

Quantifying the Impact of Accessibility on Preventive Healthcare in Sub-Saharan Africa Using Mobile Phone Data

Amy Wesolowski,^{a,b,c} Wendy Prudhomme O'Meara,^d Andrew J. Tatem,^{e,f} Samson Ndege,^g
Nathan Eagle,^{b,h} and Caroline O. Buckee^{b,c}

Background: Poor physical access to health facilities has been identified as an important contributor to reduced uptake of preventive health services and is likely to be most critical in low-income settings. However, the relation among physical access, travel behavior, and the uptake of healthcare is difficult to quantify.

Methods: Using anonymized mobile phone data from 2008 to 2009, we analyze individual and spatially aggregated travel patterns of 14,816,521 subscribers across Kenya and compare these measures to (1) estimated travel times to health facilities and (2) data on the uptake of 2 preventive healthcare interventions in an area of western Kenya: childhood immunizations and antenatal care.

Results: We document that long travel times to health facilities are strongly correlated with increased mobility in geographically isolated areas. Furthermore, we found that in areas with equal physical access to healthcare, mobile phone-derived measures of mobility predict which regions are lacking preventive care.

Conclusions: Routinely collected mobile phone data provide a simple and low-cost approach to mapping the uptake of preventive healthcare in low-income settings.

(*Epidemiology* 2015;26: 223–228)

Submitted 23 July 2014; accepted 14 October 2014.

From the ^aDepartment of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA; ^bDepartment of Epidemiology, Harvard School of Public Health, Boston, MA; ^cCenter for Communicable Disease Dynamics, Harvard School of Public Health, Boston, MA; ^dDepartment of Medicine, Duke University and Duke Global Health Institute, Durham, NC; ^eDepartment of Geography and Environment, University of Southampton, Highfield, Southampton, United Kingdom; ^fFogarty International Center, National Institutes of Health, Bethesda, MD; ^gSchool of Public Health Moi University, Eldoret, Kenya; and ^hDepartment of Computer Science, Northeastern University, Boston, MA.

A.W. was supported by the NSF (#0750271) and the James S. McDonnell Foundation. A.J.T. acknowledges support from Bill & Melinda Gates Foundation (#49446 and #OPP1032350), NIH/NIAID (U19AI089674), and the RAPIDD program of the Science & Technology Directorate, Department of Homeland Security, and the Fogarty International Center, National Institutes of Health. C.O.B. was supported by the Models of Infectious Disease Agent Study program (cooperative agreement 1U54GM088558). The other authors have no conflicts of interest to report. The Home-based Counseling and Testing program was supported by grants from Abbott Laboratories, the Purpleville Foundation, and the Global Business Coalition. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com). This content is not peer-reviewed or copy-edited; it is the sole responsibility of the authors.

Correspondence: Caroline O. Buckee, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115. E-mail: cbuckee@hsph.harvard.edu.

Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives 3.0 License, where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially.

ISSN: 1044-3983/15/2602-0223

DOI: 10.1097/EDE.0000000000000239

In Sub-Saharan Africa, a large number of deaths in children under 5 years of age result from a small number of common causes, such as diarrhea, malaria, and poor prenatal care.¹ These deaths are largely preventable through health services that are often available at local health facilities, including childhood immunizations and antenatal care for pregnant women.^{2–4} Variable uptake of available health interventions can undermine healthcare programs, however, and plays an important role in child and maternal mortality, particularly in Sub-Saharan Africa.^{5–9}

Poor physical access to health facilities is thought to be a critical determinant in these outcomes. Households further from health facilities have higher incidence of child mortality and malaria, for example, and are less likely to seek treatment for fevers, although these relationships are heterogeneous and appear to be stronger in very rural areas.^{5–17} The relation between the travel time to healthcare facilities and actual travel behavior of a community is unclear, however, and 2 contradicting hypotheses have been suggested. First, a broad spatial analysis of the determinants of poverty in Kenya found a positive correlation between access to health facilities, frequency of travel, and wealth, suggesting that more wealthy populations are both more able to travel frequently and have easy access to health facilities.¹⁸ However, it has also been suggested that the inverse is true; poor physical access imposes a higher travel burden on people to reach the resources they need, whether economic or health related. Persons in areas with poor access to health services, which are associated with poor health outcomes, may need to travel frequently and relatively far.¹⁹

Human mobility—the spatial range and frequency of individual travel—reflects a range of basic needs, social incentives, and economic constraints and inherently encompasses multiple modalities. Therefore, it remains unclear how measures of mobility, which reflect more than just geographic location and physical infrastructure for travel, correlate with access to healthcare or resulting health outcomes.¹⁶ Disentangling the relationships among healthcare access, travel behavior, and health outcomes is challenging. Household surveys and travel diaries are often used to measure variation in travel times to public resources such as health facilities,^{20–22} but these are limited in scope and hard to generalize across regions. Geospatial techniques have been used to derive maps of approximate distances and travel times to nearby health facilities,²⁰ but these do not provide insights into how physical access affects human behavior. Due to the conflicting hypotheses described above and the lack of evidence available, identifying the factors underlying poor healthcare uptake and outcomes hinders the ability of policy makers to design programs that help those most in need.

The availability of de-identified mobile phone call detail records provides a new source of data on human mobility on unprecedented scales.^{23,24} Mobile phone operators log cell tower locations whenever a person uses a mobile phone. These call detail records, once anonymized, provide detailed, longitudinal information about the travel patterns of millions of subscribers and offer unique opportunities for directly measuring the impact of geographic isolation on human travel behavior. Here, we combine an analysis of the geographic variation in mobility patterns of 14,816,521 anonymous mobile phone subscribers in Kenya with modeled estimates of travel times to health facilities, generated through standard cost–distance-based spatial analysis methods on various scales. We compare these with direct, geocoded measures of childhood immunization and antenatal care uptake in Western Kenya, measured by household surveys during a home-based HIV counseling and testing program.²⁵ This provides a direct measure of the impact of physical access to healthcare on mobility and health outcomes and a cost-effective new approach to mapping human behaviors underlying health outcomes.

METHODS

Mobile Phone Data

Mobile phone data were provided by the incumbent mobile phone provider during the time of data collection. Anonymized call detail records included the cell tower location (from one of 11,920 routing tower locations), sender, receiver, and timing of calls or SMS communications from 14,816,512 subscribers from June 2008 through June 2009 (with February 2009 missing from the data set). Subscribers were represented as unique anonymized hashed IDs. In total, over 12 billion mobile phone communications were recorded. From these data, we modeled the longitudinal travel behavior

of each subscriber. We used a previously developed measure of individual mobility, the radius of gyration, to measure the mobility of each subscriber. This measure encompasses both the range and frequency traveled by a person into a single measure, with higher values characterizing more mobile subscribers and lower values resulting from subscribers who do not travel as frequently or as far. We calculated a primary location (center of mass) for each subscriber using the tower location of each mobile phone communication, based on the most common location over the course of the year. For each of the subscriber's records in the call detail records, we then consider the frequency and distance to other towers. We calculated average radius of gyration from subscribers assigned a primary location at each mobile phone tower (eAppendix, <http://links.lww.com/EDE/A870>).^{23,24} Explicitly, radius of gyration is calculated as follows^{23,26}:

$$r_g^a(t) = \sqrt{\frac{1}{n_c^a(t)} \sum_{i=1}^{n_c^a} (\bar{r}_i^a - \bar{r}_{cm}^a)^2}$$

where \bar{r}_i^a represents the $i = 1, \dots, n_c^a(t)$ tower locations recorded for subscriber a at time t and $\bar{r}_{cm}^a = \frac{1}{n_c^a} \sum_{i=1}^{n_c^a} \bar{r}_i^a$ represents the subscriber's center of mass trajectory. Yearly radius-of-gyration values for subscribers were aggregated to provide distributions of mobility for populations on 2 spatial scales: the county level and the individual cell tower level (11,920 total towers).

Travel-time Data

Geospatial techniques were used to estimate the travel time to the nearest health facility. The population distribution was calculated using land cover and census data, as was done by Linard and colleagues.²⁰ In brief, travel times between pairs of locations was calculated using a cost–distance algorithm that computes the “cost” of traveling on a regular raster grid based on information about the transportation network (obtained from the Kenyan National Bureau of Statistics), land cover data (<http://www.africover.org>), and topography data (<http://srtm.csi.cgiar.org/>).²⁰ The locations of all national health facilities were obtained from Noor et al.²⁷

Survey Data

Survey data were collected from 5 districts in western Kenya that are part of the catchment area for the Academic Model Providing Access to Healthcare.²⁵ Between 2009 and 2011, a door-to-door HIV counseling and testing campaign was conducted in western Kenya. Most individuals (95%) agreed to be counseled and tested across 5 districts. The program collected geocoded information about pregnancy, attendance at antenatal clinics, immunization of children, and basic socioeconomic information. Individual surveys from these study sites ($n = 273,213$) were aggregated to the household level ($n = 78,882$). Households were then aggregated to

sublocations (n = 89) where the mean mobility variable and travel time to a health facility were calculated.

We considered 2 primary outcome variables related to the uptake of preventive healthcare interventions—completed childhood immunizations and antenatal care for pregnant women (see eTable S1; <http://links.lww.com/EDE/A870>). For a child to be fully immunized, they should have received Bacillus Calmette–Guérin vaccine at birth and measles at 9 months. A household was considered to be missing immunizations if there was at least 1 child over 1 year who had not had any of these immunizations. Households were considered to be missing antenatal care if there was a pregnant woman who reported not receiving antenatal care. For each sublocation, we calculated the percentage of eligible households (ie, households with children or pregnant women) who were lacking either immunizations or antenatal care.

We aggregated household-level point estimates of both mobility and travel times to obtain sublocation averages. To aggregate mobility values to sublocations for comparison, each household was assigned the mobility value of the nearest mobile phone tower and a corresponding travel-time estimate.

RESULTS

On large spatial scales, people living in counties with poor physical access to healthcare were also the most mobile, in terms of both distance and frequency. Figure 1A and B shows the Kenyan road network (see eFigure 1; <http://links.lww.com/>

EDE/A870 for continuous maps), colored by the observed mobility of local populations, and the relationship between mobility and travel times to health facilities aggregated at the county level (eFigures 1 and 2; <http://links.lww.com/EDE/A870> for similar analyses for population centers and schools). Radius-of-gyration values decreased nonlinearly as county population density increased and travel times to reach resources were reduced (eFigure 2; <http://links.lww.com/EDE/A870>: fit for population density; adjusted $R^2 = 0.5$, sum of squared errors = 3.3, root-mean-square error = 0.29; correlation with average travel time to the nearest health facility, adjusted $R^2 = 0.5$). eFigure 3 (<http://links.lww.com/EDE/A870>) shows the impact of mobile phone tower density on radius gyration values. This finding is consistent with the hypothesis that remote communities experience an increased burden of travel.

Substantial heterogeneity in the relationship between radius of gyration and travel time to health facilities was observed on the smaller spatial scale of the individual cell tower, however. For example, Figure 2A and B compares observed mobility with estimated travel times to the nearest health facility for a region of Western Kenya (eFigure 4; <http://links.lww.com/EDE/A870> for similar analyses comparing health facilities to schools). Here, we observed that populations with high and low average mobility are adjacent to each other, despite similar estimated travel times to the nearest health facility. Thus, while communities can be defined as equally remote mobility, patterns are locally heterogeneous.

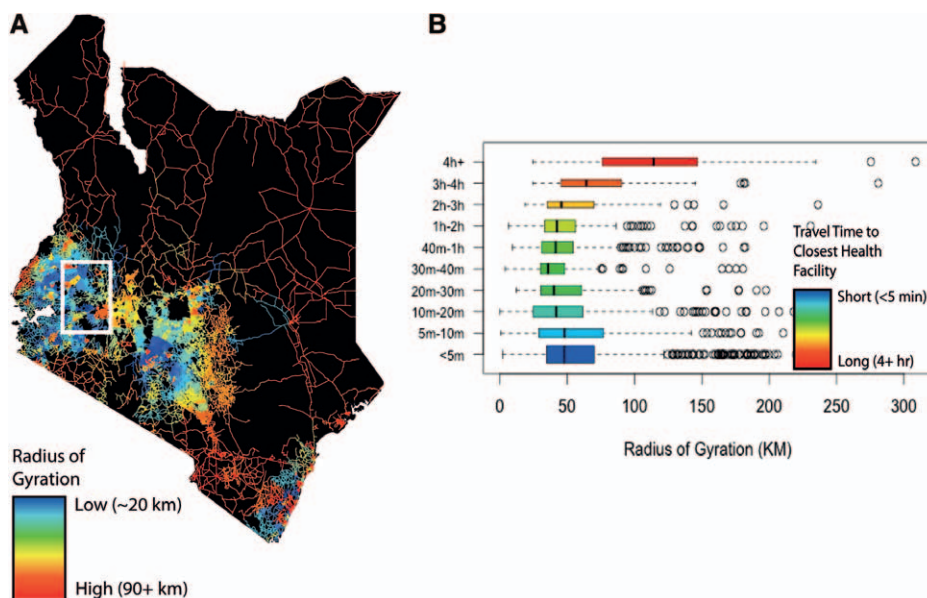


FIGURE 1. The relationship between radius of gyration and the travel time to the nearest health facility. A, Roads in Kenya are colored according to the radius-of-gyration value (median of resident values) for the nearest mobile phone tower (broken into 30 quantiles). Red roads have the highest radius-of-gyration values (90+ km), whereas blue roads have the lowest values (~20 km) (see eAppendix and eFigure 4; <http://links.lww.com/EDE/A870> for the impact of varying mobile phone tower density on radius-of-gyration estimates). B, The relationship between travel time to the nearest health facility and radius of gyration is shown grouped by travel time. For subscriber estimates aggregated to the tower level, the median radius of gyration is generally lower for towers with shorter travel times to the nearest health facility than for towers with longer travel times.

To examine whether these heterogeneities in mobility were associated with health outcomes, we analyzed income and health data from 5 districts in western Kenya (78,882 households). We identified households missing preventive care—either incomplete immunization among children age 1–5 years or a lack of antenatal care among pregnant women. Most children completed their immunization schedule in the majority of eligible households (2% of eligible households had children missing immunizations, $n = 686$). Almost 40% of households with pregnant women were missing antenatal care (39%, $n = 1,374$). The households missing care were

geographically heterogeneous, with higher percentages found in 2 of the study districts (Figure 3A). We aggregated household data to sublocations ($n = 89$) to compare with the mobile phone data. (On average, 1% sublocations were missing immunization, and 31% were missing antenatal care.)

Mobility estimates derived from mobile phone data were better able to predict the percentage of households in a sublocation missing immunizations or antenatal care than were estimates of travel times to health facilities (Figure 3B and eTable S2; <http://links.lww.com/EDE/A870>). The correlation coefficients for immunizations and health-facility

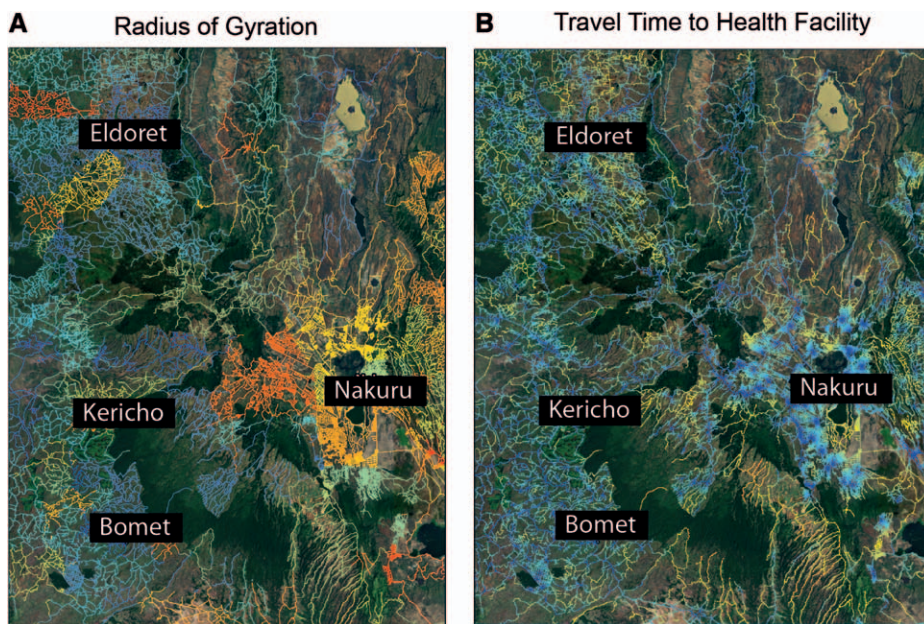


FIGURE 2. The radius-of-gyration values in a localized area. An area of Western Kenya showing radius-of-gyration values (A) and estimated travel times to the nearest health facility (B) is shown highlighting the substantial differences in travel behaviors that exist across similar levels of geographic access to health services.

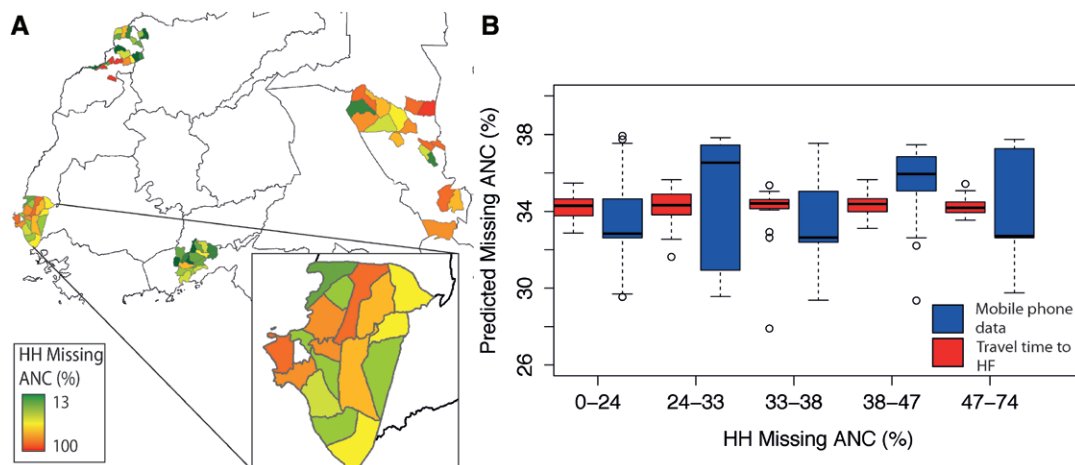


FIGURE 3. The relationship between mobility, travel times, and households missing preventive healthcare. A, The percentage of households (HHs) within each area's sublocation who reported missing antenatal care (ANC). The entire study area and a sample study site are shown. B, The predicted percentage of eligible HHs missing ANC per sublocation using travel times to the nearest health facility (HF) (red) or the mobility values from the mobile phone data (blue) (reduction in deviance from the mobility model, 3%; from the travel-time model, 0.41%). The travel-time data would predict that nearly every sublocation is missing the same percentage, whereas the mobile phone data provides more accurate estimates over a wider range.

travel time was 0.06 ($P = 0.6$), whereas average radius of gyration was correlated with a coefficient of -0.23 ($P = 0.03$). For antenatal care, there was no relationship between either measure of travel and uptake, although the relationship with mobility was somewhat stronger (Figure 2B) (health facility travel time, -0.06 [$P = 0.56$] and mobility, -0.14 [$P = 0.2$]). In sublocations of similar physical access to health facilities, increased mobility was associated with a higher percentage of households accessing preventive healthcare. (For sublocations within 30 minutes of a health facility, correlation coefficient = -0.21 [$P = 0.24$]; for sublocations between 40 minutes and 2 hours of a health facility, correlation coefficient = -0.38 [$P = 0.067$].) The strength of this relationship was not related to the average household wage hours and was stronger in areas further from health facilities, suggesting that the impact of access and mobility is greater in more remote communities. Thus, mobility can help explain heterogeneities in accessing care for populations with comparable physical travel times to health facilities, highlighting the ability of new data sources to describe variable uptake of preventive healthcare.

DISCUSSION

We propose that mobility measured by mobile phone data may reflect more about economic conditions and behavior than travel-time estimates, making it a more comprehensive measure that takes into account various aspects of health care access in addition to geography. Our results are consistent with the hypothesis that poor physical access places a high travel burden on persons living in remote regions, but that the most vulnerable among them are the least mobile. We propose that mobile phone call records provide valuable insights into human travel behaviors associated with poverty and access to healthcare and could be used to identify vulnerable populations where households are at risk for missing basic preventive care. Measuring and understanding the relationship between geographic isolation and travel is key to the targeting of development assistance that aims to alleviate health disparities and identify the role of travel and physical accessibility to health resources.¹⁶

Although mobile phone data have inherent biases, analytic tools for adjusting estimates are improving, and in this case, they may not have a major impact on estimates.^{26,28} We have previously shown that mobile phone ownership is generally still biased toward wealthier, urban-dwelling males. However, we observed mobile phone ownership in all income brackets, and these biases are rapidly diminishing.²⁶ Mobile phone data are increasingly being analyzed for epidemiologic studies through individual agreements between the operator and researchers, with anonymization and aggregation often taking place before sharing. The value of these approaches is likely to be most profound in low-income settings where data are scarce, but methods of analysis are still in their infancy, as are effective protocols for sharing the data more widely.^{29–31}

Despite the obstacles to the routine use of mobile phone data in public health research, however, we believe they represent a valuable and inexpensive source of information that can be used to identify areas facing substantial travel burden and to identify local variability in access that can exceed the spatial resolution of conventional access measures.

REFERENCES

- Rajaratnam JK, Marcus JR, Flaxman AD, et al. Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970–2010: a systematic analysis of progress towards Millennium Development Goal 4. *Lancet*. 2010;375:1988–2008.
- Black R, Cousens S, Johnson HL, Lawn JE, Rudan I. Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet*. 2010;375:1969–1987.
- Claeson M, Gillespie D, Mshinda H, Troedsson H, Victora CG; Bellagio Study Group on Child Survival. Knowledge into action for child survival. *Lancet*. 2003;362:323–327.
- Demicheli V, Rivetti A, Debalini MG, Di Pietrantonj C. Vaccines for measles, mumps and rubella in children. *Cochrane Database Syst Rev*. 2012;2:CD004407.
- Feikin DR, Nguyen LM, Adazu K, et al. The impact of distance of residence from a peripheral health facility on pediatric health utilisation in rural western Kenya. *Trop Med Int Health*. 2009;14:54–61.
- Gething PW, Noor AM, Zurovac D, et al. Empirical modelling of government health service use by children with fevers in Kenya. *Acta Trop*. 2004;91:227–237.
- Rutherford ME, Mulholland K, Hill PC. How access to health care relates to under-five mortality in sub-Saharan Africa: systematic review. *Trop Med Int Health*. 2010;15:508–519.
- Rutherford ME, Dockerty JD, Jasseh M, et al. Access to health care and mortality of children under 5 years of age in the Gambia: a case-control study. *Bull World Health Organ*. 2009;87:216–224.
- Schoeps A, Gabrysch S, Niamba L, Sié A, Becher H. The effect of distance to health-care facilities on childhood mortality in rural Burkina Faso. *Am J Epidemiol*. 2011;173:492–498.
- Gabrysch S, Cousens S, Cox J, Campbell OM. The influence of distance and level of care on delivery place in rural Zambia: a study of linked national data in a geographic information system. *PLoS Med*. 2011;8:e1000394.
- Gething PW, Johnson FA, Frempong-Ainguah F, et al. Geographic access to care at birth in Ghana: a barrier to safe motherhood. *BMC Public Health*. 2012;12:991.
- Guagliardo MF. Spatial accessibility of primary care: concepts, methods and challenges. *Int J Health Geogr*. 2004;3:3.
- Huerta Munoz U, Källestål C. Geographical accessibility and spatial coverage modeling of the primary health care network in the Western Province of Rwanda. *Int J Health Geogr*. 2012;11:40.
- Moisi JC, Gatakaa H, Noor AM, et al. Geographic access to care is not a determinant of child mortality in a rural Kenyan setting with high health facility density. *BMC Public Health*. 2010;10:142.
- O'Meara WP, Noor A, Gatakaa H, Tsofa B, McKenzie FE, Marsh K. The impact of primary health care on malaria morbidity—defining access by disease burden. *Trop Med Int Health*. 2009;14:29–35.
- Porter G. Transport services and their impact on poverty and growth in rural Sub-Saharan Africa: a review of recent and future research needs. *Transport Rev* 2014;34:25–45.
- Tanser F, Gijsbertsen B, Herbst K. Modelling and understanding primary health care accessibility and utilization in rural South Africa: an exploration using a geographical information system. *Soc Sci Med*. 2006;63:691–705.
- Okwi PO, Ndeng'e G, Kristjanson P, et al. Spatial determinants of poverty in rural Kenya. *Proc Natl Acad Sci U S A*. 2007;104:16769–16774.
- Bryceson DF, Maunder DAC, Mbara TC, Kibombo R, Davis ASC, Howe JDGF. *Sustainable Livelihoods, Mobility, and Access Needs*. TRL Report TRL544. 2003.
- Linard C, Gilbert M, Snow RW, Noor AM, Tatem AJ. Population distribution, settlement patterns and accessibility across Africa in 2010. *PLoS One*. 2012;7:e31743.

21. Kenya National Bureau of Statistics. *Kenya Demographic and Health Survey 2008–2009*. 2010. Available at <http://dhsprogram.com/pubs/pdf/FR229/FR229.pdf>.
22. Schlich R, Axhausen KW. Habitual travel behavior: evidence from a six-week travel diary. *Transportation*. 2003;30:13–36.
23. González MC, Hidalgo CA, Barabási AL. Understanding individual human mobility patterns. *Nature*. 2008;453:779–782.
24. Wesolowski A, Eagle N, Tatem AJ, et al. Quantifying the impact of human mobility on malaria. *Science*. 2012;338:267–270.
25. Kimaiyo S, Were MC, Shen C, et al. Home-based HIV counselling and testing in western Kenya. *East Afr Med J*. 2010;87:100–108.
26. Wesolowski A, Eagle N, Noor AM, Snow RW, Buckee CO. The impact of biases in mobile phone ownership on estimates of human mobility. *J R Soc Interface*. 2013;10:20120986.
27. Noor AM, Alegana VA, Gething PW, Snow RW. A spatial national health facility database for public health sector planning in Kenya in 2008. *Int J Health Geogr*. 2009;8:13.
28. Wesolowski A, Eagle N, Noor AM, Snow RW, Buckee CO. Heterogeneous mobile phone ownership and usage patterns in Kenya. *PLoS One*. 2012;7:e35319.
29. Blondel VD, Esch M, Chan C, et al. Data for development: the D4D challenge on mobile phone data. Available at: <http://arxiv.org/abs/1210.0137>. Accessed September 1, 2014.
30. de Montjoye YA, Smoreda Z, Trinquart R, Ziemlicki C, Blondel VD. D4D-Senegal: the second mobile phone data for development challenge. Available at: <http://arxiv.org/abs/1407.4885>. Accessed September 1, 2014.
31. de Montjoye YA, Hidalgo CA, Verleysen M, Blondel VD. Unique in the crowd: the privacy bounds of human mobility. *Sci Rep*. 2013;3:1376.