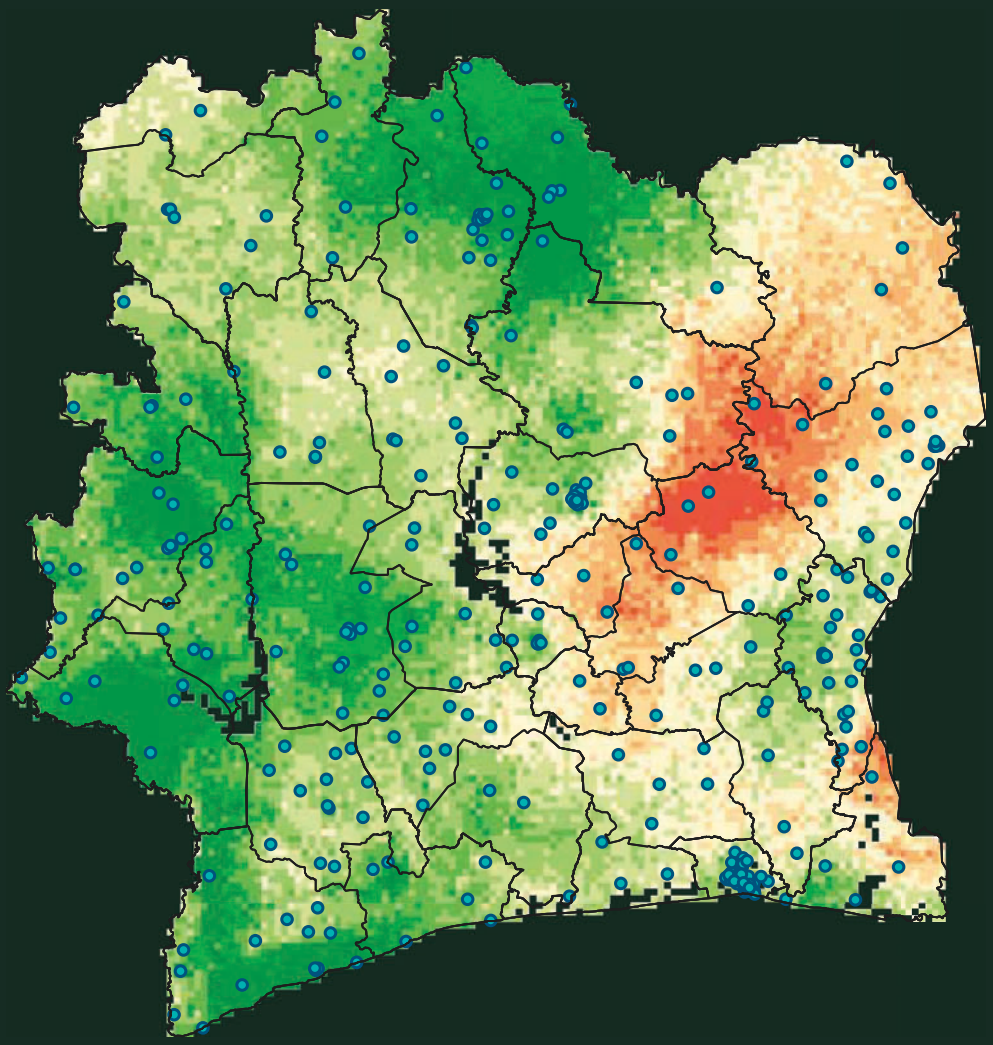
Reducing the Costs and Barriers to Evaluations Using Geospatial Data:

New Methods with an Application for HIV/AIDS in Côte d'Ivoire

White Paper for USAID | PEPFAR | Côte d'Ivoire
AidData \ William & Mary

Ariel BenYishay and Katherine Nolan
July 2020



USAID
FROM THE AMERICAN PEOPLE

AIDDATA
A Research Lab at William & Mary

Reducing the Costs and Barriers to Evaluations Using Geospatial Data: New Methods with an Application for HIV/AIDS in Côte d'Ivoire

White Paper for USAID | PEPFAR | Côte d'Ivoire

AidData \ William & Mary

Ariel BenYishay and Katherine Nolan

July 2020

ABSTRACT

We aim to help evaluators, funders, and project implementers understand their options for combining multiple rounds of surveys and spatial data to evaluate projects, particularly those in the health sector. We describe new geospatial methods that allow one to use geospatially interpolated ("predicted") data in place of one or more rounds of primary survey data. We simulate the statistical power under the most common cases that entail alternative configurations of these methods. As an application of these methods, we focus on the case of HIV/AIDS indicators in Côte d'Ivoire, derived from geo-located DHS and MICS data, as well as interpolated layers from IHME. We find that combining baseline predicted surface data with follow-up survey data provides the most statistical precision, allowing evaluators to detect even small treatment effects. Such configurations are feasible as retrospectively designed evaluations and thus are usable in a wide array of real world contexts and can help evaluators overcome barriers to running impact evaluations such as cost constraints that inhibit the collection of project-specific data.

Table of Contents

Tables and Graphs

Acronym List

CDC - Center for Disease Control
DPT - Diphtheria, Pertussis, and Tetanus
DHIS2 - District Health Information System
SIGDEP 2.0 - Système d'Information de Gestion des Dossier Electronique des Patients
eLIMS - Logistics Management Information System
OpenELIS - Open Laboratory Information System
OVC - Orphans and Vulnerable Children
ENSEA - Ecole Nationale Supérieure de Statistique et d'Economie Appliquée
MICS - Multiple Indicator Cluster Surveys
FFP - Food for Peace
DHS - Demographic Health Surveys
IHME - Institute for Health Metrics and Evaluation
WALA - The Wellness and Agriculture for Life Advancement program
HIV/AIDS - Human Immunodeficiency Virus and Acquired Immune Deficiency Syndrome
UNAIDS - the Joint United National Programme on HIV/AIDS
GEF- Global Financing Facility

Acknowledgements

This project was funded by USAID Côte d'Ivoire through a cooperative agreement (AID-OAA-A-12-00096) between USAID's Global Development Lab and AidData at the College of William & Mary under the Higher Education Solutions Network (HESN) Program. We greatly appreciate their support. We'd also like to thank William Ouattara and Anastase Mian for their hard work and contributions towards this paper.

Introduction

To understand living conditions and assess the impacts of programs aimed at improving those conditions, governments and development partners have often relied on household quantitative and qualitative surveys of samples of the population. Sometimes, this sample focuses on specific types of individuals, while at other times the sample is meant to be representative of the whole population. These surveys are seen as the best way to discover both the overall well-being of the population and the potential impacts of a certain project.

Development aid resources are finite, and while these surveys are important, they can be very expensive. For example, the cost of conducting a standard round of the Demographic & Health Surveys (DHS) in even a medium-sized country averages approximately $1M; in countries with large populations, such as Côte d'Ivoire, these costs typically exceed $1.2M (Jerven 2017). The high cost of these surveys means they are conducted only infrequently (typically only every 3-5 years) for nationally representative samples. Moreover, these costs can be prohibitive for potential impact evaluations of development programs, leading only a small share of such programs to be covered by survey-based data collections. Finding ways to economize on survey costs while still gathering high quality data from the sought-after population of people may thus be a crucial

One way to potentially decrease survey costs is to build on existing (secondary) geolocated data that are increasingly available, including national or project household surveys, administrative data, and remotely sensed data (from satellites and other sensors). These can be used to improve the precision of estimates from new primary data, thereby reducing the sample sizes required in these new data collection efforts and therefore the associated cost and effort. Moreover, in evaluation settings, these additional data sources can sometimes be used in place of primary data collection efforts if they do not exist, especially as baseline data that help draw comparison groups that are similar to the treated groups.

There are a number of challenges in potentially using existing secondary data for these purposes. First, high-quality surveys, such as the DHS and Multiple Indicator Cluster Surveys (MICS), are available only for samples of the population, and these are often sparsely distributed over a given country or area and only representative at national scales. As a result, most locations in a country are *not* directly covered by a DHS round. Second, administrative data (such as health facility data) and remotely sensed data (such as satellite imagery) may not capture the exact outcomes of interest in a given study, particularly in the health, education, and related social sectors. However, recent progress against both challenges now puts workable, general solutions within reach. For example, newly produced interpolated layers from the Institute for Health Metrics and Evaluation (IHME) provide estimates of key health measures at small subnational scales, allowing us to potentially compensate for missing populations not covered in sample surveys.

We aim to help evaluators, funders, and project implementers understand their options for combining multiple rounds of surveys and spatial data to evaluate projects. We describe new geospatial methods that allow these diverse datasets to be geospatially interpolated and joined to primary data collection survey data. We simulate the statistical power under the most common cases that entail alternative configurations of these methods. We concentrate on cases where new, primary data collection for both baseline and follow-up rounds is not feasible or within-budget. We then provide both an overall assessment and a comparison of the statistical power afforded by the alternative configurations. As an application of these methods, we focus on the case of HIV/AIDS indicators in Côte d'Ivoire, derived from geo-located DHS and MICS data, as well as interpolated layers from IHME. This setting is particularly salient for our application because it offers both a rich potential set of data configurations and an array of development efforts aimed at combating HIV/AIDS in a large population.

Our simulation results suggest that combining baseline outcome measures from predicted surfaces with follow-up measures from newly collected primary data provides the greatest statistical precision. Notably, such designs are more feasible than alternatives that rely on

baseline survey data collection, opening the door for many more retrospectively conducted evaluation designs. These results also highlight how newly created predicted surfaces can be used to help recover "missing baseline" scenarios (where no baseline survey data was conducted in both treated and comparison locations)--an important feature of geospatial impact evaluations.

HIV/AIDS Data Landscape

Over the past two decades in Côte d'Ivoire, a variety of detailed data on the health sector have been collected and compiled into a number of data rich platforms. These local data sources include DHIS 2, SIGDEP 2.0, eLIMS, OpenELIS, and the OVC Database and they cover a wide range of topics, including HIV/AIDS, maternal health, health clinic information, health drug stocks, information on orphans and vulnerable children, and more. There are also private organizations collecting household or project data across the country. Despite the large amount of data that is collected, data quality remains a concern (there are significant challenges getting the data entered, correct, and verified) and much of the data remains difficult to access for both programming and evaluation uses.

Système d'Information de Gestion des Dossier Electronique des Patients (SIGDEP 2.0), Logistics Management Information System (eLMIS), Open Laboratory Information System (OpenELIS), and the Orphans and Vulnerable Children (OVC) database face issues with (a) low capacity and time to collect data at the facilities; (b) ad-hoc collection processes for non-clinical data; (c) inefficient and costly validation procedures, especially at data-validation workshops; and (d) lack of IT in remote areas (Desai, 2018b). Given these challenges, this wealth of data can be difficult for evaluators to access because of concerns surrounding that releasing inaccurate/incomplete data could create backlash and may jeopardize general welfare (Desai, 2018b). Other barriers to accessing the data include privacy concerns for patients, doctors, and clinics (Desai, 2018b).

The main actors in the health data system include several ministries, private institutions, and Universities, such as Ecole Nationale Supérieure de Statistique et d'Economie Appliquée (ENSEA) which helps build capacity of health professionals and government ministries on basic data literacy. All major development partners are also investing resources towards strengthening health systems and developing new tools (i.e. CDC and openEllis, UNAIDS and the Situation Room, or the World Bank and the *carte sanitaire*) (Desai, 2018b). Funding is expected to increase for data production activities, in light of recent concept notes to the Global Fund and the onset of the Global Financing Facility (GFF) for Côte D'Ivoire (Desai, 2018b).

Despite the number of actors, increased funding and investments in Côte d'Ivoire's health data systems, growing momentum for creating open data sources, and Côte d'Ivoire's open data policy, significant challenges exist for organizations seeking to make health datasets public and for evaluators to gain access to them (Desai, 2018b). The current usage is also limited due to gaps in data quality, timeliness, and usability of geospatial information (Barlow et al., 2017). Overall, Côte d'Ivoire has a wealth of health data that could be useful for health project evaluators but, despite these rich data collection activities, there are still relatively few options for evaluators who seek to run impact evaluations on HIV/AIDS programs.

Related Literature

Over the past five years, predictive (interpolated) surfaces have become increasingly widespread, both in terms of the countries they cover and the different health and demographic outcomes they reflect (Paige et al., 2019). These surfaces now include population estimates (Wardrop et al., 2018; Lloyd et al., 2019), child mortality (Gething et al., 201; Akachi et al. 2011; Golding et al., 2017; Utazi et al., 2018; Bird et al., 2016), travel times to and utilization of health services (Broer et al., 2018; Ruktanonchai et al. 2018; Ruktanonchai et al., 2016), birth and pregnancy (James et al., 2018), child health inequality (Youkavitch et al., 2018), DHS indicators (Gething et al., 2015), child growth failure (Osgood-Zimmerman et al.,

2018), educational attainment (Graetz et al., 2018), diseases (Giorgi et al., 2018), vaccines (Utazi et al. 2019) and other general public health indicators (Diggle and Giorgi, 2019; Diggle and Giorgi, 2016). There has also been rich discussion across multiple disciplines about the history (Mahy, et al. 2017), methods (Aids, 2016), accuracy (Dube et al. 2019), and strengths and weaknesses (Carter, 2018; Leyk et al. 2019) of these surfaces.

More narrowly focusing on HIV/AIDS, there have also been a number of papers that look at subnational HIV prevalence variation in specific countries or sets of countries (Anderson et al., 2014; Kleinschmidt et al., 2007; Kandala et al., 2012; Larmarange and Bendaud, 2014; Okana and Blowe, 2016; Carrel et al., 2016; Coburn et al., 2017; Cuadros et al., 2017; Meyer-Rath et al., 2018; TB et al., 2016). In addition to this country-specific work, researchers at IHME developed predictive layers for a comprehensive set of indicators for all of sub-saharan Africa (Dwyer-Lindgren et al., 2019).

Beyond creating surfaces and using them for tracking specific health indicators, there has not been much use of predictive surfaces in evaluations, specifically in the HIV/AIDS sector. One potential use of predictive surfaces in evaluation is utilizing them to retrospectively draw comparison groups. For example, in Southern Malawi, BenYishay et al. (2019) utilized a predictive surface that was created based on DHS estimates to select villages that did not receive The Wellness and Agriculture for Life Advancement (WALA) program funded by Food For Peace (FFP) but had similar characteristics at baseline as villages that did receive the program. The full sample of these comparison and treated villages were then surveyed five years after the program ended to understand the lasting impact of the program. Beyond these uses, there have been limited evaluations utilizing predictive surfaces to look at program outcomes. Given the length of time evaluations take and the recent surge in both available predictive surfaces and related geospatial evaluation methods, this could change in the future. There is a related, growing field devoted to geospatial impact evaluation methodologies, which use geospatial outcome measurement tied to intervention data to recover quasi-experimental estimates of causal impacts (c.f. BenYishay et al 2017). As this paper explores,

there are several options for evaluators to use the surfaces either alone or combined with survey data to look at project outcomes.

Data Overview

In order to answer these questions, we used a combination of MICS 2015 data, DHS 2012 data, and IHME data for Côte d'Ivoire. We used these data sources because they are health focused, geolocated, and publicly available. We concentrated on a set of common measures available across multiple sources, including the HIV prevalence rate found in the DHS data as well as the first diphtheria, pertussis, and tetanus (DPT 1) vaccine rates and HIV questions in the MICs data. Both the DPT 1 vaccine estimates and the HIV estimates were utilized from the IHME datasets for years 2012 and 2015.

The DHS data was collected in 2011/2012 by the Demographic Health Survey Program. This survey covered a wide range of health topics, such as vaccines, domestic violence, HIV knowledge and testing, and child mortality. This survey was geo-referenced at the cluster level. In order to prevent the re identification of survey respondents, the survey administrators randomly displace the GPS latitude/longitude positions for all of the surveys. In order to utilize the HIV data, the team created a cluster-level number of HIV positive individuals. Similarly, we calculated a DPT1 vaccination percentage at the cluster level.[1]

The MICS data was collected in 2015/16 in lieu of the DHS survey and was geocoded to the cluster level. This survey also contained a wide range of health questions from child nutrition and vaccines to maternal and household health practices. While they did not conduct any HIV testing as part of the survey, the research team did ask a number of questions about HIV/AIDS. To estimate HIV prevalence, we utilized an index of HIV-related questions that included

[1] This was calculated by using the number of children in the household vaccinated with the DPT1 vaccine divided by the total number of children that enumerators asked about during the vaccine section. This differed from the IHME DPT1 calculations because IHME used only children who were between 12-23 months. When we calculated the the DPT1 vaccine percentage using only children aged 12-23 months, the IHME and DHS DPT1 vaccine percentage were poorly correlated (.266). The correlation increased significantly when we used all children who were included in the vaccine section (.733).

whether the respondent was married or living with a partner as married, whether they used a condom during the last sexual encounter, whether they had intercourse, and whether the respondent had multiple partners in the last year for men and for women. The IHME data used the same variables in their HIV prevalence estimates, as described below.

The third data source used in this analysis is the IHME HIV prevalence estimates and DPT1 vaccine coverage. The IHME HIV prevalence estimates are estimates of the number of adults aged 15-49 who are living with HIV at the 5x5 km grid cell level across 47 countries in Africa (IHME, 2019a). These estimates were created using a number of different sources including covariates from the 2000 Multiple Indicator Cluster Survey (MICS) microdata (UNICEF, 2000), 2005 DHS survey microdata (ICF International, 2005), the 2011-12 DHS survey microdata (ICF International, 2011-12), and the 2016 MICS survey microdata (UNICEF, 2018). The covariates used from the two DHS surveys include the prevalence of:

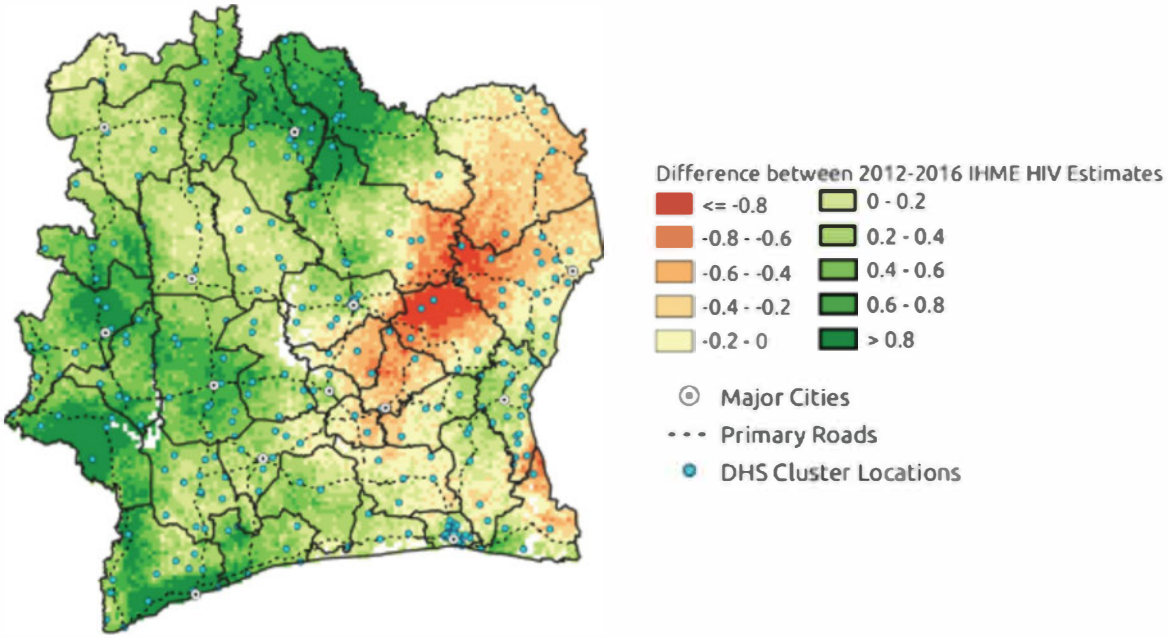
- male circumcision (all forms)
- self-reported STI symptoms
- marriage or living with a partner as married
- one's current partner living elsewhere
- condom use at last sexual encounter
- reporting ever had intercourse among young adults
- multiple partners in the last year for men
- multiple partners in the last year for women and women

The covariates used from the MICS surveys include the prevalence of:

- had intercourse
- marriage or living with a partner as married
- condom use at last sexual encounter
- reporting ever had intercourse among young adults
- multiple partners in the last year for men and women

IHME also used the 2017 ICAP, Columbia University Mailman School of Public Health survey report (ICAP, 2018), HIV estimates from the 2017 and 2018 UNAIDS estimates file (UNAIDS, 2018) that incorporates information from 2004-2005, 2008, 2013, and 2017.

Map 1: Changes in IHME HIV Estimates, 2012-2016 (with DHS locations overlaid)



IHME prevalence estimates for DPT1 coverage were based on the number of children aged 12-23 months who received the DPT1 vaccine across 52 African countries at the 5x5 km-level (IHME, 2019b). For these estimates, IHME utilized the 2011-2012 DHS survey (ICF International ,2012), the 2000 MICS survey (UNICEF, 2000), and the 2006 MICS survey (UNICEF, 2006) to create a 5x5 km-level prevalence rate. From these surveys, the prevalence estimates were constructed using the age, sex, and DPT1 coverage covariates. The researchers utilized an existing Bayesian spatiotemporal methods to analyse a geolocated database of HIV prevalence across Africa. They produced a gridded estimate of HIV prevalence for Africa. Researchers estimated grid-cell-level HIV prevalence estimates and grid-cell-level estimates of the 15-49 aged population to ultimately estimate the number of people living with HIV. After creating these estimates, they "calculated population-weighted averages of the grid-cell-level estimates to generate estimates for first-level administrative subdivisions (for example,

provinces or regions) and second-level administrative subdivisions (for example, districts or departments) in each country" (Dwyer-Lindgren, 2019).

IHME DPT1 and HIV prevalence estimate data was extracted for all of the DHS and MICS survey locations using the point extraction tool in QGIS. This data was then matched with the DHS and MICS survey data using a unique identifier. Qgis was also used to match DHS and MICS survey locations using Hub Distance Analysis which paired the closest villages to each other. Since there were more MICS locations than DHS locations, all MICS locations that were not matched were dropped from the survey sample.

Data Analysis

This paper looks at three different simulated cases that reflect common situations researchers face:

- **Case 1: Survey Baseline, Predicted Follow-up**
  Survey data are available at baseline, but no follow-up survey data are available, so researchers rely on a predictive surface to reflect follow-up conditions
- **Case 2: Predicted Baseline, Survey Follow-up**
  No survey data are available at baseline, so researchers rely on predictive surface, along with available survey data at follow-up.
- **Case 3: Predicted Baseline and Follow-up**
  No survey data are available for either round, so researchers must rely on predictive surfaces for both the endline and baseline.

Case 1 utilized DHS 2012 survey data and the 2016 IHME predictive surface to look at DPT1 vaccines. We used a simulation that looks at a hypothetical change in the number of children vaccinated with the DPT1 vaccine between 2012 (using DHS data) and 2016 (using IHME estimates based on the DHS locations) for villages that were randomly assigned to a treatment or control group. We estimated outcomes by running 20,000 simulations that randomly

assigned treatment and control groups for all locations. The simulation then created an estimated treatment effect by multiplying the standard deviation of the IHME 2016 DPT1 vaccine estimate for that location by an assigned standardized treatment effect that increased outcomes by 0.01-0.20 standard deviation across the simulations. In each iteration, we simulated the new outcome data by including the treatment effect for the randomly assigned treatment group and then regressed this outcome variable on a treatment indicator and the baseline number of children vaccinated with DPT1 in 2012 from the DHS data. We then compared the outcome coefficients across each of the 1,000 iterations for each of the 20 values of the treatment effect. Summary statistics for our main variables used are included in Table 1 below.

Table 1: Summary Statistics for Case 1 Variables

VARIABLES	N	Mean	SD	Min	Max
DHS cluster level DPT1 rates	10,060	0.693	0.192	0.0256	1
IHME 2016 DPT1 vaccine estimates*	9,488	0.945	0.0278	0.859	0.991

*Extracted from DHS cluster locations

Case 2 utilized the IHME 2012 predictive surface as a baseline and the 2016 MICs survey as an follow-up. Instead of DPT1 vaccine data used in Case 1, this simulation used an HIV index created from the 2016 MICS men's and women's data (discussed above), as well as IHME HIV prevalence estimates from 2012 that were extracted from the MICS survey locations. This simulation utilized the same analysis methods including assigning an increasing treatment effect that is measured in terms of the standard deviation of the HIV proxy variable from the 2016 MICS data to create a simulated treatment effect. We then regressed this outcome on the randomly generated treatment status and the 2012 IHME HIV estimates, again repeating this 1,000 times for each of the 20 potential values of the treatment effect (0.01-0.20).

Table 2: Summary Statistics for Case 2 Variables

VARIABLES	N	Mean	SD	Min	Max
IHME 2012 HIV Estimates*	17,823	3.473	0.762	1.196	5.093
IHME 2016 HIV Estimates*	17,823	3.199	0.796	0.896	4.733
MICS HIV Index	13,972	0.494	0.131	0.250	1

*These estimates were extracted from MICS locations

Case 3 utilized the IHME predicted surfaces for both the baseline and endline to look at the changes in HIV estimates outcomes in a hypothetical situation where some clusters were treated with a development HIV/AIDS program and some were not. The HIV proxy for the MICS data was replaced with the IHME 2016 HIV estimate and added baseline controls to the regression. These baseline controls included two variables from the DHS survey on floor material and roof material of the individual's homes. We followed the same procedure as the first two simulations and created a new outcome variable with the added estimated treatment effect. The regression included the treatment variable, IHME 2012 HIV estimates, IHME 2016 HIV estimates, and DHS and MICS control variables. The IHME estimates were extracted from the DHS locations.

Table 3: Summary Statistics for Case 3 Variables

VARIABLES	N	Mean	SD	Min	Max
IHME 2016 HIV Estimates*	994	3.171	0.754	0.896	4.733
IHME 2012 HIV Estimates*	994	3.443	0.733	1.196	5.551
Floor Quality	1,022	30.36	5.823	11	47
Roof Quality	1,022	30.09	6.645	13	88

*These estimates were extracted from MICS and DHS locations

Lessons About Data Integration

Completing these three analyses came with several challenges mainly related to data integration. Since each dataset was created by different organizations there were challenges in creating indicators that matched across the DHS, MICS, and IHME data. For example, the DPT1 vaccine rate in each cluster had to be calculated using the survey data. That calculation needed to match the calculation performed to calculate the IHME DPT1 vaccine statistics. For IHME, they calculated the DPT1 vaccine rate using the number of children aged 12-23 months that had received the vaccine. After performing that same calculation for the DHS data, there was low correlation between the DHS and IHME DPT1 vaccine rates (.266). Given this, the DHS calculation was performed using all children that were asked about the DPT1 vaccine (the section indicates this is for the last five children that were born to the respondent). That resulting calculation correlated significantly better with the IHME calculation (.733).

While each dataset had its own unique indicators and structure, in general, using a combination of Qgis and a statistical software package (in this case Stata), was an effective way to perform the analysis on three different datasets. The key is being able to link them through the geolocated points collected by each organization. Implementers and project evaluators

will have more flexibility and options for evaluating programs if they can collect coordinates for both their own data collection and their project implementation locations. These coordinates can be utilized to match with other geolocated survey data and predictive surfaces that can inform implementers and evaluators how the projects are working to improve individuals livelihoods.

Results

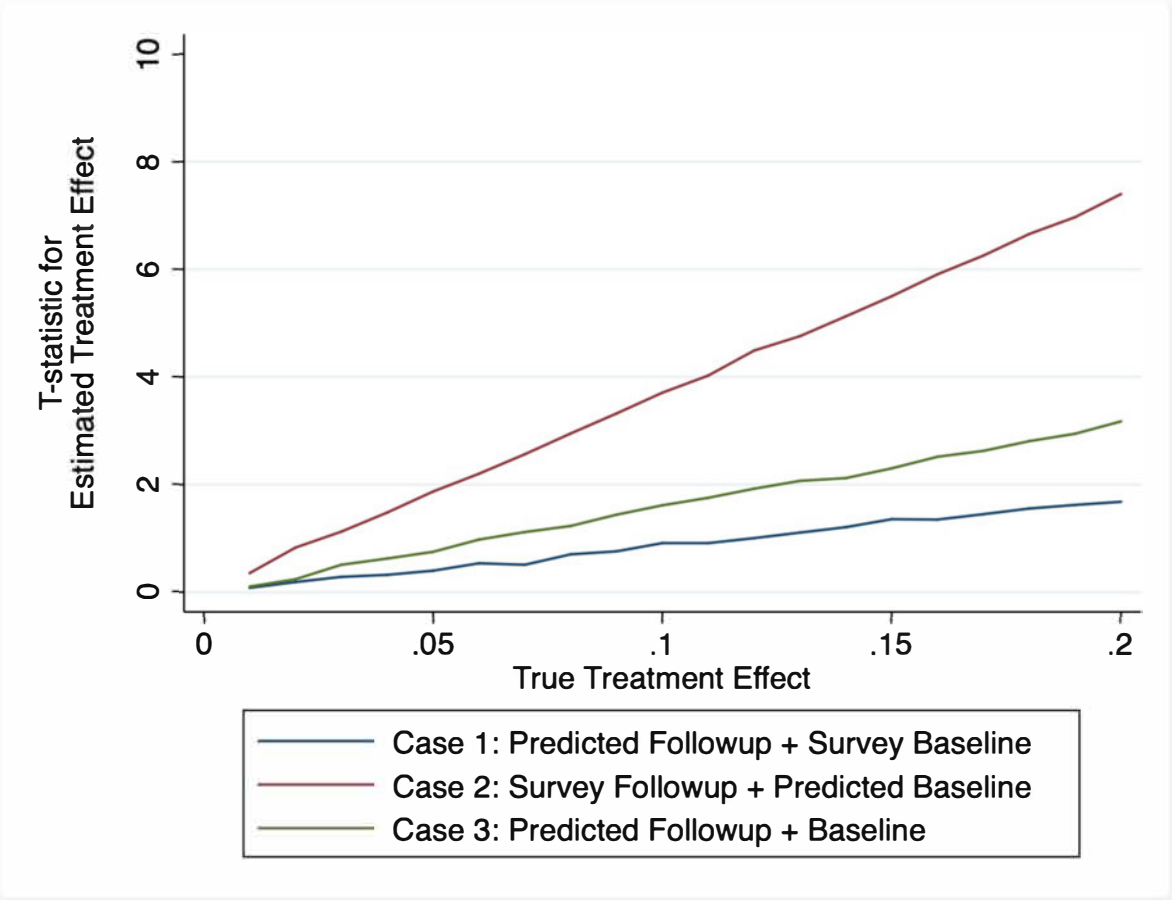
We lay out the results of the three simulated cases in Figure 1 below, which shows the t-statistic associated with each simulation over a given treatment effect size. For all treatment effect magnitudes, we find that Case 2 (Survey Follow-up + Predicted Baseline) provides more statistical significance than either of the other two cases. In fact, Case 2 can statistically detect even quite small treatment effects, with its minimum detectable effect at approximately 0.05 standard deviation units (detectable at a t-statistic of ~2, with a p=0.05). This result is driven largely by the substantial variation in outcomes reflected in the follow-up survey data, which includes variation both across clusters and within them.

Notably, Case 1 (Predicted Follow-up + Survey Baseline) fares most poorly, with no treatment effects that are statistically distinguishable from zero, even for the largest magnitude effects. This is due in part to the narrower variation offered by the predicted surface, which does not reflect any of the within-cluster variation in outcomes captured in the follow-up survey. Moreover, relying on the survey data at baseline naturally limits the sample to only the baseline cluster locations. As a result of these two constraints, the overall sample variance in outcomes limits the ability to detect treatment effects under this scenario.

Case 3 (Predicted Follow-up + Baseline) offers an intermediate performance, with medium-to-large treatment effects (those exceeding 0.15 standard deviations) statistically detectable. While Case 3 also faces limited variation in the predicted surface outcome measures, the ability to draw sample locations out of both the baseline and endline rounds of the survey data

provides a large enough overall sample to detect some treatment effects. In other words, the use of baseline outcome measures derived from the predicted surface rather than the survey data overcomes one key limitation in such contexts.

Figure 1: Results from Simulated Cases



Discussion

This paper aims to help evaluators, funders, and implementers understand how predicted surfaces drawn from geospatial health data may be used to overcome survey constraints such as budget and time issues. In this context, we find that joining predicted surfaces reflecting baseline conditions with follow-up survey data affords the greatest statistical power. Under this configuration, one can detect even quite small treatment effects (smaller than 0.05 standard deviation units). Importantly, this configuration is also easily usable in geospatial impact

evaluation--an increasingly used approach to providing rigorous impact estimates without randomized designs or baseline data collection. Using this method, evaluators could potentially utilize an interpolated surface for their baseline estimates and use their own survey for an endline, which would cut the cost of surveying in half. In fact, should there be circumstances where evaluators are faced with certain restrictions, the ability to use this type of impact evaluation methodology with only one round of primary data collection creates opportunities for learning and evaluation where none may have existed.

We specifically used the MICS and DHS data in this analysis because these health surveys are geolocated and publicly available. However, other datasets which are not currently publicly accessible (such as the DHIS2) can provide additional insights into how health facilities are functioning, what types of diseases they are treating, what drugs they use and need, and more. Organizations such as PEPFAR or other project funders or implementers may have access to these and other additional geolocated project data and could use these datasets instead of— or combined with — the publicly available surveys used herein. The methodology we describe is not limited to just the MICS, DHS, and IHME data. Moreover, digging deeper into the pathways or mechanisms leading to specific project impacts may still require this additional access to datasets beyond those used directly in our simulations.

We do not assess the ability of these datasets to draw similar comparison groups--that is, to overcome biases due to preexisting differences across treated and untreated sites (typically accomplished by limiting the comparison group to be sufficiently prior to the interventions). Because the treatment assignment in our simulation is at the location level (rather than the individual- or household-level), we expect both the predicted surfaces and the survey data to similarly compensate for such location-specific differences. Other interventions that target individuals more specifically within locations may see greater importance placed on survey data, as predicted surfaces may not be able to sufficiently overcome these differences.

Predicted surfaces are relatively new to the development world, but the estimates from these sources could open up a new suite of analyses and learning for practitioners and evaluators. These data sources are particularly useful in cases where project teams or evaluators were able to collect one round of data, but a second round was not feasible. Some cases entail projects with baselines but no follow-up survey waves, while others include only the follow-up rounds (or even no survey rounds at all). We find that statistically meaningful results are certainly achievable, even in the latter case where only predicted surfaces may be available. The greatest statistical power is achieved by the case where baseline predicted surface data is coupled with survey follow-up data.

The key trade-offs between relying on survey and predicted surface data center on the smaller samples of the surveys and the more limited variation inherent in predicted surfaces. Because they are available for entire countries (or even regions), predicted surfaces can offer much larger samples of locations. However, because the predictions behind these surfaces are only meaningful over clusters of locations (for example, 5km grid cells), they provide less individual- or household-level variation. In our application in Côte d'Ivoire, these trade-offs lead a specific blend of these two datasets to serve as the optimal configuration. It is possible that with more precise predicted surfaces available in the future, this optimal configuration could eventually shift to being entirely reliant on the predicted surfaces. The new and innovative research technique presented in this paper provides new options to evaluators that can reduce the cost of implementing an impact evaluation without reducing the rigor.

Bibliography

Anderson, S.-J. et al. Maximising the effect of combination HIV prevention through prioritisation of the people and places in greatest need: a modelling study. Lancet 384, 249–256 (2014).

Akachi Y, Atun R. Effect of investment in malaria control on child mortality in sub-Saharan Africa in 2002-2008. *PLoS One*. 2011;6(6):e21309. doi:10.1371/journal.pone.0021309

Austin Carter. Comparative Analysis of HIV Estimation Methods and Results in Locations with Case Reports and High-Quality Vital Registration Data. 2018. Thesis, University of Washington. https://digital.lib.washington.edu/researchworks/handle/1773/42899

Barlow, S., Spencer, J., and Wilkes, B. (2017). Innovations in Geographic Information System Mapping Technology. MEASURE Evaluation. https://www.measureevaluation.org/resources/publications/ws-17-41.

BenYishay, Ariel, Daniel Runfola, Rachel Trichler, Carrie Dolan, Seth Goodman, Bradley Parks, Jeffery Tanner, Silke Heuser, Geeta Batra, and Anupam Anand. 2017. A Primer on Geospatial Impact Evaluation Methods, Tools, and Applications. AidData Working Paper #44. Williamsburg, VA: AidData at William & Mary.

BenYishay, A., Velyvis, K, Nolan, K., Khatiwada L. K., Dolan, C., Guzmán, D. B., Purekal, T., Mamun, A., & Wilf, S. (2019). Long-term Impact Evaluation of the Malawi Wellness and Agriculture for Life Advancement Program. Williamsburg, VA and Washington, DC: AidData at William & Mary and Mathematica Policy Research.

Bird, Tomas & Pezzulo, Carla & Bird, Tomas & Utazi, Chigozie & Sorichetta, Alessandro & Tatem, Andrew & Yourkavitch, Jennifer & Burgert, Clara. (2016). Geospatial Modeling of Child Mortality across 27 Countries in Sub-Saharan Africa.

Broer, P. Niclas, Klug, Stefanie J., Snow, Rachel C., Okiro, Emelda A., Ouma, Paul O., Snow, Robert W., Tatem, Andrew J., Meara, John G., Alegana, Victor A. and Juran, Sabrina (2018) Geospatial mapping of access to timely essential surgery in sub-Saharan Africa. *BMJ Global Health*, 3 (4), 1-10, [e000875]. (doi:10.1136/bmjgh-2018-000875).

CDC Retro-CI, Ministry of the Fight Against AIDS (Côte d'Ivoire), National Institute of Statistics (Côte d'Ivoire), ORC Macro. Côte d'Ivoire AIDS Indicator Survey 2005. Fairfax, United States: ICF International.

Carrel, M. et al. Changing spatial patterns and increasing rurality of HIV prevalence in the Democratic Republic of the Congo between 2007 and 2013. Health Place 39, 79–85 (2016).

Coburn, B. J., Okano, J. T. & Blower, S. Using geospatial mapping to design HIV elimination strategies for sub-Saharan Africa. Sci. Transl. Med. 9, eaag0019 (2017).

“Côte d’Ivoire.” PHIA Project. ICAP at Columbia University. 2018. https://phia.icap.columbia.edu/countries/cote-divoire/. Accessed on December 2019.

Cuadros, D. F. et al. Mapping the spatial variability of HIV infection in subSaharan Africa: effective information for localized HIV prevention and control. Sci. Rep. 7, 9093 (2017).

Desai, Harsh (2018a). Open DCH Project Synthesis Report.

Desai, Harsh (2018b). Trip Report - Cote D’Ivoire.

Diggle, P. and E. Giorgi (2016). Model-based geostatistics for prevalence mapping in low-resource settings. Journal of the American Statistical Association 111, 1096– 1120.

Diggle, P. J. and E. Giorgi (2019). Model-based Geostatistics for Global Public Health: Methods and Applications. Boca-Raton: Chapman and Hall/CRC

Dube, Yolisa Prudence, Ruktanonchai, Corrine W, Sacoor, Charfudin, Tatem, Andrew, Munguambe, Khatia, Boene, Helena, Vilanculo, Faustino Carlos, Sevene, Esperanca, Matthews, Zoe, von Dadelszen, Peter and Makanga, Prestige T (2019) How accurate are modelled birth and pregnancy estimates? Comparison of four models using high resolution maternal health census data in southern Mozambique. BMJ Global Health, 4 (5), 1-12. (doi:10.1136/bmjgh-2018-000894).

Dwyer-Lindgren, L, Cork MA, Sligar A, Steuben KM, Wilson KF, Provost NR, Mayala BK, VanderHeide JD, Collison ML, Hall JB, Biehl MH, Carter A, Frank T, Douwes-Schultz D, Burstein R, Casey DC, Deshpande A, Earl L, El Bcheraoui C, Farag TH, Henry NJ, Kinyoki D, Marczak LB, Nixon MR, Osgood-Zimmerman A, Pigott D, Reiner RC, Ross JM, Schaeffer LE, Smith DL, Weaver ND, Wiens KE, Eaton JW, Justman JE, Opio A, Sartorius B, Tanser F, Wabiri N, Piot P, Murray CJL, Hay SI. Mapping HIV prevalence in sub-Saharan Africa between 2000 and 2017. Nature. 15 May 2019. doi:10.1038/s41586-019-1200-9.

Gething, P. W., D. C. Casey, D. J. Weiss, D. Bisanzio, S. Bhatt, E. Cameron, K. E. Battle, U. Dalrymple, J. Rozier, P. C. Rao, M. Kutz, R. Barber, C. Huynh, K. Shackleford, M. Coates, G. Nguyen, M. Fraser, R. Kulikoff, H. Wang, M. Naghavi, D. Smith, C. Murray, S. Hay, and S. Lim (2016). Mapping plasmodium falciparum mortality in Africa between 1990 and 2015. New England Journal of Medicine 375, 2435–2445.

Gething, P., A. Tatem, T. Bird, and C. Burgert-Brucker (2015). Creating spatial interpolation surfaces with DHS data. Technical report, ICF International. DHS Spatial Analysis Reports No. 11.

Giorgi, E., P. J. Diggle, R. W. Snow, and A. M. Noor (2018). Geostatistical methods for disease mapping and visualization using data from spatio-temporally referenced prevalence surveys. International Statistical Review 86, 571–597.

Graetz, N., J. Friedman, A. Osgood-Zimmerman, R. Burstein, M. H. Biehl, C. Shields, J. F. Mosser, D. C. Casey, A. Deshpande, L. Earl, R. Reiner, S. Ray, N. Fullman, A. Levine, R. Stubbs, B. Mayala, J. Longbottom, A. Browne, S. Bhatt, D. Weiss, P. Gething, A. Mokdad, S. Lim, C. Murray, E. Gakidou, and S. Hay (2018). Mapping local variation in educational attainment across Africa. Nature 555, 48

Golding, N., R. Burstein, J. Longbottom, A. Browne, N. Fullman, A. OsgoodZimmerman, L. Earl, S. Bhatt, E. Cameron, D. Casey, L. Dwyer-Lindgren, T. Farag, A. Flaxman, M. Fraser, P. Gething, H. Gibson, N. Graetz, L. Krause, X. Kulikoff, S. Lim, B. Mappin, C. Morozoff, R. Reiner, A. Sligar, D. Smith, H. Wang, D. Weiss, C. Murray, C. Moyes, and S. Hay (2017). Mapping under-5 and neontal mortality in Africa, 2000–15: a baseline analysis for the Sustainable Development Goals. The Lancet 390, 2171–2182.

Harsh Desai, Open DCH Project Synthesis Report, 2018a

Harsh Desai, Trip Report - Côte D'Ivoire, Jan 2018b

ICF International, Ministry of the Fight Against AIDS (Côte d'Ivoire), National Institute of Statistics (Côte d'Ivoire). Côte d'Ivoire Demographic and Health Survey 2011-2012. Fairfax, United States of America: ICF International.

Institute for Health Metrics and Evaluation (IHME). Africa HIV Prevalence Geospatial Estimates 2000- 2017. Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2019a.

Institute for Health Metrics and Evaluation (IHME). Africa Diphtheria-Pertussis-Tetanus Vaccine Coverage Geospatial Estimates 2000-2016. Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2019b.

James, William H.M., Tejedor Garavito, Natalia, Hanspal, Sophie E., Sutton, Andrew, Hornby, Graeme, Pezzulo, Carla, Nilsen, Kristine, Sorichetta, Alessandro, Ruktanonchai, Corrine W., Carioli, Alessandra, Kerr, David, Matthews, Zoe and Tatem, Andrew (2018) Gridded birth and pregnancy datasets for Africa, Latin America and the Caribbean. *Scientific Data*, 5, [180090]. (doi:10.1038/sdata.2018.90).

Kandala, N.-B., Campbell, E. K., Rakgoasi, S. D., Madi-Segwagwe, B. C. & Fako, T. T. The geography of HIV/AIDS prevalence rates in Botswana. HIV AIDS 4, 95–102 (2012).

Kleinschmidt, I., Pettifor, A., Morris, N., MacPhail, C. & Rees, H. Geographic distribution of human immunodeficiency virus in South Africa. Am. J. Trop. Med. Hyg. 77, 1163–1169 (2007)

Larmarange, J. & Bendaud, V. HIV estimates at second subnational level from national population-based surveys. AIDS 28, S469–S476 (2014).

Leyk, Stefan, Gaughan, Andrea E., Adamo, Susana B., de Sherbinin, Alex, Balk, Deborah, Freire, Sergio, Rose, Amy, Stevens, Forrest R., Blankespoor, Brian, Frye, Charlie, Comenetz, Joshua, Sorichetta, Alessandro, MacManus, Kytt, Pistolesi, Linda, Levy, Marc, Tatem, Andrew and Pesaresi, Martino (2019) The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use. *Earth System Science Data*, 11 (3). (doi:10.5194/essd-11-1385-2019).

Lloyd, Christopher T., Chamberlain, Heather, Kerr, David, Yetman, Greg, Pistolesi, Linda, Stevens, Forrest R., Gaughan, Andrew E., Nieves, Jeremiah, Joseph, Hornby, Graeme, MacManus, Kytt, Sinha, Parmanand, Bondarenko, Maksym, Sorichetta, Alessandro and Tatem, Andrew (2019) Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. *Big Earth Data*, 3 (2), 108-139. (doi:10.1080/20964471.2019.1625151).

Maina, J., Ouma, P.O., Macharia, P.M. et al. A spatial database of health facilities managed by the public health sector in sub Saharan Africa. Sci Data 6, 134 (2019). https://doi.org/10.1038/s41597-019-0142-2

Mary Mahy, Tim Brown, John Stover, Neff Walker, Karen Stanecki, Wilford Kirungi, Txema Garcia-Calleja & Peter D. Ghys (2017) Producing HIV estimates: from global advocacy to country planning and impact measurement, Global Health Action, 10:sup1, DOI: 10.1080/16549716.2017.1291169

Meyer-Rath, G. et al. Targeting the right interventions to the right people and places: the role of geospatial analysis in HIV program planning. AIDS 32, 957–963 (2018).

Morten Jerven (2017) How Much Will a Data Revolution in Development Cost?, Forum for Development Studies, 44:1, 31-50, DOI: 10.1080/08039410.2016.1260050

"National HIV estimates file." UNAIDS. 2018. https://www.unaids.org/en/dataanalysis/datatools/spectrum-epp. Accessed on January 2020.

National School for Statistics and Economics Applied (ENSEA), United Nations Children's Fund (UNICEF), United Nations Educational, Scientific and Cultural Organization (UNESCO). Côte d'Ivoire Multiple Indicator Cluster Survey 2000. New York, United States of America: United Nations Children's Fund (UNICEF).

National Institute of Statistics (Côte d'Ivoire), United Nations Children's Fund (UNICEF). Côte d'Ivoire Multiple Indicator Cluster Survey 2016. New York, United States: United Nations Children's Fund (UNICEF), 2018.

Okano, J. T. & Blower, S. Sex-specifc maps of HIV epidemics in sub-Saharan Africa. Lancet Infect. Dis. 16, 1320–1322 (2016)

Osgood-Zimmerman, A., A. I. Millear, R. W. Stubbs, C. Shields, B. V. Pickering, L. Earl, N. Graetz, D. K. Kinyoki, S. E. Ray, S. Bhatt, A. Browne, R. Burstein, E. Cameron, D. Casey, A. Deshpande, N. Fullman, P. Gething, H. Gibson, N. Henry, M. Herrero, L. Krause, I. Letourneau, A. Levine, P. Liu, J. Longbottom, B. Mayala, J. Mosser, A. Noor, D. Pigott, E. Piwoz, P. Rao, R. Rawat, R. Reiner, D. Smith, D. Weiss, K. Wiens, A. Mokdad, L. S.S., C. Murray, N. Kassebaum, and S. Hay (2018). Mapping child growth failure in Africa between 2000 and 2015. Nature 555, 41.

Paige, John & Fuglstad, Geir-Arne & Riebler, Andrea & Wakefield, Jon. (2019). Design- and Model-Based Approaches to Small-Area Estimation in a Low and Middle Income Country Context: Comparisons and Recommendations.

Pezzulo, C., T.Bird, C. Edson, C. Utazi, A. Sorichetta, A. Tatem, J. Yourkavitch, and C. Burgert-Brucker (2016). Geospatial modeling of child mortality across 27 countries in sub-Saharan Africa. Technical report, ICF International. DHS Spatial Analysis Reports No. 13.

Ruktanonchai, Corrine, Warren, Nilsen, Kristine, Alegana, Victor, Bosco, Claudio, Ayiko, Rogers, Kajeguka, Andrew C. Seven, Matthews, Zoe and Tatem, Andrew (2018) Temporal trends in spatial inequalities of maternal and newborn health services among four East African countries, 1999 - 2015. *BMC Public Health*, 18, [1339]. (doi:10.1186/s12889-018-6241-8).

Ruktanonchai, Corrine W., Ruktanonchai, Nick W., Nove, Andrea, Lopes, Sofia, Pezzulo, Carla, Bosco, Claudio, Alegana, Victor A., Burgert, Clara R., Ayiko, Rogers, Charles, Andrew S.E.K., Lambert, Nkurunziza, Msechu, Esther, Kathini, Esther, Matthews, Zoe and Tatem, Andrew J. (2016) Equality in maternal and newborn health: modelling geographic disparities in utilisation of care in five East African countries. *PLoS ONE*, 11 (8), 1-17. (doi:10.1371/journal.pone.0162006).

Subnational Estimates Working Group of the HIV Modelling Consortium. Evaluation of geospatial methods to generate subnational HIV prevalence estimates for local level planning. Aids (London, England). 2016 Jun;30(9):1467-1474. DOI: 10.1097/QAD.0000000000001075.

TB, Hallett & S-J, Anderson & Asante, Cynthia & N, Bartlett & Bendaud, Victoria & Bhatt, Samir & Burgert, Clara & Cuadros, Diego & Dzangare, Janet & Fecht, Daniela & Gething, P. & Ghys, Peter & JM, Guwani & Heard, Nathan & Kalipeni, Ezekiel & Kandala, Ngianga-Bakwin & AA,

Kim & ID, Kwao & Larmarange, Joseph & Zulu, Leo. (2016). Evaluation of geospatial methods to generate subnational HIV prevalence estimates for local level planning. AIDS.

United Nations Children's Fund (UNICEF), National Institute of Statistics (Côte d'Ivoire). Côte d'Ivoire Multiple Indicator Cluster Survey 2006. New York, United States: United Nations Children's Fund (UNICEF).

Utazi, C. Edson, Thorley, Julia, Alegana, Victor A., Ferrari, Matthew J., Takahashi, Saki, Metcalf, C. Jessica E., Lessler, Justin, Cutts, Felicity T. and Tatem, Andrew J. (2019) Mapping vaccination coverage to explore the effects of delivery mechanisms and inform vaccination strategies. *Nature Communications*, 10 (1), [1633]. (doi:10.1038/s41467-019-09611-1).

Wardrop, N., W. Jochem, T. Bird, H. Chamberlain, D. Clarke, D. Kerr, L. Bengtsson, S. Juran, V. Seaman, and A. Tatem (2018). Spatially disaggregated population estimates in the absence of national population and housing census data. Proceedings of the National Academy of Sciences 115, 3529–3537.

Yourkavitch, Jennifer & Burgert, Clara & Assaf, Shireen & Delgado, Stephen. (2018). Using geographical analysis to identify child health inequality in sub-Saharan Africa. PLoS ONE. 13. e0201870. 10.1371/journal.pone.0201870.