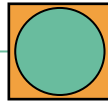


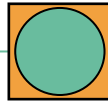
Video Tutorial: Interrupted Time Series

Data for Impact



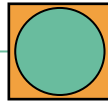
Overview

- Part 1:
 - Interrupted time series overview
 - Important considerations
 - Preparing the dataset
- Part 2 & 3:
 - Interrupted time series analysis
 - Single group
 - Multiple group



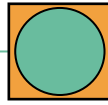
D4I works to:

- Generate strong evidence for program and policy decision making
- Build individual and organizational capacity
- Enhance the use of data for global health programs and policies



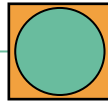
Video tutorial objective

- The team was tasked with creating video tutorials on how to use specific methods and commands to analyze routine data for evaluation and research.
 - This video is not intended for beginners but for those who have some experience with data analysis and STATA
- This tutorial focuses on interrupted time series and is part of a part of a series that outlines the use of two statistical methods in the evaluation of routine data
- For information on different methods used to analyze routine data, as well as the materials used in this presentation see <https://www.data4impactproject.org/>



Interrupted time series (ITS)

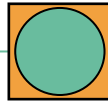
- Compares the level and mean trend in outcome indicators before and after a breakpoint.
- Data collected at multiple points before and after an intervention to determine the impact of the intervention.
- Identifies
 - level change (immediate effect)
 - slope change (sustained effect)
 - or both the level and slope change



Interrupted time series

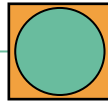
It can be used to

- Compare pre event/policy/interruption to the post
- Compare trends before and after a policy change/event
- Evaluate trends
- Evaluate variance



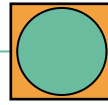
Benefits of ITS

- ITS is one of the strongest quasi experimental study designs when a randomized control trial is not possible
- It detects changes that are delayed or intermittent.
- Retrospective data can be used
- Controls for secular trends
- Controls for long term trends in the data



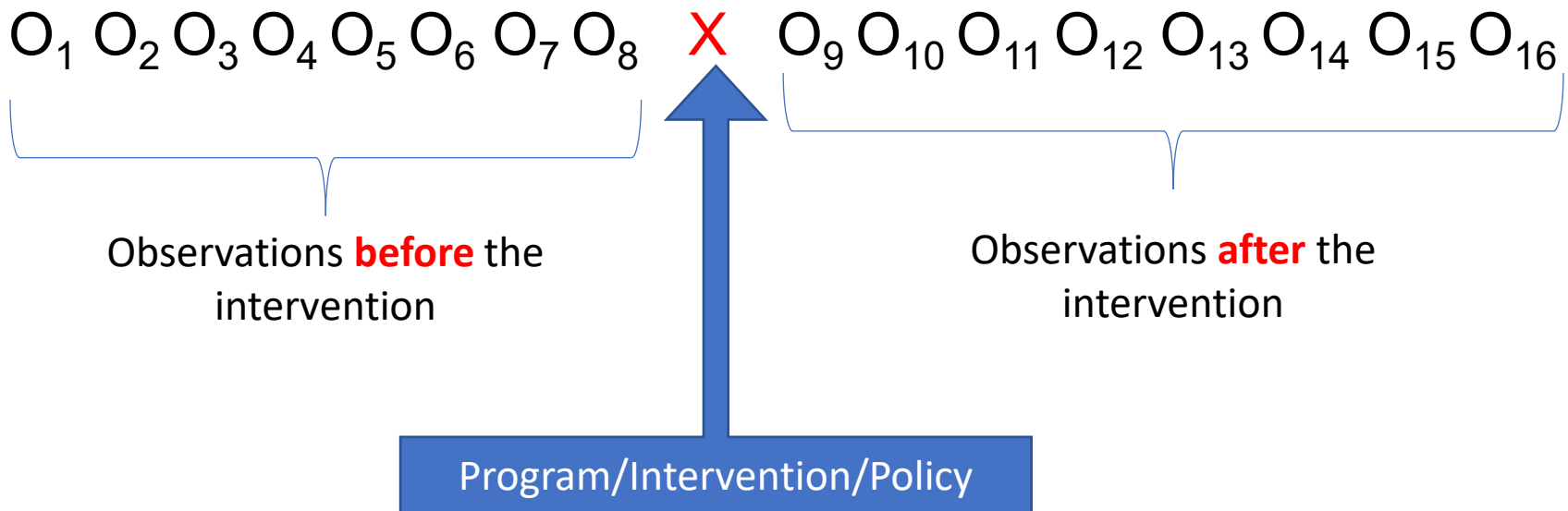
Limitations of ITS

- Multiple data entries needed
 - Minimum of 8 periods before and 8 after
- Difficulty in analysis when separate components
 - Difficult to differentiate the effects of several components if they coincide or are implemented close together in time
- Requires population level data
 - Inferences can't be made at the individual level



ITS Study design

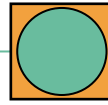
Single Group ITS: Without a control group



Legend

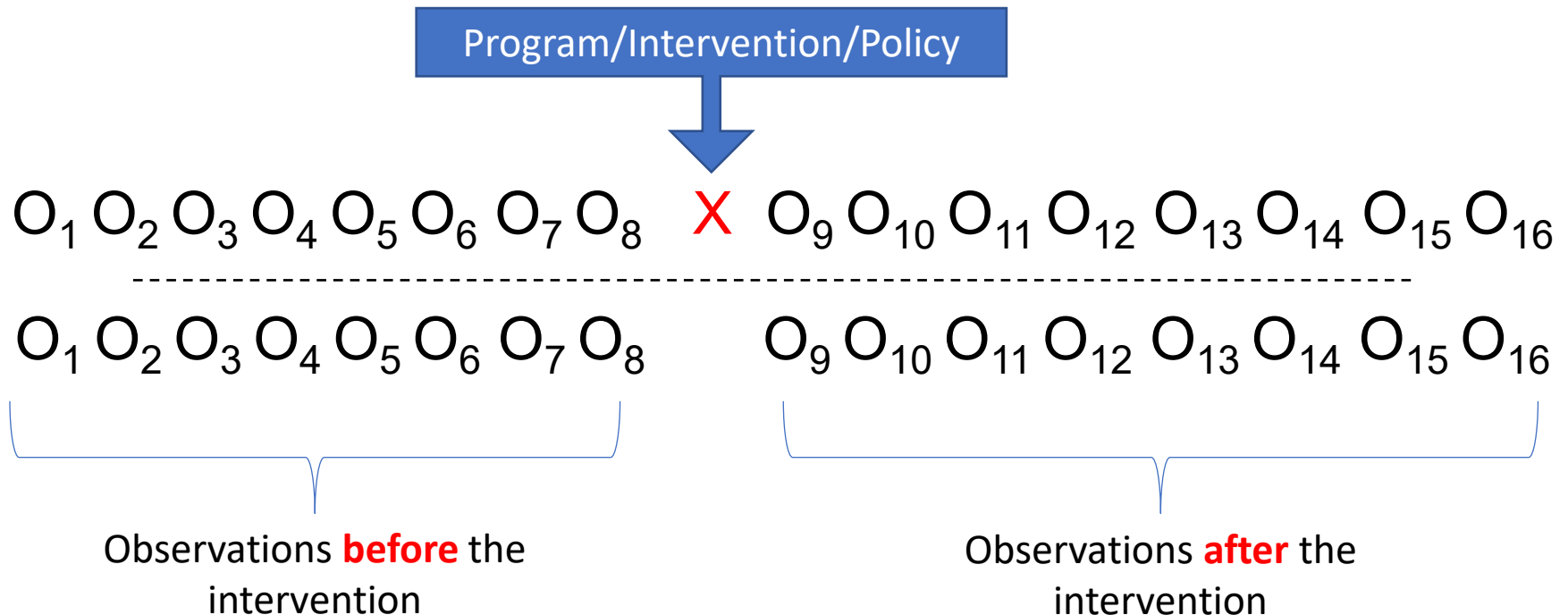
O - Observation

X – Program/Intervention/Policy



ITS Study design

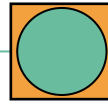
Multiple Group ITS: With a control group



Legend

O - Observation

X – Program/Intervention/Policy



Examples of studies using ITS

Use of Routine Health Information System Data to Evaluate Impact of Malaria Control Interventions in Zanzibar, Tanzania from 2000 to 2015

Ruth A. Ashton ^{a,*,1}, Adam Bennett ^{a,b,1}, Abdul-Wahid Al-Mafazy ^c, Ali K. Abass ^{c,2}, Mwinyi I. Msellem ^d, Peter McElroy ^e, S. Patrick Kachur ^{f,g}, Abdullah S. Ali ^c, Joshua Yukich ^a, Thomas P. Eisele ^a, Achuyt Bhattarai ^e

^a Center for Applied Malaria Research and Evaluation, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, USA

^b Malaria Elimination Initiative, Global Health Group, University of California, San Francisco, CA, USA

^c Zanzibar Malaria Elimination Programme, Ministry of Health, Zanzibar, United Republic of Tanzania

^d Mnazi Mmoja Hospital, Zanzibar, United Republic of Tanzania

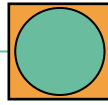
^e U.S. President's Malaria Initiative, Malaria Branch, U.S. Centers for Disease Control and Prevention, Atlanta, GA, USA

^f Malaria Branch, U.S. Centers for Disease Control and Prevention, Atlanta, GA, USA

^g Mailman School of Public Health, Columbia University, NY, New York, USA

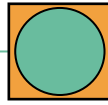
Impact of a free care policy on the utilisation of health services during an Ebola outbreak in the Democratic Republic of Congo: an interrupted time-series analysis

Yuen W Hung,¹ Michael R Law ,² Lucy Cheng,² Sharon Abramowitz ,³ Lys Alcayna-Stevens ,⁴ Grégoire Lurton ,⁵ Serge Manitu Mayaka,⁶ Romain Olekhnovitch ,⁵ Gabriel Kyomba ,⁶ Hinda Ruton ,^{2,7} Sylvain Yuma Ramazani,⁸ Karen A Grépin ,^{1,9}



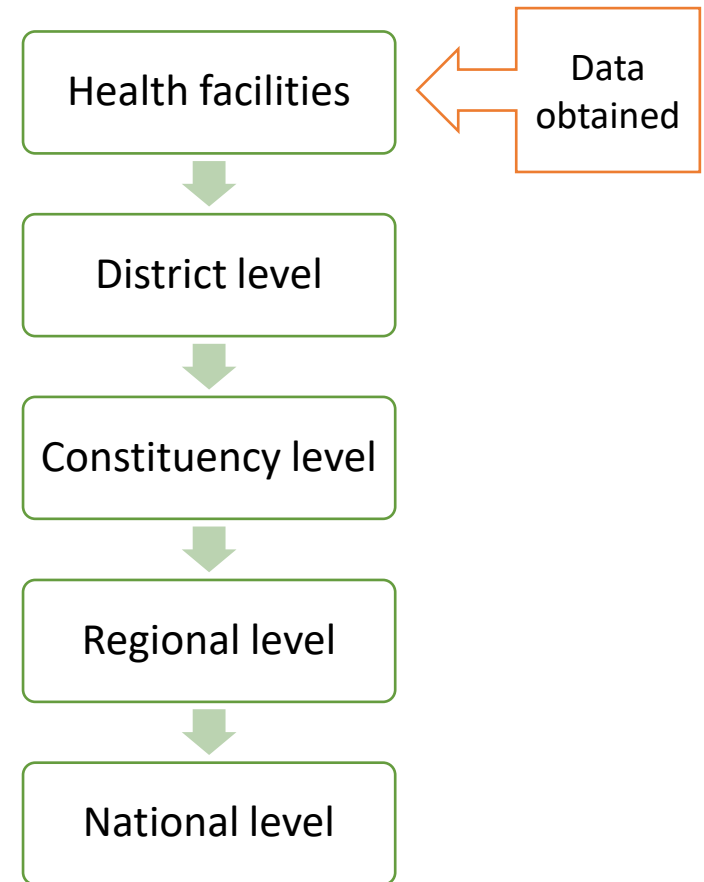
Steps needed to use ITS

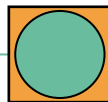
1. Determine time period(s)
2. Select the analytic cohort, if applicable
3. Determine outcomes of interest
4. Setup the data and visually inspect data
5. Preliminary analysis
6. Address autocorrelation, if applicable
7. Run final model
8. Plot results
9. Predict relative and absolute effects



Tutorial Example

- Sample data is from a fictional country called Exile.
- In November 2017, Exile introduced a policy that eliminated fees at health facilities associated with maternity care services, with the aim of increasing utilization of health services.
- It was piloted in two constituencies, Gold-Coast IX and Gold-Coast XV. The policy was not piloted in North Troy XXVI.
- Dataset: Health facility data on the number of women who attended their 4th antenatal care (ANC) visits from 2017 to 2020





Tutorial Example

Downloaded

Health
facility data

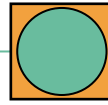


Analysis

Constituency
level data

Research Questions

1. What was the impact of the policy on the utilization of ANC, specifically the 4th ANC visit of pregnant women, in Gold Coast IX?
2. What was the impact of the policy on the utilization of ANC, specifically the 4th ANC visit of pregnant women, in the intervention constituency, Gold Coast IX, compared to the control constituency, North Troy XXVI?



Downloaded dataset

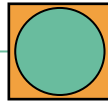
Region	Constituency	District	Facility	ANC4Jan2017	ANC4Feb2017	ANC4Mar2017
Gold Coast Region	Gold-C. I	GC1 Baku	GC1 Baku Health Center	6	23	36
Gold Coast Region	Gold-C. I	GC1 Baku	GC1 Baku Referral Facility	11	10	
Gold Coast Region	Gold-C. I	GC1 Aire	GC1 Aire Health Center	22	31	30
Gold Coast Region	Gold-C. I	GC1 King	GC1 King Health Center	28	30	31
Gold Coast Region	Gold-C. I	GC1 King	GC1 King Health Post			
Gold Coast Region	Gold-C. I	GC1 Orleans	GC1 Orleans Health Center	20	23	20
Gold Coast Region	Gold-C. I	GC1 Jefferson	GC1 Jefferson Health Center	10	18	18
Gold Coast Region	Gold-C. I	GC1 Jefferson	GC1 Jefferson Health Post	15	17	19
Gold Coast Region	Gold-C. I	GC1 Zion	GC1 Zion Health Center			
Gold Coast Region	Gold-C. I	GC1 Zion	GC1 Zion Health Post	18	24	23
Gold Coast Region	Gold-C. I	GC1 Hectar	GC1 Hectar Health Post			
Gold Coast Region	Gold-C. I	GC1 Hectar	GC1 Hectar Referral Center	22	19	24
Gold Coast Region	Gold-C. I	GC1 Gide	GC1 Gide Health Center	20	15	14
Gold Coast Region	Gold-C. I	GC1 Water	GC1 Water Health Center			
Gold Coast Region	Gold-C. I	GC1 Nectar	GC1 Nectar Health Center	20	23	25
Gold Coast Region	Gold-C. I	GC1 Nine	GC1 Nine Health Post	5	16	16
Gold Coast Region	Gold-C. I	GC1 Nine	GC1 Nine Health Center		23	23
Gold Coast Region	Gold-C. I	GC1 Mane	GC1 Mane Health Center			
Gold Coast Region	Gold-C. I	GC1 June	GC1 June Health Center	10	22	20
Gold Coast Region	Gold-C. I	GC1 June	GC1 June 1 Health Post			
Gold Coast Region	Gold-C. I	GC1 June	GC1 June 2 Health Post			

- **Variable names must be consistent**

Why?

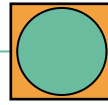
Merging files, if analyzing more than one outcome

Reshaping files



Routine Data Preparation

1. Assess missing data
2. Aggregate data
3. Reshape data
4. Date format



Routine Data Preparation: Data aggregation

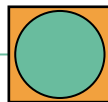
- Downloaded data may be at the health facility level.
 - If it is, aggregate the data to the level of interest.
- WAIT
 - Are there any duplicate health facilities?
 - **STATA Command:** `duplicates tag ***var list***, gen(newvar)`

STATA command: `duplicates tag facility, gen(dup1)`

dup1	Freq.	Percent	Cum.
0	17,492	93.53	93.53
1	952	5.09	98.62
2	108	0.58	99.19
3	80	0.43	99.62
4	25	0.13	99.75
5	30	0.16	99.91
6	7	0.04	99.95
8	9	0.05	100.00
Total	18,703	100.00	

STATA command: `duplicates tag regional constituency district facility, gen(duphf)`

duphf	Freq.	Percent	Cum.
0	18,703	100.00	100.00
Total	18,703	100.00	



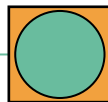
Routine Data Preparation: Data aggregation

- Sample dataset: aggregate at constituency level
 - Total number of 4th ANC visits in each constituency

STATA Command: *egen newvar= total(outcome var), by(var at level of interest)*
egen conANC4Jan2017_1= total(ANC4Jan2017), by(constituency)

Shortcut: *When you have multiple months, use loop command*

```
foreach var in Jan2017 Feb2017 Mar2017 Apr2017 May2017 Jun2017  
    Jul2017 Aug2017 Sept2017 Oct2017 Nov2017 Dec2017 {  
    egen conANC4`var'_1=total(ANC4`var'), by(constituency)  
}
```

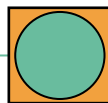


Routine Data Preparation: Data aggregation

- Calculate rates of ANC coverage
$$\frac{\text{\# of 4th ANC visits}}{\text{\# of women of reproductive age}}$$
- **STATA command:** *bysort id region constituency: gen conANC4Jan2017=conANC4Jan2017_1/CON_2017popWomRepAge*

Shortcut: When you have multiple months, use loop command

```
foreach var in Jan2017 Feb2017 Mar2017 Apr2017 May2017 Jun2017 Jul2017 Aug2017  
Sept2017 Oct2017 Nov2017 Dec2017 {  
  bysort id region constituency: gen conANC4`var`=conANC4`var`_1/con_2017popWomRepAge  
}
```



Aggregated dataset

region	constituency	nfac_cons	conANC4Jan~7	conANC4Feb~7	conANC4Mar~7
Gold Coast Region	Gold-C. IX	17	.0022075	.0022075	.0024835
Gold Coast Region	Gold-C. XV	32	.0018762	.002261	.0021407
Caliga Region	Caliga IX	12	.0022487	.0024361	.0021238
Caliga Region	Caliga X	76	.0071103	.0066354	.0058348
Spine Region	Spine XXV	56	.0030288	.0030288	.0031686
North Troy Region	N-Troy XXVI	49	.0015474	.0018204	.0021585

Variable

Definition

region

Region

constituency

Constituency

nfac_cons

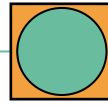
number of facilities in each constituency

conANC4Jan2017

rate of the 4th ANC visit in January 2017 at the constituency level

conANC4Feb2017

rate of the 4th ANC visit in February 2017 at the constituency level



Routine Data Preparation: Reshape data

Reshape from wide to long

STATA command: reshape long conANC4, i(id) j(smdate) string

region	constituency	id	nfac_cons	conANC4Jan~7	conANC4Feb~7	conANC4Mar~7
Gold Coast Region	Gold-C. IX	20	17	.0022075	.0022075	.0024835
Gold Coast Region	Gold-C. XV	26	32	.0018762	.002261	.0021407
North Troy Region	N-Troy XXVI	235	49	.0015474	.0018204	.0021585

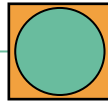
Data is wide

```
. reshape long conANC4, i(id) j(smdate) string
(note: j = Apr2017 Apr2018 Apr2019 Apr2020 Aug2017 Aug2018 Aug2019 Aug2020 Dec20
> 7 Feb2018 Feb2019 Feb2020 Jan2017 Jan2018 Jan2019 Jan2020 Jul2017 Jul2018 Jul
> 18 Jun2019 Jun2020 Mar2017 Mar2018 Mar2019 Mar2020 May2017 May2018 May2019 Ma
> 019 Oct2017 Oct2018 Oct2019 Sept2017 Sept2018 Sept2019)
```

Data	wide	->	long
Number of obs.	3	->	132
Number of variables	49	->	7
j variable (44 values)		->	smdate
xij variables:			

id	smdate	region	constituency	nfac_cons	conANC4	ndis_con
20	Apr2017	Gold Coast Region	Gold-C. IX	17	.0013797	17
20	Apr2018	Gold Coast Region	Gold-C. IX	17	.0056047	17
20	Apr2019	Gold Coast Region	Gold-C. IX	17	.0103648	17
20	Apr2020	Gold Coast Region	Gold-C. IX	17	.0088009	17
20	Aug2017	Gold Coast Region	Gold-C. IX	17	.0013797	17
20	Aug2018	Gold Coast Region	Gold-C. IX	17	.0072951	17
20	Aug2019	Gold Coast Region	Gold-C. IX	17	.0086657	17
20	Aug2020	Gold Coast Region	Gold-C. IX	17	.0072568	17
20	Dec2017	Gold Coast Region	Gold-C. IX	17	.0016557	17
20	Dec2018	Gold Coast Region	Gold-C. IX	17	.0072061	17
20	Dec2019	Gold Coast Region	Gold-C. IX	17	.0088356	17

Data is long



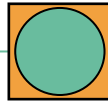
Routine Data Preparation: Date format

- Yearly
- Quarterly
- Monthly
- Weekly



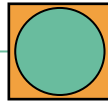
Format for sample dataset

Tip : search "help datetime" for more information on date styles



Reshape and date format illustration

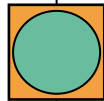




Summary

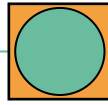
- ITS is a strong quasi experimental design that uses data collected at multiple points before and after an intervention to determine the impact of the intervention.
- ITS can identify the immediate effect, sustained effect or both.
- To prepare our routine data for ITS analysis, we need to assess for missing data, if necessary aggregate data, reshape data and properly format the dates.

Resources



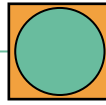
- Background information sheet
 - Overview of time series and interrupted time series
 - List of useful resources
 - Summary of studies using interrupted time series analysis
- PowerPoint presentation
- Sample dataset*
- Sample do file*

* Resources are available for Part 2 and 3 of the ITS video tutorial only



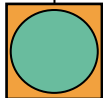
Looking ahead

- Part 1:
 - Interrupted time series overview
 - Important considerations
 - Preparing the dataset
- Part 2:
 - Interrupted time series analysis: Single group
- Part 3:
 - Interrupted time series analysis: Multiple group



References

- Baum, C.F., Schaffer, M.E. 2013. `actest`: Stata module to perform Cumby-Huizinga general test for autocorrelation in time series. <http://ideas.repec.org/c/boc/bocode/s457668.html>
- Bernal, J. L., Cummins, S., & Gasparrini, A. (2017). Interrupted time series regression for the evaluation of public health interventions: A tutorial. *International Journal of Epidemiology*, 46(1), 348–355. <https://doi.org/10.1093/ije/dyw098>
- Cochrane Effective Practice and Organisation of Care (EPOC). (2017). Interrupted time series (ITS) analyses: EPOC Resources for review authors. <https://www.coursehero.com/file/74328642/interrupted-time-series-analysesdocx/>
- Law, M. Policy Analysis Using Interrupted Time Series. University of British Columbia. <https://www.edx.org/course/policy-analysis-using-interrupted-time-series>
- Linden, A. (2015). Conducting interrupted time-series analysis for single- and multiple-group comparisons. *The Stata Journal*, 15(2), 480-500. <https://www.stata-journal.com/article.html?article=st0389>
- Penfold, R. B., & Zhang, F. (2013). Use of interrupted time series analysis in evaluating health care quality improvements. *Academic pediatrics*, 13(6 Suppl), S38–S44. <https://doi.org/10.1016/j.acap.2013.08.002>



This presentation 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.

www.data4impactproject.org

