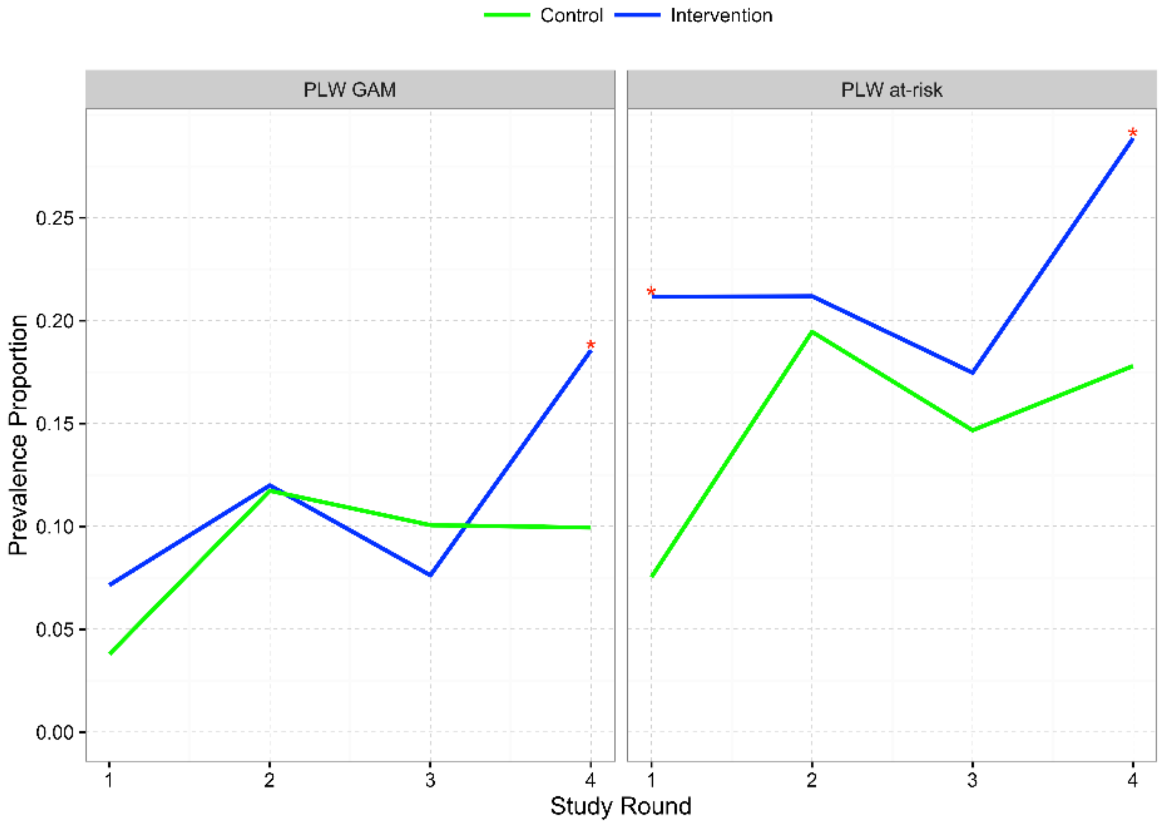
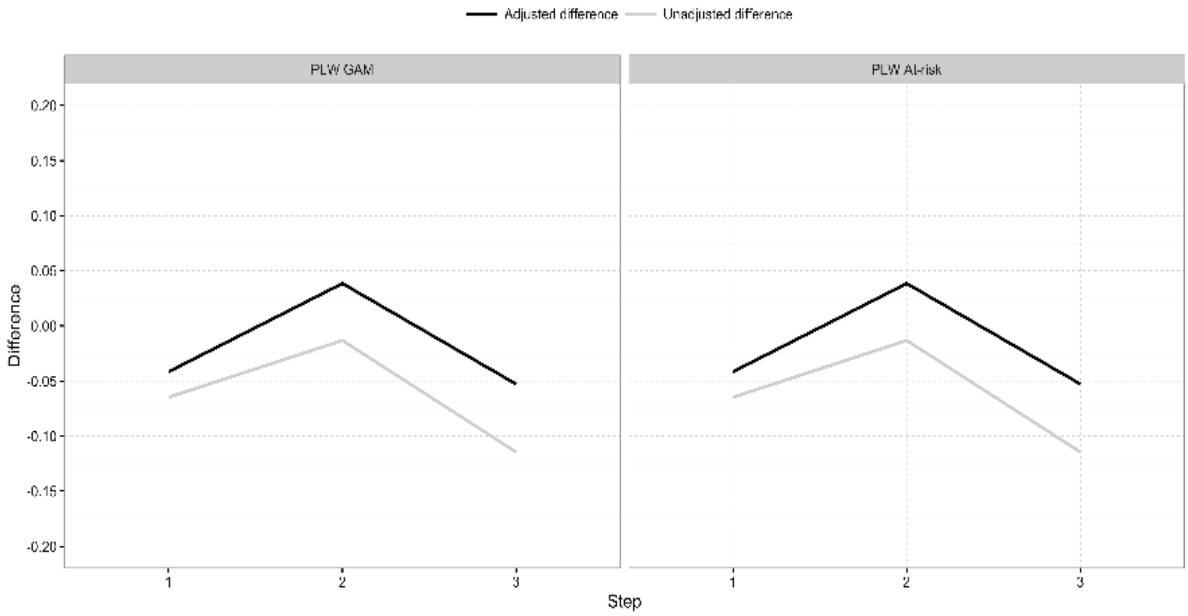


Figure 17: Difference (adjusted and unadjusted) in PLW acute malnutrition prevalence using an intention-to-treat analysis



6.4.2 Delivery of service

We assessed the performance of the MAM treatment and prevention programmes in delivering services using the prescribed standard performance indicators against routine programme monitoring data. We assessed the performance of the TSFP to determine whether the pre-existing intervention, to which FBMAM was added, was functioning at the same level across all study clusters. If TSFP performance varied across clusters, this could potentially confound the impact results. We undertook a similar assessment for FBMAM based on the same rationale that programme performance (as an indicator of how the intervention was implemented) can explain any change or difference noted.

TSFP admissions over time and defaulters over time

Admissions over time and defaulters over time are time-series analyses used to assess the trend in TSFP admissions and defaulters. The TSFP is responsive to need when the trend of admissions- and defaulters over time shows an appropriate upward or downward change due to events in the same time period affecting service access or retention and compliance with treatment. For example, events such as outbreaks of acute watery diarrhoea are known to increase the prevalence of wasting. Therefore, with a responsive TSFP we would expect to see an increase in the number of admissions at the time of the outbreak (or immediately after). Defaulting, on the other hand, may show patterns of increase when events such as floods impede access. However a responsive programme will be able to minimise the impact of such events and maintain low levels of defaulting even in times of stress.

Figure 18 and 19 below present child admissions and defaulters over a one-year period (2016).

Figure 18: TSFP admissions over time – children

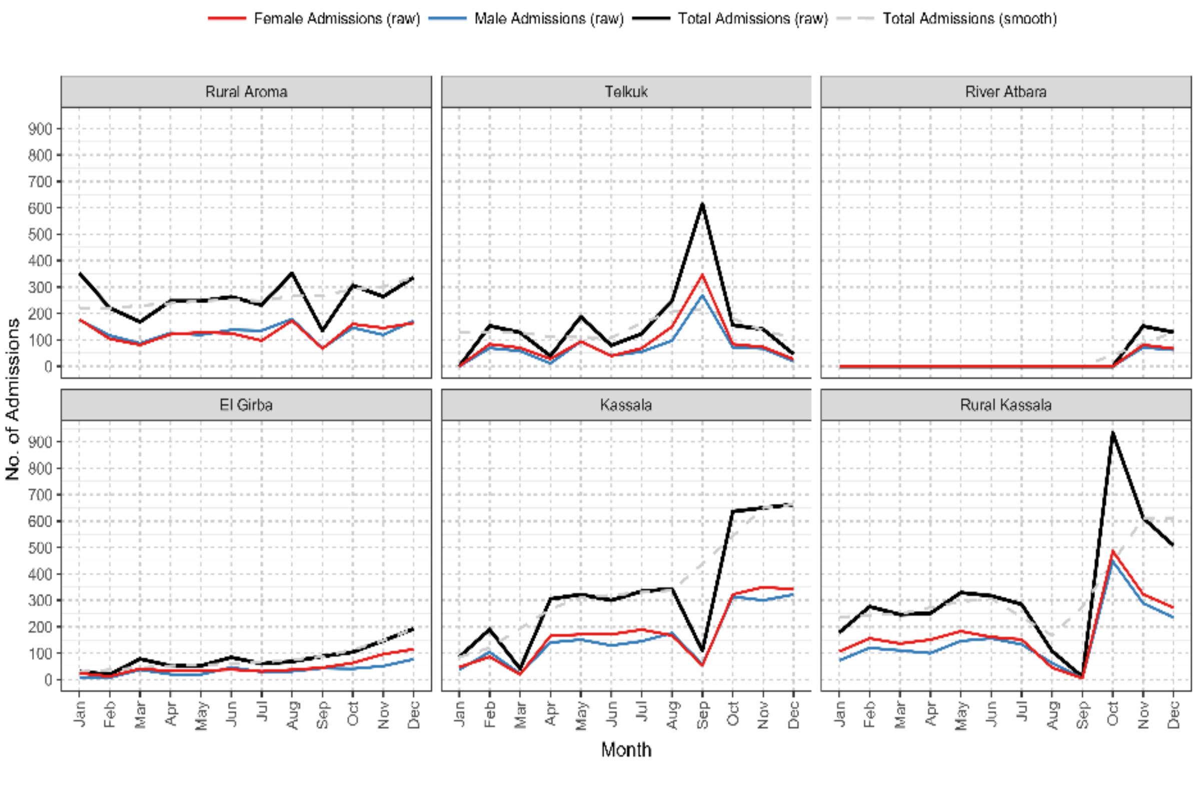


Figure 19: TSFP defaulters over time – children

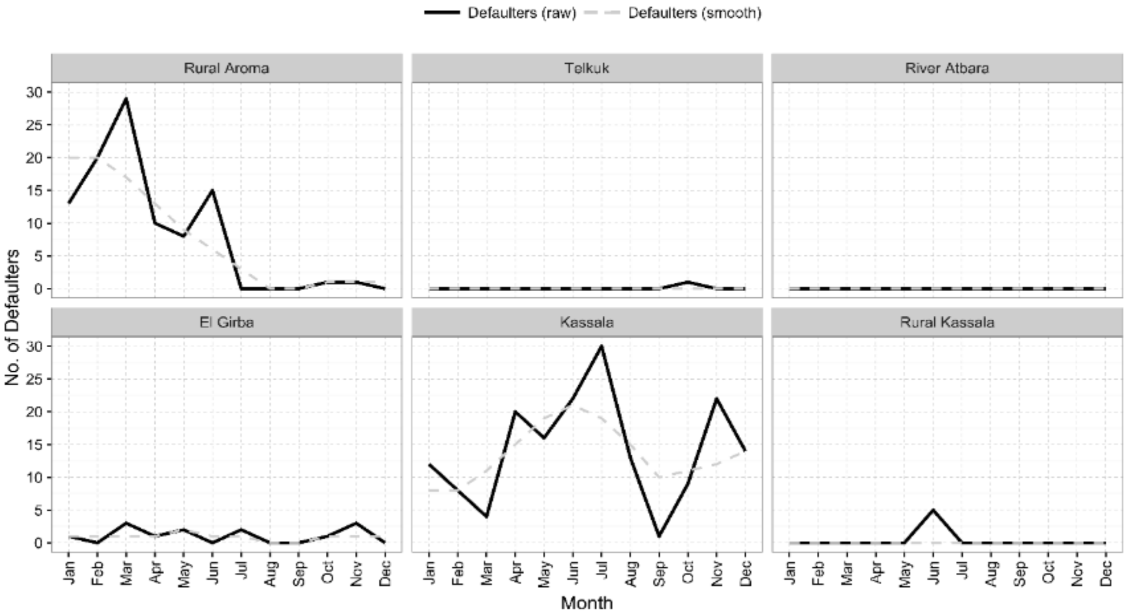


Figure 18 presents varying patterns of children's MAM admissions per cluster. In Telkuk, Kassala and rural Kassala, there are significant peaks around September and October – immediately after heavy rains, flooding and episodes of acute watery diarrhoea in the state. This also directly followed a mass screening undertaken by the SMoH across the whole of the state and supported by UNICEF (in September 2016). However rural Aroma seems static in pattern/trend, with only a slight increase in admissions toward the last few months of 2016. There is a similar pattern in El Girba, but with lower admission numbers. River Atbara is different from the rest, as it appears to show no admissions until November 2016. This indicates that there was no TSFP at all from January to October 2016.

The patterns of admissions presented in each cluster appear to indicate a less-than-responsive programme compared to need. Triggers to increase admission seem connected to one-off screening efforts rather than a consistent and ongoing effort to register all or nearly all prevalent and incident cases routinely and regularly. We also observed no difference in trend of admissions over time between males and females.

There is also a varying pattern per cluster in Figure 19 for defaulters over time. Rural Aroma and Kassala, which have the highest number of admissions, also have the most defaulters (up to about 30% in certain periods). However, there is no notable pattern of defaulting that can be associated with events known to affect attendance. During periods of heavy rain and flooding, levels of defaulting remained stable. Telkuk, River Atbara and rural Kassala show near-zero defaulters. This is particularly striking for rural Aroma and rural Kassala, given that they also have high levels of admissions. El Girba has a low number of defaulters that remains steady over time. In relation to the number of admissions, this seems a consistent pattern. Based on these trends, the programme seems to be very responsive in terms of defaulting by ensuring that those admitted to the programme remain in the programme until qualifying for discharge. This is an important aspect of programme quality and is a factor that could support good programme coverage.

Figure 20: TSFP admissions over time – PLW

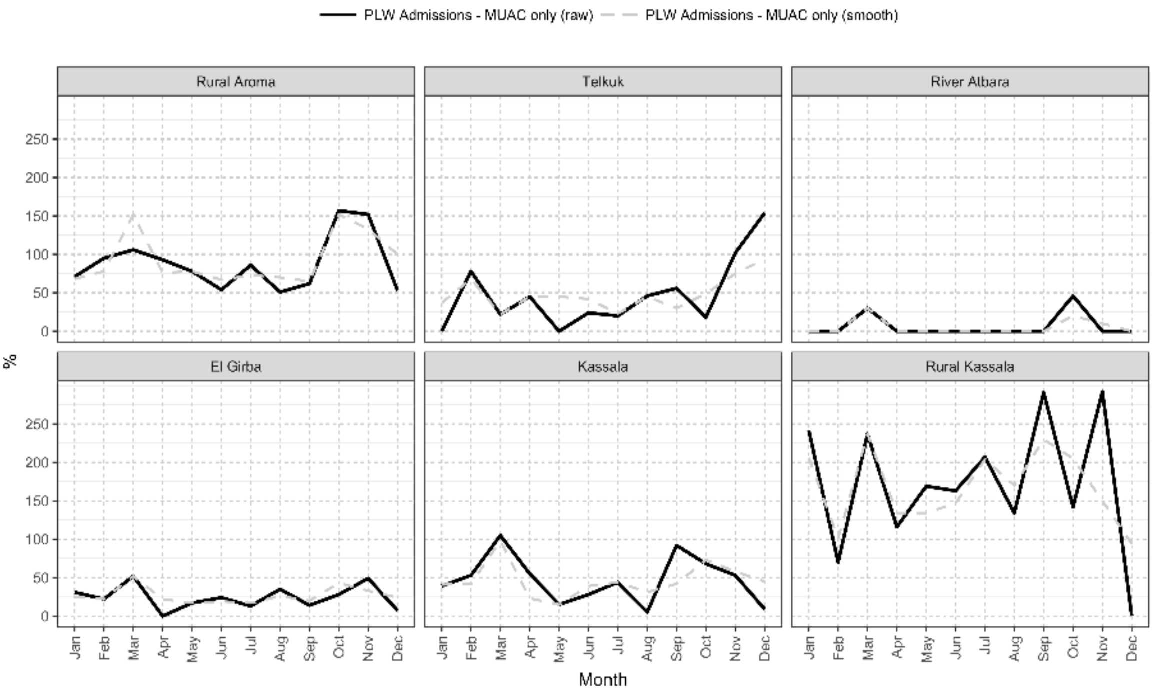


Figure 21: TSFP defaulters over time – PLW

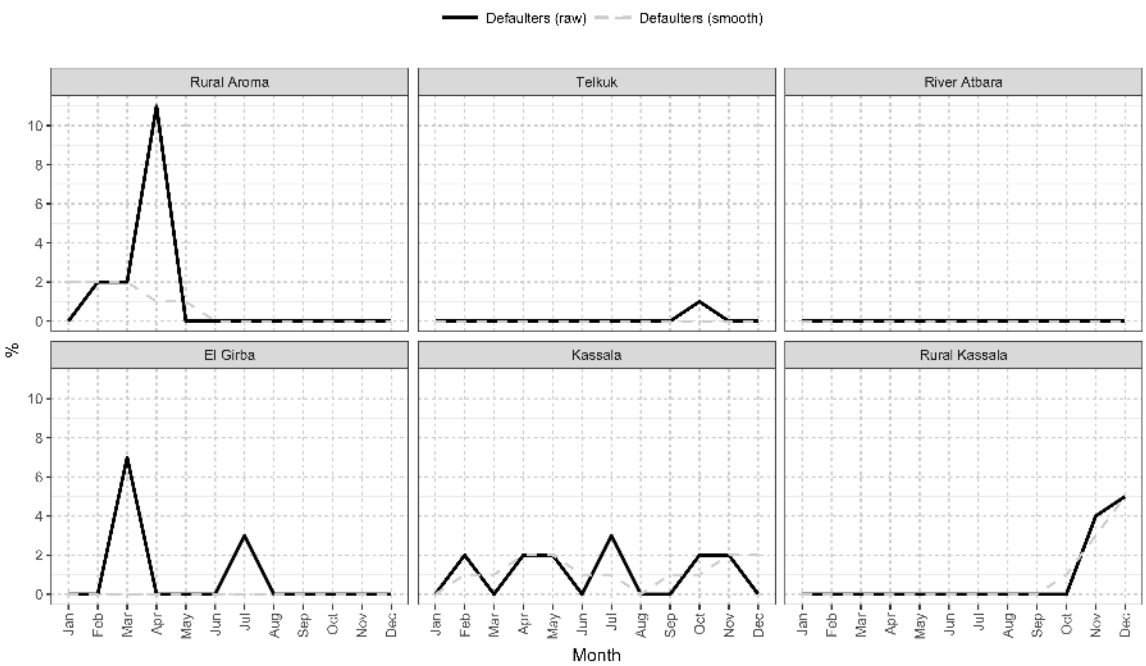


Figure 20 presents trends of PLW admissions over time. Unlike child admissions, we did not observe any obvious increasing or decreasing pattern in PLW admissions, other than an up-and-down fluctuation between months, which when smoothed indicates a generally flat trend over time. This would be indicative of erratic and inconsistent case-finding. Figure 21, on the other hand, presents a similar pattern of defaulting in PLW as with the children, with peaks of defaults early in the programme eventually tapering off in the latter period.

TSFP performance

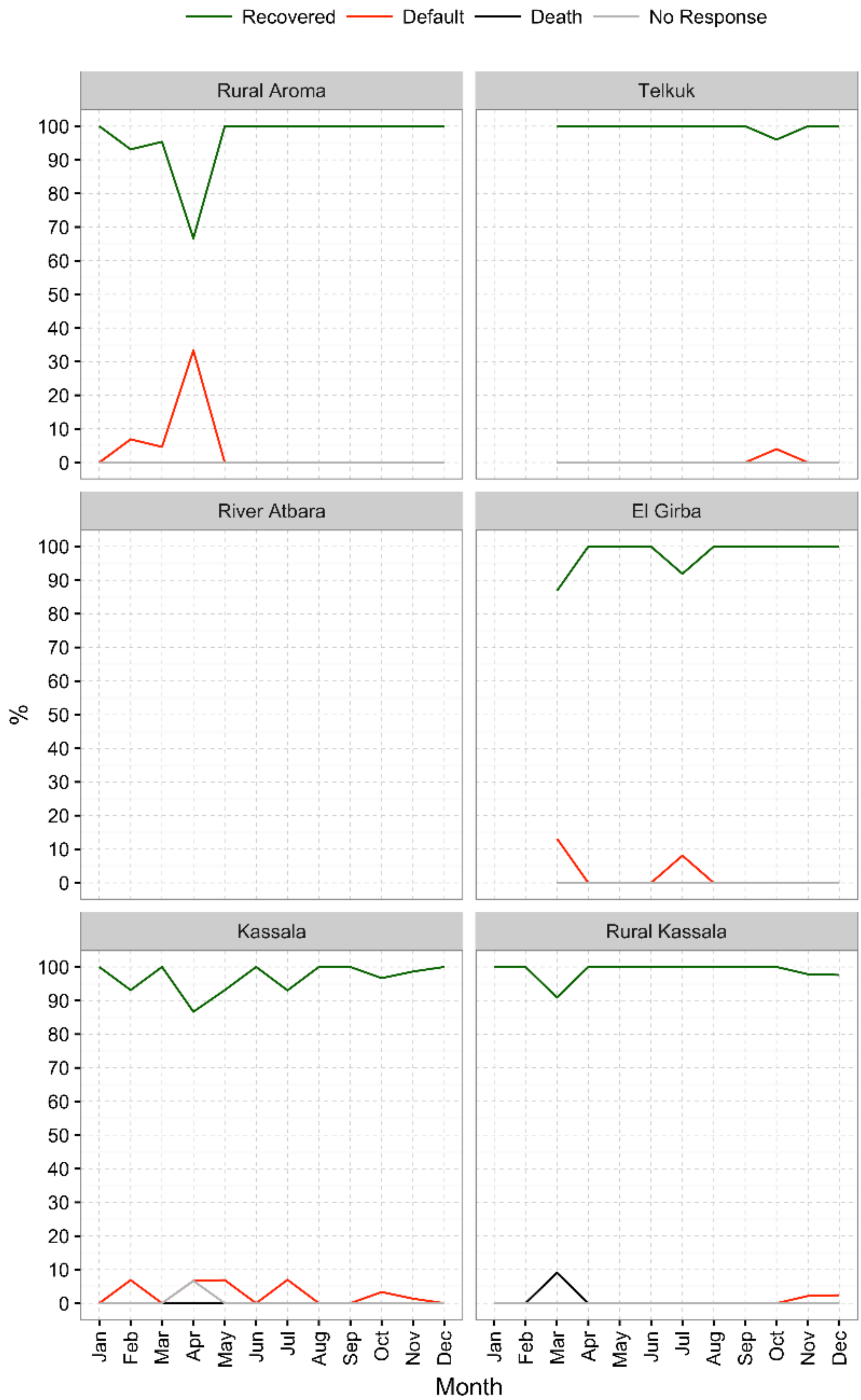
TSFP performance is assessed according to levels of exit categories such as recovery rate, default rate, death rate and non-response rate. An effective programme maximises recovery and minimises defaulting and other adverse outcomes. Therefore when performance indicators are plotted over time, the expectation is that recovery rates are high and defaulting, death and non-response rates are low. The current Sphere benchmark for assessing the acceptability of recovery rates in an emergency TSFP is 75% (The Sphere Project 2011). Figure 22 below shows the trend of recovery rates and other discharge outcomes over time for each cluster. Recovery rates across the clusters[29] are above 75% throughout the year, with defaulting staying below 15% most of the time. Based on this data, programme performance appears to be good and consistent across the different localities.

We observed a similar performance pattern for PLW (see Figure 23), with cure rates remaining high, defaulting low, no deaths and few non-responders.

[29] With the exception of River Atbara, which has not reported any discharges given that the programme appears to have begun only in November 2016.

Figure 22: TSFP performance – children

Figure 23: TSFP performance – PLW



FBMAM performance
Given the staged rollout of the intervention in each of the clusters, the routine programme monitoring data for FBMAM is patchy, with most data coming from the first 2 clusters that started as intervention (rural Aroma and Telkuk). Most discharges from the FBMAM during the study period came from these clusters, whilst there were a few in the December period for the other 4 clusters. Therefore, we have focused this assessment on admissions and defaulters over time to assess responsiveness to need, which is presented in Figure 21 and Figure 22 below.

When comparing rural Aroma and Telkuk, it seems that rural Aroma has performed well in taking on admissions, specifically in the round 3 period. Kassala and rural Kassala show peaks in admissions as soon as they turned into intervention clusters. It is important to note that earlier in the year (January and February period) admissions were made to the programme before the study had started. This was due to an initial FBMAM pilot in late 2015 that spilled over to early 2016 prior to the study. This pilot was stopped a month before the study started. This may be treated as a possible confounder, especially for the clusters that were assigned to intervention early on. These results aside, it is hard to fully assess programme implementation and delivery of service given the limited time period of implementation.

Figure 24: FBMAM admissions over time

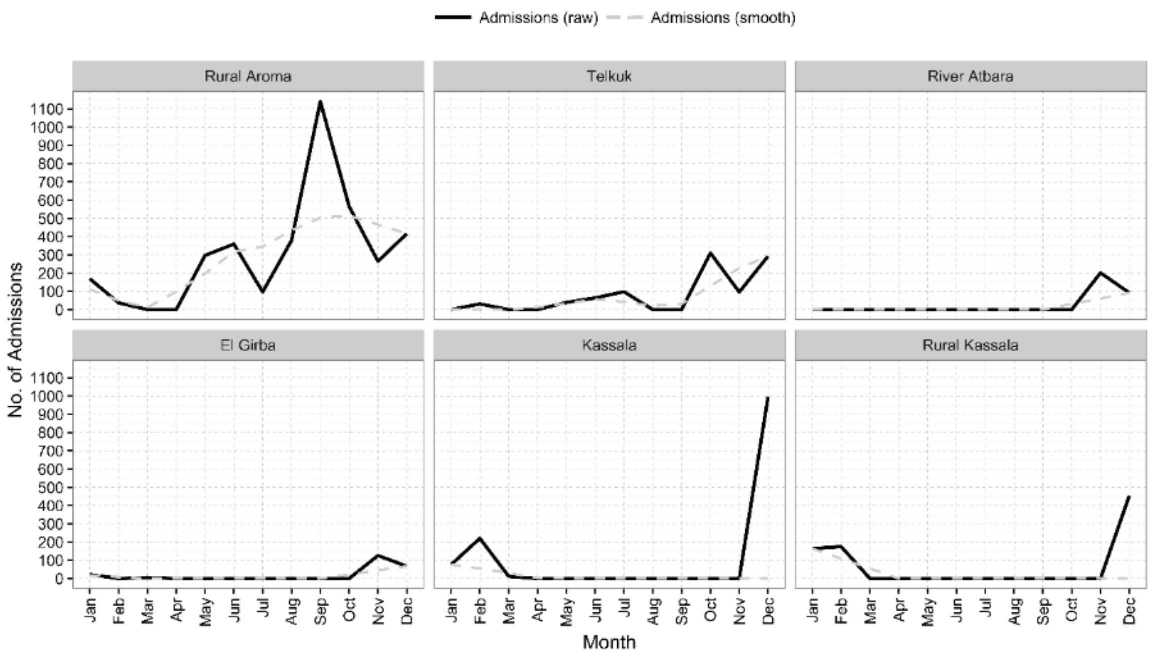
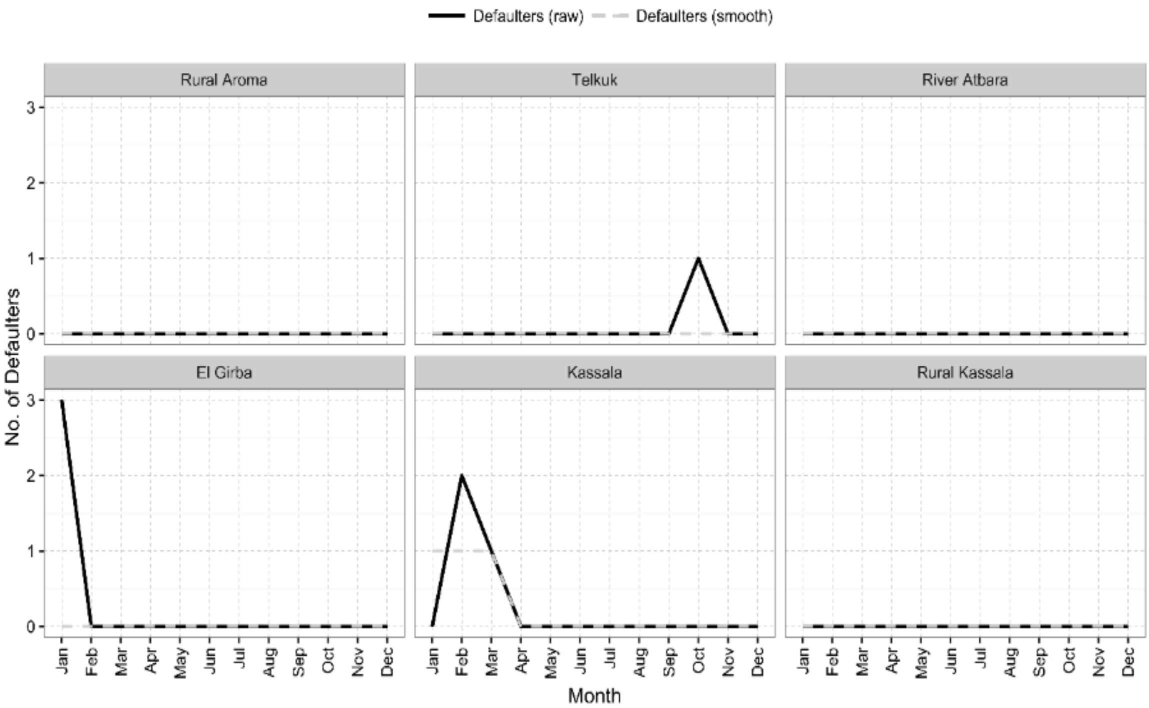


Figure 25: FBMAM defaulters over time



6.4.3 Programme coverage

TSFP coverage over time

The study collected information that would allow the assessment of programme coverage based on direct estimation methods developed for CMAM programming, which includes TSFP (Myatt et al. 2012). Figure 26 presents TSFP coverage over time by cluster (and per study round).

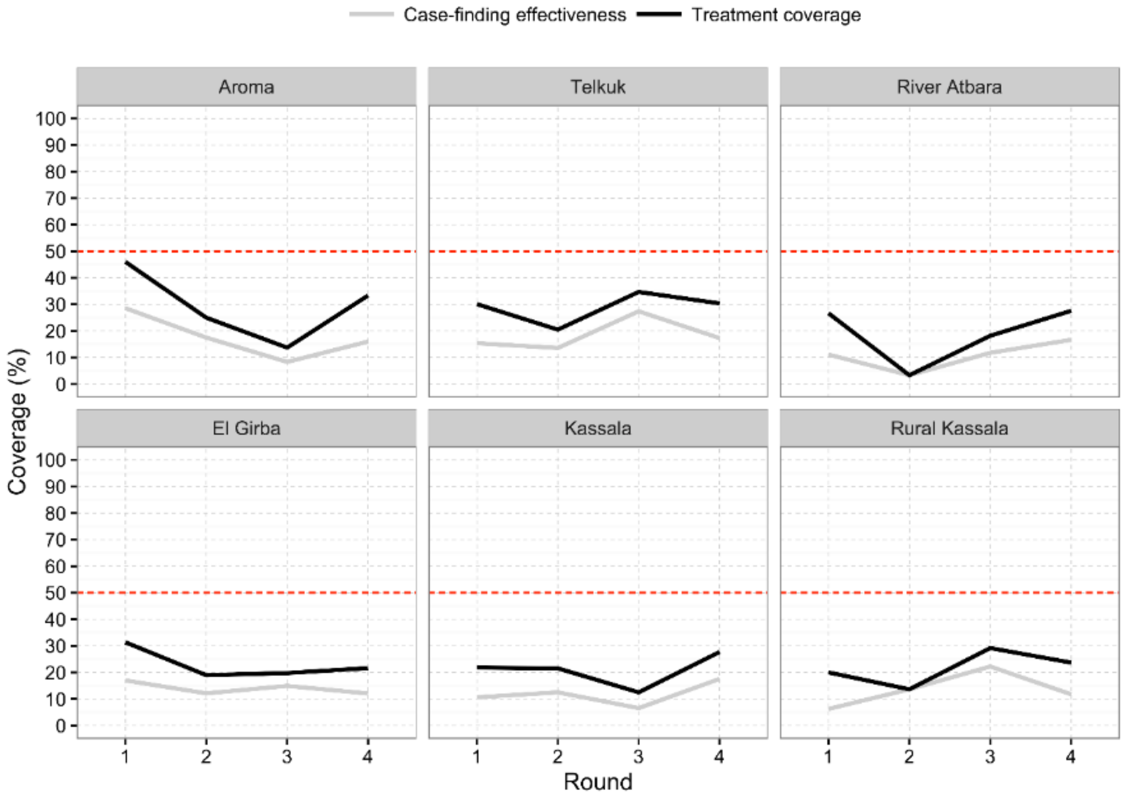
At each round of the study, coverage levels across all localities were similar. Case-finding effectiveness (grey line) indicates how effective the programme is at finding MAM cases. High case-finding effectiveness is usually a result of routine and regular active case-finding at the community level. This is often made possible by outreach workers and/or volunteers who routinely visit villages and use a MUAC tape to assess children and refer them on accordingly. A high case-finding effectiveness measure also usually indicates that programme coverage will be high. Case-finding effectiveness at each round of the study ranges from 9–12 per cent, which is indicative of a programme with a weak case-finding approach.

Treatment coverage is the programme coverage indicator, which determines whether a programme is able to find and retain MAM cases until recovery. Treatment coverage will always be higher (but not significantly) than case-finding effectiveness. Overall, treatment coverage at each round of the study ranges from 15–28 per cent (see Table 9). This is not very high, but relative to the coverage for TSFP programmes in similar countries, it is a comparable (and in some cases better) result (Shoham et al. 2013; Guevarra et al. 2015).

Table 9: Overall TSFP case-finding effectiveness and treatment coverage for children per data collection round

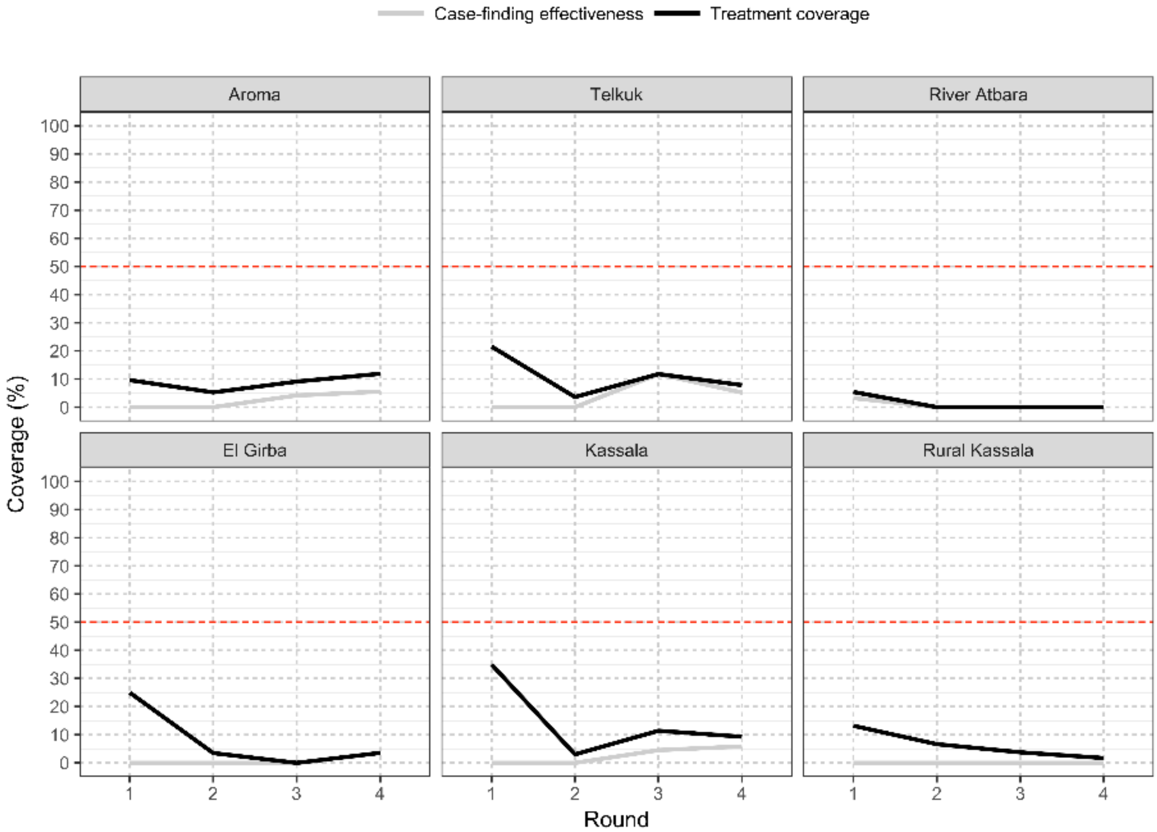
Study Round	Case-finding effectiveness			Treatment coverage		
	Estimate	95% LCL	95% UCL	Estimate	95% LCL	95% UCL
Round 1	13.4%	8.4%	18.5%	27.6%	23.6%	31.6%
Round 2	10.3%	7.0%	13.6%	15.3%	12.1%	18.5%
Round 3	15.0%	9.7%	20.3%	21.1%	15.7%	26.5%
Round 4	15.4%	9.6%	21.3%	26.4%	21.0%	31.8%

Figure 26: TSFP coverage over time – children



The temporal pattern of coverage across all localities is that of good coverage at the start of the study (round 1 period), then a dip at round 2, and a subsequent progressive increase from round 3 to round 4. Rural Aroma and Telkuk show a relatively higher level of coverage. We also observed differences in coverage achieved at the locality level compared to the overall average, with areas such as Aroma reaching nearly 50 per cent coverage at round 1.

Figure 27: TSFP coverage over time – PLW



We found a similar temporal pattern for PLW as with TSFP coverage for children: a relatively higher coverage in the start of the programme declining over time (see Figure 27). However, the magnitude of coverage (both case-finding effectiveness and treatment coverage) was lower for PLW than for children, with coverage levels far below 50%;[30] the highest coverage was slightly over 30% in Kassala during round 1 of data collection.

TSFP spatial distribution of coverage

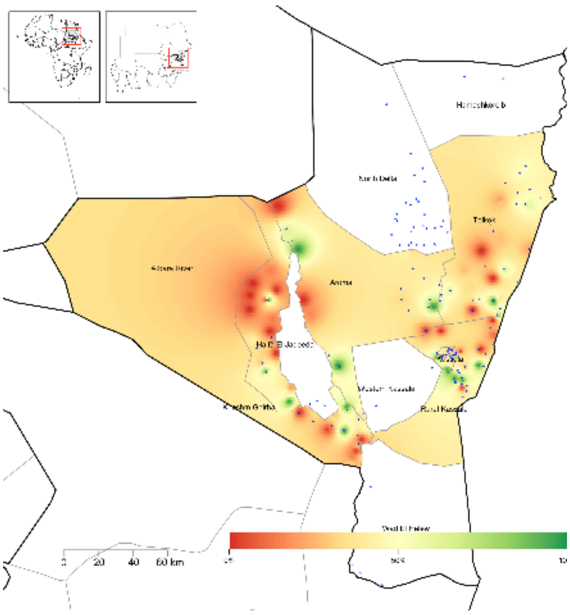
Figure 28 shows spatial variation in the coverage of children achieved by the TSFP programme across the six study localities.

The temporal pattern of coverage described earlier is consistent here, with relatively good coverage evenly distributed spatially at round 1, and poor coverage almost evenly distributed spatially at round 2. Coverage began picking up again by round 3 and round 4. Good coverage was predominantly at, or near, distribution sites/health centres providing TSFP.

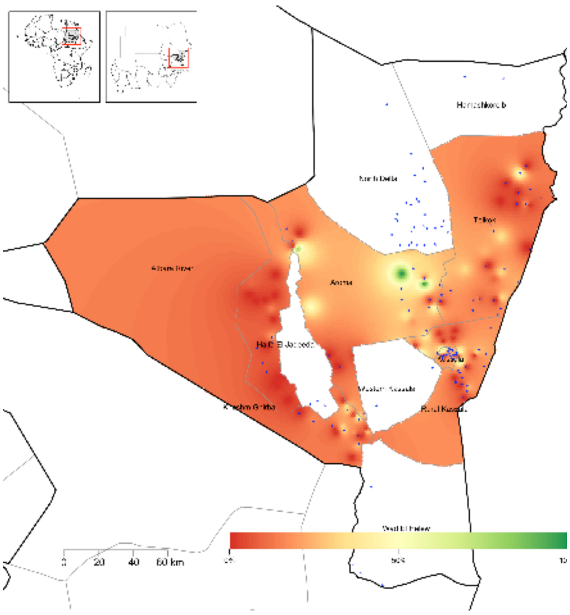
[30] Whilst there is no set Sphere standard for targeted feeding programmes for PLW, 50% is an acceptable benchmark based on experience with CMAM programming for children.

Figure 28: Spatial distribution of TSFP coverage of children by study round

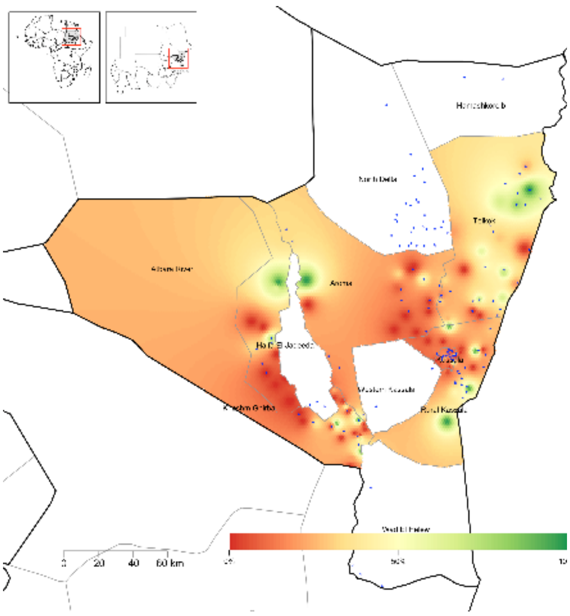
Round 1



Round 2



Round 3



Round 4

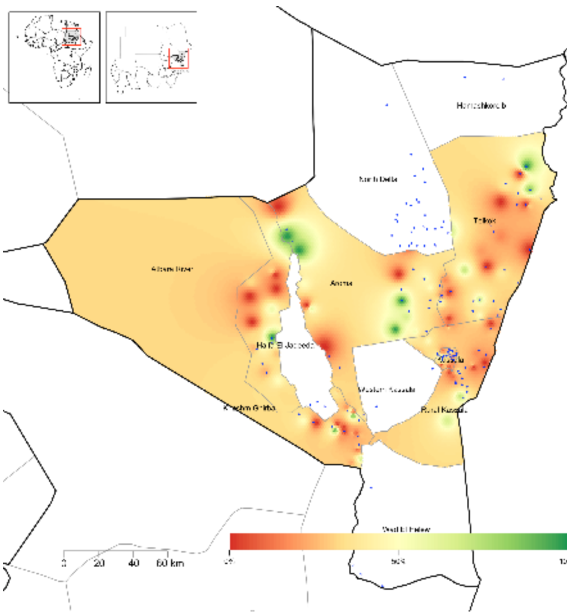


Figure 29: Spatial distribution of TSFP coverage of PLW by study round

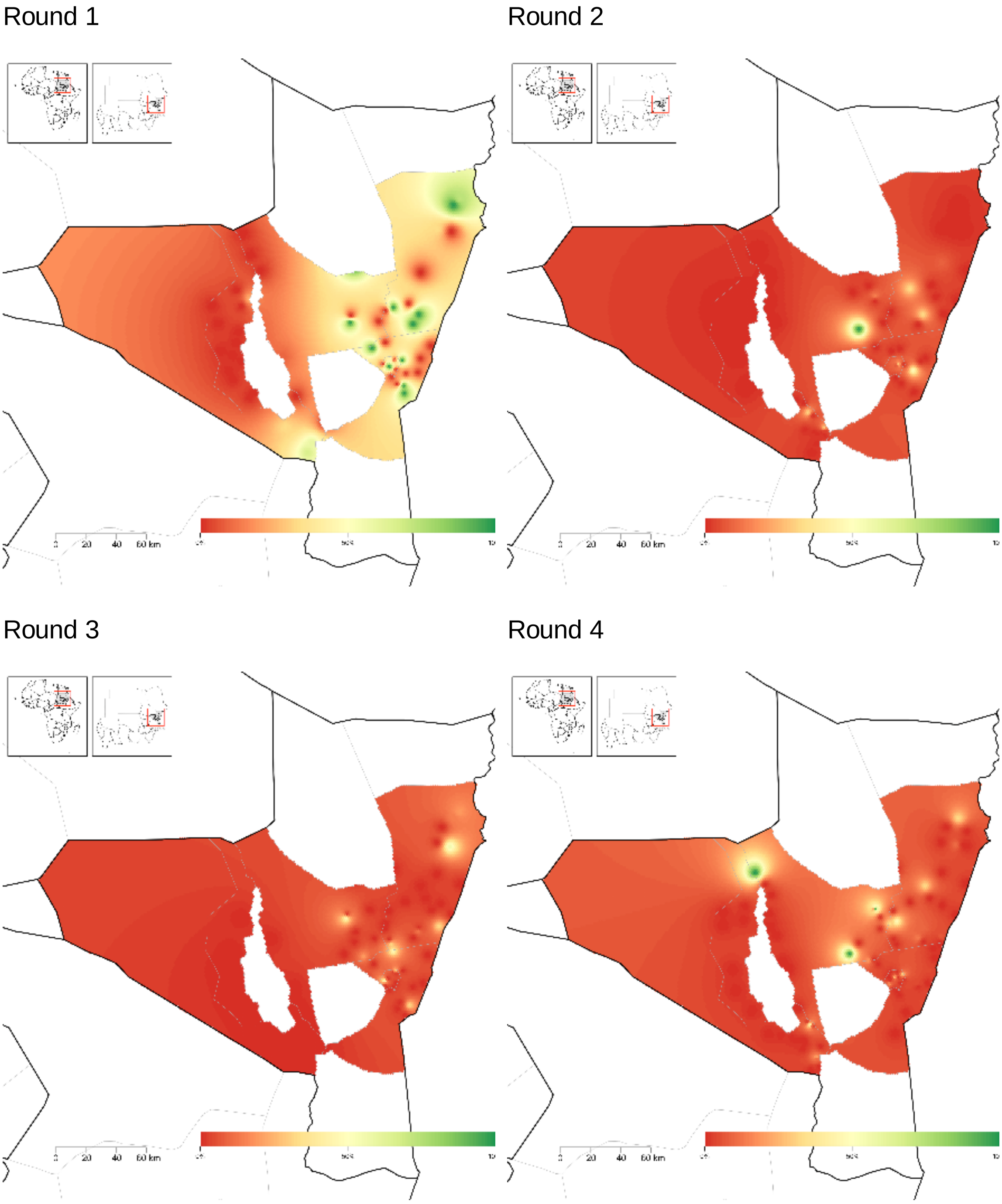
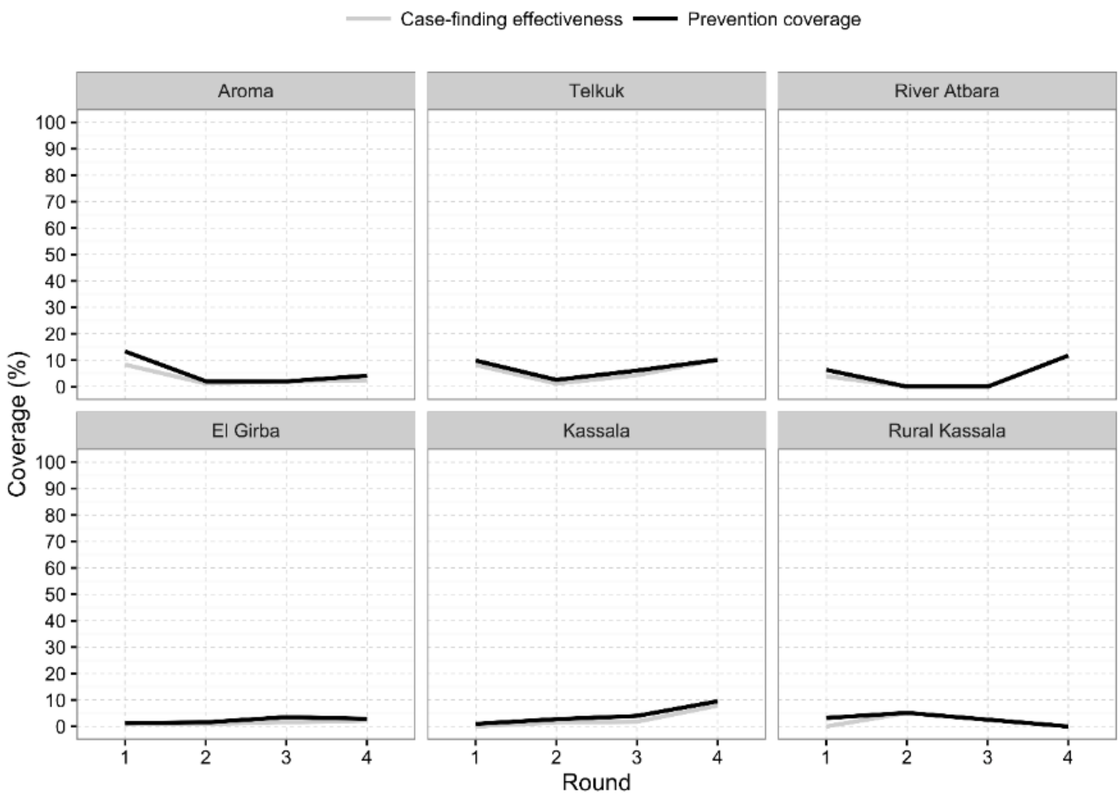


Figure 29 presents the spatial pattern of coverage of PLW by the TSFP programme. We again note a consistent trend as with the temporal pattern of coverage, in which coverage was relatively higher in the first round of the study (specifically in parts of Telkuk, rural Kassal and Kassala) and dropped to low levels by rounds 2, 3 and 4.

FBMAM coverage over time

Figure 30 shows coverage of children by the FBMAM programme over time (at each round) by cluster.

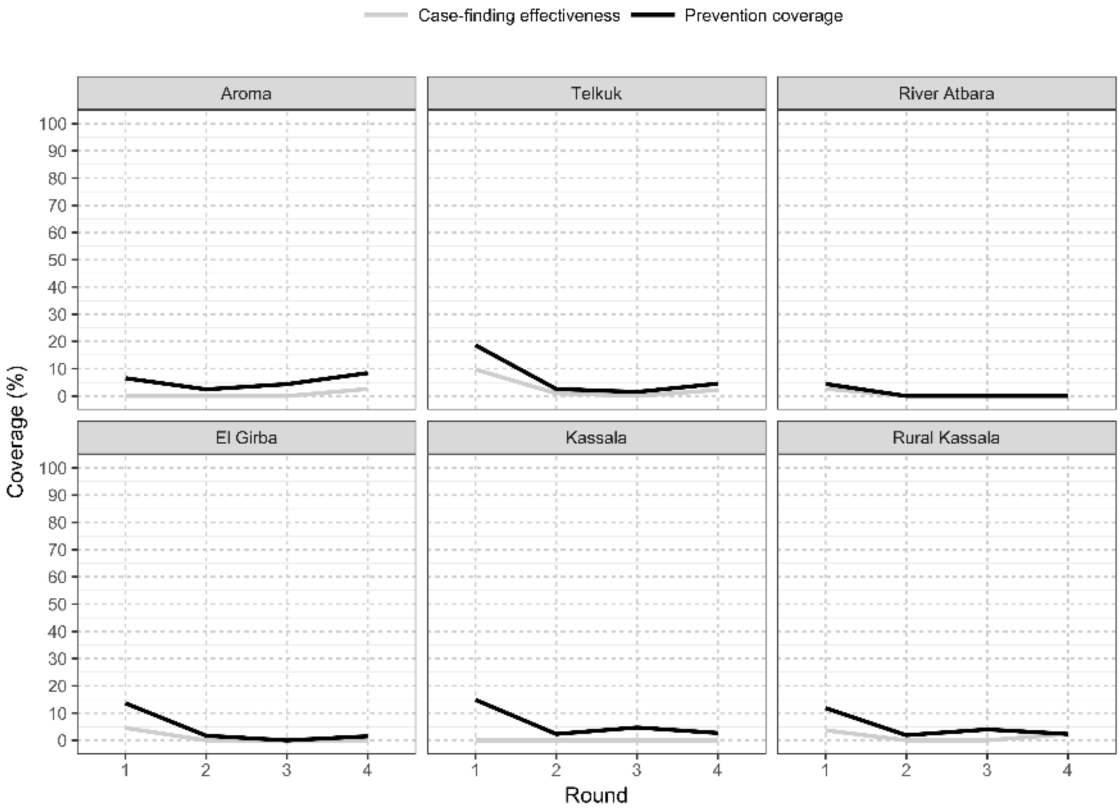
Figure 30: FBMAM coverage over time – children



The main clusters to focus on are rural Aroma and Telkuk, which have had the longest period as intervention clusters. Coverage in both clusters started off close to 10%, dipped in subsequent rounds, and returned to near the 10% mark in round 4.

Interestingly, even without implementation in the earlier rounds in River Atbara, El Girba, Kassala and rural Kassala, there are a few cases of at-risk children in the programme during this period. By Round 4, most of the remaining clusters are approaching the coverage levels of Aroma and Telkuk.

Figure 31: FBMAM coverage over time – PLW

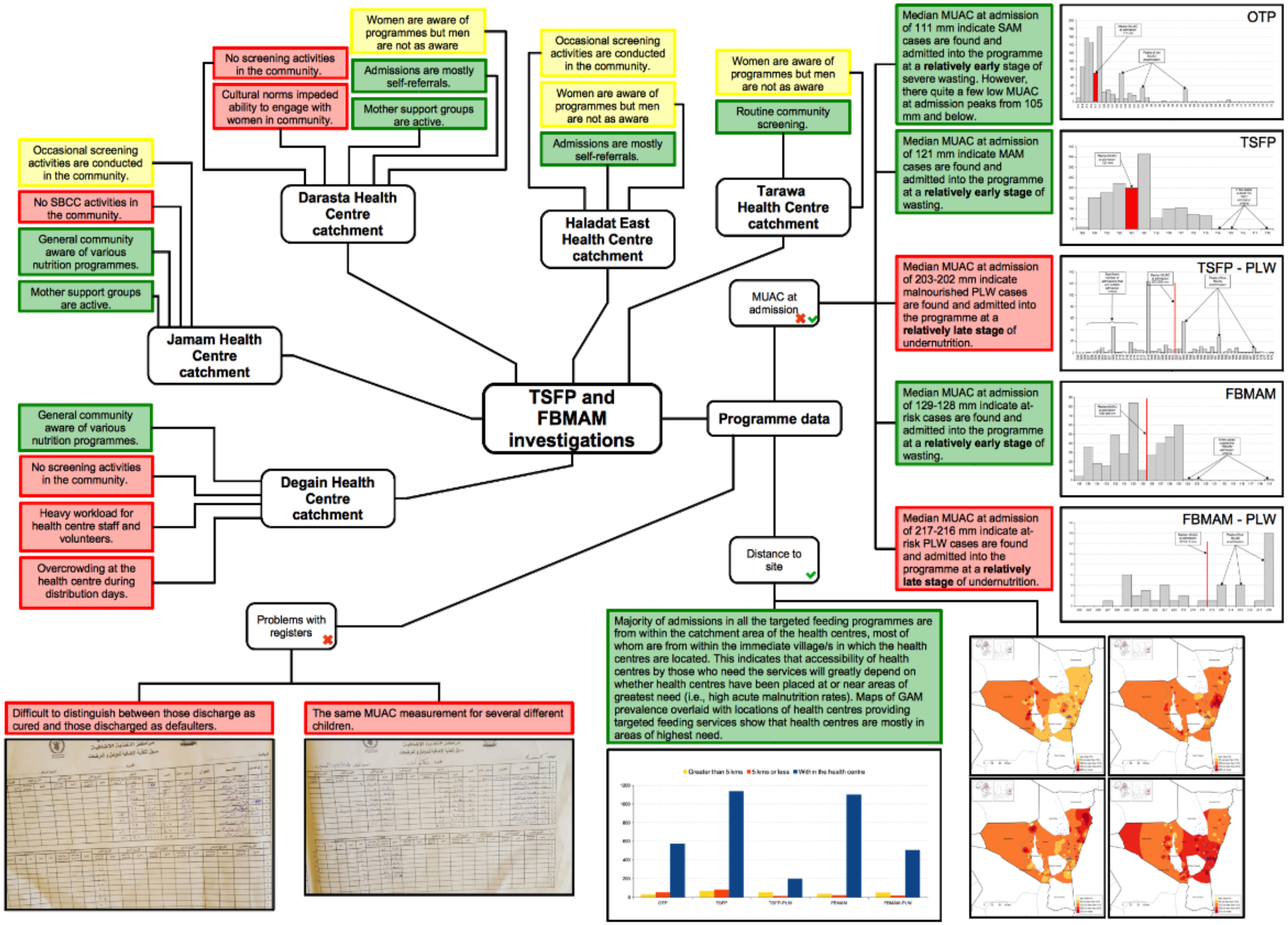


We noted the same temporal coverage of FBPM when assessing PLW, with a generally similar magnitude of coverage (low) for PLW and children (see Figure 31).

Factors impacting coverage

To further contextualise the coverage results of TSFP and FBMAM, we summarised the findings of the qualitative study using a mind map shown in Figure 26 below. Mind-mapping is a graphical way of storing and organising data and ideas. A mind map organises findings using tree structures organised around a central theme.

Figure 32: Mind map of qualitative study investigating factors of coverage for TSFP and FBMAM



The mind map points to three key issues that affect the programmes' ability to achieve good coverage:

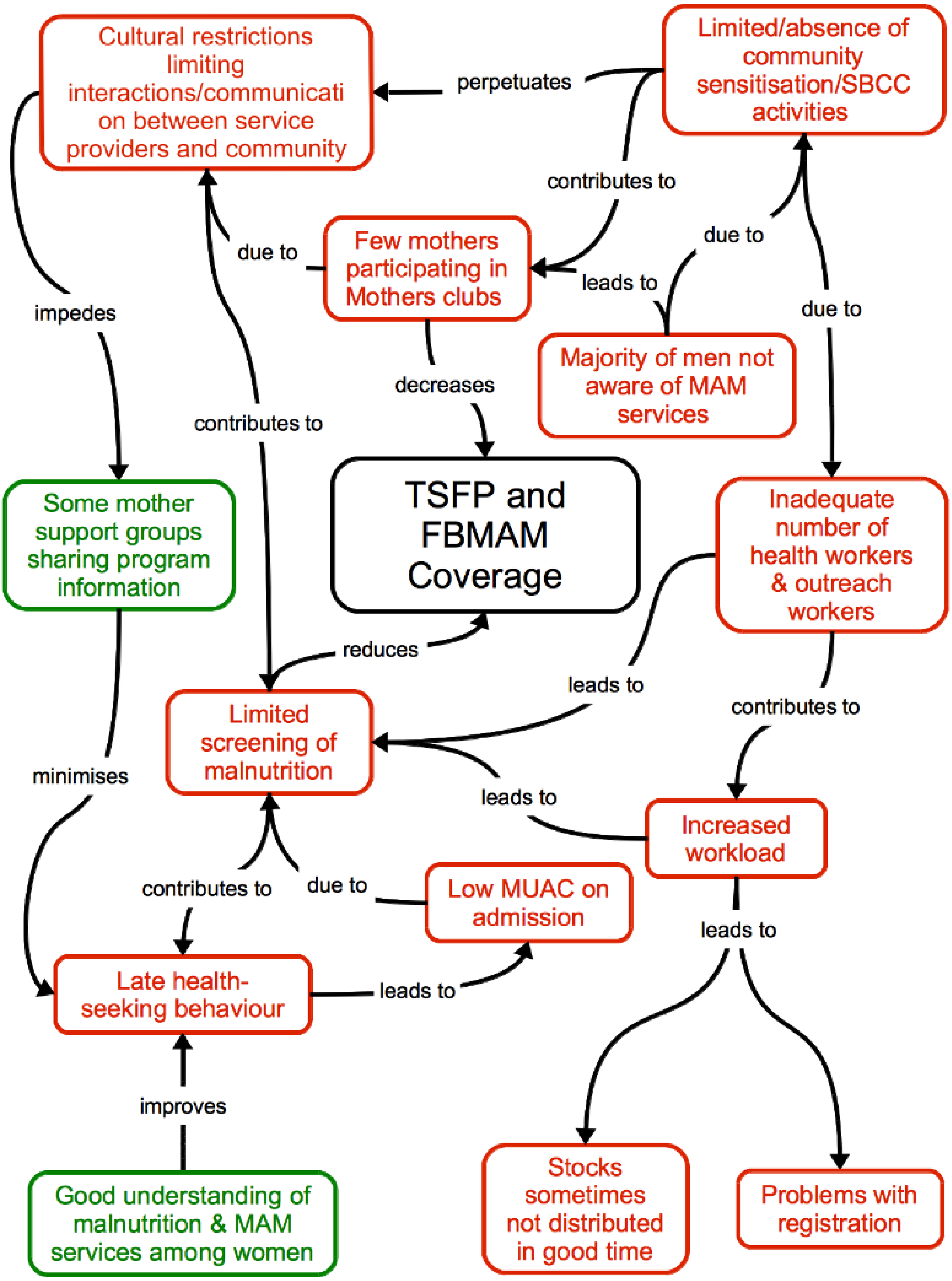
- Limited routine community-based screening to find cases of MAM and those at risk;
- Issues with record-keeping and maintenance of case registers at the health centre level; and
- Programme staff and volunteers with heavy workloads, which limit their efforts in clinic-based activities.

We further analysed the qualitative study results using a concept map[31] (see Figure 27) as a means of illustrating interaction between different factors affecting coverage. Despite factors such as the good geographical scope of health centres that provide TSFP and FBMAM – which allows for relatively easy geographical access in most cases – and generally high community awareness about the programme and acute malnutrition, the programme does not seem to have a consistent link with its communities. This difficulty would ideally be mediated by community volunteers through local sensitisation activities (as part of the SBCC activities; see Section 6.5) and routine screening.

We also determined that SBCC delivery was fraught with delays and implementation challenges. This was mainly due to a huge workload, specifically in delivering TSFP and FBMAM, which required community volunteers to work alongside programme staff at the clinic to support routine distribution of food products. It is very likely that this workload also impacted the ability of programme staff to keep registers and records correctly filled out and regularly updated. These issues with record-keeping impact on the ability of programme staff and community volunteers to track cases and identify those missing follow-up and eventually defaulting.

[31] Concept mapping is a graphical data-analysis technique that is useful for representing relationships between findings. Concept maps show findings and the connections (relationships) between findings.

Figure 33: Interactions between factors affecting coverage of TSFP and FBMAM

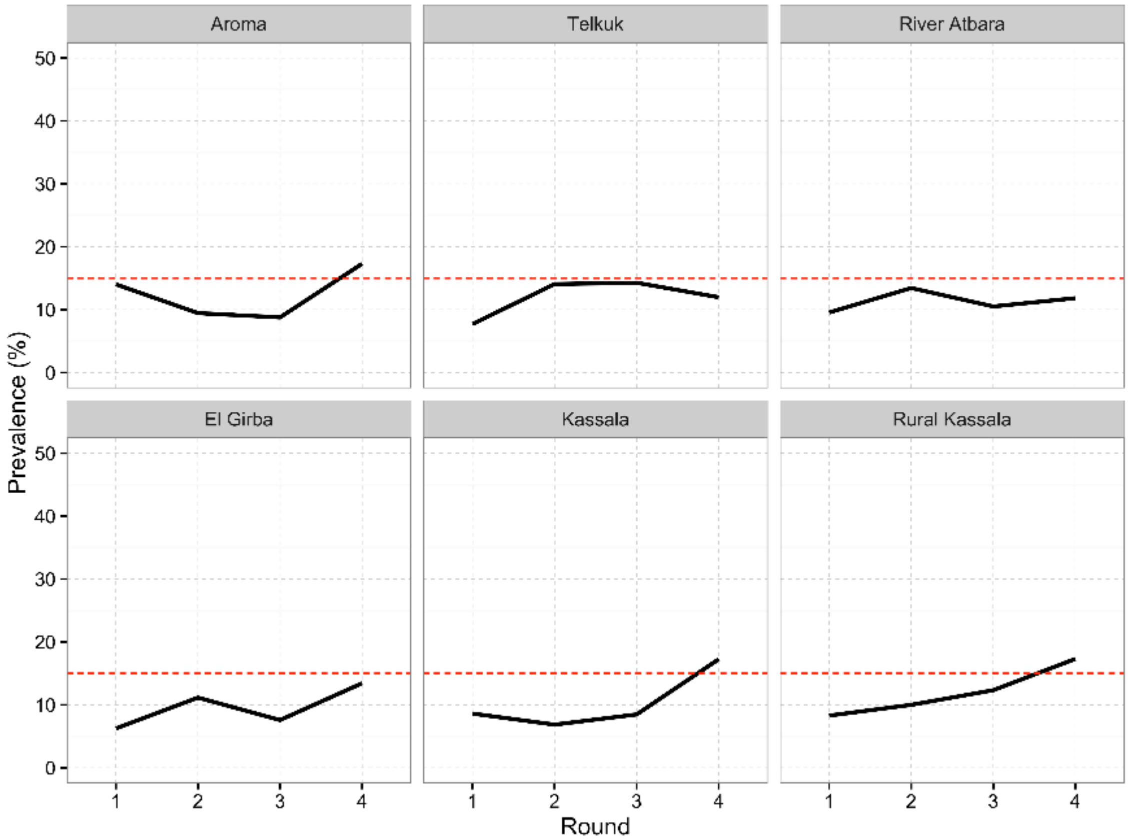


6.5 Sub-question 3: How timely and effective is an eBSFP?

To be able to assess the timeliness of any rapid response intervention, it is necessary to establish whether the response is needed in the first place. The current WFP criterion for deciding whether a rapid response is needed is based on GAM prevalence rates. If GAM prevalence reaches a level of 15 per cent or more, an eBSFP should be initiated.

Figure 28 presents the trend over time (at each round) for GAM prevalence by cluster. The 15% emergency cut-off has been marked. GAM rates were below 15% in rounds 1 to 3 of the study in all localities. In rounds 2 and 3, GAM rates in Telkuk already noticeably approach 15% but decline by round 4. By round 4, GAM rates peaked at over 15% in Aroma, Kassala and rural Kassala. Based on programme guidelines, we would have expected an emergency to be called and an appropriate response mounted– particularly an eBSFP – in Aroma, Kassala and rural Kassala during or immediately after the round 4 period. Although GAM rates were just below 15% in Telkuk during data collection in rounds 2 and 3, we can argue that there was reason to call an emergency based on heavy rains and flooding at that time, and the subsequent probability of limited access to food supplies for the locality. However, based on the timeline presented in Section 4, an eBSFP response was not implemented during this time period. Therefore, in terms of timeliness, opportunities to intervene were missed in some localities experiencing an emergency because WFP and partners were only working on assumed and historical GAM rates rather than GAM rates obtained through routine nutritional surveillance. Given that there was no emergency response mounted, effectiveness could not be assessed.

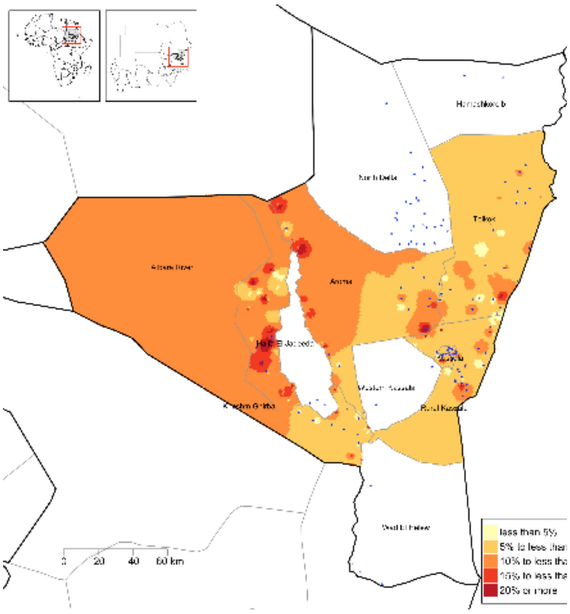
Figure 34: GAM prevalence over time



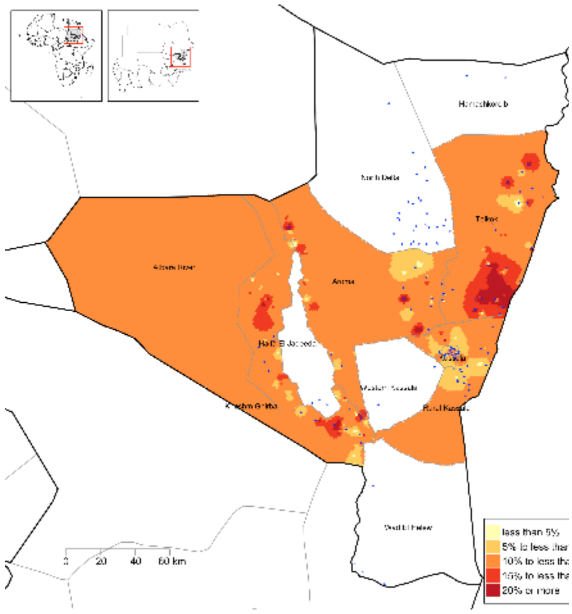
35 shows the spatial distribution within localities for GAM prevalence per round. The red areas are those areas with estimated GAM prevalence of 15% or more.

Figure 35: Spatial distribution of GAM prevalence

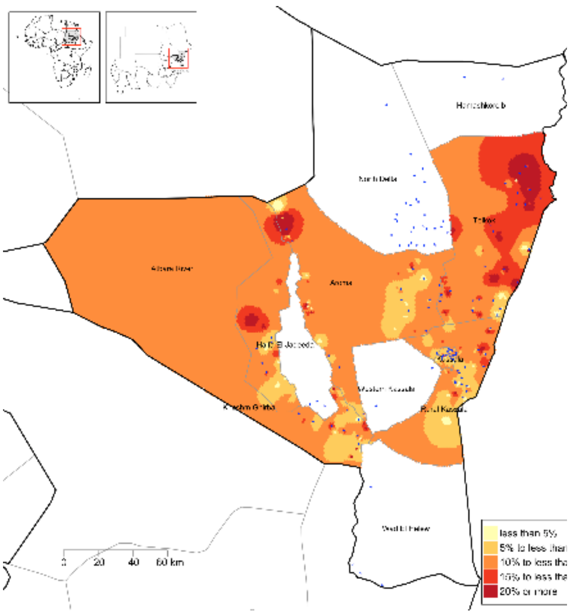
Round 1



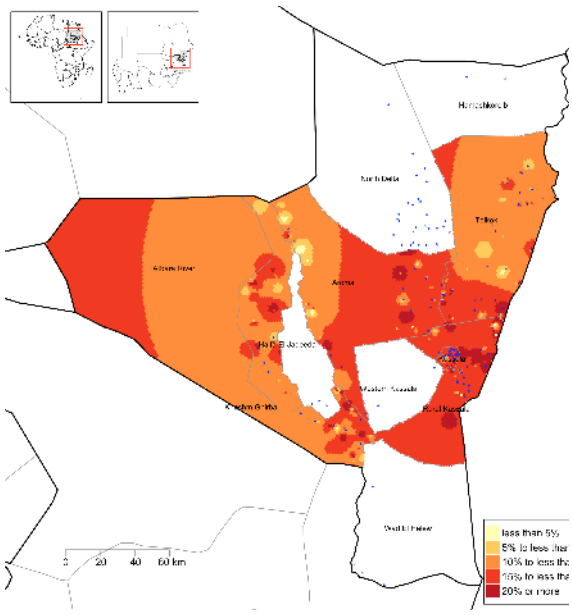
Round 2



Round 3



Round 4



In round 1, the general spatial pattern for GAM is between 10–15% across most areas, with small pockets having between 15–20%, and specific locations with 20% or higher GAM prevalence. The areas reaching the 15% threshold may have been too small to require an emergency response. However, as time passed and prevalence increased, there were relatively bigger pockets with 15% or more GAM prevalence, specifically in Telkuk, where the number of affected children was most likely higher. Meanwhile, no specific emergency response was put in place for these areas during rounds 2 and 3. By round 4, GAM prevalence was still increasing, with larger areas showing GAM prevalence at the 15% level or higher. Again, no emergency response was declared during this period; however, it can be argued that there was indeed a basis for doing so given the ever-increasing prevalence as shown spatially. The main finding here is that the only way an emergency response can be triggered appropriately is based on prevalence levels. However, during the study period, there was insufficient information available on GAM prevalence to enable any decision on an emergency response at locality or sub-locality level.

6.6 Sub-question 4: How does the inclusion of SBCC impact effectiveness?

We evaluated the impact of SBCC on effectiveness using a hierarchical coverage assessment approach similar to that described by Tanahashi (1978). First we present the level of awareness of SBCC activities in the community at large. We called this indicator a message coverage indicator. We determined whether a mother had heard about specific community sensitisation activities. Table 10 illustrates message coverage throughout the study clusters by round.

Table 10: **Mothers who have heard of community sensitisation activities**

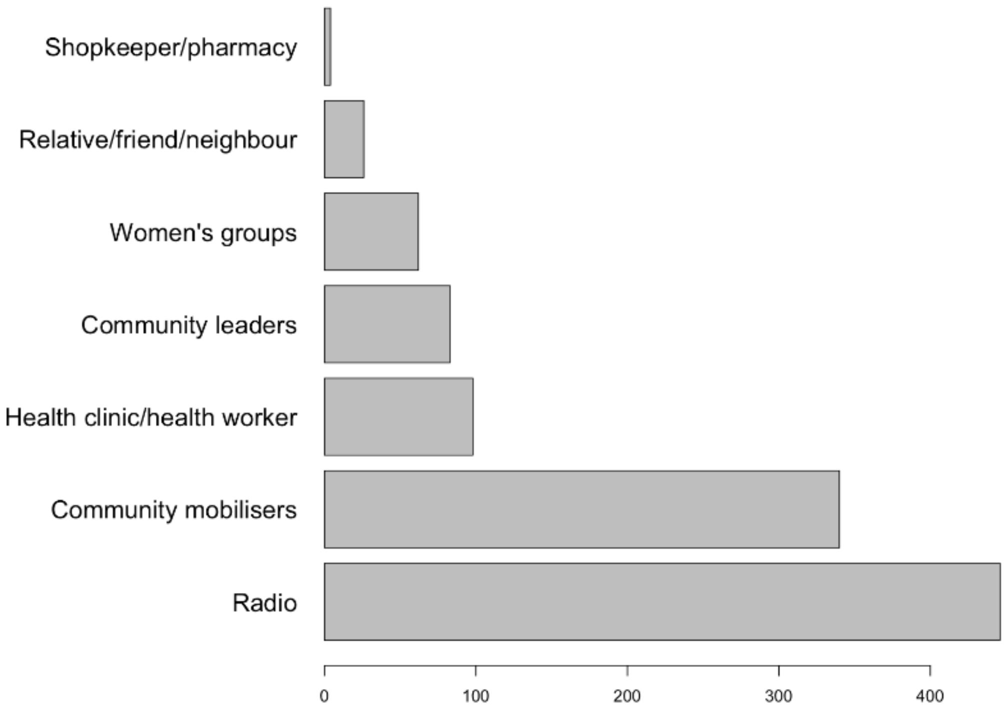
Round	Sample size	Estimate	95% LCL	95% UCL
Round 1	4638	11.4%	6.6%	13.4%
Round 2	9264	10.9%	9.4%	12.7%
Round 3	8677	10.08%	8.5%	11.7%
Round 4	7311	11.2%	9.5%	12.9%

At best, SBCC messages were heard by up to 12% of the target audience,[32] with message coverage roughly staying around the same level (between 10-12%). This is a lower-level indicator for coverage, but it also determines the highest possible level of coverage that can be achieved by SBCC activities, as only those who have heard the message about community sensitisation will eventually participate.

We then asked those who had heard about community sensitisation activities where or from whom they had received this information. Radio and community mobilisers were reported to be the most common source of information on community sensitisation activities (see Figure 30).

[32] SBCC was meant to cover all members of the community in which the programme is implemented. There are no eligibility criteria for this programme other than being within its catchment area.

Figure 36: Sources of information on community sensitisation activities



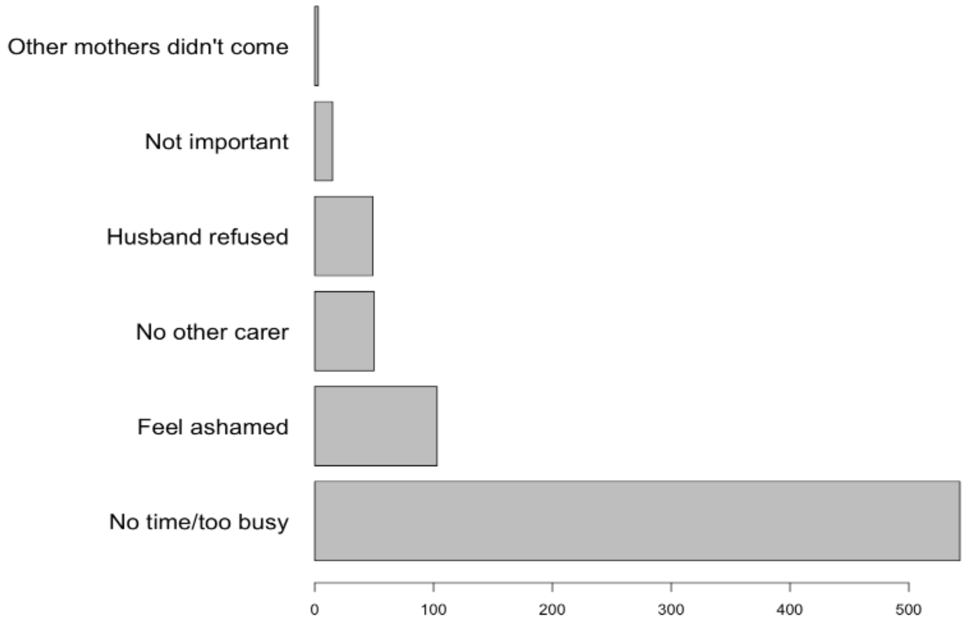
Second, we present the next level of coverage indicator, which is the proportion of mothers who heard about community sensitisation activities and then participated in them. We call this contact coverage, and it assesses those who had direct contact with the SBCC intervention.

Table 11: Mothers who have participated in community sensitisation activities

Round	Sample size	Estimate	95% LCL	95% UCL
Round 1	610	37.1%	18.5%	37.1%
Round 2	848	42.8%	34.5%	51.3%
Round 3	909	22.2%	17.7%	26.7%
Round 4	867	11.8%	8.5%	15.7%

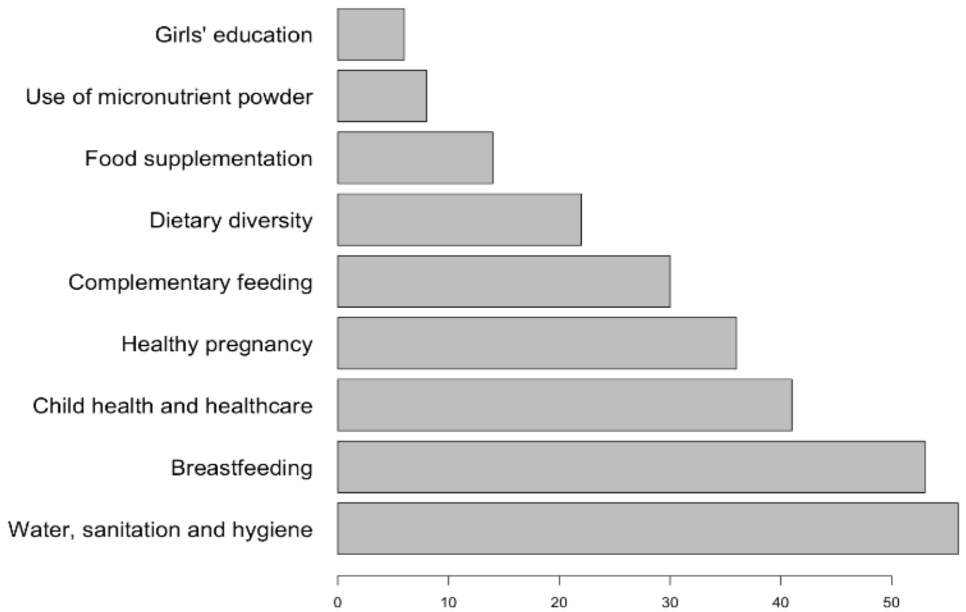
This indicator in Table 11 above shows that of those who heard of the activities, only a little over 40% (at best) participated, and this dwindled to as low as 12% in the latter period of the programme. The majority of those who said they did not attend the sensitisation activities reported that they did not have time or were too busy (see Figure 31).

Figure 37: Barriers to participation in SBCC activities (Round 4)



We also collected information about what participants learned in community sensitisation activities. Figure 38 shows that topics on WASH and breastfeeding were most frequently and easily recalled by mothers. The most likely reason is that most study areas were affected by an acute watery diarrhoea outbreak during round 3 and towards round 4 of data collection. During this period, health messaging (either related to the programme or in general), focused on WASH as a means of stemming the incidence of acute watery diarrhoea. Breastfeeding, on the other hand, is always a key topic that underpins most nutrition education and is most likely frequently emphasised during community sensitisation.

Figure 38: Topics that mothers reported to have learned in SBCC activities (Round 4)



Finally, we assessed the effectiveness of coverage for SBCC activities by determining whether there had been any change in the behaviour and practices of mothers on a specific topic that the SBCC activities were meant to address. Specifically, we chose IYCF as the behaviour to assess, as this topic is given the greatest importance in the SBCC intervention relating to acute malnutrition prevention.

Figure 39: Comparison of control and intervention groups per study round for each category of infant and young child feeding practices and behaviours

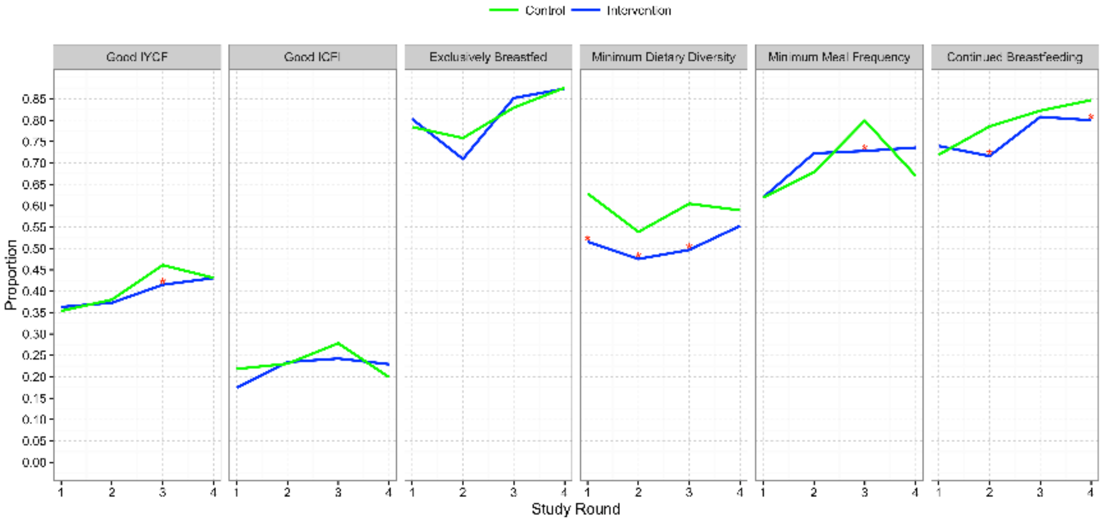


Figure 39 presents a comparison of control and intervention groups with regard to the six key infant and young child feeding practices that we assessed. The overall IYCF indicator (good IYCF) showed very little difference between control and intervention groups, except in round 3 where intervention groups had a significantly lower rate of appropriate infant and young child feeding practices as compared to the control group.

The IYCF indicator components of “Good ICFI” (Infant and Child Feeding Index) and “Exclusively Breastfed” indicate that the most likely reason for this difference in round 3 was due to ICFI practices in the 6–23 month age group. Exclusive breastfeeding was similar between control and intervention groups at each round, and rates were impressively high in both. We further disaggregated the good ICFI into its three components (minimum dietary diversity, minimum meal frequency, continued breastfeeding) and found that minimum dietary diversity in the intervention group was significantly lower than the control group from rounds 1 to 3 of the study. This seems to be the greatest driver of low ICFI in both control and intervention.

These results indicate that there has been no change in infant and young child feeding behaviours and practices attributable to the intervention. Given the delay in implementation, it is likely that the short period of exposure to SBCC activities was an important reason for the absence of change, but coverage is also a critical consideration. Message coverage within target populations was 12% at best, and therefore unlikely to support any change in the indicators reported here. In addition, the drop in contact coverage over time (from 40–12%) and the finding that the most common barrier to participation was ‘no time/too busy’, suggests a re-examination of the relevance of messages (and how they are delivered) for target communities and the opportunity costs linked to participation in this set of interventions. We discuss this further in chapter 8.

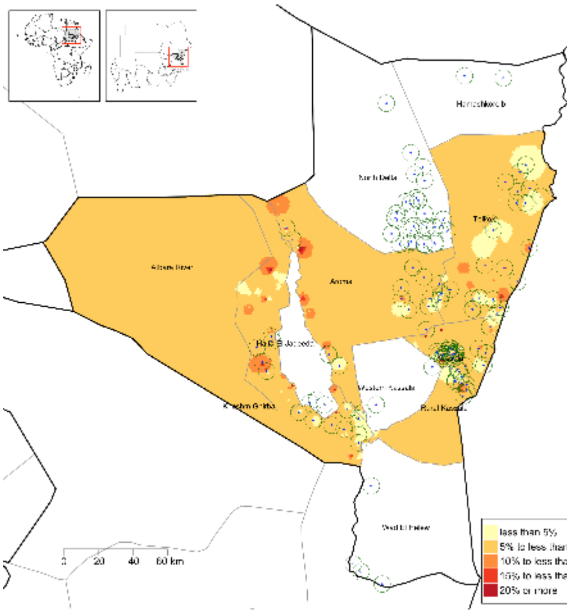
6.7 Sub-question 5: How appropriate are geographical targeting criteria for each intervention?

Both TSFP and FBMAM were implemented through the same health centres in each cluster. The selection of health centres was guided by the known prevalence of GAM and the population size of the catchment area. The catchment area of each health centre is defined as a 5km radius around it.

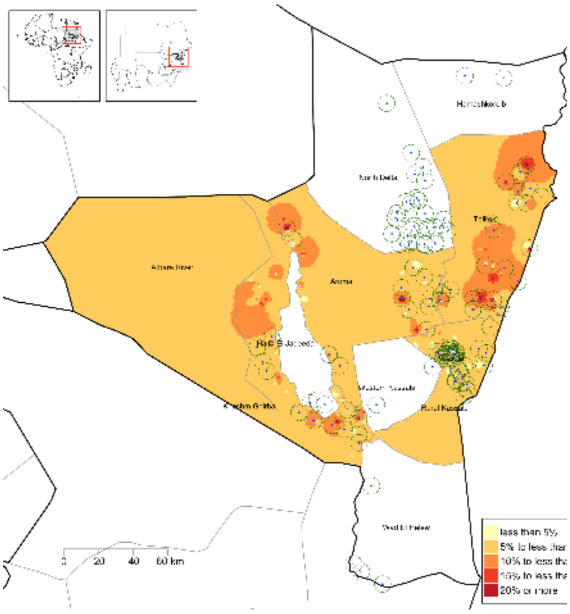
To assess the appropriateness of geographical targeting for both TSFP and FBMAM, we created interpolated maps of MAM prevalence. We then mapped the locations of each health centre implementing the programmes and added this as an overlay onto the malnutrition maps. Finally, we created a map of the catchment areas around each health centre and added this as another overlay. Figure 34 presents the maps created. The selected health centres covered areas with higher MAM prevalence than other locations. This indicates a generally appropriate geographic targeting by the programme.

Figure 40: Spatial distribution of MAM prevalence overlaid with catchment area map of health centres providing TSFP and FBMAM

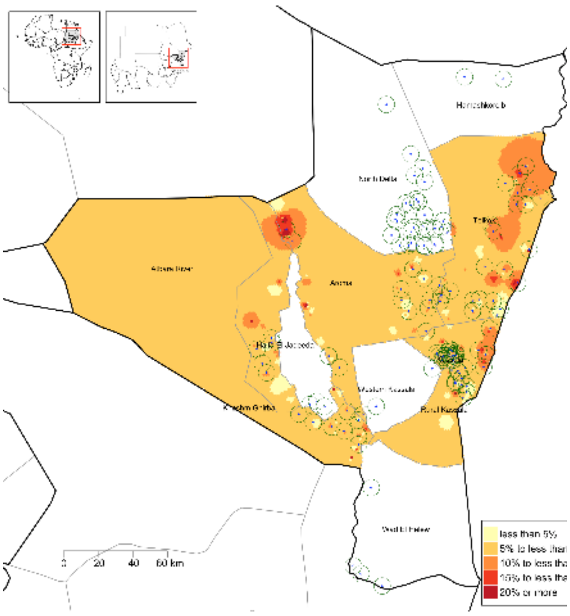
Round 1



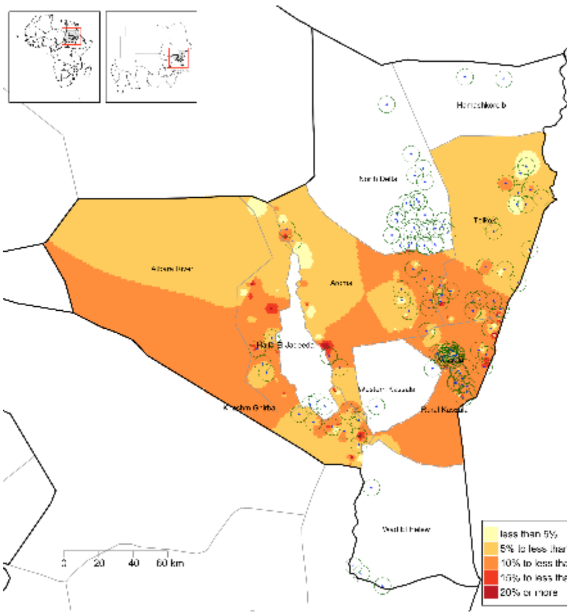
Round 2



Round 3



Round 4



6.8 What is the cost effectiveness of the different packages from a WFP perspective?

To analyse the data, we considered both the cost information shared by WFP and the impact data collected during programme monitoring. Given that the costing information does not distinguish between children under five and PLW, the analysis consolidated both types of beneficiaries into a single total for the treatment programme (for the FBMAM only data for children under five was available).

Tables 12 and 13 show results for the different programmes by locality. The cost component includes:

- The actual cost of the commodity per metric ton: SC+ from April to September and RUSF from October to December;
- External transport to move the commodity to the main port of the country;
- All in-country related costs such as transportation, shipping and handling;
- Operational costs from WFP related to partners, such as training, materials and agreements;
- WFP staff-related costs;
- Administrative costs related to WFP headquarters.

From the tables, it can be observed that many data were not-available (NA). The particularities of the information for each one of the localities are explained below.

Table 12: Cost for TSFP per beneficiary by locality

Locality	Treatment				
	Cost (USD)	**Admitted**	**Cured**	**Cost/admit**	**Cost/cured**
Aroma	NA	4183	3366	NA	NA
Telkuk	6430.2	955	350	**6.73**	**18.37**
El Girba	20030.2	1135	961	**17.65**	**20.84**
River Atbara	7785.8	357	0	**21.81**	NA
Kassala City	175157.3	4153	2690	**42.18**	**65.11**
Rural Kassala	87494.3	4997	3930	**17.51**	**22.26**

Table 13: Cost for FBMAM programme for children under five by locality

Locality	FBMAM				
	Cost (USD)	**Admitted**	**Cured**	**Cost/admit**	**Cost/cured**
Aroma	NA	3724	1314	NA	NA
Telkuk	NA	931	298	NA	NA
El Girba	NA	193	0	NA	NA
River Atbara	14206.9	92	0	**154.42**	NA
Kassala City	84292.7	996	20	**84.63**	**4214.63**
Rural Kassala	9212.8	454	0	**20.29**	NA

Aroma: Information about programme monitoring was available for the treatment and FBMAM programmes from January to December. However, the costing information shared by WFP did not include any data for these programmes. Therefore, the computations for the CEA were not possible for this locality.

Telkuk: Information about programme monitoring was available for the treatment and FBMAM programmes from February to December. However, the costing information shared by WFP only included treatment costing for April, July and September, and there was no information for FBMAM. Therefore, the computations for the CEA were only possible for three months of the treatment programme.

El Girba: Information about the programme monitoring was available for the treatment programme from January to December, while there was costing information from April to December. Therefore, the data summarised in Table 9 for El Girba includes this last range of time.

Monitoring data for the FBMAM programme included both November and December, but there was no costing information available, and therefore no CEA calculations were possible.

River Atbara: Costing information was available for the treatment programme from April to December. However, monitoring data was only available for November and December. The calculations in Table 9 include only these two months. FBMAM information on costs and monitoring data was available for November and December. Monitoring data only included admissions.

Kassala City: Information about programme monitoring was available for the treatment programme from January to December, while costing information was available from April to December. Therefore, the data summarised in Table 9 for Kassala City includes this last range of time.

The information for Kassala City initially looked reliable. However, the cost data for the treatment programme in December (USD 59,327.25), accounts for 34% of the total yearly cost, which is very high for only one month. Table 9 includes this information but should be interpreted carefully.

FBMAM information on costs was available from October to December. However, monitoring data was available only for December.[33] Therefore, only December was considered in the calculations of Table 10.

Rural Kassala: Information about programme monitoring was available for the treatment programme from January to December, while there was costing information from April to December. Therefore, the data summarised in Table 1 for Rural Kassala includes this last range of time. FBMAM information both for costs and monitoring data were available only for December.

In general, the most reliable data in Table 12 (El Girba and Rural Kassala) show that the costs incurred by WFP for the treatment programme are approximately $18 USD per beneficiary admitted and $22 USD per beneficiary cured. These estimates are consistent with those made by WFP with cost per treatment of MAM at $17 USD for programmes

[33] Monitoring data for FBMAM in Kassala City and rural Kassala were also available from January to March. However, according to the 3ie timeline and the costing information, the FBMAM programme only took place from October to December. Therefore, the data from the beginning of the year were excluded from the analysis.

providing SC, $15 USD for those providing SC+ and $18 USD for programmes providing RUSF (World Food Programme 2012b).

Unfortunately, a similar estimate cannot be reached for the FBMAM programme (Table 13), where the range for cost per child admitted goes from $20 USD to $155 USD, and there is only one estimate for the cost per beneficiary cured. These high variations are due to the lack of data, because all current computations rely only on one or two data points. Consequently, the costs for the FBMAM as calculated here cannot be relied upon. The monitoring data for both Aroma and Telkuk is very rich, and therefore including analyses for these localities would have provided more reliable estimates. However, as mentioned above, there was no costing information available for the programme in these two localities.

Given the constraints related to the availability and reliability of the data, particularly for the FBMAM programme, it is not realistic to compare cost effectiveness of the treatment to prevention programmes at this stage.

6.9 What are the wider impacts, positive or negative, of the packages at household, community or institutional level (opportunity, social, economic, environmental)?

To address this question, we reflected on the theory of change presented in Section 2.2 to assess the potential wider impact of the programme. Applying a causal pathway perspective, we assumed that the most likely determinant of whether the programme can provide a wider impact at these levels is its ability to increase knowledge on various nutrition-specific and nutrition-related behaviours and practices. As noted in our graphical presentation of the theory of change, change in behaviour and practices underpins the more clinical (curative and preventative) aspects of the programme, and is the key factor that will determine whether any impact observed on these clinical parameters can potentially be sustained.

As presented in the previous sections, the SBCC component of the intervention package has not been fully implemented in a manner and to a degree that allows for change in behaviour and practice to be possible. As such, we were unable to assess any wider impact at the household or community level. We also acknowledge that change in behaviours and practices generally takes much longer than the nine months' duration of the programme examined here; therefore this evaluation question would be more relevant for a programme that has been sustained for much longer.

7. Discussion

The overarching research question of this impact evaluation was as follows:

> *What is the impact on the incidence and prevalence of MAM and SAM in children under five and Pregnant and Lactating Women (PLW) of different MAM treatment and prevention interventions in Sudan (i.e. Targeted Supplementary Feeding Programme [TSFP] for the treatment of MAM; Targeted Food-Based prevention of MAM [FBMAM]; emergency Blanket Supplementary Feeding Programme [eBSFP] as a rapid response to crisis for the prevention of MAM; Home Fortification [HF] for prevention of MAM; and Social and Behaviour Change Communication* [*SBCC] for prevention of MAM)?*

The study results demonstrate that, in this context, the addition of a FBMAM programme to a TSFP as a package intervention for the treatment and prevention of MAM has decreased the prevalence of at-risk children but not the incidence or prevalence of MAM, SAM and GAM. We found no difference in outcomes between male and female beneficiaries. In this case, our prevalence of at risk and incidence findings seem to contradict each other, given that we expected to see a significant reduction in numbers of at-risk children to translate into a reduction in MAM prevalence secondary to reduced MAM incidence. For the PLW group, we found a similar pattern of decreasing but non-significant MAM prevalence over time and a decreasing significance of at-risk prevalence over time.

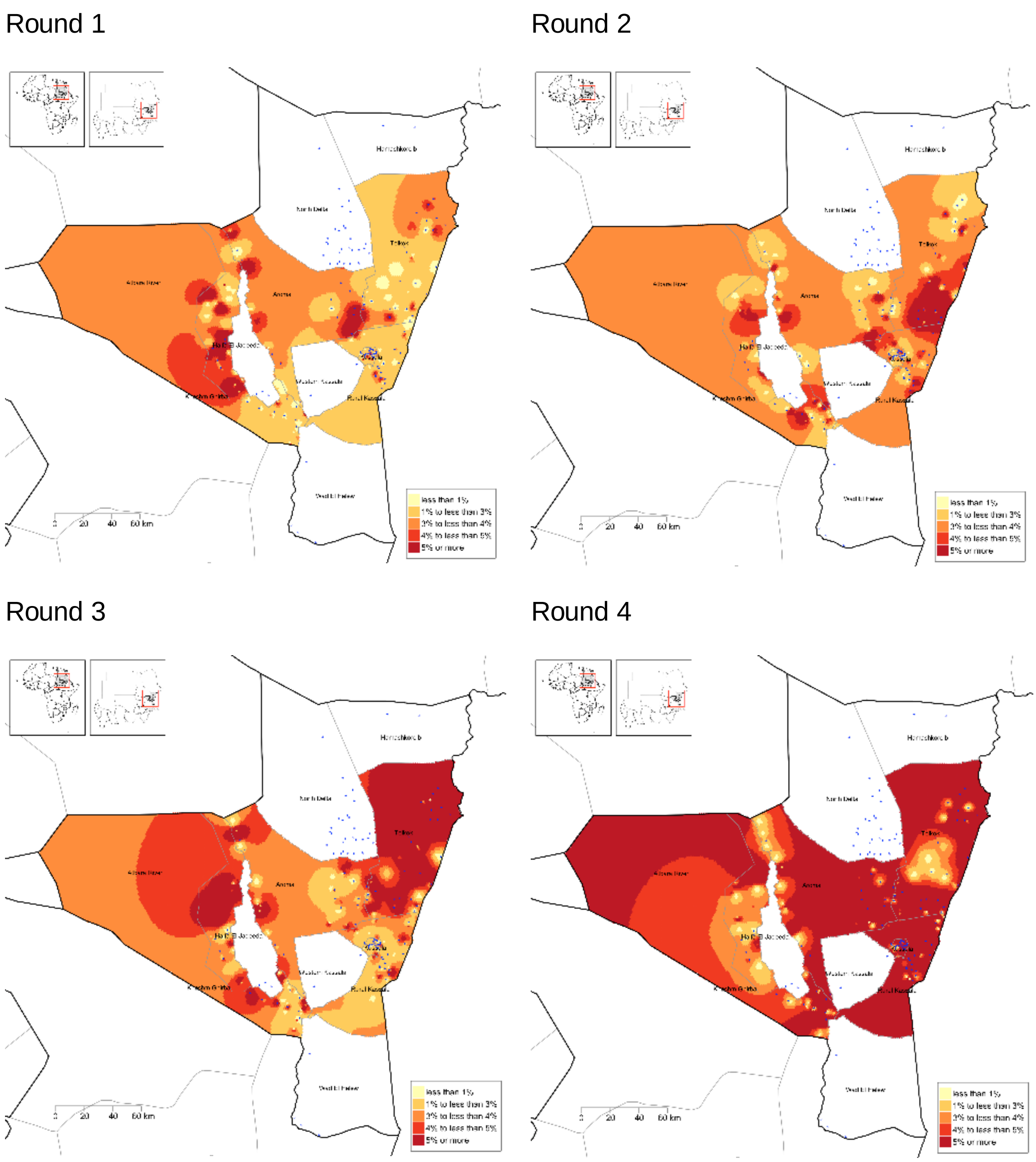
However, we think there are plausible explanations to this seeming contradiction. Possible reasons for the absence of effect on prevalent and incident cases despite the observed decrease in numbers 'at risk' are: (1) there is a time lag between at-risk reduction and prevalence reduction (as mediated by incidence reduction) and this time lag is linked (at least in part) to low coverage of prevention interventions; (2) incident MAM cases may come from previous SAM cases discharged from SAM treatment; and (3) the level of FBMAM programme coverage is too low to support any change in prevalence or incidence at a population level, especially over a relatively short-duration intervention, as assessed during this evaluation.

The temporal trend of at-risk prevalence reduction shows a continuing decrease over time in the presence of the FBMAM programme. It is possible that over a longer period of observation and exposure to the prevention programme, particularly if longer duration coincides with improving coverage (see below), at-risk prevalence would continue to decrease and eventually manifest in a reduction in MAM prevalence secondary to a reduction in MAM incidence.

A time lag between at-risk reduction and MAM prevalence and incidence reduction was shown in a previous study undertaken by Ruel and others (2008) on age-based preventive targeting of food assistance in Haiti. It found that food supplementation of non-wasted children aged 6–23 months, similar to the FBMAM programme, had wasting rates four percentage points lower than those only receiving treatment. An important feature of this study was that it was conducted over a three-year period, which allowed the authors to determine that this reduction effect was greater in children exposed to the prevention programme for a longer period of time.

We observed up to a 12% reduction in at-risk prevalence by round 4 of the study but with no effect on MAM prevalence and MAM incidence. Based on feedback and discussion with country stakeholders running treatment programmes for SAM within Kassala, it is likely that some incident MAM cases come not only from those who are not currently malnourished, but also from SAM cases that have recovered from severe wasting and have been discharged from OTP with anthropometric measurements, which classify them as MAM. Kassala is considered a priority area for SAM treatment given its historically high levels of SAM prevalence (according to personal communication with a UNICEF nutrition officer). This is supported by SAM prevalence results obtained from the study (see Figure 41), which show an increase in SAM prevalence over time.

Figure 41: Spatial distribution of SAM prevalence over time



Conceptually, we can visualise the relationship of SAM, MAM and GAM incidence and prevalence based on a compartment model approach similar to what Isanaka and colleagues have described (Isanaka et al. 2011) as shown in Figure 42.

Figure 42: Compartment model of relationship between SAM, MAM and GAM incidence and prevalence

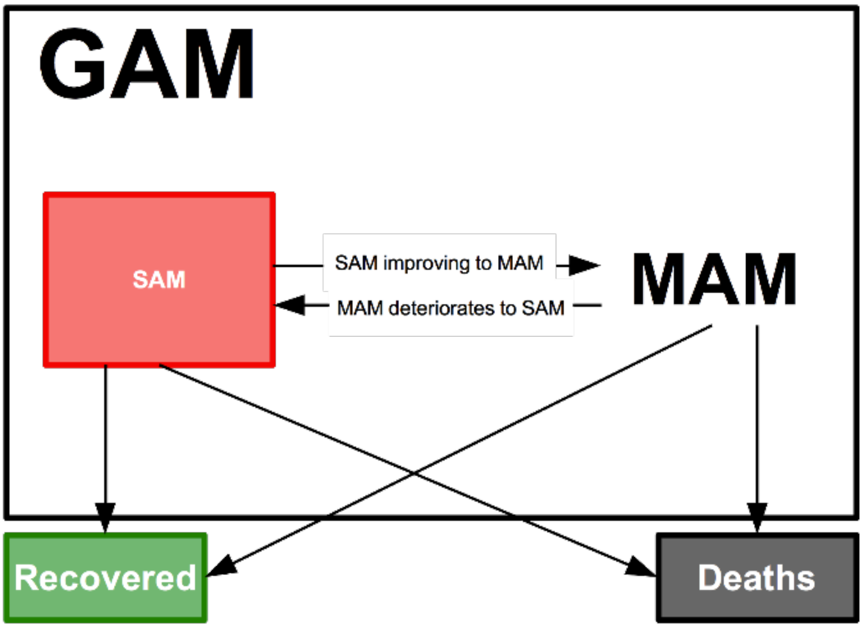
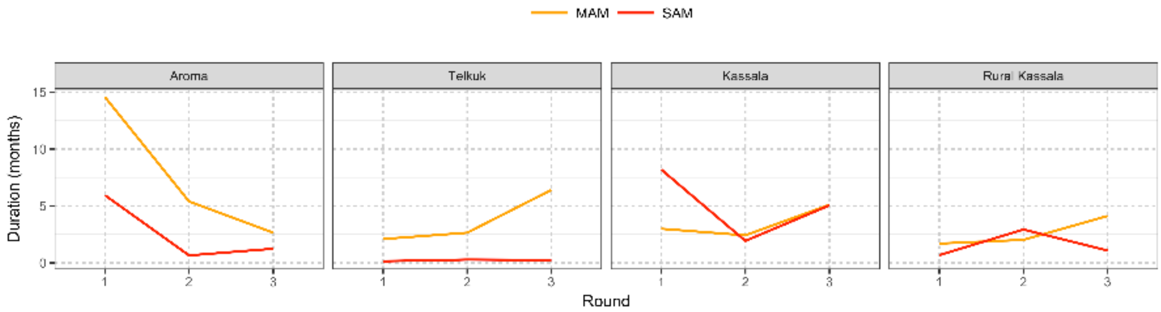


Figure 42 demonstrates how a decrease in GAM prevalence and incidence can only be affected if the prevalence and incidence of SAM or MAM (or both) decrease via the pathway of recovered cases (cases that are neither SAM nor MAM) or deaths. The deaths pathway, however, is often negligible (Isanaka et al. 2011). Given this, any decrease in SAM prevalence and incidence due to improvement of cases to MAM or any decrease in MAM prevalence and incidence due to deterioration of MAM cases to SAM would not change GAM prevalence and incidence to any significant or tangible degree.

If we were to pose the hypothesis that GAM incidence and prevalence in our study did not significantly decrease partly due to SAM cases improving to become MAM cases, we would need to show that the rate by which the programme is able to assist in MAM recovery is slower than the rate by which SAM cases improve and recover to become MAM. This requires an assessment of the effectiveness of the MAM treatment component of the programme (TSFP). Factors that contribute to this effectiveness are the programme's cure rate (see Figure 22) which, based on routine monitoring data, is above average (between 80% to 100% cure rate over time across localities) and programme coverage (see Figure 26) which, based on the study results, is quite poor. To further test this hypothesis empirically, we used routine programme data on admissions of MAM cases (from TSFP) and admissions of SAM cases (from OTP), along with the MAM and SAM prevalence estimates we obtained from the study, to calculate duration of MAM and SAM in the four study localities where we conducted the incidence study using the mathematical model used by Dale and others (2017). Figure 43 shows the results of our estimates.

For the combined TSFP and FBMAM programme to reduce GAM incidence and prevalence, the duration of MAM episodes should be as short as, or shorter than, the duration of SAM to be able to offset the recovery of SAM cases into MAM cases (among others). Our estimates show that for some of the localities (Aroma and Telkuk, who had the longest exposure to the treatment), duration of MAM was much higher than SAM duration.

Figure 43: Duration of MAM and SAM episodes



Given this, it is very likely that the programme's overall effectiveness in curing incidence and prevalence of MAM cases is insufficient to reduce GAM incidence and prevalence, particularly where there is recovery of SAM to MAM occurring in the same vicinity.

This could, in part, explain why incidence has not changed and why it doesn't appear to correlate with the level of inputs provided by the TSFP and FBMAM programmes. Further examination of SAM treatment performance and coverage may provide data that could explain this. Anecdotal information shows that Kassala is known as one of the best-performing SAM treatment programmes in Sudan, with good performance indicators and programme coverage (according to personal communication with a UNICEF nutrition officer). If that is the case, it is possible that the level of discharge of MAM incident cases from the SAM treatment programme increased over the duration of the study and is significant enough to offset any potential reduction in MAM incidence created by the reduction in at-risk prevalence.

Furthermore, although it has not yet been possible to provide unequivocal evidence as to the most effective intervention modality, service delivery and quality are clearly important issues. As highlighted by the 2014 review on MAM management (Rogers and Guerrero 2014), implementation issues have often been linked to poor impact on programme objectives, and outcomes have not been a focus to date. In particular, the quality of FBMAM programme implementation (including SBCC, which plays a critical role in the projects theory of change, noted above) could be a key factor in ensuring improvement in the decrease of at-risk prevalence observed in the study. Performance and coverage are the main elements to consider here. Currently, a programme that is achieving only approximately 10% coverage and has only recently initiated many of the SBCC actions at the community level, has had an observed effect of approximately 12% reduction in the at-risk category, with no effect on MAM prevalence. Increasing coverage could contribute significantly to a greater reduction in at-risk prevalence and MAM incidence at population level, and subsequently lead to a decrease in MAM prevalence. Further evaluation and research is critical to provide a robust evidence base for the link between coverage of these types of interventions and outcomes such as incidence and prevalence of MAM and GAM.

Findings from the qualitative component of the study provide a picture of a programme that is still in its infancy, with some teething problems requiring resolution. Key issues noted during the investigation were: (1) a need for more effective case-finding of MAM and at-risk cases; (2) a need for improved record-keeping at the clinic level for admissions and defaulters; (3) community mobilisers are inundated by multiple tasks and roles that limit their ability to perform more community-orientated tasks, including community sensitisation, as part of SBCC interventions; and (4) due to very low participation, there is a need to re-examine the relevance of SBCC actions (and how they are delivered) for target communities, as well as the opportunity costs linked to participation in this set of interventions.

The multi-ethnic context of Kassala state must also be considered. Despite years of settlement and acculturation among the differing tribes forming the state population, these groups still practice and pass on variants of their historical ways of life, which include perceptions and approaches to child care and feeding. Such differences may contribute to the results that we have observed. However, as noted in the IYCF result comparisons (see Section 6.6, Figure 39), no difference in IYCF practices was observed between control and intervention groups at baseline, which indicates relatively similar patterns of behaviours and practices related to infant and child feeding. However, it would be important to consider this multi-ethnic context in relation to the programme's SBCC intervention. This may require a tailored SBCC content and delivery by group that will address any infant and young child feeding beliefs and practices that are inconsistent with promoting healthy child growth.

The findings presented here were generated using a study design that is relatively more flexible to changes in programme implementation. The programme went through a number of modifications during the study period, such as a change in food products provided for treatment and prevention and some alterations to distribution modalities. These changes were all accounted for in the analysis, and this is reflected in the round-by-round and step-by-step analysis we performed over time. Given the sample sizes achieved for each round, a significant amount of data is available, allowing for robust analysis at temporal and spatial units of disaggregation.

River Atbara and rural Kassala were exposed to blanket FBMAM at specific points, although they were allocated as the control. However, based on the low level of programme coverage for these areas at certain points indicated in the timeline, it is unlikely that exposure to the intervention would have had any significant impact on results. Similarly, based on the low levels of coverage in intervention areas, the probability of spill-over to control areas is very low and will most likely have had little effect on the results.

The study has been designed and implemented in Kassala State, where the nutritional situation can deteriorate at certain periods during the year without it necessarily being considered an emergency and/or humanitarian situation in the Sudan context. Other states such as the Darfurs have poorer nutrition and are often in periods of nutritional emergency due to very high levels of GAM prevalence over considerable lengths of time.

However, an emergency/humanitarian context mainly impacts programme performance and coverage, and thus the effectiveness of its service delivery. The results of the study show that, even with a programme that is not at peak effectiveness (i.e. good performance but low coverage), reduction in at-risk prevalence is possible, though most likely of limited magnitude, which therefore decreases its potential ability to reduce MAM, and consequently GAM prevalence.

It is possible that the same package of treatment and prevention designed for Kassala, but with improved performance (specifically with regard to coverage) and sustained over at least a year or more, will produce similar effects on at-risk prevalence reduction, and potentially on MAM prevalence reduction, in other states in Sudan (or other countries). Differences in effects and impact may be influenced by programme quality, including the level of performance and coverage it achieves and how this is sustained over time. The links between all these factors, and nutrition outcomes such as prevalence of GAM, MAM and SAM, require further evaluation in different contexts to provide stronger evidence that can guide programme and policy decisions.

Our findings, and the points we discuss above, should be considered alongside the limitations of our study and its design. First, we acknowledge that our study was non-randomised. Programming needs were prioritised and specific areas were selected to receive intervention at the start. This could indicate inherent differences between intervention and control areas (such as a higher prevalence of acute malnutrition), and therefore liability to selection bias. However, as noted in Figure 9, there were no significant differences in rates of acute malnutrition in the control and intervention groups. We also reason that selection bias likely played a minimal factor in the incidence study, given that we performed exhaustive sampling in all areas covered by the programme in selected clusters.

Second, we acknowledge the limitations of the difference in differences analytical approach used in this study. A key assumption of this type of analysis is that the differences between intervention and control groups remain constant in the absence of treatment. The plausibility of this assumption can be tested when there are more than two periods being compared, and if there is baseline or pre-intervention period data in which no groups are receiving the intervention. We had originally planned for a baseline data collection period specifically for this purpose; however, this was not possible due to start-up delays. Whilst there is no overt reason to suggest that this assumption cannot be true, we are unable to empirically show this to be the case. Another issue common in the difference in differences approach is the Ashenfelter's dip (Heckman and Smith 1999), which affects studies where the decision to receive intervention is selected by the participants. This is similar to selection bias and is a plausible limitation for a non-randomised study such as this. However, other than the purposeful selection of intervention clusters at the start, individuals were selected for the programme based on clear, well-defined and measurable anthropometric criteria.

Finally, the wider external application of this study's findings should consider the various contextual factors of running the programme in Sudan's Kassala state. These include the area's chronically high rates of acute and chronic malnutrition; poor socio-economic situation; periodic crises linked to natural disasters or insecurity; operational challenges caused by state- and country-level socio-political structures affecting supply chains, logistics and finance systems; and the still-evolving CNIP, which is continually adjusted and organised to be relevant to state- and country-level contexts. These considerations likely make the results reported here very specific to a context with transitory food insecurity linked to seasonal or other fluctuating factors, and to settings with a highly mature and evolved community-based targeted feeding programme.

8. Specific findings for policy and practice

The prevention intervention did not show an effect on the prevalence and incidence of moderate or global acute malnutrition. However, the intervention did show a significant reduction in the prevalence of children at risk of acute malnutrition. Despite its lack of effect on the main outcome measure (prevalence of MAM, and consequently that of GAM), significant reduction in the prevalence of children at risk of acute malnutrition indicates the potential role for FBMAM to reduce malnutrition in a setting where treatment programmes for the acutely malnourished are already available and performing well.

It is possible that two key factors contributed to the intervention's lack of demonstrable effect on the prevalence of acute malnutrition: First, the duration of the prevention programme on which the study was based was relatively short (approximately eight months). Given the trajectory of acute malnutrition prevalence over time in the intervention group, it is possible that longer exposure to both treatment and prevention programmes would support an improved effect on incidence and prevalence, particularly if paired with improved programme performance (see below). This is important to note for WFP and others who run acute malnutrition programmes and are considering, or intending to run, prevention programmes. A sustained package of treatment and prevention could potentially have a positive effect on incidence and prevalence; this would therefore be a useful focus for further evaluation.

Currently, prevention programmes are initiated alongside treatment programmes in other states of Sudan and various countries where WFP operates. When resources are limited, treatment is often prioritised over prevention. Further research in the context of this evaluation is needed before conclusions can be drawn about the most effective and/or cost effective balance between FBMAM programmes and treatment in areas where acute malnutrition is not yet at high levels. Lower MAM caseloads could potentially facilitate concurrent implementation of prevention programmes.

There are examples of intervention modalities focused on acute malnutrition prevention instead of treatment that have shown a significant decrease in acute malnutrition incidence. Isanaka and others (2009) used a cluster randomised trial in Niger to demonstrate how short-term supplementation with SNF (specifically ready-to-use therapeutic food) among non-malnourished 6–60 month old children has reduced the incidence of wasting and severe wasting within a period of eight months. Talley and others (2012) showed similar results in a quasi-experimental study performed in Darfur, where they provided children 6–36 months of age (malnourished or non-malnourished) with either a Lipid-based Nutrient Supplement (LNS)[34] or an improved dry ration (similar

[34] Ready-to-use foods such as ready-to-use therapeutic food and RUSF are considered LNS products.

to SuperCereal+). The group given LNS noted a significantly higher mean weight-for-height z-score as compared to those receiving only improved dry rations, thereby prompting the proposal of LNS as an option for the prevention of acute malnutrition. Finally, another study in Niger by Langendorf and others (2014) provided various combinations of preventative interventions that included either provision of supplementary foods (LNS or SuperCereal+), cash, or a combination of both to households with at least one child 6–60 months (with or without malnutrition), and tested whether any of these modalities decreased incidence of MAM and SAM. They were able to demonstrate that supplementary food combined with cash transfers showed a more significant decrease in both MAM and SAM incidence.

Second, increased programme coverage has the potential to further reduce at-risk prevalence, which could lead to a demonstrable effect on the prevalence of moderate acute malnutrition. Currently, TSFP coverage levels are approaching 50% in some localities (i.e. Aroma) and up to 28% overall, which is relatively high compared to other treatment programmes. However prevention programme coverage is low at 10%. There are several areas of action identified during this evaluation that could support improved coverage of both the treatment and prevention arms of this programme:

Improve effectiveness of case-finding and referral of incident cases to food-based interventions

- Whilst ad hoc mass screening days will continue to be a useful mechanism to identify and refer both cases of MAM and those ‘at risk’, they will miss a high number of incident cases that occur between screening days. Active, regular screening should be strengthened to identify and refer these cases as they occur. The most obvious mechanism for this in WFP’s intervention areas in Sudan is MUAC screening during house-to-house visits through the existing network of community health workers/mobilisers. To enable this, however, their clinic-based workload must be reduced, and job descriptions must clearly state their responsibility for village-based house-to-house visits. The efficiency of this type of work can be supported by linking it to other community- and household-based activities (such as follow up of defaulters as noted below; delivery of information about the SBCC intervention, which is currently reaching a very low proportion of the population; and other health and nutrition tasks) and targeting households (or villages) known to be at particular risk of malnutrition.
- In lieu of (or in addition to) house-to-house screening by community workers, there is a growing appreciation of the role that mothers/carers themselves can play in the early identification and referral of malnutrition in their own children. A recent study by Blackwell and others (2015) found that mothers in Niger could classify their children into one of three colour classes on a MUAC tape, and had a sensitivity and specificity for classification of their child’s nutritional status of greater than 90% and greater than 80%, respectively, for GAM (defined by a MUAC less than 125mm). This suggests that mothers could become the focal point for improving early identification of children with, and at risk of, malnutrition in WFP’s target communities in Sudan – and that this may be a way to relieve the task-load of the community health workforce. As a first step, WFP may consider piloting this mechanism of case-finding and referral in a small area to examine its appropriateness and effectiveness for their Sudan programmes.

Improve effectiveness of community sensitisation to the programme, community understanding of malnutrition (and the programme's target groups) and participation in SBCC activities

- Community engagement and understanding of nutrition interventions has long been a known essential precursor to intervention effectiveness and coverage (Valid International 2006; Guerrero et al. 2010). The findings of this evaluation show limited knowledge and understanding among target groups of the mechanics and process of the FBMAM programme, most likely due to poor interface between the programme and the community. It also found that participation in the SBCC components of the programme is decreasing over time. One of the most effective mechanisms in addressing these issues is to support the community health workforce to spend more time at the village level doing house-to-house visits, and sensitising families, village leaders and stakeholders for health/nutrition about the programme's importance and objectives for the health of target groups. This will be especially important for a prevention programme that, in the eyes of the community, is targeting children that may not appear especially sick. Actions needed here are linked closely to those recommended above under case finding. The process of greater engagement at the village level in active screening, and involving mothers in monitoring the nutritional status of their own children, is likely to support sustained engagement, understanding and participation of beneficiaries in programme interventions.
- The most common reason given by individuals for non-participation in SBCC activities was 'no time/too busy'. Therefore, it may be relevant to review the delivery platforms for SBCC interventions with opportunity costs for participants in mind. The scale up of delivery mechanisms for this component such as TV and radio shows (which were starting as our evaluation came to an end) may support improved involvement, as the need for regular involvement in education sessions and women's health groups would no longer be required. Community consultation will be key to identify and support the most effective delivery mechanisms.
- Similar programmes have found low participation in SBCC-type interventions to be linked to a perceived (or real) irrelevance of the messages for the target audience. Whilst the sensitisation activities described above will help, a review of the design of WFP's SBCC interventions based on a context-specific causal analysis of undernutrition in target communities may improve perceived relevance, and thereby support better participation. It is also important to ensure good linkages between the intervention design and delivery to the MoH's own national SBCC package.

Improve understanding of programme performance and the nutritional situation in target areas over time

- Monitoring and record keeping: Whilst programme performance is currently reported to be very satisfactory (i.e. high recovery with low mortality and default), considerable gaps were noted in the collection and reporting of monitoring data at programme sites. This may result in misrepresentation of performance and, more importantly for coverage, a poor understanding of defaulting levels with absence of follow-up of individual defaulters. Defaulting (and understanding the reasons for defaulting) can be an extremely important indicator for programme coverage, as it represents opportunity costs families face for participation in an intervention. Having a complete picture of defaulting both over time and by individual programme sites can support timely adaptation of interventions to reduce opportunity costs for participants and keep coverage high.

- Monitoring prevalence of malnutrition at the community level: At present, the FMoH and WFP Sudan base all programme planning decisions on the results of ad hoc nutritional surveys implemented by themselves or partners. Considering that the last such survey was completed in 2013 (Sudan National S3M 2013), this approach is likely to mean that decisions are based on out-of-date information. A surveillance approach that collects data on an ongoing basis could be a more effective mechanism to ensure there are no missed opportunities to intervene in localities where GAM rates have exceeded emergency thresholds; that resources are being targeted to areas of highest need; and that the programme is responsive to need as it changes by area over time. Ensuring that interventions target areas of greatest need (prevalence) will support improved programme coverage by avoiding omission of high numbers of eligible individuals within the boundaries of specified project areas. Such surveillance approaches do not have to be high-cost methods: the community nutrition surveillance system established by UNICEF and the FMoH in Darfur (United Nations Children's Fund et al. 2011) uses a Rapid Assessment Methodology that relies on sampling as few as 200 children per locality twice a year.

Whilst the results of this evaluation have highlighted useful learning about the quality of WFP's programmes in Sudan and the impact of interventions on secondary outcomes (such as prevalence of children at risk of malnutrition), it has not been able to provide unequivocal evidence as to the effectiveness (or potential impacts) of the examined interventions on primary outcomes of prevalence/incidence of GAM, MAM and SAM. We have proposed several plausible explanations of this lack of impact on primary outcomes, including a time lag between at-risk and prevalence reduction, the effect of SAM treatment programmes on incident MAM cases, and the FBMAM programme's level of coverage. Further evaluation and research is therefore critical to providing a robust evidence base for the link between these factors and the incidence and prevalence of MAM and GAM, and to address other questions identified in recent reviews of MAM programming – including the type of product used for supplementation (LNS or SuperCereal+), timing and duration of the prevention intervention, and the effect of inclusion of additional household inputs such as cash transfers and other similar support to the prevention package. Wherever possible, future food-based prevention programmes run by WFP and others should maximise learning outputs through the inclusion of an operational research component at the design stage and/or a strong evaluation design such as that used for this project.

A missing key piece of information from this impact evaluation is a reliable comparison of the economic input needed for WFP's MAM treatment and prevention programmes and the respective cost effectiveness of different intervention packages. With limited resource availability, policy and budget allocations must be informed by programme costs and administrative complexity; therefore it is important to establish the FBMAM's cost-effectiveness to inform future decision-making within WFP and Sudan.

As explained throughout this report, it has not been possible to conduct a comprehensive CEA or basic cost analysis as part of this evaluation, despite different approaches. This is unfortunate, as economic reporting from field to HQ level is already an integral part of WFP's programme management and monitoring. Whilst provider cost data and budgets were provided to the evaluation team, it was not possible for the WFP CO team to provide data disaggregated by intervention package. This prevented the creation of a CEA to feed into programme evaluation and prioritisation. Future evaluations with a CEA component must adopt mechanisms to better ensure that the format of data collected throughout the programme cycle is well-adapted to the data needs for this type of analysis. Areas that might be actioned for future programmes and evaluations to support comprehensive cost-effectiveness analysis include:

- Cost data should be collected from the beginning of the programme, and in a format that distinguishes between individual programmes and different programme components/activities (e.g. start-up costs, personnel costs, operating costs and intervention components. A suggestion for cost centre allocation can be found in appendix H). The exercise of separating costs according to different programmes can be challenging but efforts should be made to achieve this, for example by estimating the proportional division of resources/costs. As WFP implementing partners undertake field activities related to MAM prevention that are not funded by WFP, they should be briefed on the importance of separating cost information from their different donors and throughout their programmes.
- Information required for a comprehensive cost analysis does not only include cost data, but routine monitoring data and background information on intervention and control communities that contextualise findings (e.g. geographic location, socio-economic and socio-cultural context and presence of other nutrition and health programmes). Therefore, recommendations for the improvement of routine monitoring data as presented above should be implemented, and implementers should be opportunistic in collecting data relevant to a cost-effectiveness evaluation when undertaking other activities such as community capacity assessments. This would entail gathering information for MAM prevention activities undertaken by WFP implementing partners with funding from other donors.
- In order to ensure the prioritisation of data collection required for comprehensive cost-effectiveness throughout the programme cycle, demand for cost data in appropriate formats must be present at all levels – from field to HQ – and the necessary resources be made available. Therefore, standard cost data reporting mechanisms and formats could be considered for different organisational levels within WFP and for implementing partners.

Finally, this study generated important information on research design in the context of working with operational programmes (such as those run by WFP) and working in challenging environments such as Sudan. Our use of the stepped-wedge design has beneficial applications in programme impact assessments in volatile and unpredictable settings. The design allows for potential variations and adaptations to be implemented as the need arises, especially given the challenging implementation landscape of humanitarian emergencies. As previously demonstrated, each round of data collection can be treated both as a stand-alone cross-sectional study that can provide useful and important information for programming and as a component element of an overall evaluation study that is able to show temporal trends in results. The stepped-wedge design also allows for a controlled study to be implemented without encountering ethical issues related to withholding interventions, as it follows the specified programme rollout. Our use of spatial sampling within the stepped-wedge designs also adds a spatial component to the interpretation of results, particularly on issues of programme prioritisation and targeting, which are critical decisions made during humanitarian emergencies.

9. Recommendations

Based on the evaluation's findings and conclusions, two broad recommendations have been developed with five specific areas of actions identified to improve future FBMAM interventions and the implementation of linked evaluations. These are set out in the following table, which also provides the rationale and implications for each.

<table>
<tr><th colspan="2">Recommendation: Wherever possible, future food-based prevention programmes run by WFP and others should maximise learning outputs through the inclusion of an operational research component in the design stage, and/or a strong evaluation design. This research/evaluation should be linked to interventions that are able to address some of the issues that are likely to have limited observable impacts under this evaluation, such as low intervention coverage and format of programme cost data.</th></tr>
<tr><th>Rationale</th><th>Specific Actions</th></tr>
<tr><td>This evaluation was not able to provide unequivocal evidence as to the most effective or cost-effective intervention modality. The study results did demonstrate that, in this context, the addition of a FBMAM programme to a TSFP as a package intervention for the treatment and prevention of MAM has decreased the prevalence of at-risk children but not of MAM, SAM and GAM incidence or prevalence. In this case, our prevalence of both at risk and incidence findings seem to contradict each other, given that we expected to see a significant reduction in numbers of at-risk children to translate into a reduction in MAM prevalence secondary to reduced MAM incidence.

The possible reasons for the absence of effect on prevalent and incident cases despite the observed</td><td>1. Research and evaluation of FBMAM programmes should be implemented in different contexts and should further address specifically:
• The length of exposure to both treatment and prevention programmes needed to support an improved effect on incidence and prevalence;
• The effect of SAM treatment performance and coverage on the nutrition impacts of FBMAM programmes;
• The link between coverage of these types of interventions and outcomes, such as incidence and prevalence of MAM and GAM.

2. To support the data needs for a comprehensive CEA of MAM prevention, both programmes and future evaluations should:
• Collect cost data from the beginning of the programme, in a format that distinguishes both between individual</td></tr>
</table>

decrease in numbers 'at risk' are: (1) the time lag between at-risk and prevalence reduction (as mediated by incidence reduction); (2) incident MAM cases may come from previous SAM cases discharged from SAM treatment; and (3) the level of coverage of the FBMAM programme is too low to support any change in prevalence or incidence at a population level. A missing key piece of information from this impact evaluation is a reliable comparison of the economic input needed for WFP's MAM treatment and prevention programmes, and the respective cost effectiveness of different packages of interventions. With limited resources available, policy and budget allocations must be informed by the costs and administrative complexity of programmes; therefore it is important that the FBMAM's cost-effectiveness is established to inform future decision-making within WFP and Sudan.	programmes and different programme components/activities. WFP implementing partners undertaking field activities related to MAM prevention that are not funded by WFP should be briefed on the importance of separating cost information from their different donors and throughout their programmes; • Strengthen collection of other data such as routine monitoring data and background information on intervention and control communities that contextualise findings. Recommended actions (outlined below) for the improvement of routine monitoring data will support this; • Ensure that demand for cost data in appropriate formats for CEA is present at all levels from field to HQ, and the necessary resources for its collection are made available. Review of standard cost data reporting mechanisms and formats could be considered for different organisational levels within WFP and for implementing partners.

Recommendation: Improve coverage of both treatment and prevention arms of this programme.	
Rationale	**Action Points**
It is very likely that the level of coverage of the FBMAM programme was too low at the time of this evaluation to support any change in prevalence or incidence at a population level. Whilst ad hoc mass screening days will continue to be a useful mechanism to identify and refer both cases of MAM and those at risk, they will miss a high number of incident cases that occur between screening days.	**3. To improve the effectiveness of case-finding and referral of incident cases to food-based interventions, we propose the following actions:** • Active, regular screenings should be strengthened to identify and refer cases as they occur. The most obvious mechanism in WFP's Sudan intervention areas is MUAC screening during house-to-house visits by the existing network of community health workers/mobilisers. To enable this, however, their clinic-based workload

Community engagement and understanding of nutrition interventions has long been a known essential precursor to intervention effectiveness and coverage. The findings of this evaluation show there is limited knowledge and understanding among target groups of the mechanics and process of the FBMAM programme, most likely due to poor interface between the programme and the community. It also found that participation in the SBCC components of the programme is decreasing over time. Whilst programme performance is currently reported to be very satisfactory (i.e. high recovery with low mortality and default), there were considerable gaps noted by this evaluation in the collection and reporting of monitoring data at programme sites. This may result in misrepresentation of performance and, more importantly for coverage, in a poor understanding of defaulting levels with absence of follow up with individual defaulters. Defaulting (and understanding the reasons for defaulting) can be an extremely important indicator for programme coverage, as it represents the opportunity costs for families of participation in an intervention. WFP Sudan bases all programme planning decisions on the results of ad hoc nutritional surveys implemented by themselves or partners. Considering the last such survey was completed in 2013, this approach is likely to mean that decisions are being made based on out-of-date information. More timely information is required to ensure that opportunities are not missed to intervene in localities where GAM	must be reduced, and job descriptions must clearly state their responsibility for village-based house-to-house visits. The efficiency of this type of work can be supported by linking it to other community- and household-based activities (such as follow up of defaulters as noted below; delivery of information about the SBCC intervention, which is currently reaching a very low proportion of the population; and other health and nutrition tasks), and by targeting households (or villages) known to be at particular risk of malnutrition. • In lieu of (or in addition to) house-to-house screening by community workers, there is a growing appreciation of the role that mothers/carers can play in the early identification and referral of malnutrition in their own children. As a first step, WFP may consider piloting this mechanism of case-finding and referral in a small area to examine the appropriateness and effectiveness of this approach for their Sudan programmes. **4. To improve the effectiveness of community sensitisation to the programme, as well as community understanding of malnutrition (and the programme's target groups) and participation in SBCC activities, we propose the following actions:** • A very effective mechanism to address these issues is supporting the community health workforce to spend more time at the village level undertaking house-to-house visits and sensitising families, village leaders and stakeholders for health/nutrition about the importance of the programme and its objectives for the health of target groups. This will be especially important for a prevention programme that, in the eyes of the community, is targeting

rates have exceeded emergency thresholds, that resources are being targeted to areas of highest need, and that the programme is responsive to need as it changes by area over time.	children that may not appear particularly sick. Required actions are linked closely to those recommended above under case finding. The process of greater engagement at the village level in active screening and involving mothers in monitoring the nutritional status of their own children is likely to support sustained engagement, understanding, and participation of beneficiaries in programme interventions. • The most common reason given by individuals for non-participation in SBCC activities was ‘no time/too busy’. In light of this, it may be relevant to review the delivery platforms for SBCC interventions while keeping in mind the opportunity costs for participants. The scale up of delivery mechanisms for this component intervention such as TV and radio shows (which were starting as our evaluation came to an end), may support improved involvement, as the need would no longer be required for regular involvement in education sessions and women’s health groups, which may incur substantial opportunity costs. Community consultation will be key to identify and support the most effective and practical delivery mechanisms. • Similar programmes have found low participation in SBCC-type interventions to be linked to a perceived (or real) irrelevance of the messages for the target audience. Whilst the sensitisation activities described above will be helpful, a review of the design of WFP’s SBCC interventions based on a context-specific causal analysis of undernutrition in target communities may improve perceived relevance and as such support better participation. Ensuring good linkages between intervention design and delivery to the

	MoH's own national SBCC package will also prove useful. **5. To improve understanding of programme performance and the nutritional situation in target areas over time, we propose the following actions:** • Strengthen the collection and reporting of monitoring data at programme sites through actions such as training and support for implementing partners. • Strengthen the monitoring of malnutrition prevalence at the community level through actions such as surveillance that collects data on an ongoing basis. Such approaches do not have to be high cost; methods such as the community nutrition surveillance system established by UNICEF and the MoH in Darfur (United Nations Children's Fund et al. 2011) uses a Rapid Assessment Methodology, which relies on sampling as few as 200 children per locality twice per year.

Online appendices

Note to the reader: These appendices are only available online and have been published as received from the authors. They have not been copy-edited or formatted by 3ie, and can be accessed through the links provided below:

Appendix A: Field notes and other information from formative work

http://www.3ieimpact.org/media/filer_public/2018/03/27/gfr-tw61026-appendix-a.pdf

Appendix B: Sample design

http://www.3ieimpact.org/media/filer_public/2018/03/27/gfr-tw61026-appendix-b.pdf

Appendix C: Survey instruments

http://www.3ieimpact.org/media/filer_public/2018/03/27/gfr-tw61026-appendix-c.pdf

Appendix D: Pre-analysis plan

http://www.3ieimpact.org/media/filer_public/2018/03/27/gfr-tw61026-appendix-d.pdf

Appendix E: Sample size and power calculations

http://www.3ieimpact.org/media/filer_public/2018/03/27/gfr-tw61026-appendix-e.pdf

Appendix F: Descriptive statistics

http://www.3ieimpact.org/media/filer_public/2018/03/27/gfr-tw61026-appendix-f.pdf

Appendix G: Results

http://www.3ieimpact.org/media/filer_public/2018/05/02/gfr-tw61026-appendix-g.pdf

Appendix H: Cost data for the programme implementation

http://www.3ieimpact.org/media/filer_public/2018/03/27/gfr-tw61026-appendix-h.pdf

Appendix I: Analysis scripts in R

http://www.3ieimpact.org/media/filer_public/2018/05/02/gfr-tw61026-appendix-i.pdf

Appendix J: Glossary

http://www.3ieimpact.org/media/filer_public/2018/05/02/gfr-tw61026-appendix-j.pdf

References

Abu-Saad, K and Fraser, D, 2010. Maternal nutrition and birth outcomes. *Epidemiologic reviews*, 32(1), pp. 5–25.

Acharya, P and Kenefick, E, 2012. Improving blanket supplementary feeding programme (BSFP) efficiency in Sudan. *Field Exchange*, (42), pp.1–7.

Annan, RA, Webb, P and Brown, R, 2014. *Management of moderate acute malnutrition (MAM): current knowledge and practice*. In CMAM Forum Technical Brief.

Bachmann, MO, 2009. Cost effectiveness of community-based therapeutic care for children with severe acute malnutrition in Zambia: decision tree model. *Cost effectiveness and resource allocation,* 7(1), pp.2–9.

Becic, T, Mokhtar, N, Wegner, C and Loechl, C, 2014. International symposium: understanding moderate malnutrition in children for effective interventions. *Sight and Life*, 28(2), pp.1–6.

Bhutta, Z.A., Ahmed, T., Black, R.E., Cousens, S., Dewey, K., Giugliani, E., Haider, B.A., Kirkwood, B., Morris, S.S., Sachdev, H.P.S. and Shekar, M., 2008. What works? Interventions for maternal and child undernutrition and survival. *The Lancet*, 371(9610), pp.417–440.

Bhutta, Z.A., Das, J.K., Rizvi, A., Gaffey, M.F., Walker, N., Horton, S., Webb, P., Lartey, A. and Black, R.E., 2013. Evidence-based interventions for improvement of maternal and child nutrition: what can be done and at what cost? *The Lancet*, 382(9890), pp.452–477.

Black, R.E., Allen, L.H., Bhutta, Z.A., Caulfield, L.E., De Onis, M., Ezzati, M., Mathers, C., Rivera, J. and Maternal and Child Undernutrition Study Group, 2008. Maternal and child undernutrition: global and regional exposures and health consequences. *The Lancet*, 371(9608), pp.243–260.

Black, R.E., Victora, C.G., Walker, S.P., Bhutta, Z.A., Christian, P., De Onis, M., Ezzati, M., Grantham-McGregor, S., Katz, J., Martorell, R. and Uauy, R., 2013. Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet*, 382(9890), pp.427–451.

Blackwell, N et al., 2015. Mothers Understand And Can do it (MUAC): a comparison of mothers and community health workers determining mid-upper arm circumference in 103 children aged from 6 months to 5 years. *Archives of Public Health*, 73(1), p.7.

Bryce, J., Coitinho, D., Darnton-Hill, I., Pelletier, D., Pinstrup-Andersen, P. and Maternal and Child Undernutrition Study Group, 2008. Maternal and child undernutrition: effective action at national level. *The Lancet*, 371(9611), pp.510–526.

Central Bureau of Statistics, 2008. *Fifth Sudan population and housing census*.

Dale, N.M., Myatt, M., Prudhon, C. and Briend, A., 2017. Using cross-sectional surveys to estimate the number of severely malnourished children needing to be enrolled in specific treatment programmes. *Public Health Nutrition*, 20(8), pp.1–5.

Diaconis, P and Efron, B, 1983. *Computer intensive methods in statistics*. Stanford, California: Division of Biostatistics, Stanford University.

Frankel, S, Roland, M and Makinen, M, 2015. *Costs, cost-effectiveness, and financial sustainability of community-based management of acute malnutrition in northern Nigeria*. Washington, D.C.: Results for Development Institute.

Guerrero, S, Myatt, M and Collins, S, 2010. Determinants of coverage in community-based therapeutic care programmes: towards a joint quantitative and qualitative analysis. *Disasters*, 34(2), pp.571–585.

Hayes, RJ and Bennett, S, 1999. Simple sample size calculation for cluster-randomized trials. *International Journal of Epidemiology*, 28(2), pp.319–326.

Heckman, JJ and Smith, JA, 1999. The pre-programme earnings dip and the determinants of participation in a social programme. Implications for simple programme evaluation strategies. *The Economic Journal*, 109(457), pp.313–348.

International Initiative for Impact Evaluation, 2016. Principles for impact evaluation. *International Initiative for Impact Evaluation*, pp.1–2. Available at: <http://www.3ieimpact.org/media/filer_public/2017/02/08/principles-for-ie.pdf> [Accessed 1 March 2017].

Isanaka, S., Nombela, N., Djibo, A., Poupard, M., Van Beckhoven, D., Gaboulaud, V., Guerin, P.J. and Grais, R.F., 2009. Effect of preventive supplementation with ready-to-use therapeutic food on the nutritional status, mortality, and morbidity of children aged 6 to 60 months in Niger: a cluster randomized trial. *JAMA: The Journal of the American Medical Association*, 301(3), pp.277–285.

Isanaka, S et al., 2011. Estimates of the duration of untreated acute malnutrition in children from Niger. *American Journal of Epidemiology*, 173(8), pp.932–940.

Isanaka, S., Boundy, E.O.N., Grais, R.F., Myatt, M. and Briend, A., 2016. Improving estimates of numbers of children with severe acute malnutrition using cohort and survey data. *American Journal of Epidemiology*, 184(12), pp.861–869.

James, P., Sadler, K., Wondafrash, M., Argaw, A., Luo, H., Geleta, B., Kedir, K., Getnet, Y., Belachew, T. and Bahwere, P., 2016. Children with moderate acute malnutrition with no access to supplementary feeding programmes experience high rates of deterioration and no improvement: results from a prospective cohort study in rural Ethiopia. *PloS One*, 11(4). e0153530–26.

Japan International Cooperation Agency, 2015. *Capacity development project for the provision of services for basic human needs in Kassala, the republic of Sudan: final report volume 1: synthesis report and planning cluster*. Japan International Cooperation Agency.

Jimenez, M and Stone-Jimenez, M, 2014. Preventing moderate acute malnutrition (MAM) through nutrition-specific interventions. *CMAM Forum Technical Brief*, pp.1–40.

Kaiser, R., Woodruff, B.A., Bilukha, O., Spiegel, P.B. and Salama, P., 2006. Using design effects from previous cluster surveys to guide sample size calculation in emergency settings. *Disasters*, 30(2), pp.199–211.

Katz, J., Lee, A.C., Kozuki, N., Lawn, J.E., Cousens, S., Blencowe, H., Ezzati, M., Bhutta, Z.A., Marchant, T., Willey, B.A. and Adair, L., 2013. Mortality risk in preterm and small-for-gestational-age infants in low-income and middle-income countries: a pooled country analysis. *The Lancet*, 382(9890), pp.417–425.

Langendorf, C., Roederer, T., de Pee, S., Brown, D., Doyon, S., Mamaty, A.A., Touré, L.W.M., Manzo, M.L. and Grais, R.F., 2014. Preventing acute malnutrition among young children in crises: a prospective intervention study in Niger. *PLoS medicine*, *11*(9), p.e1001714.

Miller, C, 2005. Power, land and ethnicity in the Kassala-Gedaref states. Land, ethnicity and political legitimacy in eastern Sudan. Cairo, Egypt: Cedej, pp. 3–58.

Milne, A, 1959. The centric systematic area-sample treated as a random sample. *Biometrics*, 15(2), pp.270–297.

Morris, SS, Cogill, B and Uauy, R, 2008. Effective international action against undernutrition: why has it proven so difficult and what can be done to accelerate progress? *The Lancet*, 371(9612), pp.608–621.

Mucha, N, 2014. Preventing moderate acute malnutrition (MAM) through nutrition-sensitive interventions. *CMAM Forum Technical Brief*, pp.1–49.

Myatt, M., Feleke, T., Sadler, K. and Collins, S., 2005. A field trial of a survey method for estimating the coverage of selective feeding programmes. *Bulletin of the World Health Organization*, 83(1), pp.20–26.

Myatt, M, 2012. *How do we estimate case load for pregnant and lactating women 6 months postpartum in a given time period?* CMAM Forum.

Puett, C., Sadler, K., Alderman, H., Coates, J., Fiedler, J.L. and Myatt, M., 2012. Cost-effectiveness of the community-based management of severe acute malnutrition by community health workers in southern Bangladesh. *Health Policy and Planning*, 28(4), pp.386–399.

Rogers, E and Guerrero, S, 2014. *Access for All, Volume 3: What can community-based SAM treatment learn from other public health interventions to improve access and coverage?* London: Coverage Monitoring Network. Available at: <http://www.coverage-monitoring.org/wp-content/uploads/2014/04/Access-for-All-Volume-3.pdf> [Accessed on 28 June 2018].

Ruel, M.T., Menon, P., Habicht, J.P., Loechl, C., Bergeron, G., Pelto, G., Arimond, M., Maluccio, J., Michaud, L. and Hankebo, B, 2008. Age-based preventive targeting of food assistance and behaviour change and communication for reduction of childhood undernutrition in Haiti: a cluster randomised trial. *The Lancet*, 371(9612), pp.588–595.

Sudan Federal Ministry of Health, 2013. *The Republic of Sudan National Health Sector Strategic Plan II (2012-2016)*. Khartoum: Sudan Federal Ministry of Health.

Sudan Federal Ministry of Health, 2014. *National Nutrition Strategic Plan 2014-2018*. Khartoum: Sudan Federal Ministry of Health.

Sudan National S3M, 2013. *Report of a simple spatial surveying method (s3m) survey in Sudan*. Khartoum: Sudan Federal Ministry of Health.

Talley, L, 2012. Prevention of acute malnutrition during the lean season: comparison of a lipid-based nutrient supplement and an improved dry ration, south Darfur, Sudan. A quasi-experimental study. *Journal of Nutritional Disorders and Therapy*, 02(03), pp.1–5.

Tanahashi, T, 1978. Health service coverage and its evaluation. *Bulletin of the World Health Organization*, 56(2), pp.295–303.

Tekeste, A., Wondafrash, M., Azene, G. and Deribe, K, 2012. Cost effectiveness of community-based and in-patient therapeutic feeding programs to treat severe acute malnutrition in Ethiopia. *Cost effectiveness and resource allocation,* 10(1), p.4.

The Sphere Project, 2011. *The Sphere Project: humanitarian charter and minimum standards in humanitarian response*.

United Nations Children's Fund, 1997. *The State of the World's Children 1997*, Oxford: Oxford University Press.

United Nations Children's Fund, 2011. *Darfur Nutrition Update*, Available at: https://www.unicef.org/sudan/Darfur_NutritionUpdate30OND2011.pdf.

Valid International, 2006. *Community-based Therapeutic Care (CTC): A Field Manual*. Oxford: Valid International.

Victora, C.G., Adair, L., Fall, C., Hallal, P.C., Martorell, R., Richter, L., Sachdev, H.S, 2008. Maternal and child undernutrition: consequences for adult health and human capital. *The Lancet*, 371(9609), pp.340–357.

Wilford, R, Golden, K and Walker, DG, 2012. Cost-effectiveness of community-based management of acute malnutrition in Malawi. *Health Policy and Planning*, 27(2), pp.127–137.

Woertman, W., de Hoop, E., Moerbeek, M., Zuidema, S.U., Gerritsen, D.L. and Teerenstra, S, 2013. Stepped wedge designs could reduce the required sample size in cluster randomized trials. *Journal of Clinical Epidemiology*, 66(7), pp.752–758.

World Health Organization, 2003. *Making choices in health: WHO guide to cost-effectiveness analysis* T. Baltussen, R.M., Adam, T., Tan-Torres Edejer, T., Hutubessy, R.C., Acharya, A., Evans, D.B., Murray, C.J., eds., Geneva: World Health Organization.

World Health Organization, 2017. WHO | Challenges. Available at: <http://www.who.int/nutrition/challenges/en/> [Accessed 20 February 2017].

World Food Programme, 2012a. *A Comprehensive Food Security Assessment in Kassala State, Sudan*. Kassala: World Food Programme.

World Food Programme, 2012b. WFP nutrition policy, nutrition programs and food supplements: Updated with new evidence, scientific knowledge and global partnerships. Available at: <http://documents.wfp.org/stellent/groups/public/documents/resources/wfp247204.pdf> [Accessed 28 June 2018].

World Food Programme, 2015. *Community-based Nutrition Integrated Programme (CNIP)* 2nd ed., World Food Programme.

World Food Programme Sudan, 2015. *UN World Food Programme Sudan Country Strategy*, Khartoum: World Food Programme Sudan.

Zerfu, TA and Ayele, HT, 2013. Micronutrients and pregnancy; effect of supplementation on pregnancy and pregnancyoutcomes: a systematic review. *Nutrition Journal*, 12(1), p.1.

Other publications in the 3ie Impact Evaluation Report Series

The following reports are available from http://www.3ieimpact.org/en/publications/3ie-impact-evaluation-reports/3ie-impact-evaluations/

Impact evaluation of WFP's programs targeting moderate acute malnutrition in humanitarian situations in Chad. 3ie Impact Evaluation Report 78. Saboya, M, Rudiger, J, Frize, J, Ruegenberg, D, Rodríguez Seco, A and McMillon, C, 2018.

Improving midday meal delivery and encouraging micronutrient fortification among children in India, 3ie Impact Evaluation Report 77. Shastry, GK, Berry, J, Mukherjee, P, Mehta, S and Ruebeck, H, 2018.

Evaluation of infant development centres: an early years intervention in Colombia, 3ie Impact Evaluation Report 76. Andrew, A, Attanasio, O, Bernal, R, Cordona, L, Krutikova, S, Heredia, DM, Medina, C, Peña, X, Rubio-Codina, M and Vera-Hernandez, M, 2018.

Can the wounds of war be healed? Experimental evidence on reconciliation in Sierra Leone. 3ie Impact Evaluation Report 75. Cilliers, J, Dube, O and Siddiqi, B, 2018.

Impact evaluation of the Menabe and Melaky development programme in Madagascar, 3ie Impact Evaluation Report 74. Ring, H, Morey, M, Kavanagh, E, Kamto, K, McCarthy, N, Brubaker, J and Rakotondrafara, C, 2018.

Impact evaluation of the Smallholder Dairy Commercialization Programme in Kenya, 3ie Impact Evaluation Report 73. Bonilla, J, McCarthy, N, Mugatha, S, Rai, N, Coombes, A and Brubaker, J, 2018.

Impact and adoption of risk-reducing drought-tolerant rice in India, 3ie Impact Evaluation Report 72. Yamano, T, Dar, MH, Panda, A, Gupta, I, Malabayabas, ML and Kelly, E, 2018.

Poverty and empowerment impacts of the Bihar Rural Livelihoods Project in India, 3ie Impact Evaluation Report 71. Hoffmann, V, Rao, V, Datta, U, Sanyal, P, Surendra, V and Majumdar, S 2018.

How should Tanzania use its natural gas? Citizens' views from a nationwide Deliberative Poll, 3ie Impact Evaluation Report 70. Birdsall, N, Fishkin, J, Haqqi, F, Kinyondo, A, Moyo, M, Richmond, J and Sandefur, J, 2018.

Impact evaluation of the conditional cash transfer program for secondary school attendance in Macedonia, 3ie Impact Evaluation Report 69. Armand, A and Carneiro, P, 2018.

Age at marriage, women's education, and mother and child outcomes in Bangladesh, 3ie Impact Evaluation Report 68. Field, E, Glennerster, R, Nazneen, S, Pimkina, S, Sen, I and Buchmann, N, 2018.

Evaluating agricultural information dissemination in western Kenya, *3ie Impact Evaluation Report 67.* Fabregas, R, Kremer, M, Robinson, J and Schilbach, F, 2017.

General equilibrium impact assessment of the Productive Safety Net Program in Ethiopia, 3ie Impact Evaluation Report 66. Filipski, M, Taylor, JE, Abegaz, GA, Ferede, T, Taffesse, AS and Diao, X, 2017.

Impact of the Uddeepan programme on child health and nutrition in India, 3ie Impact Evaluation Report 65. Kochar, A, Sharma, A and Sharma, A, 2017.

Evaluating oral HIV self-testing to increase HIV testing uptake among truck drivers in Kenya, 3ie Impact Evaluation Report 64. Kelvin, EA, Mwai, E, Romo, ML, George, G, Govender, K, Mantell, JE, Strauss, M, Nyaga, EN and Odhiambo, JO, 2017.

Integration of EPI and paediatric HIV services for improved ART initiation in Zimbabwe, *3ie Impact Evaluation Report 63.* Prescott, M, Boeke, C, Gotora, T, Mafaune, HW, Motsi, W, Graves, J, Mangwiro, A and McCarthy, E, 2017.

Increasing male partner HIV testing using self-test kits in Kenya, 3ie Impact Evaluation Report 62. Gichangi, A, Korte, JE, Wambua, J, Vrana, C and Stevens, D, 2017.

Evaluating the impact of community health worker integration into prevention of mother-to-child transmission of HIV services in Tanzania, 3ie Impact Evaluation Report 61. Nance, N, McCoy, S, Ngilangwa, D, Masanja, J, Njau, P and Noronha, R, 2017.

Using HIV self-testing to promote male partner and couples testing in Kenya, 3ie Impact Evaluation Report 60. Thirumurthy, H, Omanga, E, Obonyo, B, Masters, S and Agot, K, 2017.

Increasing male partner HIV self-testing at antenatal care clinics in Kenya, 3ie Impact Evaluation Report 59. Gichangi, A, Korte, JE, Wambua, J, Vrana, C and Stevens, D, 2017.

Impact of free availability of public childcare on labour supply and child development in Brazil, *3ie Impact Evaluation Report 58*. Attanasio, O, Paes de Barros, R, Carneiro, P, Evans, D, Lima, L, Olinto, P and Schady, N, 2017.

Estimating the effects of a low-cost early stimulation and parenting education programme in Mexico, 3ie Impact Evaluation Report 57. Cardenas, S, Evans, D and Holland, P, 2017.

The Better Obstetrics in Rural Nigeria study: an impact evaluation of the Nigerian Midwives Service Scheme, 3ie Impact Evaluation Report 56. Okeke, E, Glick, P, Abubakar, IS, Chari, AV, Pitchforth, E, Exley, J, Bashir, U, Setodji, C, Gu, K and Onwujekwe, O, 2017.

The Productive Safety Net Programme in Ethiopia: impacts on children's schooling, labour and nutritional status, *3ie Impact Evaluation Report 55.* Berhane, G, Hoddinott, J, Kumar, N and Margolies, A, 2016.

The impact of youth skills training on the financial behaviour, employability and educational choice in Morocco, 3ie Impact Evaluation Report 54. Bausch, J, Dyer, P, Gardiner, D, Kluve, J and Mizrokhi, E, 2016.

Using advertisements to create demand for voluntary medical male circumcision in South Africa, 3ie Impact Evaluation Report 53. Frade, S, Friedman, W, Rech, D and Wilson, N, 2016.

The use of peer referral incentives to increase demand for voluntary medical male circumcision in Zambia, 3ie Impact Evaluation Report 52. Zanolini, A, Bolton, C, Lyabola, LL, Phiri, G, Samona, A, Kaonga, A and Harsha Thirumurthy, H, 2016.

Using smartphone raffles to increase demand for voluntary medical male circumcision in Tanzania, 3ie Impact Evaluation Report 51. Mahler, H and Bazant, E, 2016.

Voluntary medical male circumcision uptake through soccer in Zimbabwe, 3ie Impact Evaluation Report 50. DeCelles, J, Kaufman, Z, Bhauti, K, Hershow, R, Weiss, H, Chaibva, C, Moyo, N, Braunschweig, E, Mantula, F, Hatzold, K and Ross, D, 2016.

Measuring the impact of SMS-based interventions on uptake of voluntary medical male circumcision in Zambia, 3ie Impact Evaluation Report 49. Leiby, K, Connor, A, Tsague, L, Sapele, C, Koanga, A, Kakaire, J and Wang, P, 2016.

Assessing the impact of delivering messages through intimate partners to create demand for voluntary medical male circumcision in Uganda, 3ie Impact Evaluation Report 48. Semeere, AS, Bbaale, DS, Castelnuovo, B, Kiragga, A, Kigozi, J, Muganzi, A, Kambugu, A and Coutinho, AG, 2016.

Optimising the use of economic interventions to increase demand for voluntary medical male circumcision in Kenya, 3ie Impact Evaluation Report 47. Thirumurthy, H, Omanga, E, Rao, SO, Murray, K, Masters, S and Agot, K, 2016.

The impact of earned and windfall cash transfers on livelihoods and conservation in Sierra Leone, 3ie Impact Evaluation Report 46. Bulte, E, Conteh, B, Kontoleon, A, List, J, Mokuwa, E, Richards, P, Turley, T and Voors, M, 2016.

Property tax experiment in Pakistan: Incentivising tax collection and improving performance, 3ie Impact Evaluation Report 45. Khan, A, Khwaja, A and Olken, B, 2016.

Impact of mobile message reminders on tuberculosis treatment outcomes in Pakistan, 3ie Impact Evaluation Report 44. Mohammed, S, Glennerster, R and Khan, A, 2016.

Making networks work for policy: Evidence from agricultural technology adoption in Malawi, 3ie Impact Evaluation Report 43. Beaman, L, BenYishay, A, Fatch, P, Magruder, J and Mobarak, AM, 2016.

Estimating the impact and cost-effectiveness of expanding access to secondary education in Ghana, 3ie Impact Evaluation Report 42. Dupas, P, Duflo, E and Kremer, M, 2016.

Evaluating the effectiveness of computers as tutors in China, 3ie Impact Evaluation Report 41. Mo, D, Bai, Y, Boswell, M and Rozelle, S, 2016.

Micro entrepreneurship support programme in Chile, 3ie Impact Evaluation Report 40. Martínez, CA, Puentes, EE and Ruiz-Tagle, JV, 2016.

Thirty-five years later: evaluating the impacts of a child health and family planning programme in Bangladesh, 3ie Impact Evaluation Report 39. Barham, T, Kuhn, R, Menken, J and Razzaque, A, 2016.

Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection and malnutrition in India, 3ie Impact Evaluation Report 38. Clasen, T, Boisson, S, Routray, P, Torondel, B, Bell, M, Cumming, O, Ensink, J, Freeman, M and Jenkins, M, 2016.

Evaluating the impact of vocational education vouchers on out-of-school youth in Kenya, 3ie Impact Evaluation Report 37. Hicks, JH, Kremer, M, Mbiti, I and Miguel, E, 2016.

Removing barriers to higher education in Chile: evaluation of peer effects and scholarships for test preparation, 3ie Impact Evaluation Report 36. Banerjee, A, Duflo E and Gallego, F, 2016.

Sustainability of impact: dimensions of decline and persistence in adopting a biofortified crop in Uganda, 3ie Impact Evaluation Report 35. McNiven, S, Gilligan, DO and Hotz, C 2016.

A triple win? The impact of Tanzania's Joint Forest Management programme on livelihoods, governance and forests, *3ie Impact Evaluation Report 34.* Persha, L and Meshack, C, 2016.

The effect of conditional transfers on intimate partner violence: evidence from Northern Ecuador, 3ie Impact Evaluation Report 33. Hidrobo, M, Peterman, A and Heise, L, 2016.

The effect of transfers and preschool on children's cognitive development in Uganda, 3*ie Impact Evaluation Report 32*. Gillian, DO and Roy, S, 2016.

Can egovernance reduce capture of public programmes? Experimental evidence from India's employment guarantee, 3ie Impact Evaluation Report 31. Banerjee, A, Duflo, E, Imbert, C, Mathew, S and Pande, R, 2015.

Improving maternal and child health in India: evaluating demand and supply strategies, 3ie Impact Evaluation Report 30. Mohanan, M, Miller, G, Forgia, GL, Shekhar, S and Singh, K, 2016.

Smallholder access to weather securities in India: demand and impact on production decisions, 3ie Impact Evaluation Report 28. Ceballos, F, Manuel, I, Robles, M and Butler, A, 2015.

What happens once the intervention ends? The medium-term impacts of a cash transfer programme in Malawi, 3ie Impact Evaluation Report 27. Baird, S, Chirwa, E, McIntosh, C and Özler, B, 2015.

Validation of hearing screening procedures in Ecuadorian schools, 3ie Impact Evaluation Report 26. Muñoz, K, White, K, Callow-Heusser, C and Ortiz, E, 2015.

Assessing the impact of farmer field schools on fertilizer use in China, 3ie Impact Evaluation Report 25. Burger, N, Fu, M, Gu, K, Jia, X, Kumar, KB and Mingliang, G, 2015.

The SASA! study: a cluster randomised trial to assess the impact of a violence and HIV prevention programme in Kampala, Uganda, *3ie Impact Evaluation Report 24.* Watts, C, Devries, K, Kiss, L, Abramsky, T, Kyegombe, N and Michau, L, 2014.

Enhancing food production and food security through improved inputs: an evaluation of Tanzania's National Agricultural Input Voucher Scheme with a focus on gender impacts, 3ie Impact Evaluation Report 23. Gine, X, Patel, S, Cuellar-Martinez, C, McCoy, S and Lauren, R, 2015.

A wide angle view of learning: evaluation of the CCE and LEP programmes in Haryana, 3ie Impact Evaluation Report 22. Duflo, E, Berry, J, Mukerji, S and Shotland, M, 2015.

Shelter from the storm: upgrading housing infrastructure in Latin American slums, *3ie Impact Evaluation Report 21.* Galiani, S, Gertler, P, Cooper, R, Martinez, S, Ross, A and Undurraga, R, 2015.

Environmental and socioeconomic impacts of Mexico's payments for ecosystem services programme, *3ie Impact Evaluation Report 20.* Alix-Garcia, J, Aronson, G, Radeloff, V, Ramirez-Reyes, C, Shapiro, E, Sims, K and Yañez-Pagans, P, 2015.

A randomised evaluation of the effects of an agricultural insurance programme on rural households' behaviour: evidence from China, 3ie Impact Evaluation Report 19. Cai, J, de Janvry, A and Sadoulet, E, 2014.

Impact of malaria control and enhanced literacy instruction on educational outcomes among school children in Kenya: a multi-sectoral, prospective, randomised evaluation, 3ie Impact Evaluation Report 18. Brooker, S and Halliday, K, 2015.

Assessing long-term impacts of conditional cash transfers on children and young adults in rural Nicaragua, *3ie Impact Evaluation Report 17.* Barham, T, Macours, K, Maluccio, JA, Regalia, F, Aguilera, V and Moncada, ME, 2014.

The impact of mother literacy and participation programmes on child learning: evidence from a randomised evaluation in India, 3ie Impact Evaluation Report 16. Banerji, R, Berry, J and Shortland, M, 2014.

A youth wage subsidy experiment for South Africa, 3ie Impact Evaluation Report 15. Levinsohn, J, Rankin, N, Roberts, G and Schöer, V, 2014.

Providing collateral and improving product market access for smallholder farmers: a randomised evaluation of inventory credit in Sierra Leone, 3ie Impact Evaluation Report 14. Casaburi, L, Glennerster, R, Suri, T and Kamara, S, 2014.

Scaling up male circumcision service provision: results from a randomised evaluation in Malawi, 3ie Impact Evaluation Report 13. Thornton, R, Chinkhumba, J, Godlonton, S and Pierotti, R, 2014.

Targeting the poor: evidence from a field experiment in Indonesia, 3ie Impact Evaluation Report 12. Atlas, V, Banerjee, A, Hanna, R, Olken, B, Wai-poi, M and Purnamasari, R, 2014.

An impact evaluation of information disclosure on elected representatives' performance: evidence from rural and urban India, 3ie Impact Evaluation Report 11. Banerjee, A, Duflo, E, Imbert, C, Pande, R, Walton, M and Mahapatra, B, 2014.

Truth-telling by third-party audits and the response of polluting firms: Experimental evidence from India, 3ie Impact Evaluation Report 10. Duflo, E, Greenstone, M, Pande, R and Ryan, N, 2013.

No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia, *3ie Impact Evaluation Report 9.* Ashraf, N, Bandiera, O and Jack, K, 2013.

Paying for performance in China's battle against anaemia, *3ie Impact Evaluation Report 8.* Zhang, L, Rozelle, S and Shi, Y, 2013.

Social and economic impacts of Tuungane: final report on the effects of a community-driven reconstruction programme in the Democratic Republic of Congo, *3ie Impact Evaluation Report 7*. Humphreys, M, Sanchez de la Sierra, R and van der Windt, P, 2013.

The impact of daycare on maternal labour supply and child development in Mexico, *3ie Impact Evaluation Report 6*. Angeles, G, Gadsden, P, Galiani, S, Gertler, P, Herrera, A, Kariger, P and Seira, E, 2014.

Impact evaluation of the non-contributory social pension programme 70 y más in Mexico, *3ie Impact Evaluation Report 5*. Rodríguez, A, Espinoza, B, Tamayo, K, Pereda, P, Góngora, V, Tagliaferro, G and Solís, M, 2014.

Does marginal cost pricing of electricity affect groundwater pumping behaviour of farmers? Evidence from India, 3ie Impact Evaluation Report 4. Meenakshi, JV, Banerji, A, Mukherji, A and Gupta, A, 2013.

The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone, *3ie Impact Evaluation Report 3*. Casey, K, Glennerster, R and Miguel, E, 2013.

A rapid assessment randomised-controlled trial of improved cookstoves in rural Ghana, *3ie Impact Evaluation Report 2.* Burwen, J and Levine, DI, 2012.

The promise of preschool in Africa: A randomised impact evaluation of early childhood development in rural Mozambique, *3ie Impact Evaluation Report 1.* Martinez, S, Naudeau, S and Pereira, V, 2012.

Acute malnutrition is considered among the most serious but least addressed health problems in Sudan. Guevarra and colleagues assessed the effectiveness of a food-based programme to prevent acute malnutrition in addition to a targeted supplementary feeding programme. The study findings showed that adding a prevention component to treatment programmes had no effect on prevalence or incidence on acute malnutrition in children under the age of five or among pregnant or lactating women. However, there was a significant reduction in the prevalence of children at risk of malnutrition where a food-based prevention intervention was added to the targeted supplementary feeding programme. The social and behaviour change communication intervention had no impact on feeding behaviours and practices.

Impact Evaluation Series

International Initiative for Impact Evaluation
202-203, Rectangle One
D-4, Saket District Centre
New Delhi – 110017
India

3ie@3ieimpact.org
Tel: +91 11 4989 4444
www.3ieimpact.org