ALIVE & THRIVE ETHIOPIA (A&T-ETHIOPIA)
MATERNAL NUTRITION IMPLEMENTATION RESEARCH (MNIR)

A Feasibility Study of Integrating a Package of Maternal Nutrition Interventions into Antenatal Care Services in Ethiopia: A Cluster-Randomized Evaluation

Data Analysis Plan

Prepared by:

Sunny S. Kim, Celeste Sununtnasuk, Rock Zagré, and Purnima Menon

International Food Policy Research Institute

Final version date: August 11, 2021

Trial registration number: ClinicalTrials.gov NCT04125368

Study protocol: Impact evaluation study protocol, version 2, dated June 2, 2021

Acknowledgements: This data analysis plan was prepared before commencing any analysis with the impact evaluation endline survey data. All final decisions on the statistical analyses remain with the evaluation team.

I. Study Background and Approach

In Ethiopia, Alive & Thrive (A&T) has developed and integrated a package of maternal nutrition interventions as part of the antenatal care (ANC) services provided by the government health system. Interventions are implemented in seven woredas (districts) in two regions – SNNP (4 woredas) and Somali (3 woredas). Key interventions include (1) maternal nutrition counseling (on diet quality and quantity), (2) iron and folic acid (IFA) supplementation (adequate supply and counseling), (3) weight gain monitoring (measurement and counseling), and (4) counseling on early initiation and exclusive breastfeeding practices. The interventions align with the 2016 WHO ANC guidelines (World Health Organization, 2016) and aims of the National Nutrition Program (NNP II).

1.1 Research questions

The implementation research study addresses three research questions:

Research question 1 (RQ1)	What are the program **impacts on maternal practices**: (1) consumption of diversified foods during pregnancy; (2) consumption of IFA supplements during pregnancy; and (3) early breastfeeding practices?
Research question 2 (RQ2)	Can the **coverage and utilization** of key maternal nutrition interventions [mentioned above] during ANC be improved through system strengthening approaches?
Research question 3 (RQ3)	What **factors influenced** integration and strengthening of maternal nutrition interventions into the government ANC service delivery platform?

1.2 Impact evaluation study design

The impact evaluation of A&T's interventions used a cluster-randomized design with repeated cross-sectional surveys at baseline and endline. We applied stratified random allocation to 30 health centers within seven woredas, which were assigned to either the A&T intervention (15 health centers) or control areas (15 health centers). A small baseline survey was conducted in October-November 2019 and the endline survey was conducted in August-September 2021 in the same 30 health center catchment areas, thereby creating panel data at the health facility level (not at individual level). Program implementation faced major disruptions and was paused between April and September 2020 due to the COVID-19 pandemic and related state-level restrictions. Refresher training and support for implementation of interventions were reengaged starting October 2020, for full implementation duration of approximately 10 months.

1.3 Study sample

The two main study sample groups are: 1) pregnant women (PW), as this sample allows the assessment of dietary diversity; and 2) recently delivered women (RDW) who have children less than 6 months of age, as this sample provides the best opportunity to assess the outcomes related to intervention exposure throughout pregnancy. PW and RDW were sampled separately but within the same health center catchment areas. For PW, we estimated a total sample size of 540 women (270 per arm) will detect a difference of 0.6 food groups in the mean dietary diversity score. For RDW, we estimated a total sample of 1890 women (945 per arm) to detect a difference of 13 tablets in the mean IFA tablets consumed after intervention.

Additionally, we included husbands of RDW present at the time of the survey. Outside of the two main sample groups, nurses-midwives (1 per health center) and health extension workers (HEW, 1 per health

post with up to 3 health posts[1]) were interviewed. At endline, direct observations of ANC visits (2 per health center and 1 per health post) to assess service quality, followed by exit interviews to assess service recall and satisfaction, were conducted among pregnant women attending ANC at the time of the survey.

Table 1: Expected sample sizes

		Baseline 2019		Endline 2021	
Survey respondent type		**Intervention**	**Control**	**Intervention**	**Control**
Household survey:					
1	Pregnant women	90	90	270	270
2	Recently delivered women (RDW) with children <6 months	180	180	945	945
3	Husbands of RDW	180	180	945	945
Service provider survey:					
4	Nurses-midwives (at health centers, HC)	15	15	15	15
5	Health extension workers (at health posts, HP)	45	45	45	45
Observations:					
6	ANC observation + exit interview (2 per HC, 1 per HP)	…	…	75	75
	Total:	510	510	2,295	2,295

II. Outcome Measures and Indicators

Outcome measures corresponding to the three research questions are presented below. Only some outcome measures under RQ 1 pertain to the primary outcomes of the evaluation (i.e., used to test study hypotheses and arrive at a decision on overall study impact and to serve as basis to calculate the sample size); RQs 2 and 3 focus on secondary outcomes.

2.1. Research question 1 (impact on maternal nutrition practices)

For impact estimates, outcome measures related to maternal dietary diversity will be used from the PW dataset, and outcomes to IFA consumption and early breastfeeding practices will be used from the RDW data.

Table 2: Outcome measures for RQ1

Outcome	Indicator	Data source
Maternal dietary diversity	*Primary outcome:* - Dietary diversity score (# of food groups) *Secondary outcomes:* - % PW consumed at least 5 food groups (minimum dietary diversity)	PW survey
IFA consumption	*Primary outcome:* - # of IFA tablets consumed *Secondary outcomes:* - % RDW consumed 90+ IFA tablets - # of IFA tablets received	RDW survey

[1] Primary health care is provided at primary hospitals, local health centers and rural health posts. One or more health centers/primary hospitals exist at the *woreda* (district) level. Within each health center catchment area, there are multiple health posts, usually one per *kebele* (ward/subdistrict) level.

Early breastfeeding practices	*Secondary outcomes:* - % infants <6 months breastfed within 1h of birth - % infants <6 months with no pre-lacteals fed - % infants <6 months exclusively breastfed	RDW survey

2.2 Research question 2 (coverage and utilization)

For effects on coverage and utilization of interventions during ANC visits, outcome measures will be used from the RDW survey data. In the context of the overall evaluation, outcomes under this research question are considered as secondary outcomes.

Table 3: Outcome measures for RQ2

Outcome	Indicator	Data source
ANC visits and contacts	- # of ANC visits (at health facility) - % RDW with at least 4 ANC visits - % RDW received ANC visit in first trimester of pregnancy - # of contacts outside of health facility (home visits and women's group meetings)	RDW survey
Counseling on dietary diversity and adequate intake	During ANC visits and other ANC contacts: - % RDW received counseling on maternal nutrition - % RDW received counseling on dietary diversity - % RDW received counseling on consuming adequate quantity of food	RDW survey
Counseling on IFA supplementation	During ANC visits and other ANC contacts: - % RDW received counseling on importance of IFA - % RDW received counseling on how/reminders to take IFA - % RDW received counseling on managing IFA side effects	RDW survey
Weight gain monitoring and counseling	During ANC visits and other ANC contacts: - # times weighed - % RDW weighted at least 4+ times/at each ANC visit - % RDW received counseling about weight gain during pregnancy	RDW survey
Counseling on early breastfeeding practices	During ANC visits and other ANC contacts: - % RDW received counseling on breastfeeding practices - % RDW received counseling on early initiation of breastfeeding - % RDW received counseling on not feeding pre-lacteals - % RDW received counseling on exclusive breastfeeding	RDW survey

2.3 Research question 3 (health system factors)

For assessing factors related to strengthening service delivery, measures will be used from the health facility checklist (health centers and posts) and N-M and HEW datasets. In the context of the overall evaluation, outcomes under this RQ3 count as secondary outcomes.

Table 4: Outcome measures for RQ3

Outcome	Indicator	Data source
Equipment and materials to support maternal nutrition services	- % HC/HP with maternal nutrition counseling job aids - % HC/HP with IFA supplementation job aid - % HC/HP with breastfeeding counseling job aids - % HC/HP with functional weighing scale - % HC/HP with currently stocked with IFA tablets - % HC/HP with register to monitor IFA stocks - % HC/HP reporting stock-out of IFA in past 6 months	Health facility checklist

Service providers' training and supportive supervision	- % NM/HEW received maternal nutrition training - % NM/HEW by training content - % NM/HEW received ANC supervision - % NM/HEW by supervision content	N-M survey HEW survey
Service providers' knowledge	- Knowledge scores for dietary diversity, adequate intake, IFA, and weight gain monitoring - Knowledge scores for breastfeeding	N-M survey HEW survey
Service providers' work tasks and workload perceptions	- % NM/HEW record-keeping on ANC services - % NM/HEW by content of record-keeping - % NM/HEW with increased workload in past 1y due to ANC services	N-M survey HEW survey
Service providers' provision of services	- % NM/HEW provided maternal nutrition interventions - % NM/HEW by counseling messages provided (on dietary diversity, IFA, weight gain monitoring, and breastfeeding) - % NM/HEW used job aids for maternal nutrition counseling - % HEW provided home visits to PW/number of visits in last 3 months - % HEW conducted PW conference/number of meetings in last 3 months	N-M survey HEW survey

III. Statistical Analysis Plan

3.1 General principles and methods

Data analyses will be performed using STATA version 16.0 (StataCorp LLC). All applicable statistical tests will be two-sided to allow potential findings of unexpected effects. Statistical significance will be presented at levels of $p<0.05$, $p<0.01$, and $p<0.001$.

A diagram presenting the flow of clusters and individuals through the trial, based on the Consolidation Standard of Reporting Trials (CONSORT) statement: extension to cluster randomized trials (Campbell et al., 2012; Eldridge et al., 2016), is shown as follows.

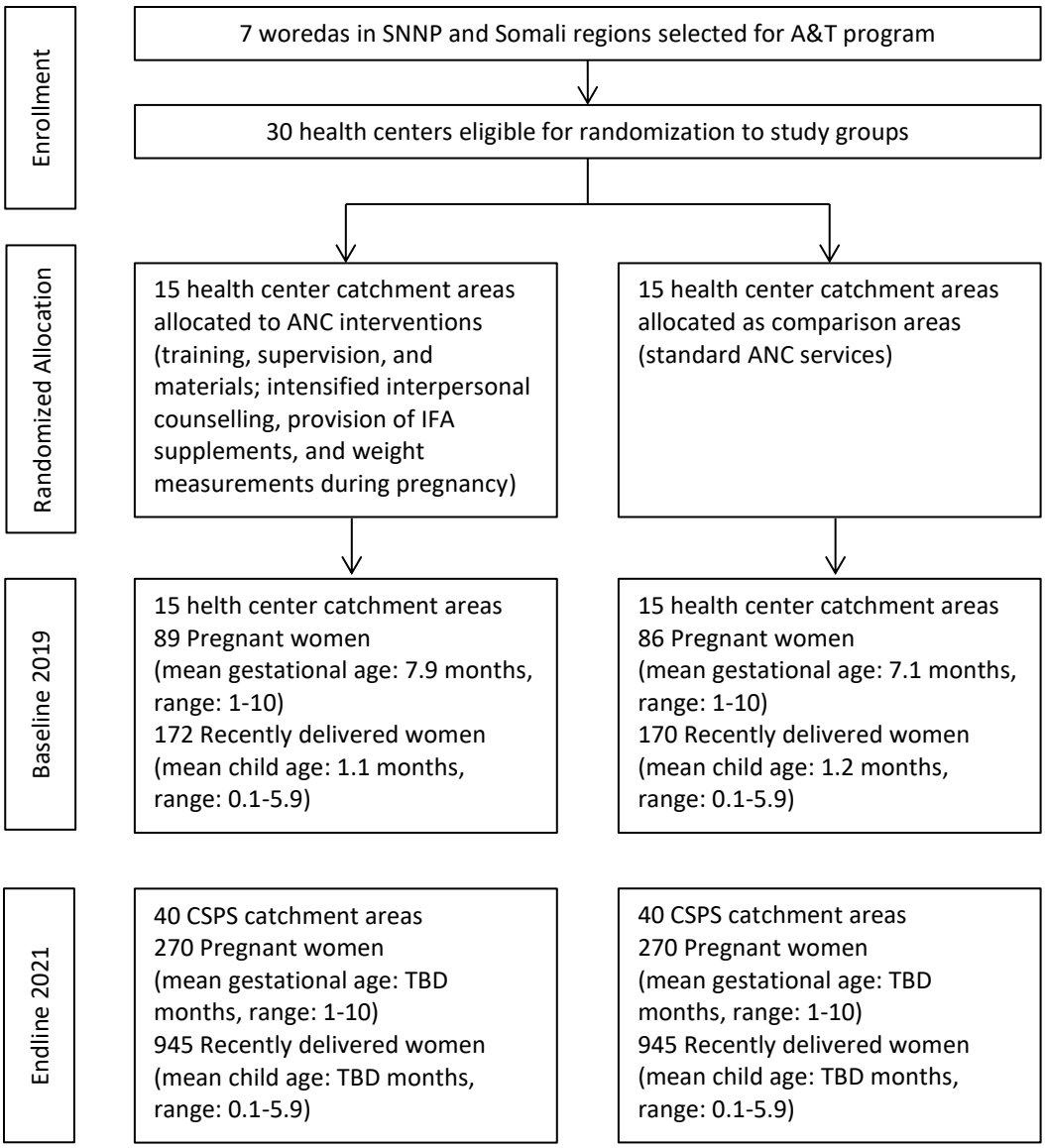

Figure 1. CONSORT flow diagram for repeated cross-sectional surveys

3.2 Sample characteristics

Baseline and endline characteristics will be reported between randomized program groups (A&T and control). For household samples, indicators of maternal characteristics (age, marital status, education, occupation, and religion), obstetric history (age of marriage, age at first birth, gravida, parity, number of living children, and trimester of pregnancy,), household composition (size, number of adults and children, and household head) and other household characteristics (household food security, livelihood, and socioeconomic status) will be reported. Binary variables will be summarized as proportions, and continuous variables will be summarized as mean values with standard deviations (when normally distributed) or as median with interquartile range (for non-normal distribution variables). T-test will be used to compare and infer significant difference between the program groups by survey round.

Table 5: Dummy table for sample characteristics

Indicator	Baseline	Endline

	A&T (N=)	Control (N=)	A&T (N=)	Control (N=)
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**
Age of respondent (years)				
	Percent	**Percent**	**Percent**	**Percent**
Marital status Education level Occupation Religion				

*p<0.05, **p<0.01, ***p<0.001

3.3 Impact estimates

The main analysis of impacts will be performed using intent-to-treat (ITT) specifications, wherein all study participants in the originally assigned program group at baseline are included in the statistical analysis and analyzed according to their program group, regardless of whether they received interventions or not. Women who refused or withdrew consent or those who are ineligible according to study protocol are excluded from ITT analysis.

The main impacts of the interventions will be estimated for: (1) dietary diversity score, (2) consumption of IFA supplements, and (3) early breastfeeding practices; secondarily, impact will be estimated for exposure to key interventions: (4) maternal nutrition counseling, (5) counseling on IFA supplementation, (6) weight gain monitoring and counseling, (7) breastfeeding counseling. The impact on dietary diversity will assessed among PW, and impact on all the remaining outcomes will be assessed among RDW.

Given that the man impact indicators with full sample sizes are collected at endline only, ITT linear regression models will be used to test the means of the outcome for estimates of group differences (intervention vs. control) at endline, with standard errors clustered at the health center level. In the ITT adjusted models, we will control for gestational age (for PW only), maternal characteristics, child age and sex (for RDW only), and other variables that may be different between study arms. Robustness tests will be conducted using difference-in-difference analysis where outcome variables exist at baseline.

Table 6: Dummy table for impact estimates

Indicator	Baseline		Endline			
	A&T (N=)	Control (N=)	A&T (N=)	Control (N=)	Unadjusted impact est.[1]	Adjusted impact est.[2]
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)		
Dietary diversity score (number of food groups) Number of IFA tablets consumed						
	Percent	**Percent**	**Percent**	**Percent**		
Minimum dietary diversity (≥5 food groups) Consumed 90+ IFA tablets						

[1] Controlling for clusters by health center

[2] Adjusted for gestational age (PW), maternal age and education, household food security and wealth, and clustering by health center
*p<0.05, **p<0.01, ***p<0.001

3.4 Plausibility analysis

In addition to the estimation of impacts, we will conduct plausibility analyses by two methods, to provide further evidence for the likelihood or strength of our impact estimates. First, we will assess whether social desirability bias may have influenced reported outcomes. Second, we will examine the intermediate outcome indicators along the program impact pathways (from service delivery to exposure and behavioral determinants) to determine whether the program resulted to the outcomes as intended by design.

3.4.1 Testing for social desirability bias

For outcome measures based on individual report, social desirability bias may play a potential role in influencing response. We applied a 13-item social desirability index, adapted from Reynolds and Gerbasi (Reynolds, 1982), to determine the extent to which respondents were likely to report behaviors based on their desire to please others, present oneself to others in a favorable way, or for social approval, i.e., "social desirability":

No.	Question item	
1	Is it sometimes hard for you to go on with your work if you are not encouraged?	NOO=1
2	Do you sometimes feel resentful when you don't get your way?	NO=1
3	Do you occasionally give up doing something because you don't think you have the ability?	NO=1
4	Do you occasionally feel like not listening to people event though you know they were right?	NO=1
5	No matter who you're talking to, are you always a good listener?	YES=1
6	Have there been occasions when you took advantage of someone?	NO=1
7	Are you always willing to admit it when you make a mistake?	YES=1
8	Do you sometimes try to get even, rather than forgive and forget?	NO=1
9	Are you always courteous, even to people who are disagreeable?	YES=1
10	Have you ever been irritated when people expressed ideas very different from your own?	NO=1
11	Have there been times when you were jealous of the good fortune of others?	NO=1
12	Are you sometimes irritated by people who ask favors of you?	NO=1
13	Have you ever deliberately said something that hurt someone's feelings?	NO=1
	Total score	13

The social desirability score (SDS) will be created by adding up the number of socially desirable answers, out of the total 13 question items. We will conduct three analyses using this score: (1) estimation of mean SDS by program group to compare differences in the level of social desirability bias between the intervention and control groups; (2) tabulation of key outcomes by SDS to assess whether the reported outcomes varied by SDS levels; and (3) regressions with each of the outcomes as dependent variables to test the interaction between SDS and intervention group, to determine whether or not social desirability bias differentially affected the impact of the A&T interventions on key outcomes.

3.4.2. Analysis of program impact pathways

The program impact pathway (PIP) was developed in collaboration with the A&T program team to map out the mechanisms through which the interventions were expected to achieve impact. The purpose of the PIP analysis is to lay out the theoretical causal links between program activities, outcomes, and impacts. We will examine key indicators along the components of pathways (addressed in part by RQ2

and RQ3), to interpret and support the impact evaluation results. We will compare differences between program groups for indicators along the pathway matched to the relevant outcomes (dietary diversity, IFA consumption, and breastfeeding practices), using linear regression models, accounting for geographic clustering.

Table 7: Measures for program impact pathways

Outcome	Key Indicators	Data source
Service providers' capacity-building and service provision:		
Training and supervision (**RQ3**)	- % NM/HEW received maternal nutrition training - % NM/HEW by training content - % NM/HEW received ANC supervision - % NM/HEW by supervision content	N-M survey HEW survey
Service providers' knowledge (**RQ3**)	- Knowledge scores for dietary diversity, adequate intake, IFA, and weight gain monitoring - Knowledge scores for breastfeeding	N-M survey HEW survey
Service provision (**RQ3**)	- % NM/HEW provided maternal nutrition interventions - % NM/HEW by counseling messages provided (on dietary diversity, IFA, weight gain monitoring, and breastfeeding) - % NM/HEW used job aids for maternal nutrition counseling - % HEW provided home visits to PW/number of visits in last 3 months - % HEW conducted PW conference/number of meetings in last 3 months	N-M survey HEW survey
Beneficiaries' exposure and behavioral determinants:		
ANC visits and contacts (**RQ2**)	- # of ANC visits (at health facility) - % RDW with at least 4 ANC visits - % RDW received ANC visit in first trimester of pregnancy - # of contacts outside of health facility (home visits and GASPAs)	PW survey RDW survey
Counseling on dietary diversity and adequate intake (**RQ2**)	During ANC visits and other ANC contacts: - % RDW received counseling on maternal nutrition - % RDW received counseling on dietary diversity - % RDW received counseling on consuming adequate quantity of food	PW survey RDW survey
Counseling on IFA supplementation (**RQ2**)	During ANC visits and other ANC contacts: - % RDW received counseling on importance of IFA - % RDW received counseling on how/reminders to take IFA - % RDW received counseling on managing IFA side effects	PW survey RDW survey
Weight gain monitoring and counseling (**RQ2**)	During ANC visits and other ANC contacts: - # times weighed - % RDW weighted at least 4+ times/ at each ANC visit - % RDW received counseling about weight gain during pregnancy	PW survey RDW survey
Counseling on early breastfeeding practices (**RQ2**)	During ANC visits and other ANC contacts: - % RDW received counseling on breastfeeding practices - % RDW received counseling on early initiation of breastfeeding - % RDW received counseling on not feeding pre-lacteals - % RDW received counseling on exclusive breastfeeding	PW survey RDW survey
Beneficiaries' knowledge and perceptions	- Knowledge scores for dietary diversity, adequate intake, IFA, and weight gain monitoring - Knowledge scores for breastfeeding - Beliefs, self-efficacy, and social norms score	PW survey RDW survey

Table 8: Dummy table for program impact pathways analysis

Indicator	Baseline		Endline	
	A&T (N=)	Control	A&T	Control
	Mean (SD)			
(see indicators in Table 7)				
	Percent			

*p<0.05, **p<0.01, ***p<0.001

IV. References

Campbell, M. K., Piaggio, G., Elbourne, D. R., & Altman, D. G. (2012). Consort 2010 statement: Extension to cluster randomised trials. *BMJ (Online)*, *345*(7881), 1–21. https://doi.org/10.1136/bmj.e5661

Eldridge, S. M., Chan, C. L., Campbell, M. J., Bond, C. M., Hopewell, S., Thabane, L., Lancaster, G. A., Altman, D., Bretz, F., Campbell, M., Cobo, E., Craig, P., Davidson, P., Groves, T., Gumedze, F., Hewison, J., Hirst, A., Hoddinott, P., Lamb, S. E., … Tugwell, P. (2016). CONSORT 2010 statement: Extension to randomised pilot and feasibility trials. *The BMJ*, *355*. https://doi.org/10.1136/bmj.i5239

Hayes, R. J., & Moulton, L. H. (2017). Cluster randomised trials, second edition. In *Cluster Randomised Trials, Second Edition*. https://doi.org/10.4324/9781315370286

Reynolds, W. M. (1982). Development of Reliable and Valid Short Forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, *38*(1), 119–125. https://doi.org/10.1002/1097-4679(198201)38:1<119::AID-JCLP2270380118>3.0.CO;2-I

World Health Organization. (2016). *WHO recommendations on antenatal care for a positive pregnancy experience*.