USAID Integrated Health Program Evaluation Report

Year 2 Impact Evaluation Results

March 2022

USAID
FROM THE AMERICAN PEOPLE

DATA FOR impact

USAID Integrated Health Program Evaluation Report

Year 2 Impact Evaluation Results

David Hotchkiss, PhD
Matt Worges, PhD
Janna Wisniewski, PhD
Paul-Samson Lusamba-Dikassa, MD, PhD
Lauren Blum, PhD
Eva Silvestre, PhD
Gael Compta, MS
Francine Wood, MPH

Data for Impact
University of North Carolina at Chapel Hill
123 West Franklin Street, Suite 330
Chapel Hill, NC 27516 USA

D4I@unc.edu
http://www.data4impactproject.org

This publication was produced with the support of the United States Agency for International Development (USAID) under the terms of the Data for Impact (D4I) associate award 7200AA18LA00008, which is implemented by the Carolina Population Center at the University of North Carolina at Chapel Hill, in partnership with Palladium International, LLC; ICF Macro, Inc.; John Snow, Inc.; and Tulane University. The views expressed in this publication do not necessarily reflect the views of USAID or the United States government. TR-21-453 D4I

D4I is committed to local partner engagement and individual and institutional strengthening. Local authorship is important, and we urge you to engage local partners in analysis and reporting.

Abstract

This report presents results from an impact evaluation conducted by Data for Impact (D4I) following two years of the United States Agency for International Development (USAID) Integrated Health Program's (IHP) implementation in nine provinces of the Democratic Republic of the Congo (DRC). The IHP focuses, in part, on increasing utilization of health facility-based maternal and child healthcare and family planning services. The impact evaluation investigates the extent to which changes in healthy behaviors and health outcomes are attributable to the USAID IHP. The evaluation used a quasi-experimental design based on a propensity score matched difference-in-differences model fit to data collected through the DRC's routine health information system (RHIS). The matching process was successful in balancing the distribution of all selected indicators across comparison and intervention sites. Restricting the propensity scores to a region of common support only led to the exclusion of 1.6 percent (57 of 3,667) of intervention facilities and 2.0 percent (113 of 5,549) of comparison facilities. The common trends assumption suggests no significant differences between the comparison and intervention groups in the preintervention time series trends for all but one of the RHIS indicators (i.e., new acceptors of modern contraceptive methods). Not satisfying the common trends assumption serves to reduce the stringency of the analysis and undermines the ability to appropriately interpret the results for implicated indicators.

Minimal and nonsignificant changes were observed in the DID estimates for all twelve indicators. The DID estimator coefficients for eight of the 12 RHIS indicators all suggest changes in the anticipated direction. One of the largest changes in rates was noted for moderate malnutrition, for which 7.06 additional cases per 1,000 children under five years were observed in the intervention health zones relative to comparison health zones although this finding was not significant and in an unanticipated direction. Although the unadjusted and Westfall-Young corrected p-values from the regressions for moderate malnutrition suggest a significant difference in the change in rate between intervention and comparison sites, the wild cluster bootstrap adjusted p-value should be used for significance interpretation. The largest although insignificant change in rate was noted for exclusive breastfeeding, with 33.07 additional cases per 1,000 children under six months.

As this impact analysis is expected to occur on an annual basis, year-over-year results need to be interpreted within the proper context. In the interim, the USAID IHP may draw its attention to those indicators that show little to no movement in the anticipated direction, such as treatment of severe malaria, diarrhea, and pneumonia in children, and consider whether any course adjustments are warranted. Findings from this portion of the evaluation will be triangulated with those from the health system surveys and the qualitative data collection, and D4I will make final recommendations in the midline evaluation report.

Acknowledgments

The authors wish to thank the United States Agency for International Development (USAID) Mission based in the Democratic Republic of the Congo as well as the Congolese Ministry of Health for supporting this work.

We acknowledge our colleagues at Tulane University (Dr. Josh Yukich and Dr. Charles Stoecker) and the University of North Carolina at Chapel Hill (Dr. Gustavo Angeles, Dr. Kristen Brugh, and Dr. Sean Sylvia) who provided technical guidance for the impact evaluation analysis.

Last, we thank the knowledge management team of the Data for Impact (D4I) project for editorial, design, and production services.

Suggested citation

Data for Impact. (2022). USAID Integrated Health Program Midline Evaluation: Year 2 Impact Evaluation Results. Chapel Hill, NC, USA: Data for Impact, University of North Carolina.

Contents

Figures

Tables

Abbreviations

ANC	Antenatal care
ASMD	Absolute standardized mean difference
CI	Confidence interval
D4I	Data for Impact
DID	Difference-in-differences
DRC	Democratic Republic of the Congo
EVD	Ebola virus disease
HZ	Health zone
IHP	Integrated Health Program
IRBs	Institutional Review Boards
NDVI	Normalized difference vegetation index
PSM	Propensity score matching
RHIS	Routine Health Information System
USAID	United States Agency for International Development
WCB	Wild cluster bootstrap
WRA	Women of reproductive age

Program Background

As part of its strategy to improve health outcomes in the Democratic Republic of the Congo (DRC), the United States Agency for International Development (USAID) funded the Integrated Health Program (IHP) in 2018. The program began operations in July 2018 and is being implemented by Abt Associates and several partner organizations. The purpose of USAID IHP is to strengthen the capacity of Congolese institutions and communities to deliver quality, integrated health services to sustainably improve the health status of the Congolese population. The project focuses on the following specific health, population, and nutrition areas: maternal health; neonatal, infant, and child health; tuberculosis; malaria; child nutrition; water, sanitation, and hygiene; and family planning.

USAID IHP seeks to reach its goal through achievement of the following overall performance objectives:

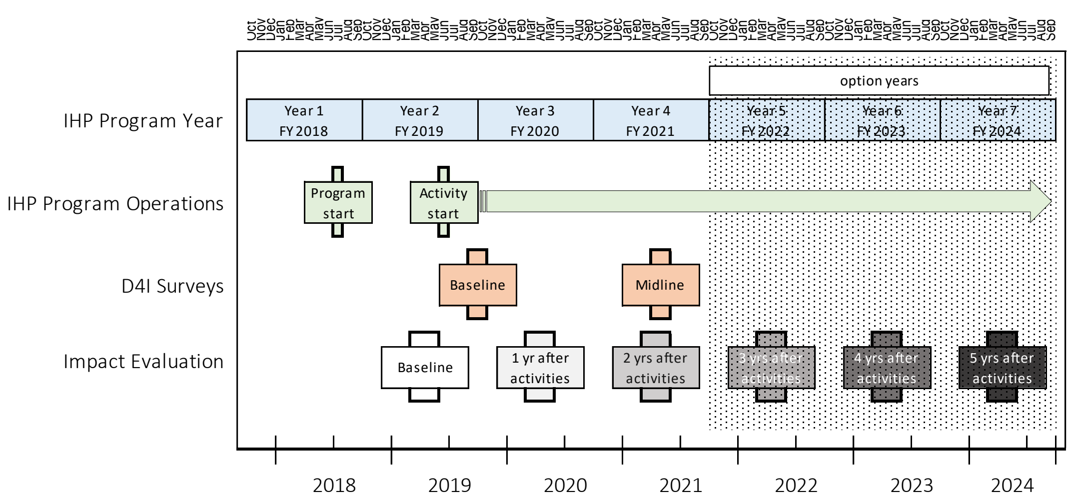
- Strengthen health systems, governance, and leadership at the provincial, health zone, and facility levels in target health zones (Objective 1)
- Increase access to quality, integrated health services in target health zones (Objective 2)
- Increase adoption of healthy behaviors, including use of health services, in target health zones (Objective 3)

USAID IHP works in nine contextually diverse provinces in the regions of Eastern Congo, Katanga, and Kasai and will include a wide array of interventions.

Given the breadth and depth of IHP's interventions, the USAID/DRC Mission requested that Data for Impact (D4I) conduct an independent third-party evaluation of the performance and impact of USAID IHP on key health systems-related outcomes: the uptake of FP and health care services; health systems functioning (i.e., improved disease surveillance, the availability of essential commodities, and health worker motivation); and the practice of key healthy behaviors.

Figure 1 below shows the timing of program rollout and D4I evaluation time points. The nearly one-year gap between program start and activity implementation allowed the IHP to staff up across the nine provinces and begin program implementation after administrative delays with the donor and ministry of health had been resolved.

Figure 1. Timeline of IHP rollout and D4I survey and evaluation time points



Methods and Limitations

Methods

Data for Impact (D4I) is carrying out two types of evaluation components for this study: a performance evaluation and an impact evaluation. Performance evaluations incorporate before and after comparisons, but generally lack a rigorously defined counterfactual to control for factors other than the project or intervention that might account for the observed change. Impact evaluations assess the extent to which changes in health outcomes or service use over time are attributable to an intervention. The specific research questions that will be addressed in the evaluation are the following:

1. Did the expected changes in outcomes and impacts occur?
    a. Strengthen health systems, governance, and leadership at provincial, health zone (HZ), and facility levels in target HZs.
    b. Increase access to quality, integrated health services in target HZs.
    c. Increase adoption of healthy behaviors, including health service use, in target HZs.
2. If there were changes in healthy behaviors over the course of the study period, to what extent were these attributable to USAID IHP?
3. Did the project contribute to gender equity in health services and within the health system?
4. What factors enabled or limited the success of USAID IHP?

The performance evaluation aspect of the study addresses Research Questions 1, 3, and 4. Data for this component of the study are collected from multiple sources, including: the DRC's routine health information system (RHIS); household surveys; surveys of healthcare facilities, health zone offices, and provincial health offices; and key informant and in-depth interviews, observations of patient-health worker interactions, and focus group discussions. The impact evaluation aspect of the study—the focus of this report—addresses Research Question 2. The impact evaluation used a quasi-experimental design based on a propensity score matched difference-in-differences model fit to RHIS data. Ethical approval for this work was given by the Institutional Review Boards (IRBs) of Tulane University and the Kinshasa School of Public Health.

Analysis of Impact Using a Difference-in-Differences Model Two Years After IHP Program Implementation

The PSM-DID method is a quasi-experimental approach that attempts to mimic an experimental research design. The PSM-DID method may facilitate causal inference even when randomization is not possible. The approach compares changes in outcomes between populations located in areas undergoing an intervention (the intervention group) and similar populations located in areas without the intervention (the comparison group) using time points before and after the start of the intervention. Simply put, the DID analysis first calculates the before-after difference of an outcome in the intervention group and then calculates the before-after difference of the same outcome for the same period in the comparison group. Next, the difference noted in the comparison group is subtracted from the difference noted in the intervention group (i.e., difference-in-differences), which provides an impact estimation of the IHP intervention.

A doubly robust model that combines PSM with a DID model was used to estimate the impact[1] of the USAID IHP on the provision of maternal and child healthcare and family planning services (as identified in Research Question 2). The data used for this analysis came from the in-country RHIS which collects health facility data monthly. The unit of analysis is the facility and as the program was implemented at the province level, a random effect was included to account for this in regression modeling. Additionally, due to the low number of provinces included in the analysis (24 total; 9 intervention and 15 comparison) and high variation in the number of health facilities per province (a low of 204 and a high of 1,298), wild cluster bootstrapping was used. Given the multiple outcomes we tested, all of which are hypothesized to be an effect of IHP activity implementation, we adjusted DID regression p-values for the number of hypothesis tests performed using the Westfall-Young multiple hypothesis testing procedure which allowed for the inclusion of propensity score weights.

The intervention arm includes facilities from within health zones in USAID IHP provinces (including both hospitals and health centers) in the preintervention (June 2018–May 2019) and postintervention (June 2019–September 2023) periods. All facilities from within IHP-targeted provinces are exposed to IHP activities. The comparison arm includes facilities within comparable and non-excluded health zones from provinces not receiving USAID IHP support. Excluded health zones were those that were a part of a previous and intensive health systems strengthening project (the Access to Primary Health Care program going by the French acronym *ASSP*) which was active in Kasaï, Nord Ubangi, and Sud Ubangi provinces. Additionally, health zones that experienced Ebola virus disease (EVD) outbreaks (N=33) were also removed from consideration as certain policies were enacted in these affected health zones that were designed to increase health service utilization. Only one of the 33 health zones affected by EVD is in an IHP target province. Nearly 85% of the EVD-affected health zones were in Ituri and North Kivu provinces. We conducted an analysis of a free care policy on service volumes in North Kivu province which showed that the enacted policy dramatically increased total clinic visits for an extended period driven in large part by malaria and pneumonia cases, which encompass two of our outcomes of interest. A previous analysis showed similar findings following the enactment of a free care policy across EVD-affected health zones in Equateur province.

Additional data cleaning included removal of anomalous data points from each individual health facility time series if these values exceeded ±4.5 standard deviations from the median facility value. Missing data in health facility time series were managed using a flexible interpolation process that took seasonality into account where it was detected; otherwise, simple linear interpolation was conducted to fill in gaps. Specifically, the ‘na.interp’ function from the forecast package (Hyndman et al, 2008; Hyndman et al, 2022) of R (R Core Team, 2021) was used to interpolate health facility time series. Note that each successive analysis takes advantage of a longer time series than predecessor analyses. This scenario may help to establish a more robust seasonal component to individual health facility time series trends, which is exploited during the interpolation process. Health facility time series with gaps of seven or more

[1] In health evaluation research literature, the term “impact” typically refers to the effects on health outcomes, such as lives saved or disability-adjusted life years averted. In health systems strengthening (HSS) evaluation literature, “impact” is also often used to refer to the effects on service delivery or other aspects of health systems functioning (Adams, et al., 2012). For the purposes of this evaluation, we use “impact” in the latter sense unless otherwise noted.

missing values in a row were dropped from analysis. See **Table 1** for a summary of the data cleaning process which details the extent of missing data and anomalous data points.

Following the data cleaning process, rates per 1,000 population were calculated for each data element of interest. We pulled population statistics for health areas and health zones directly from the RHIS. As these were overall population counts, we calculated sub-populations (women of reproductive age [15–49 years], children 6–59 months, and children <6 months) from population pyramids estimated by the Population Division of the Department of Economic and Social Affairs within the United Nations. Hospital-based rates were calculated using their respective health zone as the catchment population. All other health facility-based rates were calculated using their respective health area as the catchment population.

Baseline measures included in the DID analyses reflect three-month averages for March 2019 through May 2019, which were compared to averages spanning the same three months in 2021. These three months were specifically chosen as the immediately precede the onset of IHP activity implementation (June 2019). Due to the variable nature of data availability within the RHIS and taking into consideration the interpolation procedure used to establish fuller health facility-level time series, a three-month average was taken.

Comparison facilities were identified through PSM, coupled with the use of a gradient boosted model. Propensity score methods are used to adjust for observed confounders to produce more valid causal effect estimates. This covariate balancing is the degree to which the distribution of covariates is similar across intervention assignment. Covariate balancing requires proper model specification to avoid biased estimates. Researchers may swap covariates in and out of their logistic models or manually add polynomial and/or interaction terms to these models to balance covariates, but this process can be tedious and inefficient. Use of gradient boosted models involves a machine learning process which captures the flexible and nonlinear relationships between intervention assignment and the preintervention covariates in an automated fashion, do not need to exclude collinear or insignificant covariates, and can automatically assign polynomial and interaction terms without overfitting the data. McCaffrey et al (2004) provide supplemental materials with annotated R code for estimating propensity scores with boosted regression which we used to build our gradient boosted model.

The objective of the PSM process was to match a pool of health facilities from the comparison provinces to those in the intervention provinces such that the distributions of selected RHIS indicators were similar between the two groups. This process helps to ensure similar baseline characteristics between these two pools of facilities. The 13 selected RHIS indicators used in the PSM process were the following:

1) Total clinic visits
2) New cases
3) Other new cases
4) Suspected malaria
5) Simple diarrhea
6) New pregnancies
7) Sulfadoxine/pyrimethamine dose 1

8) Antenatal care clinic visits 1, 2, 3, and 4
9) Births
10) Live births

An underlying assumption of the DID analysis is the common trends assumption. This was initially explored for each outcome indicator of interest by plotting the overall average case incidence per 1,000 target population for the preintervention time series for both the comparison and intervention areas. Satisfying this assumption infers that the comparison units provide the appropriate counterfactual trend that the treated units would have followed if they had not been exposed to the intervention (i.e., in the absence of the IHP package of interventions, the two groups of health facilities would have had similar trends). Additionally, a formal statistical test was used to assess equality of trends using a 12-month preintervention period.

Before the DID analyses were run, it was necessary to compile covariates that would be useful in the DID linear regression models. Remote sensing data from the National Aeronautics and Space Administration and the National Oceanic and Atmospheric Administration were leveraged to create three-month average measures of the normalized difference vegetation index—a measure of greenness in the environment—and monthly rainfall. The vegetation index and rainfall measures were pulled for the three-month periods in 2019 and 2021 that corresponded to the timeframe for which average estimates were taken for the outcome indicators identified for use in the DID analyses. Health zone-level estimates were obtained for 2015 prevalence of improved housing and for 2014 educational attainment of women of reproductive age, the last time such published measures were estimated. Health zone values for educational attainment and prevalence of improved housing were held constant across all time points. A binary variable describing urbanicity (rural and urban) was also assigned to each health zone.

The DID linear regression model can be defined as follows:

$$\mathbf{Yit = \alpha + \beta Ti + \gamma At + \delta(Ti \times At) + COVSit + COVSi + \epsilon it}$$

where Yit is the outcome of interest for facility i at time t, Ti indexes health facilities in the intervention health zones, At distinguishes between pre and postintervention values, $COVSit$ represents time-varying covariates, $COVSi$ represents time-invariant covariates, and ϵit is a normal random variable with mean zero. The Greek letters are the parameters to be estimated. The null hypothesis $\delta=0$ is tested to determine whether the IHP intervention had an effect as δ represents the change in the intervention group from pre to postintervention relative to the comparison group. Inverse probability weights calculated from the propensity scores (ps/(1-ps)) for comparison facilities were also used in these models (note: intervention facilities were ascribed a weight of one).

2021 Nurses' Strike

The ongoing nurses' strike in DRC may unduly influence the results of the impact evaluation. While the time point used for Year 2 outcome values was before the nurses' strike started, an extended time series was used in the interpolation process to lend a more robust structure to the seasonal decomposition step. The time series used for the updated PSM-DID analysis extends through September 2021, when the RHIS reporting rate for three-quarters of all DRC provinces was still ≥90%. The most affected province, as far as reporting rate, appears to be Tanganyika where rates dropped to between 25–40 percent across August, September, and October 2021 (**Figure 7** of Appendix 1). The month of October 2021, excluded from the interpolation process, saw a drop in reporting rates below 90 percent for just over half of all provinces.

Limitations

There are several limitations of the evaluation and threats to carrying it out as planned.

First, the impact evaluation component of the study investigates only the impact of the USAID IHP on proxy indicators related to service provision, including treatment of childhood illnesses, contraceptive use, vaccinations, and antenatal care (ANC). Because data on health outcomes, service quality, and health systems governance and leadership are not available from non-project areas, impacts on these aspects cannot be rigorously assessed. However, to descriptively explore these aspects, a performance evaluation is being carried out using both quantitative and qualitative data collected in the nine USAID IHP provinces[2] to explore changes in proxy indicators for three USAID IHP objectives—health systems strengthening, quality, integrated health services, and healthy behaviors—and the factors that enabled or limited the success of the project.

Second, the impact evaluation of the study is based on routine data from the RHIS. Although it is expected that using a research design based on these data (i.e., numerous, repeated health facility observations over extended periods and the real-time indicators of service coverage) provides power and cost advantages over a research design based on intermittent population-based surveys, poor data quality remains a threat due to inaccurate data on counts of services provided (numerators) and the populations that are served (denominators). These disadvantages could lead to two consequences. First, poor data quality could add spurious variability to the dependent variable. If it is a random measurement error, it will add to the variation of the random error in the model, with the consequence of larger standard errors in the estimated coefficients increasing the chances of not finding significant effects when there is impact. This is a common problem in research studies based on data from health management information systems, and there is little that the evaluator can do to address the issue. Second, measurement error could potentially evolve over time as data quality improves concurrently with increasing rigorousness and completeness of data reporting. Because USAID IHP aims to improve RHIS data quality, these improvements could be different in intervention and comparison areas, which will create a type of endogeneity[3] in the program variable of the model that varies over time, so it is not controlled by the fixed effects. This heteroskedasticity[4] will be accounted for through the estimation of robust standard errors. Additionally, the RHIS indicators used as outcomes for this evaluation are not sex-disaggregated, making an assessment of gender-related differences impossible to perform. However, despite the lack of data on sex and gender, several indicators included in the analysis are focused on women's experiences as they relate to women's and children's wellbeing, such as new modern contraceptive acceptors, antenatal care clinic visits, birth rate, low birth weight births, and exclusive breastfeeding.

Third, when adjusting p-values for the number of hypothesis tests performed, the Westfall-Young procedure is unable to account for wild cluster bootstrapping and the corresponding p-value is relative to

[2] Qualitative data will be collected in three provinces and in Kinshasa.
[3] A model in which the explanatory variable is correlated with the error term.
[4] A case in which the standard errors of a variable are not constant over time.

the unadjusted p-value where wild clusters are not taken into consideration. The Romano-Wolf procedure for multiple hypothesis testing was not used as it does not allow for the inclusion of propensity score weights.

Fourth, the DRC is an unstable environment and there is a possibility that both the implementation of USAID IHP and D4I and USAID IHP survey activities could be affected by political and social unrest during the project period. This is not likely to affect the impact evaluation component of the study, which relies on RHIS data, but it could affect future surveys by precluding travel to sampled provinces, health zones, and facilities.

Results

Data Processing

Table 1 on the next page shows the effects of the data cleaning process. Just over 20,000 health facility records were pulled from the RHIS – a complete take for the entire country. After removing those provinces and health zones that previously received intensive health systems strengthening support or experienced policies meant to boost overall clinic volumes in the face of EVD epidemics, a loss of 2,884 health facility records was noted. An additional 739 health facility records were removed because they were completely blank across all selected data elements. Across the remaining records, the degree of data missingness varied with an average missingness of 63.0% (i.e., on average, each data set for each individual data element had 63.0% of its data missing/blank). The data set for treatment of severe diarrhea/dehydration was nearly 95% blank. The most well-reported data element among the 12 selected as outcomes was live births with only about 30% of its data missing/blank. Few data points were identified (and removed) as outliers comprising, on average, only about 0.25% of available, non-missing values. We arbitrarily set a decision rule to remove health facility records with seven or more consecutive missing values, which allows for a liberal run of missing values. Final data sets ranges from 735 health facility records (treatment of severe diarrhea/dehydration) to 11,725 records (live births). Data availability across these final data sets averaged at about 81% with a high of 90.7% (live births) to a low of 65.4% (live births <2,500g). The final data sets were those across which interpolation was conducted.

Table 1. Data processing steps and loss of health facility records

Routine Health Information System Data Element	Data Cleaning Steps: Total health facility record count	Count after removal of ASSP and EVD-affected areas	Count after removal of records with no data	Percent of missing data points	Percent of non-missing data points removed as outliers	Percent of records removed with ≥7 consecutive missing values	Interpolation Step: Available health facility records (final)	Percent of non-missing data points (final)
New acceptors of modern contraceptive methods	20,451	17,567	16,828	61.3%	0.37%	65.8%	5,760	85.1%
Attendance at the fourth ANC visit	20,451	17,567	16,828	38.0%	0.12%	39.0%	10,266	90.0%
Insecticide-treated bed net distribution during ANC visits	20,451	17,567	16,828	54.6%	0.16%	60.5%	6,641	75.0%
Live births	20,451	17,567	16,828	29.8%	0.14%	30.3%	11,725	90.7%
Live births <2,500g	20,451	17,567	16,828	86.4%	0.34%	91.4%	1,445	64.5%
Exclusive breastfeeding	20,451	17,567	16,828	52.7%	0.24%	54.5%	7,654	86.0%
Measles vaccination	20,451	17,567	16,828	51.6%	0.43%	51.3%	8,203	87.9%
Pentavalent vaccination	20,451	17,567	16,828	50.6%	0.39%	50.7%	8,296	89.0%
Moderate malnutrition	20,451	17,567	16,828	71.9%	0.35%	76.3%	3,982	81.3%
Severe malaria treatment	20,451	17,567	16,828	78.4%	0.13%	83.5%	2,783	81.8%
Severe diarrhea/dehydration treatment	20,451	17,567	16,828	92.5%	0.10%	95.6%	735	70.9%
Severe pneumonia treatment	20,451	17,567	16,828	88.2%	0.12%	92.2%	1,305	74.9%

Covariate Balance

The gradient boosted model and PSM process were successful in balancing the distribution of all selected indicators across comparison and intervention sites. The comparison of absolute standardized mean difference values before and after the gradient boosted model and PSM process showed that matching significantly reduced the absolute standardized mean difference for each indicator to below the standard threshold of 10 percent. **Figure 4** of Appendix 1 shows that each of the propensity score weighted absolute standardized mean difference values was below the five percent threshold as well. Restricting the propensity scores to a region of common support only led to the exclusion of 1.6 percent (57 of 3,667) of intervention facilities and 2.0 percent (113 of 5,549) of comparison facilities (**Figure 5** of Appendix 1).

Common Trends

The common trends assumption suggests no significant differences between the comparison and intervention groups in the preintervention time series trends for all but one of the RHIS indicators (i.e., new acceptors of modern contraceptive methods). Note that attempts to satisfy the common trends assumption between the comparison and intervention group in the preintervention period for new acceptors of modern contraceptive methods proved unsuccessful. To formally test the assumption of common trends, we separately regressed each of the 12 selected RHIS indicators transformed to rates on a linear measure for time trend in the preintervention period interacted with a dummy variable for the intervention group. Time-invariant covariates of urban/rural health zone status, hospital facility, prevalence of improved housing (health zone level), and educational attainment of women of reproductive age (health zone level) were included in each linear regression model. Inverse probability weights calculated from the propensity scores (ps/(1-ps)) for comparison facilities were also used in these models (note that intervention facilities were ascribed a weight of 1). The p-values reported in **Figures 6a** and **6b** of Appendix 1 are for the coefficients of the interaction term, which shows whether the comparison and intervention groups demonstrated a different time trend prior to the onset of IHP activities. For all but one of the assessed RHIS indicators, the differences in trends across the comparison and intervention groups were not statistically significant at an alpha level of 0.05. Taking into consideration the success in achieving covariate balance between the comparison and intervention groups, the results of the linear regression for preintervention time trends, and a visual inspection of these trends, we can expect that the common trend assumption is plausible for each indicator, with the exception of new acceptors of modern contraceptive methods.

Difference-in-Differences

Table 2 includes a summary of the DID estimators for each of the 12 assessed RHIS indicators. In the graphs (**Figures 2a** and **2b**) and **Table 2**, minimal and nonsignificant changes were observed in the DID estimates for all twelve indicators. The DID estimator coefficients for eight of the 12 RHIS indicators all suggest changes in the anticipated direction. One of the largest changes in rates was noted for moderate malnutrition, for which 7.06 additional cases per 1,000 children under five years were observed in the intervention health zones relative to comparison health zones although this finding was not significant and in an unanticipated direction. Although the unadjusted and Westfall-Young corrected p-values from the regressions for moderate malnutrition suggest a significant difference in the change in rate between intervention and comparison sites, the wild cluster bootstrap adjusted p-value should be used for significance interpretation. The largest although insignificant change in rate was noted for exclusive breastfeeding, with 33.07 additional cases per 1,000 children under six months. Tables 3 through 14 of Appendix 1 show detailed results from the 12 DID regressions.

Figure 2a. Graphical depictions of the difference-in-differences analyses by RHIS indicator (note that 2020 results are superseded by 2021)

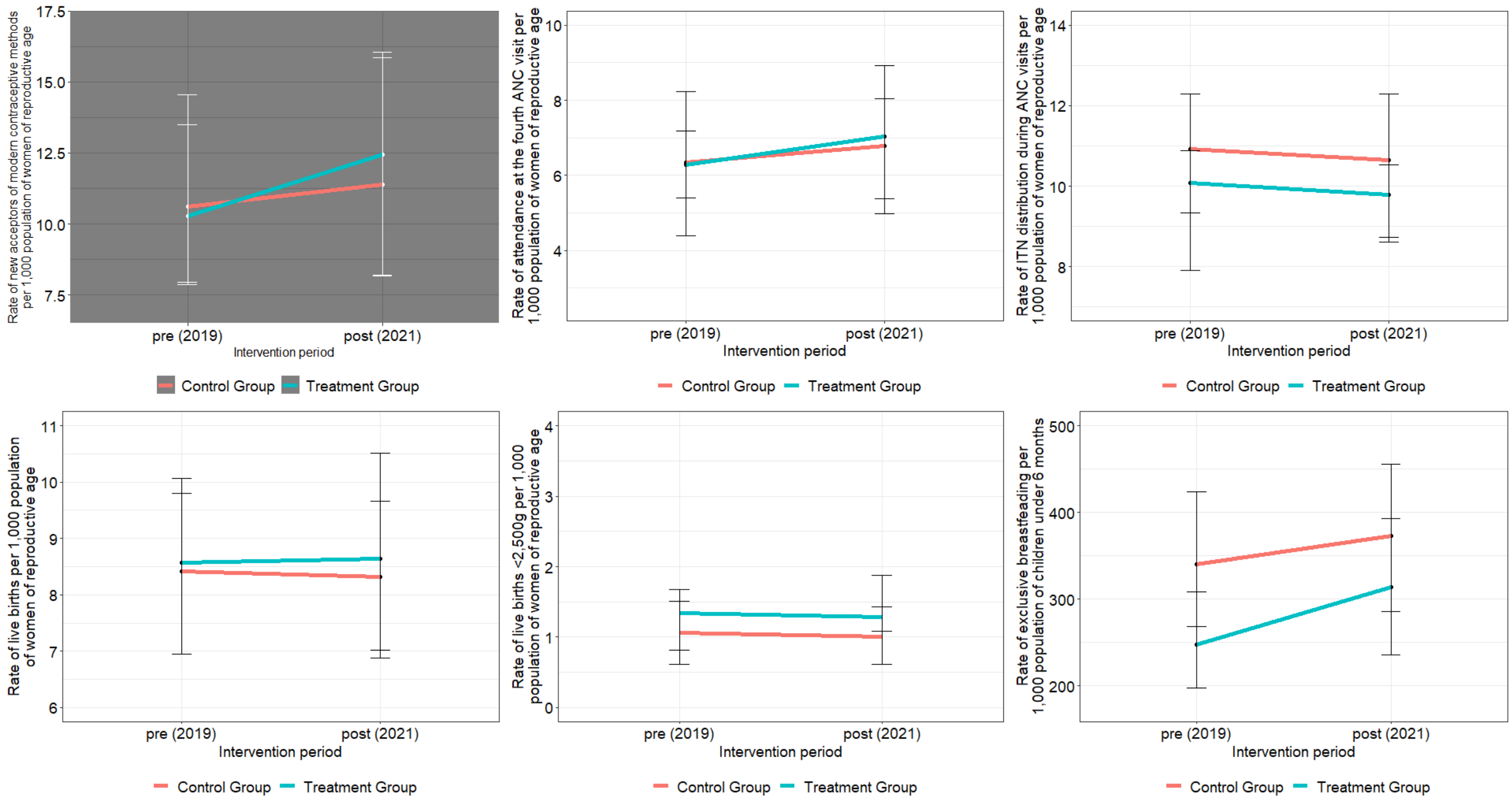


Figure 2b. Graphical depictions of the difference-in-differences analyses by RHIS indicator (note that 2020 results are superseded by 2021)

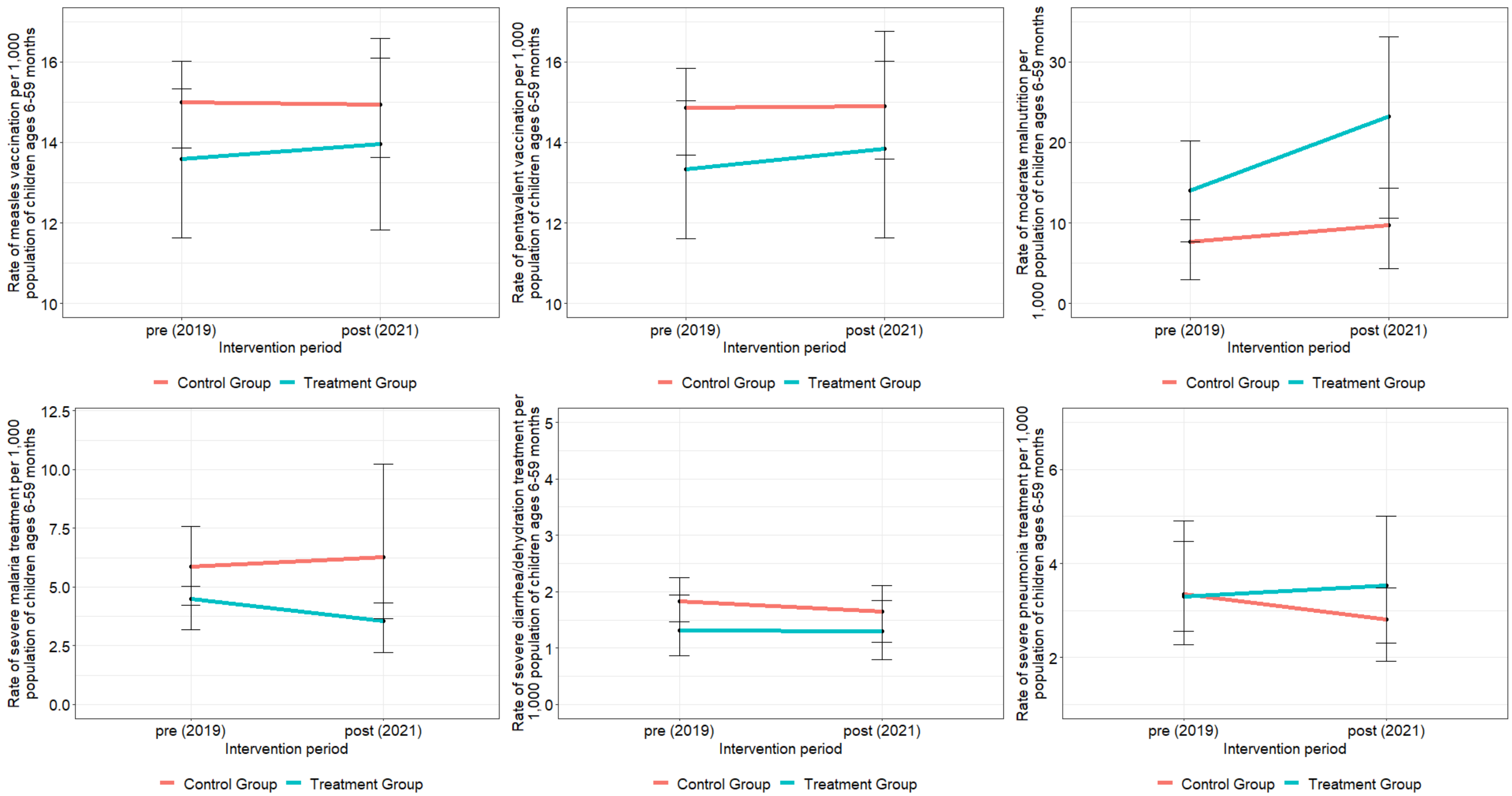


Table 2. Summary of difference-in-differences estimators by assessed RHIS indicator (note that 2020 results are superseded by 2021)

RHIS indicator	Mean difference in comparison sites (2019 vs 2021)	2019 vs 2021			
		DID estimator [WCB 95% CI]	Unadjusted p-value	WCB p-value	Westfall-Young multiple hypothesis testing adjusted p-value
New acceptors of modern contraceptive methods per 1,000 WRA*	0 .83	1.40 [-0.58, 3.52]	0.170	0.183	0.732
Attendance at the fourth ANC visit per 1,000 WRA	0 .59	0.31 [-0.26, 0.91]	0.287	0.270	0.793
Insecticide-treated bed net distribution during ANC visits per 1,000 WRA	-0.66	-0.03 [-1.04, 1.42]	0.943	0.946	0.988
Live births per 1,000 WRA	-0.21	0.15 [-0.37, 0.76]	0.568	0.560	0.988
Live births <2,500g per 1,000 WRA	-0.07	0.00 [-0.38, 0.60]	0.984	0.988	0.988
Exclusive breastfeeding per 1,000 children under-6 months	17.88	33.07 [-40.76, 107.70]	0.301	0.339	0.742
Measles vaccination per 1,000 children under-5 years	-0.22	0.43 [-0.59, 1.46]	0.368	0.443	0.793
Pentavalent vaccination per 1,000 children under-5 years	-0.07	0.47 [-0.65, 1.69]	0.362	0.460	0.764
Moderate malnutrition per 1,000 children under-5 years	2.30	7.06 [-0.18, 12.94]	***0.018***	0.055	***0.033***
Severe malaria treatment per 1,000 children under-5 years	-0.52	-1.35 [-4.19, 1.09]	0.316	0.469	0.793
Severe diarrhea/dehydration treatment per 1,000 children under-5 years	0.18	0.17 [-0.52, 0.79]	0.558	0.593	0.934
Severe pneumonia treatment per 1,000 children under-5 years	-0.81	0.76 [-0.63, 2.05]	0.177	0.270	0.793

*Common trends assumption not satisfied in the preintervention period.
Note that the Westfall-Young procedure is unable to account for wild cluster bootstrapping and the corresponding p-value is relative to the unadjusted p-value. In the case of moderate malnutrition, the WCB p-value should take precedence as it is insignificant. Also note that all WCB p-values are >0.05 indicating that no DID estimators are significant – this scenario should only result in adjusted p-values that are even closer to 1 when considering multiple hypothesis testing.

Baseline measures included in the DID analyses reflect three-month averages for March 2019 through May 2019, which were compared to averages spanning the same three months in 2021.
RHIS – routine health information system; DID – difference-in-differences; WCB – wild cluster bootstrap; *CI* – confidence interval; *WRA* –women of reproductive age; *ANC* – antenatal care

Summary and Implications

Overall, mostly small changes are noted in the assessed RHIS indicators two years into USAID IHP program implementation. As this impact analysis is expected to occur on an annual basis, year-over-year results need to be interpreted within the proper context. For example, the time required for project start-up activities before integrated health strategies began in earnest would have tempered the potential impact of the project during its first year. Nevertheless, Year 2 impact evaluation results show movement in the anticipated direction for many of the assessed indicators although not to a statistically significant degree when compared to comparison sites. Just as important, these results highlight areas that may warrant additional program focus, such as treatment of severe malaria, diarrhea, and pneumonia in children, and prevention and treatment of moderate malnutrition in children.

Additionally, it is important to remember that the impact evaluation is making use of routinely reported health facility data for which poor data quality remains an issue. Inaccuracies in data reporting could add spurious variability to the dependent variables used in the DID regression analyses, which has the potential to mask the detection of significant effects where there may be an impact. Note that our analyses for both the PSM and DID procedures made use of the most well-reported data elements in the RHIS.

Unfortunately, the RHIS indicator for new acceptors of modern contraceptive methods did not exhibit similar trends between the comparison and intervention facilities in the preintervention period and as such should be interpreted through the lens of not satisfying a basic assumption of the DID analysis. This means the comparison group does not serve as an appropriate counterfactual to the intervention group and results should be interpreted with caution.

Recommendations

The USAID IHP may draw its attention to those indicators that show little to no movement in the anticipated direction, such as treatment of severe malaria, diarrhea, and pneumonia in children, and consider whether any course adjustments are warranted.

Findings from this portion of the evaluation will be triangulated with those from the health system surveys and the qualitative data collection, and D4I will make final recommendations in the midline evaluation report.

References

Adam T, Hsu J, de Savigny D, Lavis JN, Røttingen JA, Bennett S. Evaluating health systems strengthening interventions in low-income and middle-income countries: are we asking the right questions? Health Policy Plan. 2012 Oct;27 Suppl 4:iv9–19. doi: 10.1093/heapol/czs086. PMID: 23014156.

McCaffrey DF, Ridgeway G, Morral AR. Propensity score estimation with boosted regression for evaluating causal effects in observational studies. Psychol Methods. 2004 Dec; 9(4):403–25. doi: 10.1037/1082-989X.9.4.403. PMID: 15598095.

Hyndman RJ, Khandakar Y. Automatic time series forecasting: the forecast package for R. Journal of Statistical Software, 2008; 26(3): 1–22. doi: 10.18637/jss.v027.i03 (https://doi.org/10.18637/jss.v027.i03).

Hyndman R, Athanasopoulos G, Bergmeir C, Caceres G, Chhay L, O'Hara-Wild M, Petropoulos F, Razbash S, Wang E, Yasmeen F. forecast: Forecasting functions for time series and linear models. 2022; R package version 8.16, (https://pkg.robjhyndman.com/forecast/).

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, 2021, Vienna, Austria. (https://www.R-project.org/).

Appendix 1. Additional Figures and Tables

Figure 3. Receiver operating characteristic curve for gradient boosted model used to estimate propensity scores

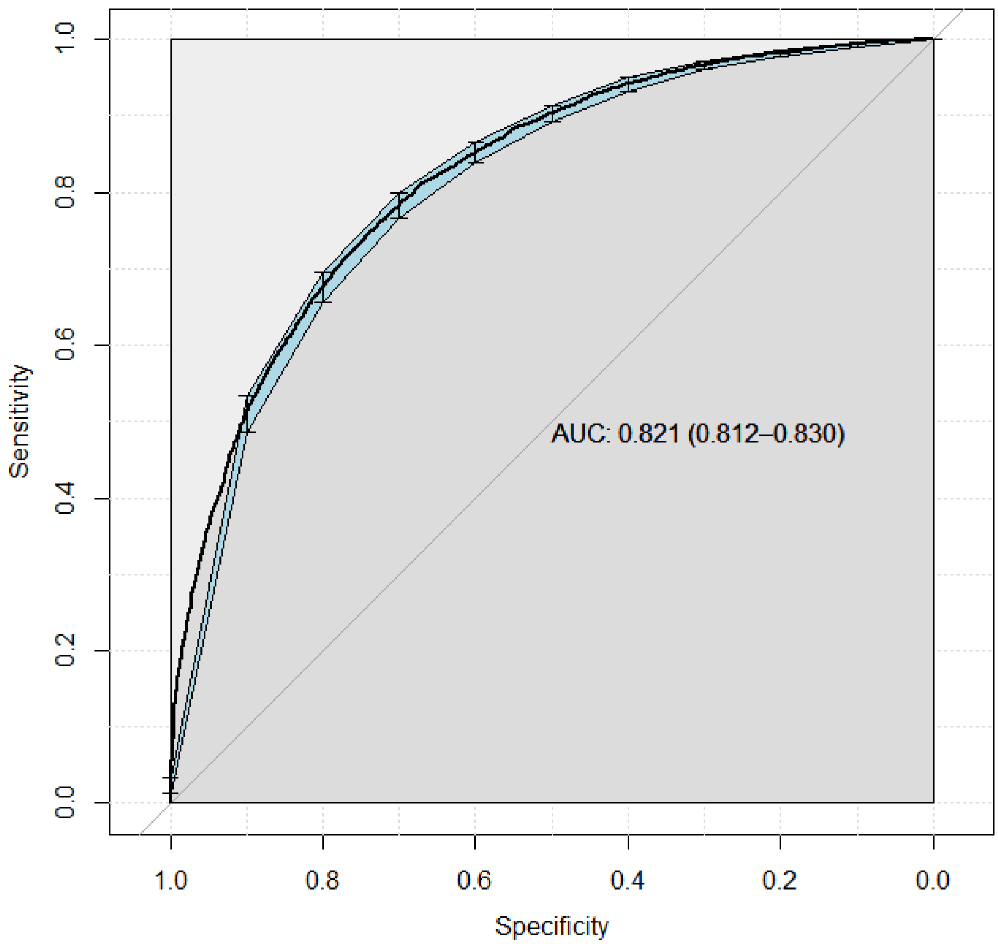
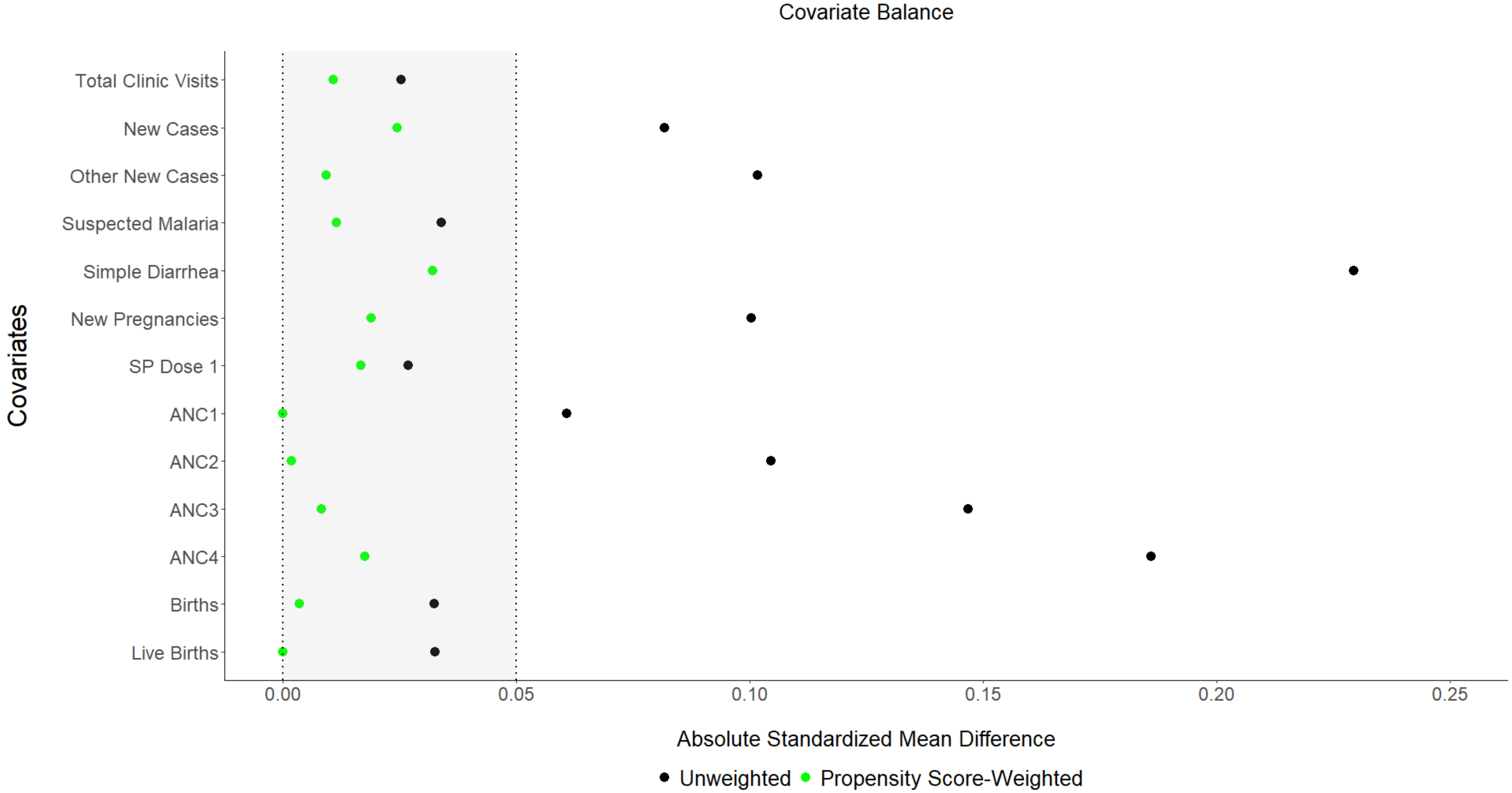


Figure 4. Love plot of covariate balance (from updated analysis)



Black points represent unadjusted absolute standardized mean difference (ASMD) between intervention and comparison values. Green points represent the adjusted ASMD between intervention and propensity score-weighted comparison values. An ASMD value less than 0.1 indicates good balance while values less than 0.05 indicate much better balance.

Figure 5. Propensity score region of common support before and after matching (from updated analysis)

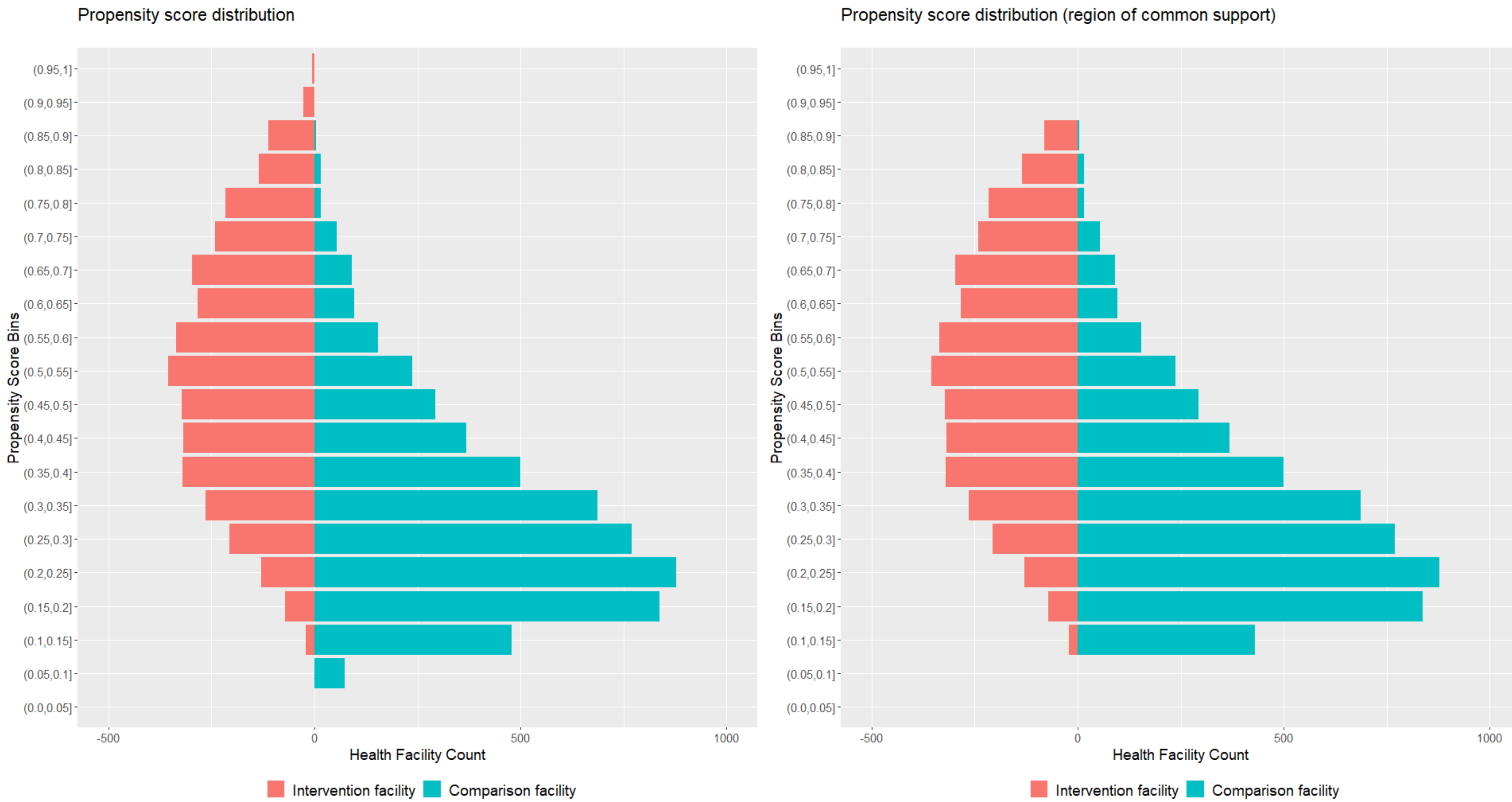


Figure 6a. Assessment of common trends assumption by RHIS indicator (from updated analysis)

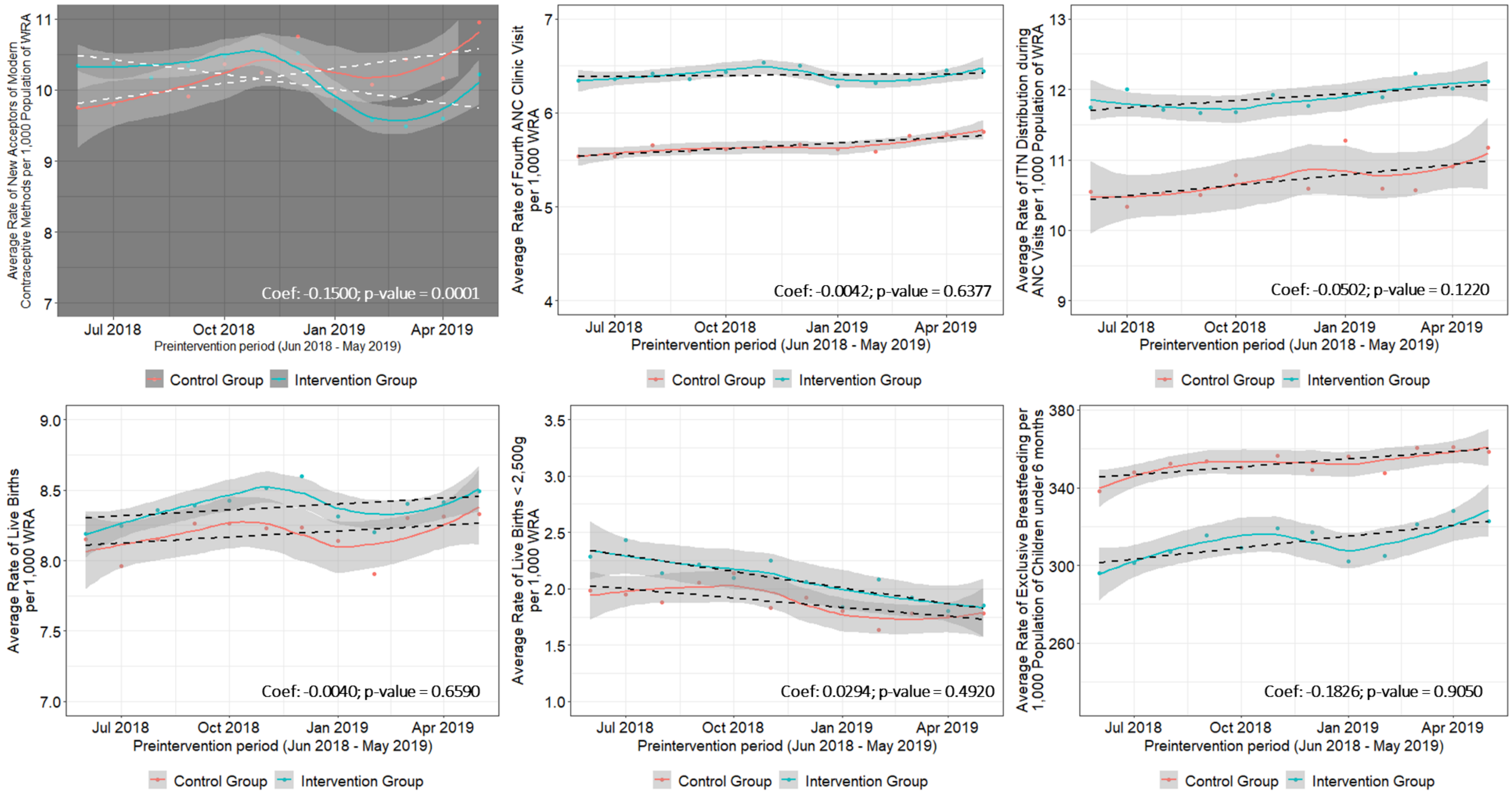


Figure 6b. Assessment of common trends assumption by RHIS indicator (from updated analysis)

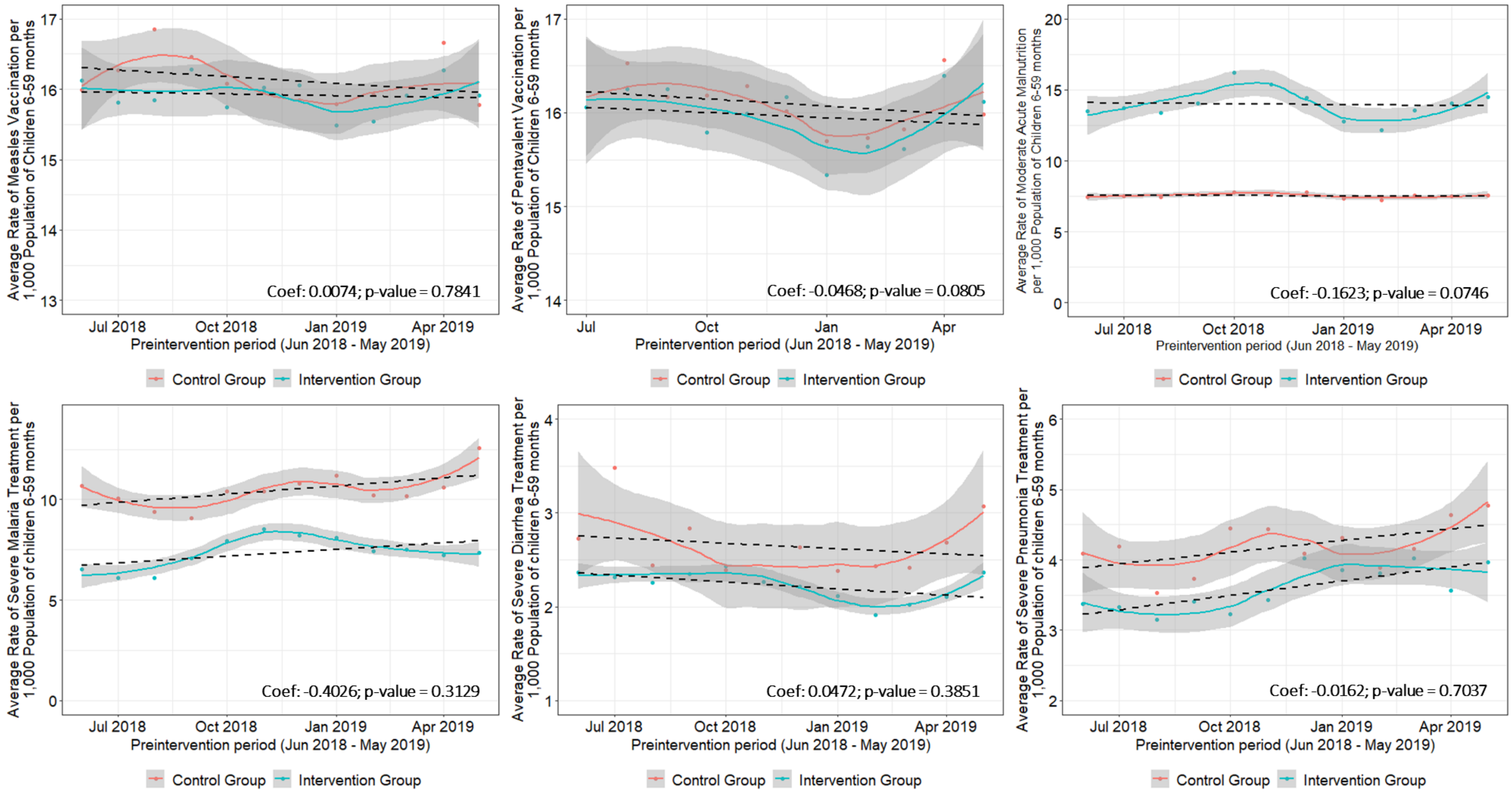


Figure 7. RHIS reporting rate of basic services by province (July 2021–November 2021)

Taux de Rapportage - Services de Base (par AS / 6 derniers mois)					
A- Services de Base Reporting rate					
	Juillet 2021	Août 2021	Septembre 2021	Octobre 2021	Novembre 2021
bu Bas Uele Province	100	100	99.4	100	100
eq Equateur Province	99.8	98.7	98.9	95.7	87.1
hk Haut Katanga Province	89.6	95	88.9	89.7	83.8
hl Haut Lomami Province	91	95.9	91.5	91.5	86.7
hu Haut Uele Province	90.6	88.8	87.6	86	76.4
it Ituri Province	99.2	95.5	95.4	88.8	80
kc Kongo Central Province	93.7	93.7	93.7	93.7	97.6
ke Kasai Oriental Province	99.8	100	99.5	97.6	94.9
kg Kwango Province	100	100	99.5	99.7	98.4
kl Kwilu Province	100	99.8	100	92.8	87.1
kn Kinshasa Province	96.6	98.1	98.5	98.1	98.5
kr Kasai Central Province	97.3	98	98.7	99.6	99.5
ks Kasai Province	97.3	99.8	99.5	99.7	98.9
ll Lualaba Province	87.6	91.6	93.3	93.8	95.5
lm Lomami Province	99.9	97	98.2	98.9	95.5
md Maindombe Province	97.8	95.2	94.7	85.5	77.8
mg Mongala Province	100	99.6	100	98.8	87.3
mn Maniema Province	100	95.6	96.3	95.8	100
nk Nord Kivu Province	97.1	96.4	96.3	81.3	67.4
nu Nord Ubangi Province	98.7	99.6	98.7	95.1	93.7
sk Sud Kivu Province	83.5	78.7	77.4	76.8	71.7
sn Sankuru Province	99.1	73.4	71.3	84.7	87.3
su Sud Ubangi Province	100	100	99.7	91.9	84.1
tn Tanganyika Province	73.3	29.5	25.8	38.2	14.9
tp Tshopo Province	99.9	99.6	100	87.4	72
tu Tshuapa Province	100	100	99.8	99.3	99.3

Extracted from the DRC RHIS January 7, 2022.

Table 3. Rate of new acceptors of modern contraceptive methods per 1,000 women of reproductive age

	estimate [WCB 95% CI]	std. err	statistic	WCB p.value
Intervention facility*Post intervention period (DID est.)	1.4 [-0.58, 3.52]	0.982884	1.42	0.1832
Intervention facility	-0.35 [-4.96, 3.96]	1.729721	-0.2	0.8689
Post intervention period	0.77 [-0.70, 2.04]	0.595835	1.3	0.2202
Educational attainment - women of reproductive age	-0.1 [-1.94, 1.48]	0.718067	-0.13	0.9129
Prevalence of improved housing	-14 [-51.37, 8.66]	11.06279	-1.27	0.3804
Rainfall: 3-month average (Mar/Apr/May)	1.05 [-1.23, 2.81]	0.786937	1.33	0.3013
NDVI: 3-month average (Mar/Apr/May)	4.51 [-24.38, 30.34]	10.2483	0.44	0.7267
Urban health zone	-0.79 [-4.07, 3.60]	1.584077	-0.5	0.6727

Note: WCB – wild cluster bootstrap; std. err. – standard error; DID est. – difference-in-differences estimate; NDVI – normalized difference vegetation index

Table 4. Rate of attendance at the fourth ANC visit per 1,000 women of reproductive age

	estimate [WCB 95% CI]	std. err	statistic	WCB p.value
Intervention facility*Post intervention period (DID est.)	0.31 [-0.26, 0.91]	0.283954	1.09	0.2703
Intervention facility	-0.06 [-2.06, 1.84]	0.784122	-0.08	0.9439
Post intervention period	0.44 [-0.09, 0.94]	0.207387	2.1	0.0791
Educational attainment - women of reproductive age	-0.56 [-1.23, 0.13]	0.278076	-2.02	0.1011
Prevalence of improved housing	0.05 [-14.47, 13.10]	3.370491	0.01	0.9950
Rainfall: 3-month average (Mar/Apr/May)	0.8 [-0.48, 1.65]	0.387855	2.07	0.1752
NDVI: 3-month average (Mar/Apr/May)	0.86 [-9.19, 10.39]	3.888809	0.22	0.8398
Urban health zone	-1.44 [-2.97, 0.60]	0.758509	-1.9	0.2322

Note: WCB – wild cluster bootstrap; std. err. – standard error; DID est. – difference-in-differences estimate; NDVI – normalized difference vegetation index

Table 5. Rate of ITN distribution during ANC visits per 1,000 women of reproductive age

	estimate [WCB 95% CI]	std. err	statistic	WCB p.value
Intervention facility*Post intervention period (DID est.)	-0.03 [-1.04, 1.42]	0.453162	-0.07	0.9459
Intervention facility	-0.84 [-3.18, 1.08]	0.828144	-1.02	0.3433
Post intervention period	-0.26 [-1.24, 0.73]	0.425388	-0.62	0.5696
Educational attainment - women of reproductive age	-1.29 [-2.15, -0.46]	0.346418	-3.71	0.0070
Prevalence of improved housing	4.85 [-11.92, 19.53]	5.851892	0.83	0.4765
Rainfall: 3-month average (Mar/Apr/May)	0.2 [-0.56, 0.73]	0.253668	0.8	0.4805
NDVI: 3-month average (Mar/Apr/May)	7.59 [-1.41, 17.23]	3.863391	1.96	0.0881
Urban health zone	-0.6 [-1.95, 0.83]	0.613314	-0.98	0.3253

Note: WCB – wild cluster bootstrap; std. err. – standard error; DID est. – difference-in-differences estimate; NDVI – normalized difference vegetation index

Table 6. Rate of live births per 1,000 women of reproductive age

	estimate [WCB 95% CI]	std. err	statistic	WCB p.value
Intervention facility*Post intervention period (DID est.)	0.15 [-0.37, 0.76]	0.265473	0.58	0.5596
Intervention facility	0.16 [-1.84, 1.98]	0.78676	0.21	0.8639
Post intervention period	-0.09 [-0.48, 0.24]	0.152168	-0.6	0.5816
Educational attainment - women of reproductive age	-0.62 [-1.47, 0.23]	0.352309	-1.77	0.1341
Prevalence of improved housing	2.45 [-14.64, 19.20]	4.490206	0.54	0.7007
Rainfall: 3-month average (Mar/Apr/May)	0.4 [-0.68, 1.11]	0.347568	1.15	0.3984
NDVI: 3-month average (Mar/Apr/May)	4.48 [-4.30, 13.77]	3.515511	1.27	0.2362
Urban health zone	-1.15 [-2.99, 0.91]	0.820875	-1.4	0.3994

Note: WCB – wild cluster bootstrap; std. err. – standard error; DID est. – difference-in-differences estimate; NDVI – normalized difference vegetation index

Table 7. Rate of live births <2,500g per 1,000 women of reproductive age

	estimate [WCB 95% CI]	std. err	statistic	WCB p.value
Intervention facility*Post intervention period (DID est.)	0 [-0.38, 0.60]	0.192233	-0.02	0.9880
Intervention facility	0.28 [-0.40, 0.86]	0.243931	1.17	0.3644
Post intervention period	-0.05 [-0.24, 0.21]	0.093768	-0.55	0.5966
Educational attainment - women of reproductive age	-0.02 [-0.29, 0.19]	0.095255	-0.18	0.8739
Prevalence of improved housing	-0.58 [-3.78, 3.92]	0.888216	-0.65	0.5856
Rainfall: 3-month average (Mar/Apr/May)	0.02 [-0.20, 0.35]	0.112597	0.17	0.8809
NDVI: 3-month average (Mar/Apr/May)	0.44 [-2.58, 5.75]	1.297538	0.34	0.7958
Urban health zone	-0.46 [-0.77, -0.16]	0.114888	-4.01	0.0060

Note: WCB – wild cluster bootstrap; std. err. – standard error; DID est. – difference-in-differences estimate; NDVI – normalized difference vegetation index

Table 8. Rate of exclusive breastfeeding per 1,000 children under 6 months

	estimate [WCB 95% CI]	std. err	statistic	WCB p.value
Intervention facility*Post intervention period (DID est.)	33.07 [-40.76, 107.70]	31.12776	1.06	0.3393
Intervention facility	-92.57 [-201.80, 16.03]	44.31327	-2.09	0.0831
Post intervention period	32.79 [-5.76, 72.33]	17.45415	1.88	0.1081
Educational attainment - women of reproductive age	-41.61 [-106.30, 9.78]	22.93694	-1.81	0.1361
Prevalence of improved housing	610.92 [-183.90, 2,560.00]	473.2394	1.29	0.3183
Rainfall: 3-month average (Mar/Apr/May)	0.66 [-25.06, 36.90]	14.08444	0.05	0.9740
NDVI: 3-month average (Mar/Apr/May)	200.9 [-686.70, 1,115.00]	348.4964	0.58	0.6887
Urban health zone	4.93 [-42.38, 62.55]	25.05686	0.2	0.8549

Note: WCB – wild cluster bootstrap; std. err. – standard error; DID est. – difference-in-differences estimate; NDVI – normalized difference vegetation index

Table 9. Rate of measles vaccination per 1,000 children ages 6–59 months

	estimate [WCB 95% CI]	std. err	statistic	WCB p.value
Intervention facility*Post intervention period (DID est.)	0.43 [-0.59, 1.46]	0.466403	0.92	0.4434
Intervention facility	-1.42 [-3.50, 0.41]	0.761055	-1.86	0.1221
Post intervention period	-0.06 [-0.74, 0.55]	0.272024	-0.22	0.8328
Educational attainment - women of reproductive age	-1.09 [-1.95, -0.24]	0.318319	-3.43	0.0090
Prevalence of improved housing	6.21 [-12.49, 24.75]	4.339244	1.43	0.4464
Rainfall: 3-month average (Mar/Apr/May)	0.34 [-0.96, 1.26]	0.391549	0.86	0.5706
NDVI: 3-month average (Mar/Apr/May)	2.91 [-8.85, 11.95]	3.885831	0.75	0.5425
Urban health zone	-1.81 [-3.95, 0.65]	0.981987	-1.84	0.2172

Note: WCB – wild cluster bootstrap; std. err. – standard error; DID est. – difference-in-differences estimate; NDVI – normalized difference vegetation index

Table 10. Rate of pentavalent vaccination per 1,000 children ages 6–59 months

	estimate [WCB 95% CI]	std. err	statistic	WCB p.value
Intervention facility*Post intervention period (DID est.)	0.47 [-0.65, 1.69]	0.501803	0.93	0.4595
Intervention facility	-1.53 [-3.36, 0.08]	0.668902	-2.28	0.0641
Post intervention period	0.03 [-0.48, 0.47]	0.20162	0.17	0.8639
Educational attainment - women of reproductive age	-1.19 [-2.04, -0.32]	0.324023	-3.66	0.0060
Prevalence of improved housing	6.85 [-11.26, 24.79]	4.233398	1.62	0.3944
Rainfall: 3-month average (Mar/Apr/May)	0.29 [-1.03, 1.23]	0.393228	0.75	0.6627
NDVI: 3-month average (Mar/Apr/May)	2.67 [-10.45, 13.52]	4.378887	0.61	0.6146
Urban health zone	-1.8 [-4.61, 0.99]	1.20684	-1.49	0.3964

Note: WCB – wild cluster bootstrap; std. err. – standard error; DID est. – difference-in-differences estimate; NDVI – normalized difference vegetation index

Table 11. Rate of moderate malnutrition per 1,000 children ages 6–59 months

	estimate [WCB 95% CI]	std. err	statistic	WCB p.value
Intervention facility*Post intervention period (DID est.)	7.06 [-0.18, 12.94]	2.73564	2.58	0.0551
Intervention facility	6.44 [-1.83, 15.38]	3.3929	1.9	0.2042
Post intervention period	2.08 [-1.03, 5.12]	1.348971	1.54	0.2262
Educational attainment - women of reproductive age	-0.88 [-4.59, 2.68]	1.35293	-0.65	0.6016
Prevalence of improved housing	-6.61 [-76.30, 76.59]	14.41929	-0.46	0.7197
Rainfall: 3-month average (Mar/Apr/May)	2.83 [-0.30, 5.19]	1.073884	2.64	0.0771
NDVI: 3-month average (Mar/Apr/May)	-16.55 [-66.98, 38.27]	22.25057	-0.74	0.5195
Urban health zone	-5.16 [-10.86, 0.88]	2.613936	-1.98	0.0991

Note: WCB – wild cluster bootstrap; std. err. – standard error; DID est. – difference-in-differences estimate; NDVI – normalized difference vegetation index

Table 12. Rate of severe malaria treatment per 1,000 children ages 6–59 months

	estimate [WCB 95% CI]	std. err	statistic	WCB p.value
Intervention facility*Post intervention period (DID est.)	-1.35 [-4.19, 1.09]	1.315504	-1.03	0.4685
Intervention facility	-1.38 [-3.99, 0.60]	0.834229	-1.66	0.1762
Post intervention period	0.4 [-1.72, 3.69]	1.342208	0.3	0.9940
Educational attainment - women of reproductive age	-0.95 [-1.94, -0.11]	0.401271	-2.37	0.0160
Prevalence of improved housing	0.27 [-17.73, 17.86]	4.83846	0.06	0.9620
Rainfall: 3-month average (Mar/Apr/May)	0.29 [-0.78, 1.34]	0.371679	0.79	0.5335
NDVI: 3-month average (Mar/Apr/May)	-2.51 [-17.03, 8.54]	4.78254	-0.53	0.6837
Urban health zone	-1.56 [-2.33, -0.40]	0.434212	-3.59	0.0100

Note: WCB – wild cluster bootstrap; std. err. – standard error; DID est. – difference-in-differences estimate; NDVI – normalized difference vegetation index

Table 13. Rate of severe diarrhea/dehydration treatment per 1,000 children ages 6–59 months

	estimate [WCB 95% CI]	std. err	statistic	WCB p.value
Intervention facility*Post intervention period (DID est.)	0.17 [-0.52, 0.79]	0.280304	0.6	0.5926
Intervention facility	-0.52 [-1.05, 0.33]	0.263169	-1.97	0.1361
Post intervention period	-0.17 [-0.73, 0.37]	0.240653	-0.72	0.5015
Educational attainment - women of reproductive age	-0.2 [-0.48, 0.12]	0.128042	-1.57	0.1862
Prevalence of improved housing	0.29 [-6.03, 5.18]	1.865816	0.15	0.8879
Rainfall: 3-month average (Mar/Apr/May)	0.12 [-0.20, 0.54]	0.135216	0.87	0.7137
NDVI: 3-month average (Mar/Apr/May)	-0.17 [-3.57, 3.38]	1.450482	-0.12	0.9109
Urban health zone	-0.5 [-1.11, 0.04]	0.239893	-2.08	0.0641

Note: WCB – wild cluster bootstrap; std. err. – standard error; DID est. – difference-in-differences estimate; NDVI – normalized difference vegetation index

Table 14. Rate of severe pneumonia treatment per 1,000 children ages 6–59 months

	estimate [WCB 95% CI]	std. err	statistic	WCB p.value
Intervention facility*Post intervention period (DID est.)	0.76 [-0.63, 2.05]	0.54429	1.4	0.2693
Intervention facility	-0.04 [-1.26, 1.41]	0.515457	-0.08	0.9439
Post intervention period	-0.53 [-1.04, 0.08]	0.245497	-2.14	0.1191
Educational attainment - women of reproductive age	-0.46 [-1.26, 0.22]	0.300326	-1.54	0.1892
Prevalence of improved housing	4.3 [-4.46, 25.90]	4.508902	0.95	0.4565
Rainfall: 3-month average (Mar/Apr/May)	1.1 [-0.14, 2.86]	0.541776	2.02	0.0841
NDVI: 3-month average (Mar/Apr/May)	4.05 [-2.25, 13.52]	3.050012	1.33	0.2603
Urban health zone	-1.53 [-2.78, -0.41]	0.464083	-3.3	0.0010

Note: WCB – wild cluster bootstrap; std. err. – standard error; DID est. – difference-in-differences estimate; NDVI – normalized difference vegetation index

Data for Impact
University of North Carolina at Chapel Hill
123 West Franklin Street, Suite 330
Chapel Hill, NC 27516 USA

http://www.data4impactproject.org

This publication was produced with the support of the United States Agency for International Development (USAID) under the terms of the Data for Impact (D4I) associate award 7200AA18LA00008, which is implemented by the Carolina Population Center at the University of North Carolina at Chapel Hill, in partnership with Palladium International, LLC; ICF Macro, Inc.; John Snow, Inc.; and Tulane University. The views expressed in this publication do not necessarily reflect the views of USAID or the United States government. TR-21-453 D4I