Am. J. Trop. Med. Hyg., 95(3), 2016, pp. 709–719
doi:10.4269/ajtmh.15-0895

Prediction of Low Community Sanitation Coverage Using Environmental and Sociodemographic Factors in Amhara Region, Ethiopia

William E. Oswald,[1]* Aisha E. P. Stewart,[2] W. Dana Flanders,[1] Michael R. Kramer,[1] Tekola Endeshaw,[3] Mulat Zerihun,[3] Birhanu Melaku,[3] Eshetu Sata,[3] Demelash Gessesse,[3] Tesfaye Teferi,[3,4] Zerihun Tadesse,[3] Birhan Guadie,[5] Jonathan D. King,[2,6] Paul M. Emerson,[2,7] Elizabeth K. Callahan,[2] Christine L. Moe,[8] and Thomas F. Clasen[9]

[1]*Department of Epidemiology, Emory University, Atlanta, Georgia;* [2]*The Carter Center, Atlanta, Georgia;* [3]*The Carter Center, Addis Ababa, Ethiopia;* [4]*International Trachoma Initiative, Addis Ababa, Ethiopia;* [5]*Amhara Regional Health Bureau, Bahir Dar, Ethiopia;* [6]*World Health Organization, Geneva, Switzerland;* [7]*International Trachoma Initiative, Atlanta, Georgia;* [8]*Hubert Department of Global Health, Center for Global Safe Water, Emory University, Atlanta, Georgia;* [9]*Department of Environmental Health, Emory University, Atlanta, Georgia*

Abstract. This study developed and validated a model for predicting the probability that communities in Amhara Region, Ethiopia, have low sanitation coverage, based on environmental and sociodemographic conditions. Community sanitation coverage was measured between 2011 and 2014 through trachoma control program evaluation surveys. Information on environmental and sociodemographic conditions was obtained from available data sources and linked with community data using a geographic information system. Logistic regression was used to identify predictors of low community sanitation coverage (< 20% versus ≥ 20%). The selected model was geographically and temporally validated. Model-predicted probabilities of low community sanitation coverage were mapped. Among 1,502 communities, 344 (22.90%) had coverage below 20%. The selected model included measures for high topsoil gravel content, an indicator for low-lying land, population density, altitude, and rainfall and had reasonable predictive discrimination (area under the curve = 0.75, 95% confidence interval = 0.72, 0.78). Measures of soil stability were strongly associated with low community sanitation coverage, controlling for community wealth, and other factors. A model using available environmental and sociodemographic data predicted low community sanitation coverage for areas across Amhara Region with fair discrimination. This approach could assist sanitation programs and trachoma control programs, scaling up or in hyperendemic areas, to target vulnerable areas with additional activities or alternate technologies.

INTRODUCTION

In 2015, an estimated 13% of the world's population, or just under 1 billion people, lacked access to any sanitation facility and defecated in the open.[1] The majority of these people live in rural areas, where 90% of all open defecation takes place.[1] This crisis remains despite progress made in the past two decades. Rates of open defecation have declined to current levels from 24% in 1990, and in 2015, 68% of the world population had access to improved sanitation facilities.[1] Similarly, despite exceptional progress in Ethiopia extending sanitation coverage over the past decade, recent estimates of household latrine ownership in rural areas of the Amhara Region range between approximately 40% and 70%, with the actual usage of latrines being still low, indicating that a large proportion of the population lacks access to a sanitation facility and practices open defecation.[2–4]

Community-led total sanitation (CLTS) is a mobilization approach aimed at ending open defecation through community-wide action. It was initially developed and used in Bangladesh, but it is now being widely implemented throughout the world.[5] In Ethiopia, a variant of CLTS, community-led total sanitation and hygiene (CLTSH) incorporating additional behavior change approaches and a focus on hygiene behaviors, has been adopted within the Federal Ministry of Health's (FMOH) national approach and is implemented through the national Health Extension Program (HEP).[6] The HEP trains extension workers to provide disease prevention and health promotion services, including sanitation, in rural communities.[7] CLTSH promotes households constructing their own basic pit latrines, using locally available materials, with the aim of providing a means to end open defecation.

Pit latrines represent the most basic form of sanitation to prevent immediate or subsequent human contact with excreta. Specifications for the construction of an adequate latrine have long been established.[8] CLTS imposes minimal emphasis on latrine design, focusing first on ending open defecation through the deposition of feces in a fixed location and subsequently on improving latrine quality.[5] As a result, the quality of built latrines depends on many factors, including the resources, skill, time, willingness, and capability of household residents tasked with building their own latrine. An evaluation of latrine promotion in Amhara Region identified several construction deficiencies with built pit latrines that could affect their acceptability and sustainability.[4] Recent work has also drawn attention to the influence of factors beyond the individual or household on behaviors related to water, sanitation, and hygiene, including the impact of contextual factors, such as time of year, land ownership, geographical conditions, and climate, that may motivate or deter positive sanitation behaviors like latrine construction and maintenance.[9,10]

Environmental conditions may have a profound effect on household uptake or maintenance of sanitation in response to a community-based sanitation intervention. Soil conditions were reported by household respondents as a barrier to latrine construction in Ghana, Tanzania, and Kenya.[11–13] Greater extent of vegetation or land cover near a community may increase available building materials, facilitating construction.[14] Alternatively, local changes related to economic development, like deforestation, increased population density, and proximity to roads, reportedly motivated increased latrine adoption in Benin and Kenya.[11,15] Just as local geographic conditions may deter or promote initial adoption of household sanitation, environmental factors may strongly influence the durability of

*Address correspondence to William E. Oswald, Department of Disease Control, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom. E-mail: william.oswald@lshtm.ac.uk

FIGURE 1. Examples of household pit latrines built in rural areas of Amhara Region, Ethiopia (Photo credit: William Oswald).

built latrines, particularly the rudimentary ones constructed in response to CLTS activities (Figure 1). Destruction of latrines through flooding during the rainy season were reported in Kenya and the Gambia and could decrease sanitation coverage.[11,16]

The relationship between environmental and sociodemographic conditions in a location and variations in pit latrine coverage has not been extensively examined.[17] Through novel use of impact assessment survey data collected by a trachoma control program, the current study aimed to develop and validate a model for estimating the probability of low community sanitation coverage based on environmental and sociodemographic conditions. This approach could then be used to identify and target vulnerable areas with additional promotional activities or alternate sanitation technologies more suitable for local conditions.

METHODS

Study area and population. Amhara Region is located in northwest Ethiopia and has a total area of approximately 150,000 km^2. Geographically, the region is centered on Lake Tana and encompasses a range of physical landscapes, characterized by rugged mountains, plateaus, valleys, and gorges.[18] Elevation ranges from 519 m in the northwestern areas to 4,420 m among mountains in the northeast. Land cover consists primarily of shrublands and croplands.[19] Based on a 2007 census, Amhara Region has a population of approximately 17 million.[20]

Trachoma control impact surveys. Impact surveys were conducted by the Amhara Regional Health Bureau to provide population-based woreda (Ethiopian administrative units equivalent to districts) estimates of the prevalence of trachoma disease and quantify uptake of the World Health Organization's SAFE (surgery; antibiotics; facial cleanliness; environmental improvements, including household water and sanitation access) strategy for trachoma control. Surveys were implemented in eligible woreda, or those that had received at least 5 years of the full SAFE strategy including annual azithromycin mass drug administration.

For the current study, data were combined from five trachoma control impact surveys conducted in distinct areas of Amhara Region between 2011 and 2014. The first survey was conducted in South Gondar zone between June and August 2011. The methods and results of this study have been described previously.[2] The second survey was conducted in North Gondar and West Gojjam zones between May and June 2012. The next three surveys were conducted in eastern Amhara from December 2012 to January 2013, in western Amhara from June to July 2013, and in eastern Amhara from January to February 2014.

All surveys used a multistage cluster random sampling methodology and were powered to estimate the prevalence of trachomatous inflammation-follicular among children aged 1–9 years in woreda. Within each eligible woreda, gotts (villages) were randomly selected from a geographically ordered line listing using probability proportional to population size and were primary sampling units. Within gotts, smaller administrative units of approximately 40 households, called development teams (DTs), were used as segments for a modified segment survey design.[21,22] DTs were listed and numbered upon arrival in the community with assistance from a designated "gott" representative, who then randomly drew a number from a hat to select the DT to be surveyed. In each DT, all households were surveyed. In gotts of 40 households or less, the entire "gott" was surveyed.

Heads of household were interviewed for demographic, socioeconomic information and knowledge and practices regarding water, sanitation, and hygiene. Visual inspections were made of household latrines. Responses were recorded electronically using tablet computers operating Swift Insights software (The Carter Center, Atlanta, GA).[23]

Geographic information. Geographic coordinates in World Geodetic System 1984 were collected using tablet computers at each household (except 2013 survey, where coordinates were only collected at two households per community). Household coordinates were averaged to provide a single point for each community, and these were projected to Universal Transverse Mercator (UTM) zone 37N.

Low community sanitation coverage. Community sanitation coverage was calculated as the proportion of households within the cluster observed to have a pit latrine. The aim of this study was to explore contextual determinants of sanitation coverage, simplistically excluding behavioral factors. Whether the latrine was in use was not considered for the current study. Community sanitation coverage was dichotomized for statistical convenience. Less than 20% was selected as low sanitation and was considered the outcome of interest, given it corresponded to the referent category used for a complementary study on the association of community sanitation usage with active trachoma.

Environmental and sociodemographic predictors. Candidate environmental and sociodemographic predictors were identified a priori based on literature review.[9–16] Table 1 lists each obtained variable, hypothesized influence on sanitation coverage, data source, and how the variable was created or transformed for analysis. Hypothesized influence may be positive or negative, recognizing the possibility that some variables may represent multiple influences on household latrine uptake and community sanitation coverage.[17]

Information on soil texture class and gravel content was obtained from the Harmonized World Soil Database (HWSD, v.1.2), which combines regional and national updates with information within the Food and Agriculture Organization–United Nations Educational, Scientific and Cultural Organization Digital Soil Map of the World.[24,25] Information for the dominant topsoil (0–30 cm) in each soil mapping unit were used.

Land surface form and topographic position, indicating low-lying land with higher moisture potential, were obtained from the U.S. Geological Survey Africa Ecosystems Mapping project.[26] This project used NASA Shuttle Radar Topographic Mission (SRTM) digital elevation data to classify seven land surface form classes (plains, irregular plains, escarpments, hills, breaks/foothills, low mountains, and high mountains/deep canyons) based on categorizations of local slope and relief. The obtained topographic position dataset had been created using 90-m SRTM elevation data and a 3 arc-second Drainage Direction dataset, to identify two classes of topographic position (uplands and lowlands/depressions), using slope measures for raster cells and contributing areas from "upstream" raster cells, which indicated potential for water to flow to a point, without considering climate or soil attributes.[27]

Elevation for the study region was obtained from SRTM data processed by the Consortium for Spatial Information of the Consultative Group for International Agricultural Research.[28] Annual average Normal Difference Vegetation Index for 2011 was obtained from the Africa Soil Information Service as a measure of land vegetation cover.[29] Annual total rainfall was calculated from interpolated surfaces with mean monthly precipitation from 1950 to 2000.[30] A shapefile representing paved, gravel, or dirt/sand roads for Amhara Region was obtained from the Global Roads Open Access (1980–2010) dataset.[31] Population density in 2011 per square kilometer was generated using the Oak Ridge National Laboratory's LandScan.[32] All raster surfaces were clipped to Amhara's geographic extent and projected to UTM zone 37N. Community coordinates were overlaid on raster surfaces in ArcMap 10.1 (ESRI, Redlands, CA), and values for predictors were extracted per community to create an analysis dataset.

A measure of community wealth was calculated as the mean of the total number of wealth indicators reported by households during interviews, including radio, television, electricity, mobile phone, and an iron roof. Two additional control variables were included for time and season of survey activities. A variable for time trend was created based on years since July 2011, the month when the first survey began, using the 15th as the reference date. An indicator of whether the survey was conducted during the rainy season (June–September) was also generated.

Statistical analyses. Logistic regression was used to identify sociodemographic and environmental factors predictive of low community sanitation coverage (< 20% versus ≥ 20%). The full dataset, combining information from all five surveys, was partitioned into training and testing datasets to allow for temporal and geographic validation of the model. The partition was based on survey year and location to include the range of landscapes across Amhara Region in both training and testing datasets (Figure 2). The training dataset included data from three surveys conducted between 2011 and 2012. The testing dataset included data from two surveys conducted between 2013 and 2014.

Collinearity of predictors was first assessed in the full dataset based on calculated condition indices and variance decomposition proportions.[33] Using the training dataset and forcing in control variables for time and season of survey activities, a model selection approach was used, fitting all possible subsets of predictors to maximize model fit based on Akaike information criteria (AIC).[34–36] The Akaike weight was calculated for each of the best-fitting models to compare its suitability among this candidate set of models.[35] The sum of Akaike weights was also calculated for each predictor from the models in which it was included to determine its relative importance.[35] Estimated coefficients from the selected model were then fit to testing data.

Because of the observed difference in frequency of low sanitation coverage between the training and testing datasets, a recalibrated model was also fit to the testing data.[37,38] Probabilities predicted by the model initially fitted to the testing data were transformed using the logit transformation. These log odds were then entered as an offset, specifying a coefficient of 1, into a new logistic model for the testing data, thereby estimating a new intercept. Outcome probability was then recalculated for each community from this model.

Finally, using the complete dataset, measures of association with low sanitation coverage were estimated for selected predictors. As a sensitivity analysis, the selected model was refit to assess its robustness using alternative thresholds for low sanitation coverage, including less than 10%, 30%, 40%, and 50%. Possible improvements in model prediction and changes in estimates from including the measure of community wealth were evaluated. The selected model was also refit using alternate community wealth measures based on the proportion of households per community with each separate indicator to compare estimation of association measures and model discrimination. Results using these alternate measures were similar to those using the mean of the total number of wealth indicators (Supplemental Table 1). Model selection and estimation of measures of association were also repeated using generalized estimating equations (GEE) with an exchangeable correlation structure to account for possible correlation of outcomes within districts.[33] Based on the quasi-likelihood under the independence model criterion including a penalty for the number of parameters (QICu), the selected GEE model included the same predictors as in the ordinary logistic model.[39] Results using GEE were similar to those from the ordinary logistic models that are presented here (Supplemental Table 2).

The Hosmer–Lemeshow statistic was calculated to assess model fit to training, testing, and complete data, and calibration was assessed by plotting predicted probabilities against observed probabilities by deciles of predicted probability.[33] Discrimination of all models was assessed using receiver-operator curves (ROC) and calculating area under the

Table 1

Factors hypothesized to influence community having low sanitation coverage in Amhara Region, Ethiopia, 2011–2014

Variable	Hypothesized influence on sanitation coverage	Reason	Data and source	Variable creation
USDA soil texture classes			HWSD, v.1.2.[24] Includes a 30 arc-second GIS raster image linked to an attribute database with characteristics of topsoil (0–30 cm) and subsoil (30–100 cm)	Observed topsoil texture classes within the study area included clay, clay loam, loam, and sandy clay loam. Indicators were created for each, excluding sandy clay loam as referent
Clay	Positive	More stable soil		
Clay loam	Positive	More stable soil		
Loam	Positive or negative	More or less stable soil		
Sandy clay loam				
High soil gravel content	Negative	Soils with higher gravel content are less stable	HWSD, v.1.2.[24] Includes a 30 arc-second GIS raster image linked to an attribute database with characteristics of topsoil (0–30 cm) and subsoil (30–100 cm)	Observed topsoil volume percentage gravel (materials > 2 mm) within study area included values of 1 and 29, 31, and 32. To create an indicator of higher gravel content, 1 was coded as 0, and 29, 31, 32 were coded as 1
Surface form	Positive or negative	Specific land forms may capture other mechanisms determining latrine pit collapse	U.S. Geological Survey Africa Ecosystems Mapping project[26]	Indicators were created for each of seven land surface form observed within study area, excluding high mountains/deep canyons as referent
Low-lying land	Negative	Low-lying areas have higher moisture potential may be more likely to flood	U.S. Geological Survey Africa Ecosystems Mapping project[26]	Indicator created for low-lying land, based on two classes of topographic position (uplands and lowlands/depressions)
Altitude	Positive or negative	Altitude may capture other mechanisms determining latrine pit collapse	NASA Shuttle Radar Topographic Mission digital elevation data processed by the Consortium for Spatial Information of the Consultative Group for International Agricultural Research[28]	Altitude divided by 100
Percent slope	Negative	Greater slope may increase soil instability	Same as altitude	Calculated in ArcMap 10.1
Annual average NDVI 2011	Positive or negative	Higher vegetation may increase availability of materials for latrine construction or decrease demand for sanitation by providing areas for open defecation	Africa Soil Information Service MODIS collection: vegetation indices, April 2014 release, Center for International Earth Science Information Network, Columbia University[29]	NDVI multiplied by 10
Total annual rainfall	Negative	Higher annual rainfall results in flooding or damage to latrines	Interpolated surfaces with mean monthly precipitation (mm) from 1950 to 2000[30] (available at: http://www.worldclim.org; accessed January 20, 2015)	Surfaces were added in ArcGIS 10.1 to get total annual rainfall. Total annual rainfall (mm) divided by 100
Population density 2011	Positive	Higher population density may increase demand for sanitation by reducing areas for open defecation and need for privacy	Oak Ridge National Laboratory's LandScan[32]	Population density $\log_{10}$ transformed
Distance to roads	Negative	Greater distance to roads may decrease demand for sanitation as it reflects lower perceived need. Lower access to roads reflects lower exposure to new ideas and markets and lack of mobility	Global Roads Open Access dataset[31]	Shapefile with paved, gravel, or dirt/sand roads used to create raster surface for Amhara Region indicating distance in kilometers to a road in ArcGIS 10.1

GIS = geographic information system; HWSD = Harmonized World Soil Database; NDVI = Normalized Difference Vegetation Index; USDA = U.S. Department of Agriculture.

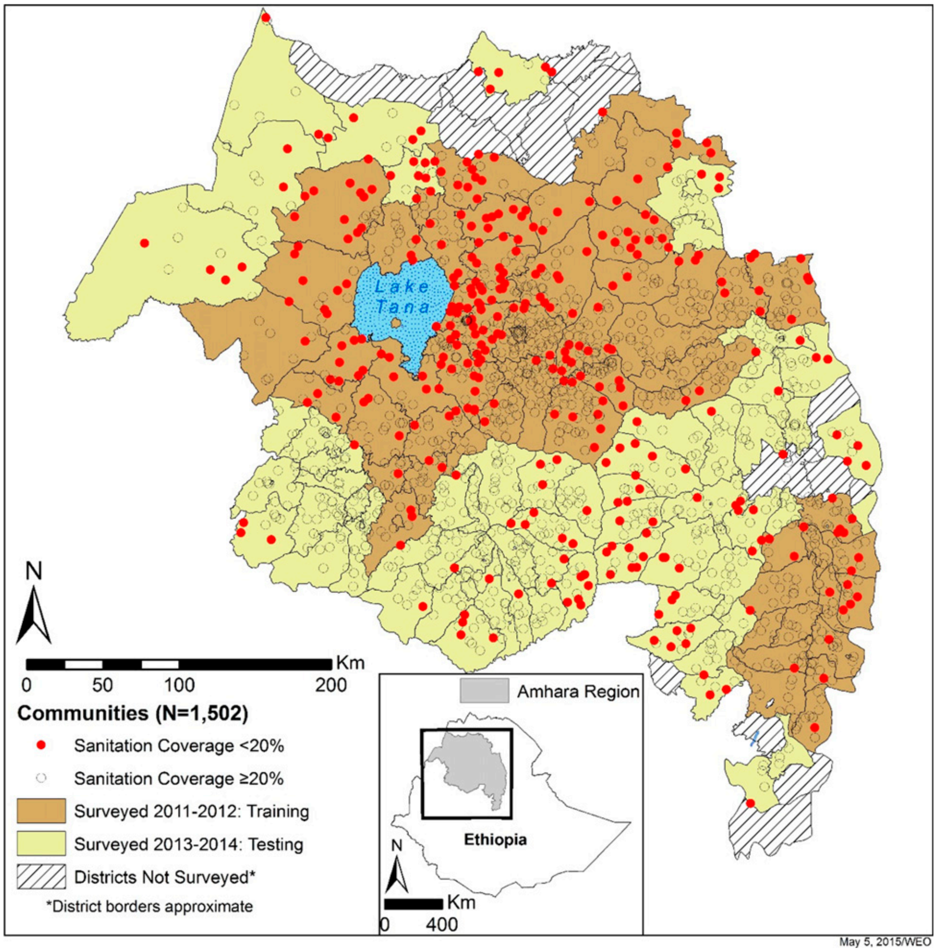

FIGURE 2. Distribution of surveyed communities (N = 1,502) and districts, by training or dataset, and low community sanitation coverage in Amhara Region, Ethiopia, 2011–2014.

curve (AUC).[33] A nonparametric approach was used to compare ROC.[40]

A model-based probability map of low community sanitation coverage was calculated using ArcMap 10.1, by applying the inverse logit function to the linear sum of the intercept and regression coefficients times their local values from raster surfaces for selected environmental and sociodemographic predictors. Analyses were performed with SAS 9.4 (SAS Institute Inc., Cary, NC).

RESULTS

Communities. Across Amhara Region, 1,510 communities were surveyed in eligible districts between 2011 and 2014. Geographic coordinates were not available for two communities (< 1%). Six communities' coordinates were placed outside the region or in areas without predictor information and were dropped (< 1%). The complete dataset contained information on 1,502 communities (Figure 2). The training dataset contained information on 876 communities, and the testing dataset contained information on 626 communities.

Table 2 presents characteristics of communities in training, testing, and complete datasets. Of 1,502 communities, 344 (22.90%) had sanitation coverage below 20% and were located throughout the region (Figure 2). The distributions of most environmental and sociodemographic conditions were similar between testing and training datasets. Clay loam was the most frequent soil texture (34.09%), and the frequency of clay (22.83% and 36.10%) and sandy clay loam (26.94% and 14.06%) differed between training and testing datasets, respectively.

Of 1,502 communities, 651 (43.34%) locations were classified as low mountainous areas. Elevation, ranging from 568 to 3,644 m, and percent slope, ranging from 0.04% to 40.54%, reflected topographic variation across study areas. Of 1,502 communities, 59 (3.93%) communities were located in low-lying areas with higher moisture potential. Approximately half of all surveyed communities were located in areas with estimated population density of less than 113.53 people/km^2 in 2011 and more than 4 km from a georeferenced paved, gravel, or dirt/sand road. Overall, within communities, households reported owning few wealth indicators.

Model development. No collinearity was detected among candidate predictors using the complete dataset. Table 3 shows eight models selected from all-possible subsets with an AIC within 2 units of the minimum AIC (AIC_{min} = 880.29), indicating little difference in estimated likelihood between models. The variable for land surface form was not selected for any model. The small ratios of Akaike weights between model 1 and other models (1.6–2.8) indicates only weak support for

TABLE 2
Environmental and sociodemographic conditions of communities in Amhara Region, Ethiopia, 2011–2014

Characteristic	Training: surveyed 2011–2012 (N = 876)		Testing: surveyed 2013–2014 (N = 626)		Complete: surveyed 2011–2014 (N = 1,502)	
	Median (IQR)	N (%)	Median (IQR)	N (%)	Median (IQR)	N (%)
Sanitation coverage < 20%		228 (26.03)		116 (18.53)		344 (22.90)
USDA soil texture classes						
Clay		200 (22.83)		226 (36.10)		426 (28.36)
Clay loam		302 (34.47)		210 (33.55)		512 (34.09)
Loam		138 (15.75)		102 (16.29)		240 (15.98)
Sandy clay loam		236 (26.94)		88 (14.06)		324 (21.57)
High soil gravel content		371 (42.35)		263 (42.01)		634 (42.21)
Annual mean, NDVI 2011	0.41 (0.35, 0.46)		0.44 (0.39, 0.49)		0.42 (0.37, 0.47)	
Surface land form category						
Smooth plains		48 (5.48)		30 (4.79)		78 (5.19)
Irregular plains		174 (19.86)		165 (26.36)		339 (22.57)
Escarpment		40 (4.57)		25 (3.99)		65 (4.33)
Hills		22 (2.51)		15 (2.40)		37 (2.46)
Breaks		147 (16.78)		90 (14.38)		237 (15.78)
Low mountains		394 (44.98)		257 (41.05)		651 (43.34)
High mountains/deep valleys		51 (5.82)		44 (7.03)		95 (6.32)
Percent slope	5.77 (2.93, 11.26)		6.08 (2.48, 12.67)		5.88 (2.72, 11.64)	
Altitude (m)	2,222 (1,917, 2,633)		2,277 (1,936, 2,598)		2,241 (1,927, 2,614)	
Low-lying land		35 (4.00)		24 (3.83)		59 (3.93)
Annual total rainfall (mm)	1,126 (1,008, 1,349)		1,155 (998, 1,421)		1,138 (1,004, 1,371)	
Distance to nearest road (km)	4.41 (1.45, 9.20)		4.10 (1.41, 7.87)		4.30 (1.43, 8.57)	
Population/km^2, 2011	116.72 (56.05, 277.85)		107.26 (55.88, 305.59)		113.53 (56.05, 289.44)	
Community mean total wealth indicators	0.73 (0.31, 1.07)		1.13 (0.83, 1.57)		0.93 (0.49, 1.30)	
Time since July 2011 (year)						
0		353 (40.30)		0 (0.00)		353 (23.50)
0.8		208 (23.74)		0 (0.00)		208 (13.85)
1.4		315 (35.96)		0 (0.00)		315 (20.97)
1.9		0 (0.00)		356 (56.87)		356 (23.70)
2.5		0 (0.00)		270 (43.13)		270 (17.98)
Surveyed during rainy season		561 (64.04)		356 (56.87)		917 (61.05)

IQR = interquartile range; NDVI = Normalized Difference Vegetation Index; USDA = U.S. Department of Agriculture.

this model, among these eight models. High gravel content and soil texture indicators were not selected together for any of the same models, indicating worse fit if modeled jointly. Soil texture was strongly associated with high gravel content (P < 0.01), and no community with topsoil classified as clay had high gravel content. The selected, best-fitting prediction model included variables for high gravel content, altitude, rainfall, population density, low-lying land, and controlled for time since July 2011 and season of survey.

Model calibration and validation. There was no evidence that the selected model did not fit training data well (Figure 3, Hosmer–Lemeshow, P = 0.63). When applied to testing data, the model fit less well and overpredicted probability of a community having low sanitation coverage (Hosmer–Lemeshow, P < 0.01; Figure 3). Figure 4 shows the ROC curves for each model. With training data, the model demonstrated reasonable discriminatory ability (AUC = 0.75, 95% confidence interval [CI] = 0.71, 0.79). Discrimination declined when the model was applied to testing data (AUC = 0.69, 95% CI = 0.63, 0.74). Model recalibration reduced overprediction (Figure 3), but tests did not indicate better fit (Hosmer–Lemeshow, P < 0.01) and discriminatory ability was unchanged (AUC = 0.69, 95% CI = 0.63, 0.74). The final model fit combined data well and maintained discriminatory ability (Hosmer–Lemeshow, P = 0.91; AUC = 0.75, 95% CI = 0.72, 0.78).

TABLE 3
Candidate training models for predicting community sanitation coverage < 20% among communities in Amhara Region, Ethiopia, 2011–2012

Model	Intercept	Altitude	Annual rainfall	High gravel	Population density	Low-lying land	Clay	Clay loam	Loam	NDVI	Distance to road	Slope	Time	Rainy	Δ_i†	ω_i
1	+X*	−X*	−X*	$+X^{NS}$	−X*	+X*	–	–	–	–	–	–	+X	+X*	0.00	0.22
2	+X*	−X*	−X*	–	−X*	+X*	$-X^{NS}$	$+X^{NS}$	$+X^{NS}$	–	–	–	+X*	+X*	0.86	0.14
3	+X*	−X*	−X*	$+X^{NS}$	−X*	+X*	–	–	–	$-X^{NS}$	–	–	+X	+X*	0.88	0.14
4	+X*	−X*	−X*	–	−X*	+X*	–	–	–	–	–	–	+X	+X*	0.91	0.14
5	+X*	−X*	−X*	–	−X*	+X*	$-X^{NS}$	$+X^{NS}$	$+X^{NS}$	$-X^{NS}$	–	–	+X*	+X*	1.61	0.10
6	+X	−X*	−X*	$+X^{NS}$	−X*	+X*	–	–	–	–	$+X^{NS}$	–	+X	+X*	1.78	0.09
7	+X*	−X*	−X*	–	−X*	+X*	–	–	–	$-X^{NS}$	–	–	+X	+X*	1.89	0.09
8	+X*	−X*	−X*	$+X^{NS}$	−X*	+X*	–	–	–	–	–	$+X^{NS}$	+X	+X*	1.93	0.08
$\sum\omega_i(j)$	1.00	1.00	1.00	0.53	1.00	1.00	0.24	0.24	0.24	0.32	0.09	0.08	1.00	1.00		

AIC = Akaike information criteria; NDVI = Normal Difference Vegetation Index; X = variable tested in model; – = variable not tested in model; − = negative association; + = positive association; NS = not significant; $\Delta_i = AIC_i - AIC_{min}$; $\omega_i = \exp(-1/2\ \Delta_i)/\sum\exp(-1/2\ \Delta_i)$; $\sum\omega_i(j)$ = sum of ω_i values from all models in which variable i was present.
* $P \leq 0.01$.
† AIC_{min} = 880.29.

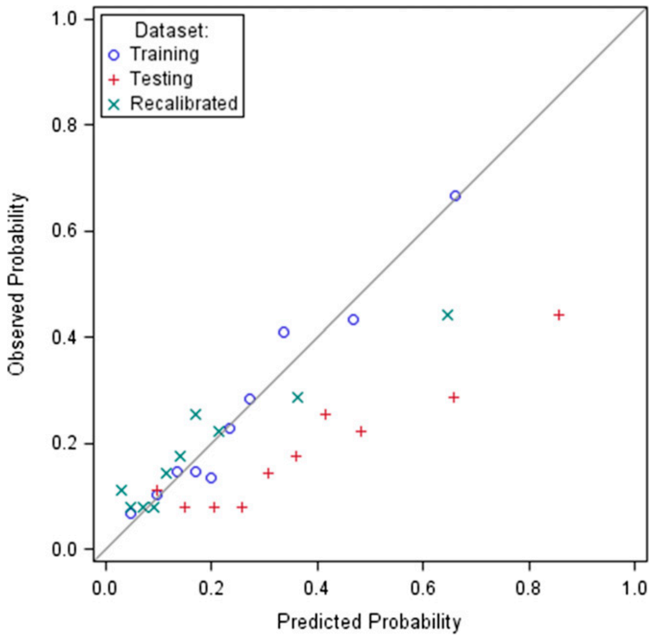

FIGURE 3. Observed vs. predicted probability for low sanitation coverage by deciles of predicted probability for models fit to training and testing datasets.

Environmental and sociodemographic conditions and low community sanitation coverage. Based on fitting the selected model to the complete dataset (Table 4), communities in areas with high topsoil gravel content had 1.76 times the odds of low sanitation coverage compared with communities in areas with low gravel content (odds ratio [OR] = 1.76, 95% CI = 1.28, 2.41), independent of other conditions. Communities in low-lying areas had 2.74 times the odds of low sanitation coverage (OR = 2.74, 95% CI = 1.49, 5.02) compared with upland areas with lower moisture potential, adjusting for other conditions. Communities in areas with higher population density, higher elevation, and higher rainfall were significantly less likely to have low sanitation coverage. Communities surveyed during the rainy season, independent of other factors, had 2.59 times the odds of low sanitation coverage (OR = 2.59, 95% CI = 1.69, 3.96). Based on the sensitivity analysis, the association of low-lying land and years since July 2011 with low sanitation coverage attenuated to null using a threshold of < 50% for low sanitation (Supplemental Table 3). There was little change in magnitude or direction of association for other predictors using alternative low sanitation thresholds, but discrimination did decrease with higher thresholds (Supplemental Table 3). In an alternate model that included a variable for community wealth, the estimated measures of association changed little (Table 4), except population density and time since 2011, which were no longer significantly associated with low sanitation coverage. An increase of one in mean total wealth indicators was associated with a 58% decrease in the odds of low community sanitation coverage (OR = 0.42, 95% CI = 0.32, 0.55), controlling for other factors. Including this variable led to a small, but significant increase, in model discrimination (AUC = 0.77, 95% CI = 0.74, 0.80; $P < 0.01$).

Figure 5 shows geographic distribution of predicted probabilities of communities having low sanitation coverage, based on estimated coefficients for environmental and sociodemographic conditions, adjusting for survey season and year. The map highlights areas in the northwestern areas of North Gondar zone, the northern half of Waghemra zone, and areas on the northern and eastern shores of Lake Tana.

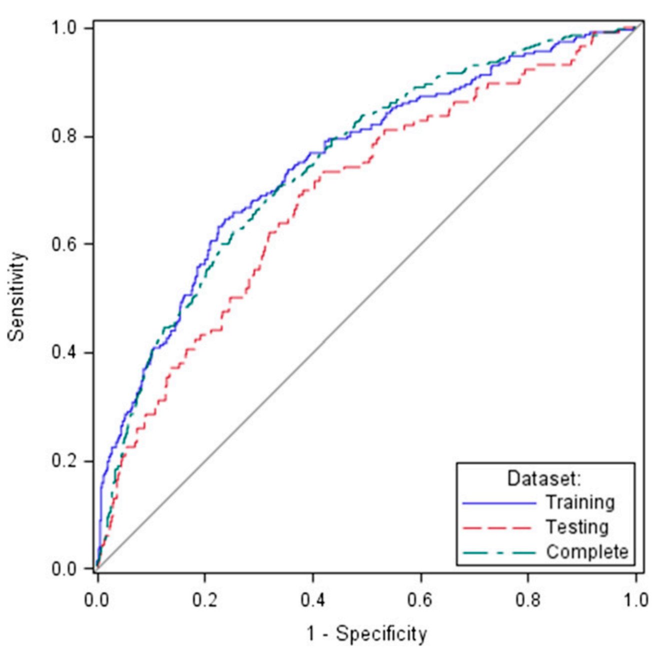

FIGURE 4. Receiver–operator curves showing sensitivity vs. 1 – specificity for models fit to training, testing, and complete datasets.

DISCUSSION

Our study used information on environmental and sociodemographic conditions from a variety of existing data sources to develop a model to predict low sanitation coverage among communities in Amhara Region. Based on the training data, the model predicted whether less than 20% of households in a community had a latrine with fair discrimination. The model's discriminatory performance decreased when applied to populations from distinct areas surveyed at later dates. We identified environmental conditions associated with low community sanitation coverage, before and after controlling for a measure of community wealth. Finally, the model was used to produce a map of the predicted probability that communities have low sanitation coverage based on local environmental and sociodemographic conditions.

Using our selected model, communities where topsoil had higher gravel content were found to have significantly higher odds of low sanitation coverage, compared with where topsoil had low gravel content, controlling for other factors. Soil characteristics and resultant mechanics represent a complex science. We based our hypotheses regarding latrine pit collapse on the relationship between soil texture and gravel content and soil cohesiveness and stability. Soils with a larger proportion of sand, silt, or clay are considered coarse, medium, or fine, respectively. Cohesive soils, the most stable, commonly have finer textures: clay, silty clay, sandy clay, clay loam, and sometimes silty or sandy clay loam.[41] Cohesive soils become only moderately stable with medium textures of silt or silt loam and least stable with coarse textures of sand or loamy sand and gravel content.[41,42] In study communities, soil texture

TABLE 4
Logistic regression models for predicting sanitation coverage < 20% among communities in Amhara Region, Ethiopia, 2011–2014

	Selected and validated model				Selected model + community wealth			
Parameters (unit of change)	Coefficient	SE	*P*	OR (95% CI)	Coefficient.	SE	*P*	OR (95% CI)
Intercept	3.54	0.52	< 0.01	–	2.84	0.53	< 0.01	–
High gravel content (yes/no)	0.56	0.16	< 0.01	1.76 (1.28, 2.41)	0.46	0.16	< 0.01	1.58 (1.15, 2.19)
Low-lying land (yes/no)	1.01	0.31	< 0.01	2.74 (1.49, 5.02)	1.04	0.31	< 0.01	2.82 (1.53, 5.20)
Population/km^2, 2011 ($\log_{10}$)	−0.55	0.11	< 0.01	0.58 (0.46, 0.73)	−0.19	0.13	0.14	0.83 (0.64, 1.07)
Altitude (100 m)	−0.08	0.02	< 0.01	0.92 (0.89, 0.95)	−0.09	0.02	< 0.01	0.91 (0.88, 0.94)
Annual total rainfall (100 mm)	−0.21	0.04	< 0.01	0.81 (0.75, 0.88)	−0.16	0.04	< 0.01	0.85 (0.79, 0.92)
Community mean total wealth indicators (+1)	–	–	–	–	−0.86	0.14	< 0.01	0.42 (0.32, 0.55)
Time since July 2011 (year)	−0.20	0.10	0.04	0.82 (0.68, 0.99)	0.05	0.10	0.66	1.05 (0.85, 1.28)
Rainy season (yes/no)	0.95	0.22	< 0.01	2.59 (1.69, 3.96)	1.10	0.23	< 0.01	3.00 (1.93, 4.66)

CI = confidence interval; OR = odds ratio; SE = standard error.

was highly associated with gravel content. Areas with clay loam, loam, and sandy clay loam had higher frequencies of high gravel content compared with areas with clay, which did not have any high gravel content. Our model selection chose high gravel content as the best soil-related predictor of low sanitation coverage in our data. The use of continuous measures for soil percentage content of sand, silt, and clay and gravel content would have allowed application of the model beyond the study region but was avoided because of concerns about accuracy. Future studies should explore the use of these alternate measures. Future applications of this approach should also consider including the major soil group classifications.[43] These classifications combine multiple aspects of soil characterization; however, the mechanistic relationship with latrine pit collapse is less clear. One soil group, Vertisols, is known to be associated with structural instability and soil collapse. Vertisols, constituting the "black cotton" soils common throughout South Sudan and parts of Ethiopia, have high clay content but are expansive in nature and have a high risk of collapse.[41] The current study might have been strengthened by including an indicator for this soil group to potentially better characterize the risk of soil collapse.

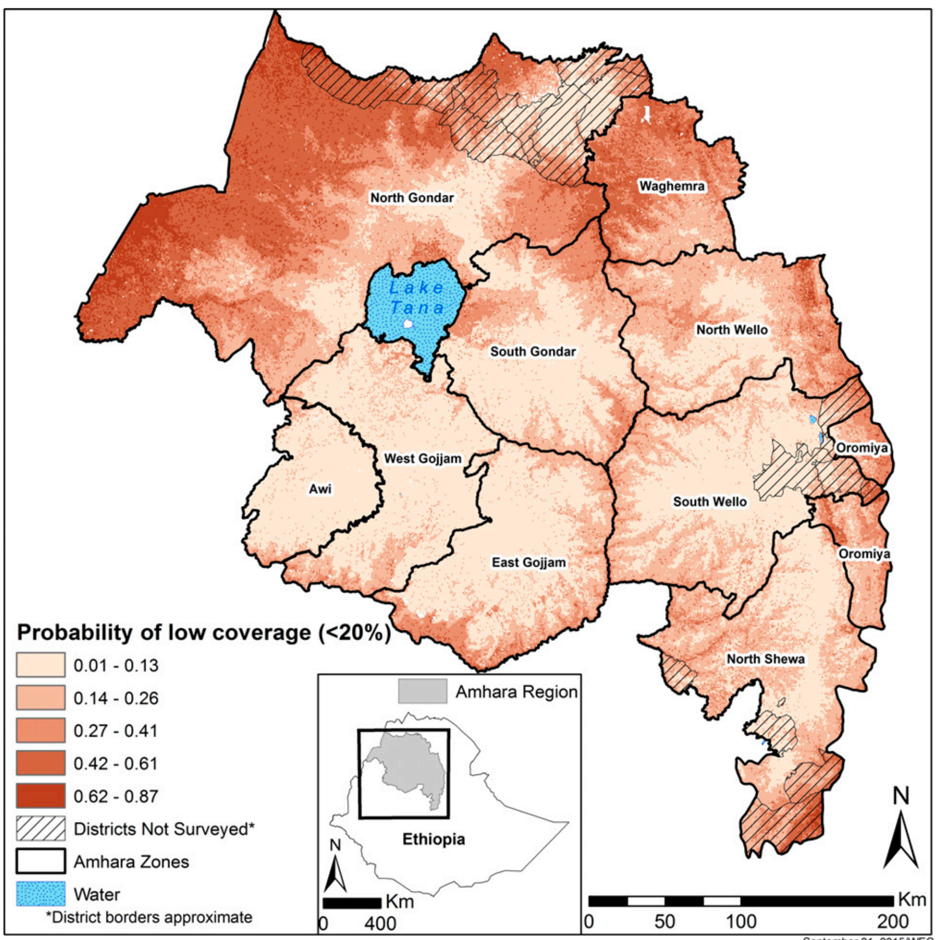

FIGURE 5. Map of model-predicted probability of low community sanitation coverage (< 20%) in Amhara Region, Ethiopia, 2011–2014, based on selected environmental and sociodemographic factors and adjusted for survey season and year.

Soil conditions related to instability and rock content were identified in a global review as challenges to both construction and durability of household latrines.[10] The recommended depth for pit latrines is approximately 2 m, though the specific pit volume needed depends on intended lifespan, the number of users, and anal cleansing materials used.[8] In their guide to onsite sanitation, Franceys and others described how cohesive soils may appear self-supporting when first excavated, but over time bonding properties of the soil may be lost, making it almost impossible to predict if or when soil may collapse.[44] At approximately 1,880 kg/m^3, soil's weight alone could exert extreme pressures on pit walls, which would be exacerbated by other factors, such as natural zones of weakness, water content, weather conditions, and the depth of excavation, that influence the stability of excavation walls.[41]

The stability of latrine pits influences both adoption and sustainability of latrines in sub-Saharan Africa. In Ghana, Jenkins and Scott described how soil conditions were an external barrier to sanitation adoption that operated late in the decision process, after households show preference and intention, by impeding construction.[12] Similarly, in Benin and Tanzania, individuals motivated to adopt sanitation reported that unsuitable soil conditions were an obstacle.[13,45] After sanitation has been adopted, collapse of latrine pits due to weak soil structure, particularly during the rainy season, was reported as a problem hindering sustainable sanitation coverage in Kenya.[11]

In addition to the influence of poor soil conditions, heavy rain and flooding can exacerbate structural weaknesses of the rudimentary onsite sanitation facilities typically found in sub-Saharan Africa.[46] A follow-up study of 666 latrines provided in the Gambia documented that 77 of the latrines collapsed over the course of two wet seasons.[16] The authors of that study described how sandy soil became liquefied, causing the latrines' cement slab and ring block to sink or collapse under their own weight. In contrast to our hypothesis, we found that areas with higher annual rainfall had significantly lower odds of low sanitation coverage, adjusting for other factors in the model. As such, annual rainfall may predict sanitation coverage, but it may not reflect mechanisms leading to low sanitation. An alternative predictor of low sanitation for future studies may be rainfall intensity. Whether the surveys in our study were conducted during the main rainy season, from June to September, was found to be strongly associated with higher odds of low sanitation coverage, which could also reflect the influence of rainfall on coverage estimates. This measure was included as a control variable. Research and evaluation activities to estimate sanitation coverage levels using household surveys should consider the season in which surveys were conducted.

The indicator for low-lying areas with higher moisture potential highlighted areas on the northeastern and northern edges of Lake Tana. Of all the predictors, this measure had the strongest association with low sanitation coverage. These areas flood during the rainy season, and inhabitants reportedly continue to reside in the area during these times. In Kenya, flooding was a major constraint to sanitation coverage in certain districts, where latrines were reported to fill up and overflow or collapse during the rainy season. After such events, residents reported preferring open defecation to the difficult and expensive repairs needed for their latrines.[11]

Flooding, high rainfall, and soil type have been described previously as challenges for sanitation.[11–13,16,45,46] Our study quantifies the association of these factors with frequency of low sanitation coverage in communities across Amhara Region. The FMOH acknowledged the influence of environmental factors like soil structure, topography, and climate in its 2005 National Hygiene and Sanitation Strategy (NHSS).[47] Therefore, what may be needed in areas that are identified to be at high risk of low sanitation coverage is information and capacity building on appropriate and affordable sanitation solutions for the environmental challenges they face. For example, it is generally recommended that all latrine pits are lined to their full depth to prevent collapse.[44] Pit collapse can be hazardous to the person excavating, disturbing to users, and can discourage households from sustaining improved sanitation practices.[41,44,47] Stabilized soil blocks are an environmentally sustainable and affordable alternative to fired brick that are gaining recognition in east Africa and could be used for lining pit latrines.[48]

Alternative approaches may also be needed to provide sanitation for populations living in flood-prone areas. A full review of solutions is beyond the scope of this discussion, but some sustainable onsite sanitation options exist for flood-prone areas.[49] Raised pit latrines have been constructed in the riverine areas of Bangladesh, where communities inhabit the small islands that are left as rivers subside and are periodically inundated when rivers rise again.[44] An intervention program there raised houses, tube wells, and pit latrines on earthen "plinths" to protect them from floodwaters and reported success extending sanitation coverage and asset protection to the residents of these communities.[50] The intervention provided a subsidy for families to raise these plinths. Flexibility for creative finance options, including subsidies where adverse ground conditions were confirmed, is also possible within the NHSS and protocol.[47,51] The model developed by our study could assist sanitation authorities and programs to identify areas where alternate interventions may be needed, based on challenging environmental conditions.

Our selected model had a reasonably high ability to classify communities as having low sanitation coverage by combining data from trachoma impact surveys and publicly available environmental data sources. This model was validated among communities distinct in both time and location. Interestingly, model discrimination was only slightly improved by inclusion of a measure of community wealth. This community wealth measure was derived from information collected at great cost during extensive population-based, household surveys. Comparing model prediction utilizing only community-level versus household-level information alongside measures of pit latrine construction quality and maintenance would be an interesting next step.

To our knowledge, a predictive model for sanitation coverage using environmental and sociodemographic conditions has not been developed previously. Care was taken to choose optimal data sources, but their accuracy at the scale used might be limited. For example, the HWSD was originally compiled from different sources with varying quality per region, though east Africa was covered by more reliable data sources.[24] Regardless, the predictive model documented here can only improve in the future as more environmental information becomes available.

The map based on the predictive model highlights spatial heterogeneity in the probability of low sanitation coverage across Amhara Region. Additional studies to examine and confirm the reasons for low coverage in these areas are warranted. A recent study described high geographic inequality for improved sanitation within countries across sub-Saharan Africa.[52] Pullan and other suggested that areas with lowest access to sanitation are likely the most challenging in terms of environmental conditions.[52] Our results demonstrate that environmental conditions, independent of community wealth, significantly predict low sanitation coverage, and that a measure of community wealth improved prediction only slightly. Areas hyperendemic for trachoma, like Amhara Region, or where interventions are being scaled up could benefit from additional tools to help prioritize and target control efforts. This prediction tool may benefit programs in Amhara Region and elsewhere by improving the targeting of information and resources to bring about practical and sustainable sanitation improvements to these areas most in need.

Received December 12, 2015. Accepted for publication May 31, 2016.

Published online July 18, 2016.

Note: Supplemental tables appear at www.ajtmh.org.

Acknowledgments: We would like to thank the Amhara National Regional Health Bureau and health offices, The Carter Center staff, field teams, and study supervisors. We are especially grateful to the residents of selected communities, who gave freely of their time to participate. We acknowledge Gonzalo Vazquez-Prokopec, Julie Clennon, Donal Bisanzio, and Cecile Janssens at Emory University for instruction and early discussions on study methods and approach.

Financial support: William Oswald was supported by the Emory University Laney Graduate School, ARCS Foundation Atlanta, and the Global 2000 program of The Carter Center. This study was supported by the Lions-Carter Center Sight-First Initiative and was made possible by the generous support of the American people through the U.S. Agency for International Development (USAID) and the ENVISION project led by RTI International in partnership with The Carter Center.

Authors' addresses: William E. Oswald, W. Dana Flanders, and Michael R. Kramer, Department of Epidemiology, Emory University, Atlanta, GA, E-mails: william.oswald@lshtm.ac.uk, wflande@emory.edu, and mkram02@emory.edu. Aisha E. P. Stewart and Elizabeth K. Callahan, Trachoma Control Program, The Carter Center, Atlanta, GA, E-mails: aisha.stewart@cartercenter.org and kelly.callahan@cartercenter.org. Tekola Endeshaw, Mulat Zerihun, Birhanu Melaku, Eshetu Sata, Demelash Gessesse, and Zerihun Tadesse, Trachoma Control Program, The Carter Center, Addis Ababa, Ethiopia, E-mails: tekola.endeshaw@cartercenter.org, mulat.zerihun@cartercenter.org, birhanu.melaku@cartercenter.org, eshetu.sata@cartercenter.org, and zerihun.tadesse@cartercenter.org. Tesfaye Teferi, Trachoma Control Program, The Carter Center, Addis Ababa, Ethiopia, and International Trachoma Initiative, Addis Ababa, Ethiopia, E-mail: tteferi@taskforce.org. Birhan Guadie, Amhara Regional Health Bureau, Bahir Dar, Ethiopia, E-mail: guadiebirhan@yahoo.com. Jonathan D. King, Trachoma Control Program, The Carter Center, Atlanta, GA, and Control of Neglected Tropical Diseases, World Health Organization, Geneva, Switzerland, E-mail: king_jond@yahoo.com. Paul M. Emerson, International Trachoma Initiative, Atlanta, GA, and Trachoma Control Program, The Carter Center, Atlanta, GA, E-mail: pemerson@taskforce.org. Christine L. Moe, Hubert Department of Global Health, Center for Global Safe Water, Emory University, Atlanta, GA, E-mail: clmoe@emory.edu. Thomas F. Clasen, Department of Environmental Health, Emory University, Atlanta, GA, E-mail: thomas.f.clasen@emory.edu.

REFERENCES

1. UNICEF/WHO, 2015. *Progress on Sanitation and Drinking-Water: 2015 Update and MDG Assessment*. World Health Organization and UNICEF, Geneva, Switzerland.
2. King JD, Endeshaw T, Escher E, Alemtaye G, Melaku S, Gelaye W, Worku A, Adugna M, Melak B, Teferi T, Zerihun M, Gesese D, Tadesse Z, Mosher AW, Odermatt P, Utzinger J, Marti H, Ngondi J, Hopkins DR, Emerson PM, 2013. Intestinal parasite prevalence in an area of Ethiopia after implementing the SAFE strategy, enhanced outreach services, and health extension program. *PLoS Negl Trop Dis 7:* e2223.
3. Ross RK, King JD, Damte M, Ayalew F, Gebre T, Cromwell EA, Teferi T, Emerson PM, 2011. Evaluation of household latrine coverage in Kewot woreda, Ethiopia, 3 years after implementing interventions to control blinding trachoma. *Int Health 3:* 251–258.
4. O'Loughlin R, Fentie G, Flannery B, Emerson PM, 2006. Follow-up of a low cost latrine promotion programme in one district of Amhara, Ethiopia: characteristics of early adopters and non-adopters. *Trop Med Int Health 11:* 1406–1415.
5. Kar K, Chambers R, 2008. *Handbook on Community-Led Total Sanitation*. London, United Kingdom: Plan UK.
6. Faris K, Rosenbaum J, 2011. *Learning by Doing: Working at Scale in Ethiopia*. Washington, DC: Water and Sanitation Program, World Bank.
7. Banteyerga H, 2011. Ethiopia's health extension program: improving health through community involvement. *MEDICC Rev 13:* 46–49.
8. Wagner EG, Lanoix JN, 1958. *Excreta Disposal for Rural Areas and Small Communities.* WHO Monograph Series. Geneva, Switzerland: World Health Organization.
9. Dreibelbis R, Winch PJ, Leontsini E, Hulland KR, Ram PK, Unicomb L, Luby SP, 2013. The integrated behavioural model for water, sanitation, and hygiene: a systematic review of behavioural models and a framework for designing and evaluating behaviour change interventions in infrastructure-restricted settings. *BMC Public Health 13:* 1015.
10. O'Connell K, 2014. *What Influences Open Defecation and Latrine Ownership in Rural Households?: Findings from a Global Review*. Washington, DC: Water and Sanitation Program, World Bank.
11. Jackson B, 2004. *Sanitation and Hygiene in Kenya: Lessons on What Drives Demand for Improved Sanitation*. Washington, DC: Water and Sanitation Program, The World Bank.
12. Jenkins MW, Scott B, 2007. Behavioral indicators of household decision-making and demand for sanitation and potential gains from social marketing in Ghana. *Soc Sci Med 64:* 2427–2442.
13. Sara S, Graham J, 2014. Ending open defecation in rural Tanzania: which factors facilitate latrine adoption? *Int J Environ Res Public Health 11:* 9854–9870.
14. Jenkins MW, Sugden S, 2006. *Rethinking Sanitation: Lessons and Innovation for Sustainability and Success in the New Millennium*. Human Development Report 2006. New York, NY: Human Development Report Office, UNDP.
15. Jenkins MW, Curtis V, 2005. Achieving the 'good life': why some people want latrines in rural Benin. *Soc Sci Med 61:* 2446–2459.
16. Simms VM, Makalo P, Bailey RL, Emerson PM, 2005. Sustainability and acceptability of latrine provision in The Gambia. *Trans R Soc Trop Med Hyg 99:* 631–637.
17. Jenkins MW, Cairncross S, 2010. Modelling latrine diffusion in Benin: towards a community typology of demand for improved sanitation in developing countries. *J Water Health 8:* 166–183.
18. Bewket W, Conway D, 2007. A note on the temporal and spatial variability of rainfall in the drought-prone Amhara region of Ethiopia. *Int J Climatol 27:* 1467–1477.
19. European Commission JRC, 2003. *Global Land Cover 2000 Database*. Available at: http://forobs.jrc.ec.europa.eu/products/glc2000/glc2000.php. Accessed July 9, 2016.

20. Central Statistical Agency (CSA), 2007. *2007 Population and Housing Census of Ethiopia*. Addis Ababa, Ethiopia: Central Statistical Agency.
21. Turner AG, Magnani RJ, Shuaib M, 1996. A not quite as quick but much cleaner alternative to the Expanded Programme on Immunization (EPI) Cluster Survey design. *Int J Epidemiol 25:* 198–203.
22. UNICEF, 2006. *Monitoring the Situation of Children and Women. Multiple Indicator Cluster Survey Manual 2005*. New York, NY: UNICEF.
23. King JD, Buolamwini J, Cromwell EA, Panfel A, Teferi T, Zerihun M, Melak B, Watson J, Tadesse Z, Vienneau D, Ngondi J, Utzinger J, Odermatt P, Emerson PM, 2013. A novel electronic data collection system for large-scale surveys of neglected tropical diseases. *PLoS One 8:* e74570.
24. FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012. *Harmonized World Soil Database*. Available at: http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/index.html?sb=1. Accessed February 10, 2015.
25. FAO, 2007. *Digital Soil Map of the World*. Rome, Italy: Food and Agriculture Organization of the United Nations.
26. USGS, 2009. *Africa Ecosystems Mapping Project*. Available at: http://rmgsc.cr.usgs.gov/outgoing/ecosystems/AfricaData/. Accessed February 27, 2015.
27. Cress JJ, Sayre R, Comer P, Warner H, 2009. *Topographic Moisture Potential of the Conterminous United States: U.S. Geological Survey Scientific Investigations Map 3086, Scale 1:5,000,000*. Available at: http://pubs.usgs.gov/sim/3086/. Accessed February 27, 2015.
28. Jarvis A, Reuter HI, Nelson A, Guevara E, 2008. *Hole-filled SRTM for the Globe Version 4*. Available at: http://www.diva-gis.org. Accessed February 10, 2015.
29. AfSIS, 2011. *Annual Normalized Difference Vegetation Index*. Available at: ftp://africagrids.net/250m/MOD13Q1/NDVI. Accessed February 4, 2015.
30. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A, 2005. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol 25:* 1965–1978.
31. Center for International Earth Science Information Network (CIESIN) Columbia University, Information Technology Outreach Services (ITOS), University of Georgia, 2013. *Global Roads Open Access Data Set, Version 1 (gROADSv1)*. Available at: http://dx.doi.org/10.7927/H4VD6WCT. Accessed January 20, 2015.
32. Bright EA, Coleman PR, Rose AN, Urban ML, 2012. *LandScan 2011*. Available at: http://www.ornl.gov/landscan/. Accessed April 17, 2015.
33. Kleinbaum DG, Klein M, 2010. *Logistic Regression: A Self-Learning Text*, 3rd edition. New York, NY: Springer.
34. Vazquez-Prokopec GM, Spillmann C, Zaidenberg M, Gurtler RE, Kitron U, 2012. Spatial heterogeneity and risk maps of community infestation by *Triatoma infestans* in rural northwestern Argentina. *PLoS Negl Trop Dis 6:* e1788.
35. Burnham KP, Anderson DR, 2002. *Model Selection and Multimodal Inference: A Practical Information-Theoretic Approach*. New York, NY: Springer.
36. Rosenberg ES, 2014. proc_EVW.sas [Program for SAS software]. Atlanta, GA: Rollins School of Public Health, Emory University.
37. Flanders WD, Tucker G, Krishnadasan A, Martin D, Honig E, McClellan WM, 1999. Validation of the pneumonia severity index. *J Gen Intern Med 14:* 333–340.
38. Janssen KJ, Moons KG, Kalkman CJ, Grobbee DE, Vergouwe Y, 2008. Updating methods improved the performance of a clinical prediction model in new patients. *J Clin Epidemiol 61:* 76–86.
39. Pan W, 2001. Akaike's information criterion in generalized estimating equations. *Biometrics 57:* 120–125.
40. DeLong ER, DeLong DM, Clarke-Pearson DL, 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics 44:* 837–845.
41. Johnson LK, 2004. Earth collapse of soil pits and trench excavations. Muckel GB, ed. *Understanding Soil Risks and Hazards: Using Soil Survey to Identify Areas with Risks and Hazards to Human Life and Property*. Lincoln, Nebraska: United States Department of Agriculture, Natural Resources Conservation Service, National Soil Survey Center.
42. USDA, 1993. *Soil Survey Manual*. Washington, DC: Soil Survey Division Staff, U.S. Department of Agriculture.
43. FAO, 1988. *FAO/UNESCO Soil Map of the World, Revised Legend, with Corrections and Updates*. World Soil Resources Report 60, FAO, Rome. Reprinted with updates as Technical Paper 20. Wageningen, Germany: ISRIC.
44. Franceys R, Pickford J, Reed R, 1992. *A Guide to the Development of On-Site Sanitation*. Geneva, Switzerland: World Health Organization.
45. Jenkins M, 2004. *Who Buys Latrines, Where and Why?* Nairobi, Kenya: Water and Sanitation Program Africa, World Bank.
46. Grimason AM, Davison K, Tembo KC, Jabu GC, Jackson MH, 2000. Problems associated with the use of pit latrines in Blantyre, Republic of Malawi. *J R Soc Promot Health 120:* 175–182.
47. FDRE/MOH, 2007. *Federal Democratic Republic of Ethiopia Ministry of Health National Hygiene and Sanitation Strategy 2005*. Washington, DC: Water and Sanitation Program, World Bank.
48. Perez A, 2009. *Interlocking Stabilised Soil Blocks: Appropriate Earth Technologies in Uganda*. Nairobi, Kenya: United Nations Human Settlements Programme (UN-Habitat).
49. Morshed G, Sobhan A, 2010. The search for appropriate latrine solutions for flood-prone areas of Bangladesh. *Waterlines 29:* 236–245.
50. O'Donnell E, 2015. *Plinth Performance Review*. Bogra, Bangladesh: Chars Livelihood Programme.
51. FDRE/MOH, 2007. *Federal Democratic Republic of Ethiopia Ministry of Health National Hygiene and "On-Site" Sanitation Protocol 2006*. Washington, DC: Water and Sanitation Program, World Bank.
52. Pullan RL, Freeman MC, Gething PW, Brooker SJ, 2014. Geographical inequalities in use of improved drinking water supply and sanitation across sub-Saharan Africa: mapping and spatial analysis of cross-sectional survey data. *PLoS Med 11:* e1001626.