

© Copyright 2016

Jessica M. Northridge

Evaluation of a Quality Improvement (QI) Intervention for
Maternal, Newborn, and Child Health (MNCH) in Colombia:
The *Mil Días* Intervention of SINERGIAS in the Department of Bolívar

Jessica M. Northridge

A thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Public Health

University of Washington

2016

Committee:

Pablo Montoya

Bernardo Hernández Prado

Ana Judith Blanco

Program Authorized to Offer Degree:
Department of Global Health

University of Washington

Abstract

Evaluation of a Quality Improvement (QI) Intervention for
Maternal, Newborn, and Child Health (MNCH) in Colombia:
The *Mil Días* Intervention of SINERGIAS in the Department of Bolívar

Jessica M. Northridge

Chair of the Supervisory Committee:

Clinical Assistant Professor Pablo Montoya
Department of Global Health

Initiated in 2012, the *Mejoramiento de la calidad de la atención a gestantes, niñas y niños en sus primeros mil días (Mil Días)* intervention is a strategy for quality improvement (QI) of maternal, child, and newborn health (MNCH) across 6 departments of Colombia. In the department of Bolívar, *Mil Días* is implemented in 25 health facilities across 12 municipalities. A time series analysis was conducted to determine whether or not contextual factors affect the rate of QI of antenatal care (ANC) among health facilities in Bolívar targeted by the *Mil Días* intervention. Multivariate linear regression analysis was used to examine relationships between predictors of interest and ANC performance over time. Important covariates included type of health facility (private or public), municipality level percentage of households with unsatisfied basic needs, number of health facilities per 100,000 municipal inhabitants, and MNCH patient volume per

health facility. Findings were that quality of ANC, as measured by mean ANC performance scores or average rates of ANC indicators, increased over the implementation of *Mil Días*. Multivariate linear regression modeling demonstrated that health facility type was a significant predictor of the rate of QI of ANC, with public facilities showing an increased rate of QI as compared to private health facilities. This evaluation of *Mil Días* provides evidence of the importance of contextual influence on QI intervention results. Future studies examining program context ought to be conducted to inform QI strategies for the implementation of MNCH and related health services.

I. Introduction

Background and Significance

Maternal, newborn, and child health (MNCH) is a priority among the health targets defined by the United Nations (UN) Millennium Development Goals (MDGs) and Sustainable Development Goals (SDGs).^{1,2} Using the UN targets as motivation, Colombia has achieved continuous decreases in the rates of under 5 year, under 1 year and maternal mortality through quality improvement (QI) strategies across the country.^{3,4} Despite documented improvements, however, national maternal mortality rates as well as maternal and infant morbidity rates failed to meet MDG targets in 2015.³ Further, large inequalities in MNCH indicators and services exist within specific regions and between populations of the country with, for example, rural and indigenous communities having greater risk of poor MNCH than urban and nonindigenous communities.⁵ In order to identify strategies to further improve MNCH, a consortium comprised of the Colombian Ministry of Health and Social Protection (MSPS), the United Nations Children's Fund (UNICEF), the United Nations Population Fund (UNFPA), and SINERGIAS – Alianzas Estratégicas para la Salud y el Desarrollo (hereafter referred to as Sinergias) sought to evaluate antenatal care (ANC) in 20 of a total of 32 departments of Colombia during 2010-2012. The purpose of this investigation was to identify the percentage of women receiving quality ANC as defined by the Colombian national guidelines (*412 Resolution of 2000*). Extensive review of medical records found that only 1% of women had received proper ANC for the early detection of complications during pregnancy.^{3,6}

Following this formative research, Sinergias, a Colombian non-governmental organization (NGO) for public health and social development, with the support of UNICEF, initiated the *Mejoramiento de la calidad de la atención a gestantes, niñas y niños en sus primeros mil días* (*Mil Días*) intervention. The focus of this QI intervention is to improve the quality of care during the first 1,000 days of life. Nutrition and access to adequate health care during this period of life, lasting from the beginning of pregnancy to approximately 24 months of age, has been shown to have lasting and profound effects on development later in life including, but not limited to, future health outcomes related and unrelated to pregnancy as well as educational attainment and economic status.⁷ Evidence supports that increasing the quality and coverage of existing family planning, antenatal, delivery, newborn, and postnatal services throughout this critical period for both mother and child is essential to reducing newborn, child and maternal mortality.⁸⁻¹⁰

Mil Días was designed as a departure from past QI strategies in the Colombian context that often led to increased burdens on health care providers at the facility level. Instead, *Mil Días* promotes active involvement of all health system stakeholders and operates as a partnership with municipal, district, and departmental Health Secretaries at Territorial Health Directorates (DTS – Direcciones Territoriales de Salud), health facilities (IPS – Institución Prestadora de Servicios), health insurance governances (EPS – Empresa Promotora de Salud), patients receiving MNCH services, and Sinergias personnel. This QI intervention is comprised of 4 general components: 1) creation and implementation of Improvement Plans for MNCH with DTS health secretaries and other key actors such as EPS representatives and IPS managers; 2) training and technical assistance for health facility personnel and DTS Health Secretaries; 3) medical records review, user satisfaction measurement, and monitoring of municipal indicators; and 4) in field support for health system stakeholders from Sinergias personnel. Over the initial 18 months of *Mil Días* implementation, intervention activities focused on the municipal level. Subsequently, the activities during the following 18 months broadened to focus on the department level. During

this latter departmental phase, the program also expanded to include new municipalities and shifted the leadership role from Sinergias to the departmental DTS. Ultimately, by June 2013, *Mil Días* was scaled up to encompass health facilities in 6 departments including La Guajira, Bolívar, Chocó, Córdoba, Putumayo, and Cauca across a total of 36 municipalities. The initial UNICEF funding period for *Mil Días* lasted through the end of 2015. Currently, *Mil Días* is implemented in the departments of Bolívar and Cauca with funding from Merck for Mothers.^{3,6}

Conceptual Model

The conceptual model of *Mil Días* embraces an ecological framework, with program components at multiple levels, from policies at the national, local, and facility levels, to health outcomes at the population and individual levels. A mapping of the program components from inputs to impact, as well as the larger context in which the intervention is implemented, is provided in Figure 1 below.

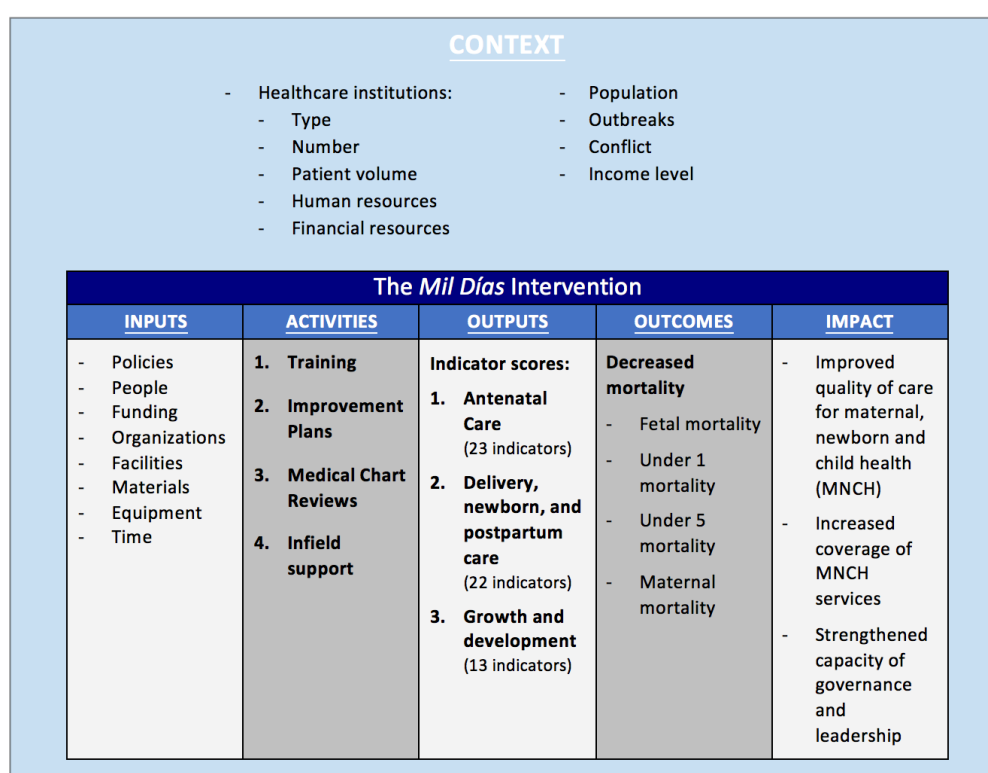


Figure 1: *Mil Días* Conceptual Model. Components of the *Mil Días* intervention can be mapped from inputs to impact. This study focuses specifically on the outputs of *Mil Días* activities, specifically those for ANC, as well as the context in which the program is implemented.

Evaluation of *Mil Días*

Routine impact evaluation of the *Mil Días* intervention is conducted through the reporting of QI indicators. Sinergias personnel collect indicator data, which quantify activities completed during MNCH visits from targeted health facility medical records at defined measurement time points. On average, measurements are completed every 3 months. For the 2015 *Mil Días* Final Report, 58 indicators were selected from the Sinergias program evaluation database to organize key QI measurements into the following 3 categories of MNCH services: 1) ANC; 2) delivery, newborn, and postpartum care; and 3) early childhood growth and development. Following each measurement, *Mil Días* summary reports are presented and reviewed at each health facility. In

addition, individual and categorical indicator scores are averaged at the departmental and municipal levels as part of a yearly *Mil Días* Final Report that is circulated by Sinergias to all intervention stakeholders.

While changes in QI indicators over time have been observed, trends in these measurements have been inconsistent from measurement to measurement. Ascertaining the reasons for any changes in indicator performance requires the examination of variables external to *Mil Días* programming.^{11–13} To date, no measured comparison of health facilities apart from indicator scores has been used to evaluate the quality of care among health facilities targeted by *Mil Días*. Contextual variables that are specific to health facilities, their geographic location, and/or the populations they serve may affect the quality of care delivered in health facilities. Examples of such measured contextual variables that are relevant to the *Mil Días* setting include type of health facility, number of health facilities providing similar services to the same population, patient volume, amount of human and financial resources at the facility level, population size of the area served, disease outbreaks, existence of political or social conflicts, and average income level among households surrounding the facility (Figure 1).^{12,14–17}

Although certain health system data exist to compare the performance of health facilities, the unavailability of MNCH-specific data across Colombian health facilities limits the strength of *Mil Días* impact evaluation.¹⁸ Comparing QI indicators among *Mil Días* health facilities to non-implementation health facilities would afford program evaluation a level of generalizability to other potential settings for QI interventions, and provide stronger evidence of effective QI strategies. Ensuring consistent data collection across settings will require the development of a universal tool for data collection. Furthermore, attempts to simplify data collection through the use of indices are warranted to promote thorough and dependable data systems that limit the burdens placed upon health care providers at the time of service delivery. Analyses that may guide the inclusion of relevant health facility variables and health indicators are needed as part of the development of an evidence-based and contextually-appropriate data collection index.¹⁹

Finally, program evaluation that considers contextual influences on quality of care will better equip future adaptations of *Mil Días* QI activities. QI for MNCH as well as other health services is dependent on a complex relationship of internal and external program characteristics.^{13,14,20,21} For this reason, effective QI strategies must be specifically tailored for their specific contexts.^{22,23} Expanding *Mil Días* evaluation to account for a larger range of influences in the attempt to explain certain program results will potentially inform the implementation of QI strategies, specifically for MNCH, in similar settings.¹¹

Specific Study Question:

What contextual factors affect the quality improvement (QI) of MNCH services delivered in health facilities targeted by the *Mil Días* intervention?

General Objectives:

1. To identify the factors external to the QI activities of *Mil Días* that influence institutional performance, and the extent of such influence, as measured by program QI indicators.
2. To contribute to the development a simpler, scalable feedback mechanism for data collection on MNCH services at the facility level.

3. To disseminate key findings from the implementation of *Mil Días* to inform future QI interventions for MNCH in similar settings.

II. Methods

A time series analysis was conducted to determine whether or not contextual factors affect the rate of quality improvement (QI) of ANC in health facilities targeted by the *Mil Días* intervention.

Study Setting

This study is an evaluation of *Mil Días* in the department of Bolívar. Bolívar is located in the Caribbean region of Colombia and encompasses a total of 46 municipalities. Among the 32 Departments of Colombia, Bolívar contains the second largest total number of people living in poverty, with approximately 1,036,131 people living in poverty and 235,897 people living in extreme poverty.⁵ Cartagena is the department's capital, and an historical site for domestic and international tourism that contributes a significant proportion of annual national profit. There is substantial geographic and demographic diversity across the department, as well as within municipalities. Both urban and rural centers are located within *Mil Días* implementation municipalities.²⁴

Selection of Health Facilities

Initial implementation municipalities for *Mil Días* were selected at the onset of the intervention, and were comprised of the most vulnerable municipalities with regard to MNCH as identified by UNICEF. These municipalities, among the 1,102 municipalities in Colombia, scored below the national average for a series of key ANC indicators. The department of Bolívar had 3 implementation municipalities at the start of *Mil Días*; an additional 9 implementation municipalities in Bolívar were included upon request from local DTSs after a concerted evaluation based on MNCH indicators.^{3,6} Implementation facilities are selected based on the number of people they serve, and in some cases, are the only health facilities located within implementation municipalities. Within the 12 implementation municipalities in Bolívar, 25 health facilities are targeted by *Mil Días*, 20 of which offer ANC. Further, among facilities offering ANC, there were 2 onsets groups for program implementation. For the first onset group, 8 facilities began implementation at measurement 0 or 1 of *Mil Días*, and continued implementation throughout a total of 6-12 program measurements, with the exception of 1 facility that was enrolled for 2 measurements (see Figure 2). A second onset group was introduced when more resources became available for program expansion. This group included an additional 12 health facilities that began implementation at measurement 8 and continued implementation throughout a total of 4 measurements, with the exception of 2 health facilities that began implementation at measurement 10 for a total of 2 implementation measurements (see Figure 2). The average time between measurements is 3 months.^{3,6}

Data Collection

This study is a secondary data analysis, using data previously collected from routine evaluation databases of Sinergias, national registries for health facilities, and Population and Housing Census data. The availability of data dictated which contextual variables (see Figure 1) could potentially be incorporated into the final analysis.

Variables of Interest

1. ANC performance score

The outcome of interest in this analysis is ANC performance score, used as a measure of quality of ANC. The number of indicators, out of a total of 23, that are completed during an individual ANC visit are recorded and then averaged to produce an ANC performance score for each health facility (Table 1). Data on ANC performance scores were obtained from the routine Sinergias database for *Mil Días*, which comprises data on a representative sample of medical charts from implementation health facilities. The number of randomly selected charts is determined in order to generate indicator scores with a 10% confidence interval (CI). Performance scores for each health facility are available for each measurement time point of *Mil Días* implementation (see Figure 2).

Table 1: ANC Performance Score. ANC indicators, or actions completed during an ANC visit, are organized with 4 composite variables that combine to produce an overall performance score for ANC in which all indicators are weighted equally. In this study, health facility ANC performance scores are the average number of ANC indicators completed during all ANC visits reviewed at a certain measurement time point, and are used as a measurement of quality of care.

Performance Score	Composite Variable	Indicators Included in the Composite Variable
Antenatal Care (23 total indicators)	Education	1. Information given on warning signs during pregnancy 2. Family planning counseling 3. Violence inquiry screening 4. Rating on biopsychosocial scale 5. Pre-HIV test counseling
	Paraclinical	1. Oral glucose tolerance test or O'Sullivan's Test 2. Record of urine culture results 3. HIV informed consent 4. HIV screening 5. Syphilis screening 6. Hepatitis B screening 7. Toxoplasmosis screening 8. Cervical cytology 9. Obstetric ultrasound
	Consultation	1. Tetanus toxoid vaccination 2. Use of uterine height chart 2. Use of weight chart 3. Obstetric risk assessment 4. Dental referral or assessment 5. Psychological referral or assessment
	Nutritional supplementation	1. Folic acid 2. Ferrous sulfate 3. Calcium

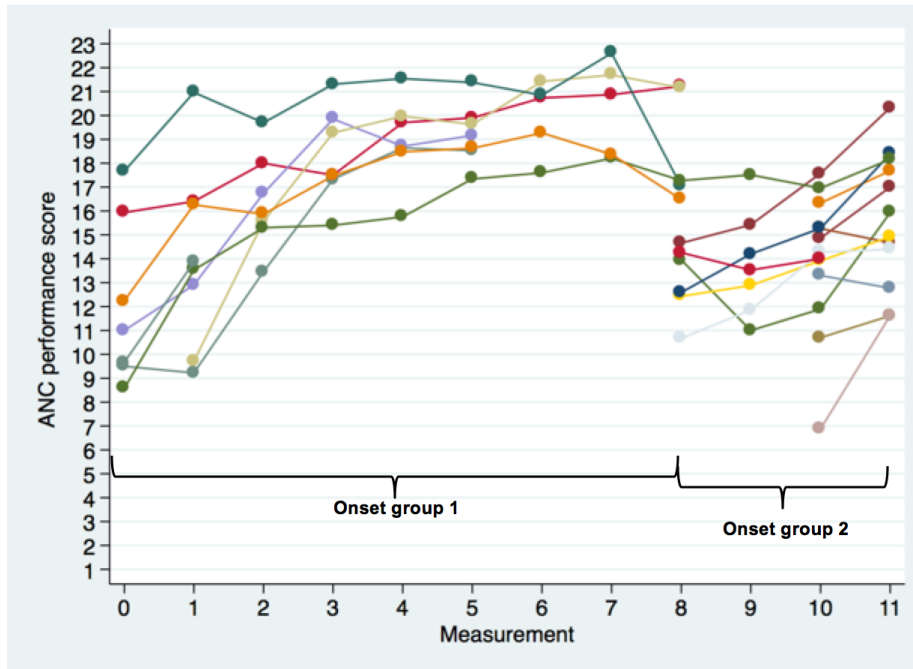


Figure 2: Health facility ANC performance scores over measurements 0-11. Each colored line in the figure represents an individual health facility. The two implementation onset groups are visible using the labels on the figure; the first onset group covers implementation from measurement 0-8, and the second onset group covers implementation from measurement 8-11. The highest possible ANC performance score is 23. ANC performance scores among all health facilities ranged from a minimum of 6.9 to a maximum of 22.6.

2. Type of health facility

Implementation health facilities are grouped into 3 main categories based on funding source: public, private, or mixed. Data on health facility type were obtained from the routine Sinergias database for *Mil Días*. Health facilities included in this study were either public or private.

3. Unsatisfied basic needs

Unsatisfied basic needs (UBN) is a proxy measure of the percentage of households living in poverty. Municipality UBN data were obtained from the 2005 Colombian Population and Data Census database, the last census published by the National Administrative Department of Statistics (DANE) of the Colombian government. Indicators used to calculate this UBN measure include inadequate housing, housing with critical overcrowding, housing with inadequate services for sanitation, households with a high number of economic dependents, and households with school-aged children who are not enrolled in school. Households are classified as poor if they do not meet a specified threshold for these indicators, and the percentage of poor households in each municipality was used in this analysis.^{5,25}

4. Number of health facilities per 100,000 inhabitants

Data on the number of health facilities per municipality were obtained from Special Registry of Health Care Services (REPS). Municipality population projections for 2012, the first year of *Mil Días* implementation, based on the 2005 Colombian Population and Data Census, were obtained from the DANE database. The number of health facilities in each municipality was calculated per 100,000 inhabitants per municipality, using 2012 population projections.

5. Patient Volume

The number of medical records reviewed at each measurement time point of *Mil Días* corresponds to a statistically representative sample of the total number of patients seen for ANC in the month previous to the measurement time point. Therefore, the number of observations for each measurement from the Sinergias database for *Mil Días* was used to represent the average patient volume at each health facility. The average number of medical records reviewed at each health facility over the first 8 measurements for the first onset group and last 4 measurements for the second onset group was used to calculate a single value for patient volume. Total patient volume was then dichotomized into a binary categorical variable, with high volume facilities having an average of 25 or more medical charts reviewed at each measurement (low volume facilities had an average of 24 or less medical charts reviewed at each measurement).

Statistical Analysis

A total of 20 health facilities were included in the overall analysis. Due to the differences in the timeframes of program implementation, separate analyses were conducted for health facilities that began enrollment before measurement 8 and those that began enrollment after measurement 8. A total of 8 facilities were included in the first enrollment group (n=8) and 12 facilities were included in the second enrollment group (n=12). The included facilities are located in 12 different municipalities in the department of Bolívar. During the first and second enrollment periods, a total of 1,224 and 1,799 medical records were used to calculate ANC performance scores, respectively. Because the purpose of this study was to test possible effects on the rate of QI, time represented by the number of measurements during *Mil Días* implementation was log-transformed for the analyses using measurements 0-8 and a quadratic term for time was used in analyses using measurements 8-11.

Multivariate linear regression analysis was used to examine relationships between predictors of interest and ANC performance over time using Stata 14 (StataCorp. 2015. Stata: Release 14. Statistical Software. College Station, TX: StataCorp LP). Two separate analyses were carried out: the first examined relationships between predictors and ANC performance over measurements 0-8, and the second examined relationships between predictors and ANC performance over measurements 8-11 of *Mil Días*. Random intercepts linear regression (RILR) and fixed-effects regression models were used to examine the effect of time on ANC performance score. Finally, because the purpose of this study was to test possible effects on the rate of QI, RILR and fixed-effects regression analyses were performed to test for interactions between covariates of interest and time. RILRs and fixed-effects regression models are appropriate for this study design, as the outcome of interest is comprised of non-independent observations, i.e., ANC performance scores are abstracted from observations of the same health facilities over time.

Ethical Considerations

This study did not require Institutional Review Board approval or exemption status because it is an evaluation of an existing approved national intervention.

III. Results

The study sample contained a total of 30 health facilities: 8 in the first onset group located in 3 municipalities and 12 in the second onset group located in 9 municipalities. Among the first onset group, a total of 6 health facilities were public and 2 were private. The average number of health facilities per 100,000 municipality inhabitants was 9.7, with a range of 4.7 to 17.8. The average UBN, reported as % of municipality households, was 59.1%. Finally, values for patient volume per health facility ranged from 8 to 35, with 4 health facilities classified as high volume and 4 classified as low volume (see Table 2). In the second onset group, all health facilities were public. The average number of health facilities per 100,000 municipality inhabitants and UBN were 10.8 and 67.5%, respectively. Patient volume was the only continuous variable of interest for which the second onset group had a lower average value than that for the first onset group (36.0 versus 26.1), with 9 facilities classified as high volume and only 1 facility classified as low volume (see Table 2).

Table 2: Descriptive statistics for predictors of interest. Municipal values, health facility values, and summary statistics are provided for each predictor of interest: number of health facilities per 100,000 municipality inhabitants, unsatisfied basic needs, health facility type, and average patient volume. The first onset group contains health facilities enrolled over measurements 0-8 of *Mil Días* implementation, while the second onset group contains health facilities enrolled over measurements 8-11 of *Mil Días* implementation.

Onset Group 1							
Municipality	Population	Total HF ¹	HF ¹ /100,000 hab. ²	UBN ³ (%)	HF ¹ Name	HF ¹ Type (Private)	Patient Volume
Arjona	68,583	5	7.3	61.4	Hospital Local de Arjona	-	43.4*
Magangué	123,312	22	17.8	49.2	Salud Especial	1	8.0
					CS Buenos Aires	-	17.0
					CS Cascajal	-	21.1
					CS Pastrana	-	23.0
					CS San Pablo	-	31.3*
					Corporación IPS Costa Atlántica	1	35.0*
Mahates	24,838	1	4.0	66.7	Hospital Local de Mahates	-	29.7*
Minimum	24,838	1	4.0	49.2			8.0
Maximum	123,312	22	17.8	66.7			35.0
Average	72,244	9	9.7	59.1			26.1
Onset Group 2							
Municipality	Population	Total HF ¹	HF ¹ /100,000 hab. ²	UBN ³ (%)	HF ¹ Name	HF ¹ Type (Private)	Patient Volume
Carmen de Bolívar	74,297	10	13.5	96.2	ESE Hospital Nuestra Señora del Carmen	-	35.5*
Cartagena	990,179	84	8.5	25.5	Centro de Salud Giovanni Cristini	-	41.5*
					CAP Pozón	-	49.5*
					CAP Arroz Barato	-	37.5*
					CAP Canapote	-	42.5*
El Guamo	7,763	1	12.9	63.9	CAB El Guamo	-	21.5
Maria la Baja	47,749	3	6.3	61.6	Hospital Local María la Baja	-	42.0*
San Estanislao	16,149	2	12.4	59.2	Hospital Local Ana María Rodríguez	-	28.7*
Santa Rosa	22,095	1	4.5	99.3	ESE Hospital Local Santa Rosa	-	36.3*
Turbaco	71,173	11	15.5	49.9	Hospital Local Turbaco	-	39.0*
Turbana	14,733	2	13.6	64.0	ESE Hospital Local Turbana	-	32.5*
Villa Nueva	19,458	2	10.3	88.3	UOL José Isabel Villa Real	-	25.3*
Minimum	7,763	1	4.5	25.5			21.5
Maximum	990,179	84	15.5	99.3			49.5
Average	140,400	13	10.8	67.5			36.0

¹Health facility

³Inhabitants

²Unsatisfied basic needs

*High volume facility

During *Mil Días* implementation, overall quality of ANC measured by program indicators increased (see Figure 3). Across all measurements, ANC performance scores for health facilities averaged 16.1 and ranged from 6.9 to 22.6, out of a possible maximum of 23 (see Table 3).

Table 3: Descriptive statistics for health facility ANC performance scores. Summary statistics for the outcome of interest, ANC performance scores, were calculated across all measurements (0-11) of *Mil Días* implementation. ANC performance scores reflect the average number of indicators completed, out of a total possible 23, during an ANC visit and are used to measure quality of care under *Mil Días* evaluation.

ANC performance score statistic	Minimum	Maximum	Mean	Standard deviation
Value	6.9	22.6	16.1	3.5

Bivariate RILR regression examined the rate of QI over time. Findings were that there was a significant change in ANC performance score across program trimesters for facilities enrolled in *Mil Días* during measurements 0-8 and those enrolled during measurements 8-11 ($p < 0.05$) (Tables 4 and A.2, crude models). Over measurements 0-8 of *Mil Días*, the rate of improvement was logarithmic, in which QI increased rapidly at first, and then proceeded to increase at an exponentially slower rate over time (Figure 3 and Table 4). Specifically, the rate of improvement among these facilities began with a performance score at measurement 1 that was 4.1 (95% CI 3.3-4.9) times greater than the score at measurement 0 (Table 4, crude model). This trend differed from the pattern of improvement among health facilities in the second implementation onset group. Among these facilities, the rate of QI measured by ANC performance exhibits a quadratic pattern, in that QI increased exponentially over time (Figure 3), starting with the ANC performance score at measurement 9 that was 0.4 (95% CI 0.0, 0.8) times greater than the score at measurement 8 (Table A.2).

Table 4: Random intercept linear regression (RILR) models using predictor and outcome of interest data over measurements 0-8. The outcome of interest for all models is ANC performance score, a measurement of quality of care. Interaction 1-4 models test predictors of interest (health facility type, unsatisfied basic needs, number of health facilities per 100,000 municipality inhabitants, and MNCH patient volume) for an effect on the rate of QI for ANC. Health facility type is the only predictor of interest shown to have an association with rate of QI, with public facilities having an increased rate of QI as compared to private health facilities [2.2 (95% CI 0.4-4.3)]. Relationships are exponential, with time log-transformed.

Coefficient	Crude	Adjusted	Interaction 1	Interaction 2	Interaction 3	Interaction 4
Intercept	11.8 (10.3 - 13.3)**	10.6 (9.2 - 11.9)**	9.8 (8.3 - 11.3)**	10.6 (9.2 - 11.9)**	10.6 (9.2 - 11.9)**	10.1 (8.6 - 11.7)**
Log time	4.1 (3.4 - 4.9)**	4.2 (3.5 - 5.0)**	4.7 (3.9 - 5.4)**	4.2 (3.5 - 5.0)**	4.2 (3.5 - 5.0)**	4.7 (3.6 - 5.9)**
HF¹ type						
Main (Private clinic)	-	2.5 (0.7 - 4.3)**	4.6 (2.0 - 7.2)**	2.5 (0.7 - 4.3)**	2.5 (0.7 - 4.3)**	2.6 (0.9 - 4.4)**
Interaction with log-time	-	-	-2.2 (-4.1 - -0.4)**	-	-	-
UBN²						
Main	-	2.1 (0.3 - 4.0)**	2.5 (0.4 - 4.5)**	2.2 (0.3 - 4.0)**	2.1 (0.3 - 4.0)**	2.0 (0.3 - 3.8)**
Interaction with log-time	-	-	-	0.0 (-0.1 - 0.1)	-	-
Number of HF¹ per 100,000 hab³						
Main	-	2.8 (0.5 - 5.0)**	3.2 (0.7 - 5.7)**	2.8 (0.5 - 5.0)**	2.8 (0.5 - 5.0)**	2.6 (0.5 - 4.8)**
Interaction with log-time	-	-	-	-	0.0 (-0.1 - 0.1)	-
Patient Volume						
Main (High volume)	-	1.0 (-0.6 - 2.5)	1.4 (-0.3 - 3.2)	1.0 (-0.6 - 2.6)	1.0 (-0.6 - 2.6)	1.9 (-0.4 - 4.1)
Interaction with log-time	-	-	-	-	-	-0.8 (-2.4 - 0.7)
ICC⁴	0.52	0.16	0.26	0.16	0.16	0.15

- Not included in model

* p<0.1

** p<0.05

¹Health facility

²Unsatisfied basic needs

³Inhabitants

⁴Intraclass Correlation Coefficient

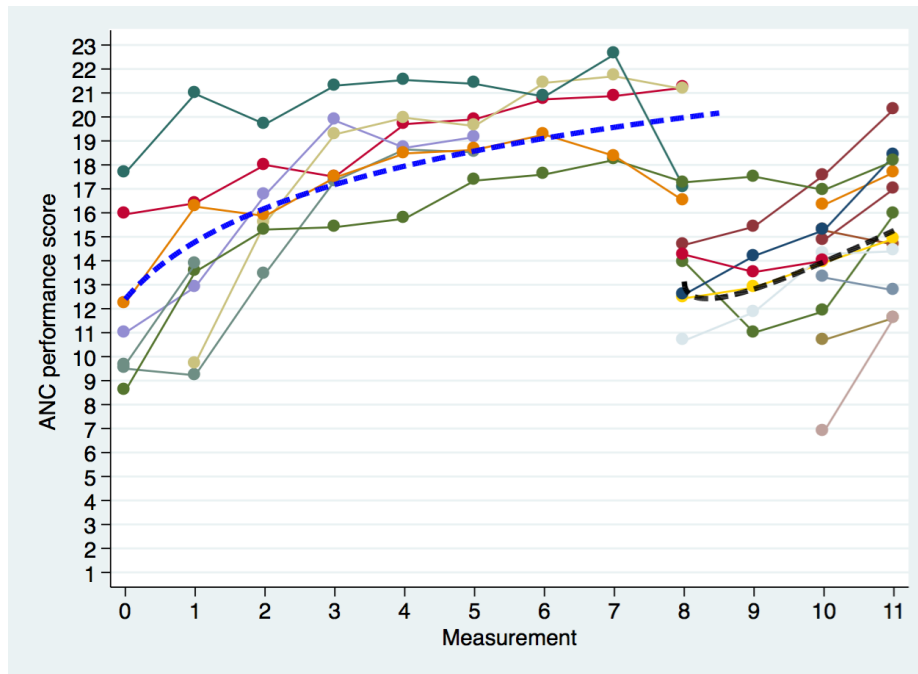


Figure 3: Rates of QI measured by ANC performance score among health facilities during measurements 0-8 and measurements 8-11. The rate of QI, measured by ANC performance score, shows a logarithmic pattern among health facilities among the first onset group, over measurements 0-8. With an overall increase in ANC performance score, the rate of improvement exponentially decreases with time. Among the second onset group, over measurements 8-11, the rate of QI shows a quadratic pattern. With an overall increase in ANC performance score, the rate of improvement exponentially increases with time.

For health facilities enrolled over measurements 0-8, the rate of QI remained similar when adjusted for covariates of interest (Table 4, adjusted model). Among this first enrollment group of health facilities, health facility type, UBN, and number of health facilities per 100,000 municipality inhabitants covariates were significantly associated with ANC performance score ($p < 0.05$) (see Table 4, adjusted model). For this group, facility type was the only covariate that exhibited a significant interaction with time to affect the rate of QI measured by ANC performance score ($p < 0.05$) (Table 4, interaction 1 model). Specifically, the rate of QI among private health facilities began as 2.2 (95% CI 0.4-4.3) times less than the rate of QI among public facilities, with the difference in QI rates between these facility types decreasing exponentially over time (Figure 4). Figure 4 also illustrates that this difference in QI rates is such that public health facilities started with a lower average ANC performance score than private health facilities at measurement 0, but ended with a higher average ANC performance score at measurement 8. Fixed-effects regression models produced similar results as RILR models (see Table A.1).

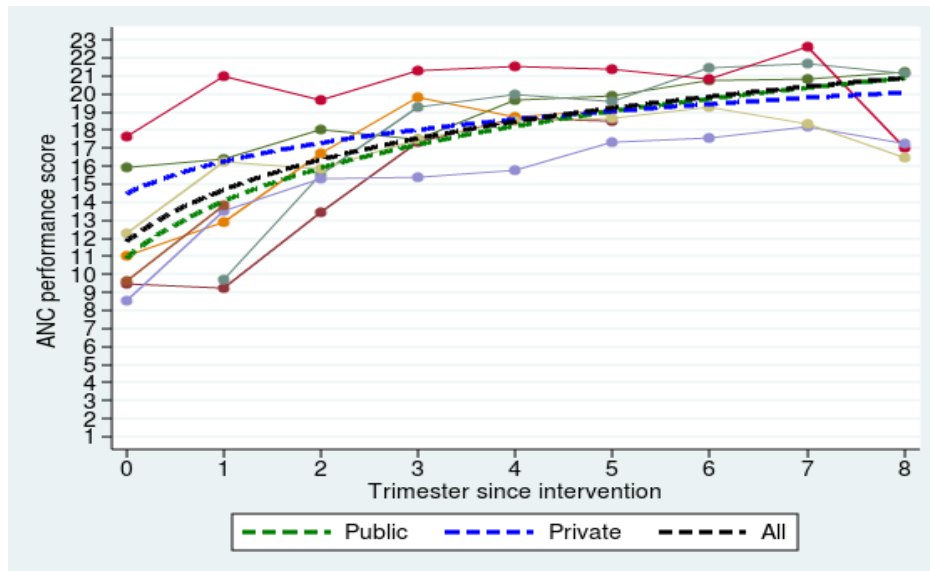


Figure 4: ANC performance score in public health facilities versus private health facilities during measurements 0-8. At baseline, or measurement 0, public health facilities on average have a lower ANC performance score than private health facilities on average. However, the rate of QI as measured by ANC performance score is greater among public facilities than among private facilities [2.2 (95% CI 0.4-4.3)]. The difference in these rates changes exponentially over time.

For health facilities enrolled over measurements 8-11, RILR analyses to test for an interaction between covariates of interest and time demonstrated that no covariate of interest had an effect on the rate of QI measured by ANC performance score (Table A.2). The same was true in fixed-effects multivariate models (Table A.3). Importantly, all of the health facilities in this group were public facilities.

IV. Discussion

This study sought to examine the influence of health system contextual factors on the QI of ANC within health facilities targeted by the *Mil Días* intervention in the department of Bolívar in Colombia. Overall, there was an increase in mean ANC performance scores among health facilities over time, demonstrating positive results during the course of program implementation for both facilities enrolled during measurements 0-8 and facilities enrolled during measurements 8-11. The major finding of the included analyses is that facility type is a determinant of the rate of QI among *Mil Días* health facilities in Bolívar, with public facilities exhibiting a greater rate of QI than private facilities. Interestingly, facility type may indicate a series of health facility characteristics that influence ability to implement *Mil Días* components and improve quality of care. Examples of such characteristics may include levels of financial and human resources and stocks of essential medicines and equipment. *Mil Días* may therefore play a role in overcoming challenges to quality MNCH services dictated by facility type, and serve to especially benefit public facilities.

It is important to note that there was a difference at measurement 0 in mean ANC performance score between facility type groups (Figure 4), which may indicate a baseline confounding of the relationship between facility type and quality of ANC. If this were the case, certain unmeasured characteristics that affect performance scores may differ significantly between public and private facilities. While the fixed-effects linear regression models may help to control this baseline

confounding by holding all other variables of interest in a fixed position, fixed-effects models are not as efficient as RILR models, as is evidenced by the larger confidence intervals obtained (Table A.1 and A.3). Nonetheless, it is important to note that both RILR and fixed-effects logistic regression models were consistent in yielding significant findings to support the relationship between facility type and rate of QI.

The statistical analyses conducted in this study failed to demonstrate a significant association between the rate of QI of ANC and other variables of interest, including UBN, number of health facilities per population unit, and patient volume. Still, results from bivariate linear regression analyses demonstrated that UBN and number of health facilities per population unit had a significant effect on ANC performance (without having a significant interaction with time).

Using these findings as a basis for future QI studies, facility type, UBN (or other economic measurements of household poverty) and number of health facilities per population unit may be valuable points of data collection to be included in indices for MNCH metrics. These conclusions are supported by other reports in the literature that likewise found that institution type,^{12,17,26,27} population characteristics and fiscal resources^{14–16} and level of competition among health care institutions^{16,27} had an effect on quality of care.

Limitations of the Study Design

Lack of other relevant contextual data was a major study limitation. Moreover, comparable data on a control group of health facilities was not available. Further limitations of the study design are a consequence of the small sample size, which was dictated by the number of included *Mil Días* intervention facilities.

Another data limitation was due to the structure of the national databases used, which were largely limited to departmental and municipality level data. In certain cases, health information systems at the facility level do indeed exist, but are underutilized by the majority of health facilities and/or are unreliable due to unclear and redundant organization, e.g., multiple ways of classifying an ANC visit. Indeed, data discrepancies in national databases were also problematic. As one example, there was imprecision in the measurement of patient volume, a predictor of interest in the present study. Of note, none of the conducted analyses demonstrated that patient volume was a significant predictor of QI of ANC. This may be due to the fact that the available data only allowed for the approximation of patient volume using the mean number of reviewed medical charts via routine evaluation by Sinergias.

Challenges with data availability and quality support the argument for methods to enable routine, baseline data collection across all levels of the Colombian health system. An index for simplified and standardized data collection is key to efforts to assess current care standards and interventions to improve care. A tool for data collection would also allow for consideration of additional contextual variables (see Figure 1) to be examined in relationship to QI strategies for MNCH as well as other health services. For example, disease outbreaks and armed conflict may have a large influence on quality of care in the Colombian context. During the course of *Mil Días* implementation, a Chikungunya outbreak affected health care providers and the prioritization of care in targeted health facilities. This outbreak, which peaked during measurements 7 and 8, may have correlated with a drop in ANC performance during this time period. Moreover, the presence of long-standing armed conflict in Colombia has resulted in profound consequences for the entire health system of Colombia that is difficult to quantify. Its aftermath will also be challenging.

This context has tremendous import for MNCH, influencing both access to and availability of care.

Implications for Future Studies

Despite these acknowledged limitations, this study adds to the knowledge base by describing various contextual factors and their effect on QI of ANC services delivered in health facilities targeted by *Mil Días*. Moving forward, the study design for program evaluation introduced here may be used in an ongoing way for measuring *Mil Días* effectiveness and informing implementation activities. In the current study, the analyses using health facilities over the measurement 8-11 enrollment period did not reveal any statistically significant results. This may have been due to the fact that there were a maximum of only 4 data points among these health facilities. Therefore, it will be important to perform a follow-up evaluation on these health facilities after more data time points are available. In addition, future execution of this program evaluation may be able to incorporate a greater number of contextual covariates depending upon changes in data availability. In Colombia, data were collected for the 2015 Population and Data Census. Once these data become available, they may provide an important update and extension of this evaluation. Notably, Sinergias has already begun to modify its data collection activities to include variables such as patient load per health care provider as part of routine impact evaluation for *Mil Días*.

Other adaptations of the study design may include expanding the outcomes of interest. For instance, mean ANC performance scores never reached the maximum value of 23 among intervention facilities in Bolívar, which would translate to an average of 100% completion of quality of care indicators. Investigating individual indicators may reveal additional effects on quality of care, such as the availability of essential materials or limited time resources of health care providers. Additionally, performance scores for other categories of QI indicators (delivery, newborn, and postpartum care; and early childhood growth and development) may be usefully examined, as may performance scores among health facilities in other implementation departments.

This program evaluation utilized a novel analysis that included contextual variables to better understand how and why *Mil Días* QI strategies improve ANC quality of care over time. The study design was intended to evaluate *Mil Días* results as a function of clinical factors as well as the broader context in which health care facilities are situated, which is essential to analyze the real world effectiveness of QI approaches.²⁸ By doing so, QI program evaluations such as this one may contribute to guiding further implementation of *Mil Días* and adaptations of its activities.¹² Future QI program evaluations that examine the intervention context are necessary to overcome implementation challenges in improving MNCH services and to sustain QI successes.²³

V. Resources

1. United Nations. *The Millennium Development Goals Report.*; 2015. doi:978-92-1-101320-7.
2. Osborn D, Cutter A, Ullah F. *Universal Sustainable Development Goals: Understanding the Transformational Challenge for Developed Countries.*; 2015.
3. SINERGIAS - Alianzas Estratégicas para la Salud y el Desarrollo Social. *Mejoramiento de La Calidad de La Atención a Gestantes, Niñas Y Niños En Sus Primeros Mil Días Mediante: Fortalecimiento de Las Capacidades Locales Del Sector Salud - Informe Final Fase III.* Bogotá, Colombia; 2015. doi:10.1017/CBO9781107415324.004.
4. Ministerio De Salud Y Protección Social. *Análisis de Situación de Salud Según Regiones Colombia.*; 2013.
5. United Nations Population Fund. *Objetivos de Desarrollo Del Milenio Colombia: Herramientas de Aproximación Al Contexto Local.* Colombia; 2016.
6. SINERGIAS - Alianzas Estratégicas para la Salud y el Desarrollo Social. *Mejoramiento de La Calidad de La Atención a Gestantes, Niñas Y Niños En Sus Primeros Mil Días Mediante El Fortalecimiento de Las Capacidades Locales Del Sector Salud - Informe Final Fase IV.* Bogotá, Colombia; 2015.
7. Hoddinott J, Behrman JR, Maluccio JA, et al. Adult consequences of growth failure in early childhood. *Am J Clin Nutr.* 2013;98(5):1170-1178. doi:10.3945/ajcn.113.064584.
8. Bhutta ZA, Das JK, Bahl R, et al. Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? *Lancet.* 2014;384(9940):347-370. doi:10.1016/S0140-6736(14)60792-3.
9. World Health Organization. *Integrated Management of Pregnancy and Childbirth: WHO Recommended Interventions for Improving Maternal and Newborn Health.* Geneva, Switzerland; 2009. doi:WHO/MPS/07.05.
10. Lassi ZS, Middleton PF, Crowther C, Bhutta ZA. Interventions to Improve Neonatal Health and Later Survival: An Overview of Systematic Reviews. *EBioMedicine.* 2015;2(8):985-1000. doi:10.1016/j.ebiom.2015.05.023.
11. Shortell SM, Bennett CL, Byck GR. Assessing the impact of continuous quality improvement on clinical practice: what it will take to accelerate progress. *Milbank Q.* 1998;76(4):593-624, 510. doi:10.1111/1468-0009.00107.
12. Ouslander JG, Lamb G, Tappen R, et al. Interventions to reduce hospitalizations from nursing homes: Evaluation of the INTERACT II collaborative quality improvement project. *J Am Geriatr Soc.* 2011;59(4):745-753. doi:10.1111/j.1532-5415.2011.03333.x.
13. Walshe K. Understanding what works-and why-in quality improvement: The need for theory-driven evaluation. *Int J Qual Heal Care.* 2007;19(2):57-59. doi:10.1093/intqhc/mzm004.
14. Davidoff F. Heterogeneity Is Not Always Noise: Lessons From Improvement. *Jama.* 2009;302(23):2580-2586. doi:10.1001/jama.2009.1845.
15. Odaga J, Henriksson DK, Nkolo C, et al. Empowering districts to target priorities for improving child health service in Uganda using change management and rapid assessment methods. *Glob Health Action.* 2016;9. doi:10.3402/gha.v9.30983.
16. Hanson C, Waiswa P, Marchant T, et al. Expanded Quality Management Using Information Power (EQUIP): protocol for a quasi-experimental study to improve maternal and newborn health in Tanzania and Uganda. *Implement Sci.* 2014;9(1):41. doi:10.1186/1748-5908-9-41.

17. Hanson C, Ronsmans C, Penfold S, et al. Health system support for childbirth care in Southern Tanzania: results from a health facility census. *BMC Res Notes*. 2013;6(1):435. doi:10.1186/1756-0500-6-435.
18. Sistema Integral de Información (SISPRO) M de S y PS. Observatorio de Calidad de la Atención en Salud. <http://calidadensalud.minsalud.gov.co/Observatorio/IndicadoresR144606/Comparaci%C3%B3ndeIPS.aspx>. Published 2014. Accessed May 8, 2016.
19. Tripathi V, Stanton C, Strobino D, Bartlett L. Development and validation of an index to measure the quality of facility-based labor and delivery care processes in sub-Saharan Africa. *PLoS One*. 2015;10(6):1-29. doi:10.1371/journal.pone.0129491.
20. Ovreteit J, Gustafson D. Evaluation of quality improvement programmes. *Qual Saf Health Care*. 2002;11(3):270-275. doi:10.1136/qhc.11.3.270.
21. Walshe K, Freeman T. Effectiveness of quality improvement: learning from evaluations. *Qual Saf Health Care*. 2002;11(1):85-87. doi:10.1136/qhc.11.1.85.
22. Dixon-Woods M, Bosk CL, Aveling EL, Goeschel CA, Pronovost PJ. Milbank quarterly. *Milbank Q*. 2011;89(2):167-205. doi:10.1111/j.1468-0009.2010.00608.x.
23. Kinney M V., Cocoman O, Dickson KE, et al. Implementation of the every Newborn Action Plan: Progress and lessons learned. *Semin Perinatol*. 2015;39(5):326-337. doi:10.1053/j.semperi.2015.06.004.
24. *Diagnóstico Departamental Bolívar*. Bolívar, Colombia.
25. DANE: Sistema Estadístico Nacional. UBN (Unsatisfied Basic Needs). <http://www.dane.gov.co/index.php/quality-of-life/unsatisfied-basic-needs>. Accessed June 24, 2016.
26. Hanson C e. A. Expanded Quality Management Using Information Power (EQUIP): protocol for a quasi-experimental study to improve maternal and newborn health in Tanzania and Uganda. *Implement Sci*. 2014;9(1):41. doi:10.1186/1748-5908-9-41.
27. García FM. Los determinantes de la calidad de la atención en salud: el caso del control prenatal. *Coyunt Económica*. 2009;XXXIX(2):61-97. http://www.fedesarrollo.org.co/wp-content/uploads/2012/03/Los-determinantes-de-la-calidad-de-la-atención-en-salud_García-F-M-C.E.-2do-semester-2009.pdf.
28. Shojania KG, Grimshaw JM. Evidence-based quality improvement: The state of the science. *Health Aff*. 2005;24(1):138-150. doi:10.1377/hlthaff.24.1.138.

VI. Appendix

Table A.1: Fixed-effects regressions using predictor and outcome of interest data over measurements 0-8. The outcome of interest for all models is ANC performance score, a measurement of quality of care. Interaction 1-4 models test predictors of interest (health facility type, unsatisfied basic needs, number of health facilities per 100,000 municipality inhabitants, and MNCH patient volume) for an effect on the rate of QI for ANC. Health facility type is the only predictor of interest shown to have an association with rate of QI, with public facilities having an increased rate of QI as compared to private health facilities [2.7 (95% CI 0.7-4.7)]. Relationships are exponential, with time log-transformed. Time-invariant, i.e. all non-interaction coefficients of predictors of interest are automatically dropped from fixed-effects regressions.

Coefficient	Crude	Adjusted	Interaction 1	Interaction 2	Interaction 3	Interaction 4
Intercept	12.0 (11.0 – 13.1)**	11.7 (2.3 – 21.1)**	9.2 (0.3 – 18.2)**	12.8 (1.9 – 23.6)**	12.8 (1.6 – 23.9)**	13.7 (3.3 – 24.1)**
Log time	4.1 (3.3 – 4.9)**	4.1 (3.0 – 5.2)**	4.8 (3.6 – 5.9)**	4.0 (2.9 – 5.2)**	4.0 (2.8 – 5.2)**	4.4 (3.1 – 5.7)**
HF¹ type						
Main (Private clinic)	-	~	~	~	~	~
Interaction with log-time	-	-	-2.7 (-4.7 – -0.7)**	-	-	-
UBN²						
Main	-	~	~	~	~	~
Interaction with log-time	-	-	-	0 (-0.2 – 0.1)	-	-
Number of HF¹ per 100,000 hab³						
Main	-	~	~	~	~	~
Interaction with log-time	-	-	-	-	0 (-0.1 – 0.2)	-
Patient Volume						
Main (High volume)	-	~	~	~	~	~
Interaction with log-time	-	-	-	-	-	-0.8 (-2.6 – 1.0)

- Not included in model

~ Time-invariant term dropped from model

* p<0.1

** p<0.05

¹Health facility

²Unsatisfied basic needs

³Inhabitants

Table A.2: Random intercept linear regressions (RILRs) using predictor and outcome of interest data over measurements 8-11. The outcome of interest for all models is ANC performance score, a measurement of quality of care. Interaction 1-3 models test predictors of interest (unsatisfied basic needs, number of health facilities per 100,000 municipality inhabitants, and MNCH patient volume) for an effect on the rate of QI for ANC. Relationships are exponential, with a quadratic term included for time.

Coefficient	Crude	Adjusted	Interaction 1	Interaction 2	Interaction 3
Intercept	12.1 (10.5 – 13.8)**	12.4 (8.4 – 16.4)**	12.2 (8.2 – 16.2)**	12.4 (8.4 – 16.4)**	13.0 (8.8 – 17.3)**
Time					
Linear term	-0.1 (-1.4 – 1.2)	-0.1 (-1.4 – 1.2)	0.1 (-1.2 – 1.3)	-0.1 (-1.4 – 1.2)	-0.5 (-2.0 – 1.0)
Quadratic term	0.4 (0.0 – 0.8)**	0.4 (0.0 – 0.8)**	0.4 (-0.0 – 0.7)*	0.4 (0.0 – 0.8)**	0.4 (0.0 – 0.8)**
UBN¹					
Main	-	0 (-0.1 – 0.0)*	0 (-0.1 – 0.1)	0 (-0.1 – 0.0)*	0 (-0.1 – 0.0)
Interaction with log-time	-	-	0 (0.0 – 0.0)		-
Number of HF² per 100,000 hab³					
Main	-	-0.2 (-0.6 – 0.2)	-0.2 (-0.5 – 0.2)	-0.2 (-0.6 – 0.2)	-0.2 (-0.6 – 0.2)
Interaction with log-time	-	-	-	0 (-0.1 – 0.1)	
Patient Volume					
Main (High volume)	-	-0.2 (-4.3 – 3.9)	-0.3 (-4.4 – 3.9)	-0.3 (-4.4 – 3.9)	-1.0 (-5.6 – 3.5)
Interaction with log-time	-	-	-	-	0.5 (-0.6 – 1.6)
ICC⁴	0.79	0.72	0.75	0.73	0.74

- Not included in model

* p<0.1

** p<0.05

¹Unsatisfied basic needs

²Health facility

³Inhabitants

⁴Intraclass Correlation Coefficient

Table A.2: Fixed-effects regressions using predictor and outcome of interest data over measurements 8-11. The outcome of interest for all models is ANC performance score, a measurement of quality of care. Interaction 1-3 models test predictors of interest (unsatisfied basic needs, number of health facilities per 100,000 municipality inhabitants, and MNCH patient volume) for an effect on the rate of QI for ANC. Relationships are exponential, with a quadratic term included for time. Time-invariant, i.e. all non-interaction coefficients of predictors of interest, are automatically dropped from fixed-effects regressions.

Coefficient	Crude	Adjusted	Interaction 1	Interaction 2	Interaction 3
Intercept	12.1 (10.5 – 13.8)**	12.3 (11.3 – 13.4)**	12.1 (11.0 – 13.2)**	12.3 (11.2 – 13.4)**	12.2 (11.1 – 13.3)**
Time					
Linear term	-0.1 (-1.4 – 1.2)	0 (-1.4 – 1.4)	0.2 (-1.3 – 1.6)	0 (-1.5 – 1.4)	-0.5 (-2.2 – 1.2)
Quadratic term	0.4 (0.0 – 0.8)**	0.4 (0.0 – 0.8)*	0.4 (-0.1 – 0.8)*	0.4 (0.0 – 0.8)*	0.4 (0.0 – 0.8)*
UBN¹					
Main	-	~	~ 0	~	~
Interaction with log-time	-	-	(0.0 – 0.0)	-	-
Number of HF² per 100,000 hab³					
Main	-	~	~	~ 0	~
Interaction with log-time	-	-	-	(-0.1 – 0.1)	-
Patient Volume					
Main (High volume)	-	~	~	~	~ 0.6
Interaction with log-time	-	-	-	-	(-0.6 – 1.8)

- Not included in model

~ Time-invariant term dropped from model

* p<0.1

** p<0.05

¹Unsatisfied basic needs

²Health facility

³Inhabitants