Data Bites Man: The Production of Malaria by Technology

PEDRO SANCHES, Dept. of Media Technology & Interaction Design, KTH Royal Institute of Technology, Sweden
BARRY BROWN, Dept. of Computer and Systems Sciences, Stockholm University, Sweden

Malaria surveillance is a practice concerned with collection and analysis of data. This paper presents an ethnographical account of an international team of researchers producing data about malaria in the Zanzibar archipelago. We show that malaria is increasingly an electronic entity, inextricably interwoven with the practices of data workers using ICT tools. The use of mobile ICT tools enables new data production practices that include situated coordination mechanisms such that 1) more people can be surveyed, including individuals not suffering from malaria but possibly carrying the parasite, 2) greater geographical areas can be covered, and 3) more data can be validated and included in malaria statistics. As electronic data, malaria builds and mobilizes diverse human, organizational, and infrastructural worlds around it, who must now be dedicated to its production, management, and care. We discuss implications for design of digital data collection tools that support the work of teams surveying malaria, as well as implications for disease surveillance methods and practices.

CCS Concepts: • **Human-centered computing → Empirical studies in collaborative and social computing**;

Additional Key Words and Phrases: data, surveillance, malaria, practices, coordination work

ACM Reference Format:
Pedro Sanches and Barry Brown. 2018. Data Bites Man: The Production of Malaria by Technology. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 153 (November 2018), 19 pages. https://doi.org/10.1145/3274422

1 INTRODUCTION

In this paper we focus on malaria monitoring. Digital data collection tools are increasingly replacing paper, and a number of them [22, 24] are specifically designed for low-resource settings. In the health literature these tools are broadly lauded as having had a positive impact on how health surveillance is conducted [12, 24]. Beside promising faster responses to detected events such as outbreaks, the introduction of information and communications technology (ICT) tools also promises more accuracy and validity in the data that is collected.

These tools are not neutral, however, and have in some important ways changed what malaria is. By changing how malaria is tracked and recorded, these tools have enabled the recording of malaria cases in settings that were previously hard to reach. By making tracking easier, cases that would have not been recorded are now recorded, resulting in an increase in malaria statistics, albeit in the context of numbers falling overall. More fundamentally, malaria becomes a socio-technical entity managed, recorded and manipulated through electronic form. While malaria, to those who suffer

Authors' addresses: Pedro Sanches, Dept. of Media Technology & Interaction Design, KTH Royal Institute of Technology, Stockholm, Sweden, sanches@kth.se; Barry Brown, Dept. of Computer and Systems Sciences, Stockholm University, Stockholm, Sweden, barry@dsv.su.se.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.
2573-0142/2018/11-ART153 $15.00
https://doi.org/10.1145/3274422

from it still manifests itself in their own or others sickness, to health professionals it increasingly also appears in electronic form. As we put it in this paper, rather than the malaria parasite becoming the center of medical practice, complex socio-technical mixes become central. As Ribes and Jackson [38] put it in their book chapter 'data bite man', data are:

> "temperamental and delicate creatures, whose existence and fraternity with one another depend on a complex assemblage of people, instruments, and practices dedicated to their production, management, and care. Like corn and flies before them, data demand and build the human, organizational, and infrastructural worlds around them."

As such, in this paper we focus not on surveillance and medical professionals, but rather the material assemblages [21] of less trained people, machines and organizations that make the measuring and counting of malaria possible. We present an ethnographic account of an international team of researchers producing data about malaria in the Zanzibar archipelago, a semi-autonomous region of Tanzania.

Our contribution is twofold. First, this paper presents a case study detailing current data production practices in the domain of active disease surveillance. These practices are described as the result of work of distributed actors working with and through mobile ICT tools. This constitutes in itself a contribution, as this domain has previously not been covered in CSCW. Second, we show that these new data production practices include situated coordination mechanisms such that 1) populations can be mobilized in real-time and more people can be surveyed, including individuals not suffering from malaria but possibly carrying the parasite, 2) more households can be included, and greater geographical areas can be covered, and 3) validation mechanisms can be made and therefore more data can be trusted and included in malaria statistics. This ethnographic account also shows that some of these practices often improvisational and contextual. We extend the literature on data production [34, 38], by showing that coordination work through mobile ICT tools, a familiar theme in CSCW [41], are essential for data to come to existence. Data on disease surveillance can be seen as a product of a socio-technical mix of formalized procedures and situated decisions by data workers and malaria researchers who must both adapt to, as well as transform the sociomaterial fabric of the people and areas and diseases they produce data on.

2 RELATED WORK

Disease surveillance denotes a practice of ongoing data collection, analysis and dissemination for the purposes of implementing and evaluating public health policies and practices [31]. Some work in disease surveillance in CSCW has been on developing systems using data sources such as mobile data to predict flu dynamics [17], evaluating the potential of public resources such as Wikipedia texts to measure global disease [37], or studying sociotechnical issues related to data collection in hospitals for surveillance purposes [25]. There is also extensive CSCW research on health care, which is focused mostly on clinical work and how it is transformed by computers [20]. When it comes to data production work related to diseases, the focus has mostly been on patient data (e.g.[27]), focusing on themes of personal engagements with data about oneself, along with narratives, personal conceptions of health and wellbeing as well as collaborative sensemaking between patients and caretakers. However, there is a lack of HCI research looking at the work of active surveillance for population-wide disease control. This type of work is mainly concerned with the act of data collection, which differs from passive disease surveillance (c.f.[25]) which makes use of data obtained in hospitals or health facilities from health-seeking patients.

This paper draws on and extends on literature that critically engages with the so called "commodity fiction" [38] of data: the belief that data and its origins, context and temporal history are separable, and therefore data in a "raw" form is out there waiting to be collected. Feinberg

[18] characterized data production as a result of design decisions in three stages: 1) conceptual infrastructure, which determines forms of measurement (e.g. parasite count in malaria) and scale (e.g. blood microliter), 2) collection processes, which implement the conceptual infrastructure, and 3) aggregation processes, which make use of a certain standard, such as time scale or a geographical unit, to integrate many individual collection acts. Rather than being neutral, each decision made, even before "raw" data is collected, enables certain possibilities for action and constrains others, making some things visible and others invisible [7, 10, 39]. Producing data is therefore not a neutral act. When data is produced by wide-governmental entities, citizens are viewed through lenses that capture high level aspects and metrics assumed to be the only ones relevant for the state interventions, sometimes disregarding everyday life context and local practices, even if the latter are often crucial for the success of the state interventions [26, 42] . Conversely, the act of producing data can be locally appropriated by activists for the purposes of enacting social change [34].

Data, particularly quantified data, is reactive [16, 34] in the way that it creates new "things", and relations between things. By doing so it causes people to think and act differently. For example, Shklovski and colleagues [44] show that mobile data infrastructures transform "location" from being a social, situational construct to a commodified object that becomes embedded in social and institutional process, and also changes them. Another example is the development of crime statistics [36] which changed the way we think about crime: from something stemming from bad human nature to a something than can be managed and predicted. Similarly, disease surveillance may change the ways we understand and act upon diseases, in a way that they become something to be pre-emptively managed as "threats" [10].

Even before data is inscribed in databases, already the work of producing data can have profound effects in the social world. The accounts by Ribes and Jackson [38] show that data can be seen as an actor that shapes individuals, organizations and technologies around producing it and maintaining it. Sustaining long-term scientific studies entails that individuals and organizations orient their practices to maintain comparability of data over the years and perform cascades of rituals which tie sample collections in the field to data points in a lab. On a larger scale, taking the example of the task of producing unique biometric-based identification numbers for over a billion people in India, Johri and Srinivasan [26] show how a focus on data production can mobilize geographically distributed and diverse actors such as ICT designers, enrollment agents, public agencies and private companies across large timespans and geographies. Over time, if not properly maintained, data can be become meaningless or lost [30]. Edwards [15] shows that a considerable amount of transformations and interventions are required for scientific data to flow across different storage devices, filesystems or disciplinary standards, a process described as overcoming "data friction". Furthermore, the context of data production, namely local organizational structures of the production sites, the practices that sustain data and are sustained by it affect how data is shared and flows across individuals and organizations [5, 48].

It is a crucial task to critically analyze data production practices to interrogate which things are represented and which are left invisible [39], and what effects these practices have on organizations that produce data and on the world they intend to represent [10]. Looking at practices in disease surveillance entails describing the material assemblage [21]: the network of people, microbes, machines and organizations that make public health surveillance possible. By describing the work of producing data using ICT tools, we unavoidably focus on familiar CSCW themes such as coordination work [41]. For example, the study of Pine and Mazmanian [35] on coordination work by healthcare workers highlights the unavoidable interplay between formal logics of action, often encoded in ICT tools [2], and emergent patterns of action. Broadly, this theme has been addressed by scholars such as Suchmann [46], showing that people adapt to the specifics of the situation at hand, using formal logics of actions as resources while also acting around them. Arguably [1] the

biggest challenge for CSCW is to design for highly flexible, nuanced, and contextualized human activities.

3 CONTEXT OF THE STUDY

The ethnographic account herein is derived from observations conducted alongside two distinct large-scale malaria survey studies, where the first author also participated as a technical consultant. These studies result from a long-term collaboration between Karolinska Institute (KI) and the Zanzibar Malaria Elimination Program (ZAMEP) through the Zanzibar Ministry of Health of over 10 years. Since its inception, ZAMEP has leveraged collaborations with international organizations and the central government to implement malaria control measures that have greatly reduced the incidence of malaria in Zanzibar. The collaboration with KI has resulted in yearly studies and interventions that assess the state of malaria in Zanzibar, and test new methods for achieving elimination combating the disease. With other collaborations (e.g. USAID), ZAMEP has also implemented surveillance systems [13] which enable public health facilities — those that provide point of care services such as clinics or hospitals — to report malaria cases, using mobile phones, to a centralized system, as well as actively following-up on detected cases.

The World Health Organization (WHO) recommends measures for countries to control and eliminate malaria that aim, at a first phase, at making malaria a problem manageable at the local government level, and then, at an elimination phase, to reduce local transmission to zero. Zanzibar is close to the elimination phase. Unlike the surrounding areas, which have high incidence of malaria all year around, malaria transmission in Zanzibar is low and seasonal, with most of the cases occurring after the rains in March-June, which provide breeding sites for the mosquitos which spread the disease. Central to malaria elimination are the notions of hotspot and hotpop [29]. Hotspots are geographically defined areas, typically smaller than 1km2, where malaria infections exceed the average level. Hotpops are high-risk population groups, typically with occupations or travel histories that make them more prone to be exposed to mosquito bites in malaria regions. Malaria elimination efforts are dependent on successfully identifying hotspots and hotpops.

Elimination phase areas have less than 1 case per population annually [32]. According to WHO, these areas are expected to devote resources to data collection and precise diagnostic methods, able to detect low-density malaria infections and provide rapid responses to outbreaks [32]. The goal is to limit the impact of the seasonal malaria by reducing and eventually eliminating local transmission. Like in many other areas of healthcare [20], ICT mobile tools are increasingly central to the daily work of health workers in development regions working with malaria surveillance and elimination. They are central to malaria elimination recommendations by researchers and international public health authorities [3], as they enable finer temporal and spatial resolutions in quantifying the disease, by e.g. enabling GPS coordinates to be collected when a malaria case is found, which was previously not possible with paper data collection forms. Other methods in the fight against malaria are the periodic screening and treating of hotspots and at-risk populations, or mass drug administration which aims at high or total coverage of distribution of antimalarial medicine to populations in hotspots [29]. These efforts are also facilitated by ICT tools which can help precisely identify people and households for re-visits or reduce time-to-analysis (paper forms needed to be entered in spreadsheets).

The analysis in this paper is derived from observations made during field work conducted from May 2015 until September 2016. The first author actively participated and observed two large-scale malaria studies. One study was a survey, conducted in June 2015, where data was collected from a random selection of households in two districts in Zanzibar, housing approximately 4.400 residents. During this survey data regarding demographics, uptake of malaria control interventions and malaria risk factors were collected. The population was tested for malaria and treatment was given

to people who were found to be infected. In addition, blood samples were collected for screening for malaria parasites with more sensitive laboratory methods. The second study, in 2016, was a mass drug administration in hotspots in Zanzibar, to further reduce malaria transmission in these areas. It covered over 20.000 residents. Again, blood samples were collected and data regarding demographics, uptake of malaria control intervention and malaria risk factors were collected.

Despite their different characteristics, both studies had a similar goal of producing epidemiological data related to the state of malaria in Zanzibar. Behavioral data was collected with a questionnaire. The questionnaire includes questions about the household, which can be observational (e.g. "what materials are used in the roof of the house"), asked to an adult specified as the head of household (e.g. "has anyone in this household suffered from malaria in the last 4 months") or asked to each individual household member. The questionnaires are either made using pen or paper, or, since 2013, with digital data collection tools such as Open Data Kit (ODK). ODK is open source software, originally devolved by the University of Washington, with the goal of helping "organizations author, field, and manage mobile data collection solutions" [1]. The software has been widely adopted for various settings in so called developing regions [22]. In the studies described herein, the digital questionnaires were made on Android tablets running version 1.4 of ODK. Other ICT involved in data production in these studies include mobile phones, spreadsheet software, statistical analysis software such as Stata, and GIS.

4 METHODS

This paper draws on participant observation from the first author of 1.5 years alongside being in the role of an external IT consultant employed by ZAMEP, in two distinct malaria studies involving ZAMEP and KI. This was a broad role that involved developing, installing and maintaining the ICT infrastructure during two studies conducted by ZAMEP and KI. Contact with first author was established by malaria researchers from KI, who were in need of IT support for the studies done in collaboration with ZAMEP. The first author was not permanently employed but was instead hired on a temporary basis whenever there was a need for developing applications for data collection or performing other supporting IT work. The analysis herein is not limited to instances of direct use of ICT since the first author occupied a particularly advantageous point that allowed observation of how decisions were made at the management level, during the planning phase of malaria studies with malaria researchers and managers at ZAMEP, as well as how digital data collection tools were used in the field.

Ethnographic materials such as field notes, photos, notes from meetings were collected continuously through the course of 1.5 years, with more intensity during the fieldwork and preparation phases of the malaria studies. The field work took four weeks for each study, and preparation period took five months for each study. After the first malaria study, materials were organized into themes related to how data workers behaved around data collection tools, and how were data from previous studies used to plan future studies. The initial thematic analysis was intended to improve the IT systems locally at ZAMEP. During the second malaria study, observations were made to confirm or disconfirm themes made during the first study, with special attention being paid to events that occurred unexpectedly during field work. The themes were expanded and refined to include the new observations. The first author did the initial analysis, and the second author helped select and refine the themes that spoke to the question of "how mobile ICT tools impact how malaria is represented". As such, in addition to notes and observations, informal and semi-structured interviews were done with the malaria researchers from KI involved in these studies, to clarify and deepen some of the

[1] https://opendatakit.org/about/

themes both during field work and during analysis. Finally, vignettes [49] were written to bring to life, in a representative way, how different actors collaborate to produce data about malaria.

By dissecting the everyday activities involved in data collection, this analysis rests on the tradition of practice theory and laboratory studies [9, 28, 48], treating scientific procedures around malaria as human, social practices. The aim is not to show failures or gaps in method. Rather otherwise, it is assumed that best procedures were followed. It also rests on a tradition of ethnography in CSCW that looks at sociality and materiality of work practices [6]. In the analysis below, different voices such as enumerators, staff at every level in ZAMEP (field supervisors, analysts, and management), and international malaria researchers were considered. It is however important to note that the main author actively participated in different management meetings where enumerators were not included, as well as international malaria researcher meetings, where many decisions about data collection and analysis were made. Observations and the understanding of malaria expressed in this analysis is therefore influenced mostly by exposure to malaria researchers, and less by clinical workers in Zanzibar or those affected by malaria. The analysis has however steered away from malaria researchers' vocabularies and framing of the problems as much as possible (although it is impossible to avoid entirely). In addition, understanding of the life in rural areas of Zanzibar is of an outsider, as they have only been visited by the first author during the malaria studies fieldwork. Understanding of these areas is based on conversations with enumerators and malaria researchers during those brief visits. In addition, the first author was not a local from Zanzibar, even though several months were spent in the island prior to the fieldwork. A limited knowledge of the local language, Swahili, hindered the understanding of some group discussions between enumerators as they were happening in real-time, but these were then often summarized by one or more enumerators when prompted by the first author, during field work.

Finally, some data production practices are not described due to space and scope restrictions. We do not follow blood samples to laboratories, nor do we describe laboratory practices such as counting parasites. These are other equally important field sites that would could possibly offer other types of insight on how malaria data is produced. The scope here is limited to the direct interactions between malaria researchers and data workers with the Zanzibar population, during the field work of malaria studies.

5 PRODUCING DATA ABOUT MALARIA

The analysis follows 9 actor groups, summarized in Table 1. Malaria researchers, sometimes referred simply as "researchers", are those employed by KI with diverse backgrounds in medicine, biology or global health. The researchers were often present in the island for some periods of time to accompany field work but were mostly active the during planning phase of the malaria studies. The first author of this paper is referred as an IT consultant, as that was the role taken during these malaria studies. The actors from ZAMEP and KI, as a group, are sometimes referred to as "data workers". Below, the following themes are unpacked: the data ecosystem on the organizational level, the data collection procedures in the field, and how the Zanzibar population is enrolled to participate in these studies. Each section starts with a vignette [49] from the field, constructed from different field observations and interviews, showing how each theme impacts different actors as they do the work of quantifying malaria.

5.1 Finding houses

In June 2016, every day, before leaving to the villages, 50 enumerators, 5 field supervisors, 3 international malaria researchers and one IT consultant meet in ZAMEP's headquarters in the city center before sunrise. Following a morning briefing where different subjects were discussed related to the future workday and feedback from previous field days, everyone gets inside two buses and

Table 1. Actor groups

Actors	Institution	Role
Malaria researchers	KI	Scientific research coordination, data analysis
IT consultant	ZAMEP	ICT development and support, also observing and collecting ethnographic materials
Data manager	ZAMEP	Responsible for surveillance data
Program manager	ZAMEP	Scientific research coordination
BCC[a] Unit	ZAMEP	Contact with the population
Enumerators	ZAMEP	Data workers who collect data
Field leaders	ZAMEP	Coordination of enumerators
Shehas	Government	Local government official
Field guides	Government	Guiding enumerator teams to households

[a] BCC stands for Behavior Change and Communication

departs to the villages, which could be more than 1 hour drive away. When the group arrives to the ward, we meet with the sheha, the local authority, responsible for the district. A shehia is the smallest administrative unit in Zanzibar. It can be composed by several villages, and it can contain several thousand people. The sheha had previously hired local guides, which are inhabitants of the villages and well acquainted with them. The sheha had also prepared a list of areas to be visited, to achieve full coverage of the district. Enumerators, in teams of two, are then sent away, each team with one guide. At this point, the local guides for each village are the ones with most control over data collection, as they are the only ones who know the field. The enumerators know that they must visit 14 households until the end of the day, as that was the information they received from the field supervisors. But as the sun sets and the day approaches the end, one of the teams has made it through 10 households and there are no more houses in the vicinity. Coordinating over mobile phones with other guides and the field supervisors, the guide of the team with missing households goes to "find houses" nearby the village center, outside of the expected area of the study, sometimes walking several kilometers.

The vignette above exemplifies a situation that occurred almost daily in the 2016 study. Let us trace back those events to the study preparation, where part of the conceptual infrastructure [18] of data production was designed. In this case, there is a mismatch between the number of houses that the guides thought to exist in an area of the village, and the number of expected houses derived compiled from a variety of data sources before the start of the study. The field work typically takes 4-6 weeks, but before the fieldwork there is a preparation period of around 4-7 months depending on the study size. This is a period where most critical decisions are made regarding the study implementation, including the which districts should be included in the study. For that, researchers need to identify the hotspots where yearly outbreaks are likely to occur. This data is taken by the researchers from malaria surveillance systems, such as the Coconut Surveillance, at the care of the data manager at ZAMEP (see Figure 1). It also relies on similar studies done in previous years. The data is found to be inconsistent and patchy [4], but it is nevertheless trusted to reveal areas where intervention is to be focused.

Malaria researchers also need to know how many households exist in each district. Most commonly new data is generated before or during each ZAMEP or field study by shehas who are asked to generate household lists in preprinted booklets. The sheha's roles [50] include keeping records of all demographics within their area, including marriages, divorces, immigration and land ownership. Data from the booklets is then entered into a spreadsheet, employing data entry staff to do this. In

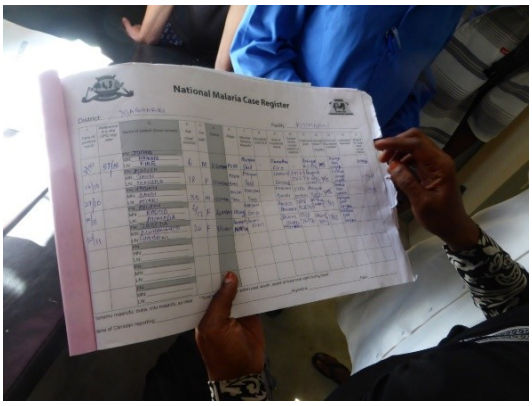

Fig. 1. Cases annotated in the National Malaria Case Register book are periodically digitized using the Coconut Surveillance System.

2016, data from a recent activity conducted by ZAMEP where households in the study areas were sprayed with insecticide was used instead, since this data already existed electronically. Unlike data from surveillance systems, this data is often found in spreadsheet files, and is retrieved from specific departments at ZAMEP, e.g. in 2016, the operational department responsible for indoor insecticide spraying.

Making sure each blood sample, each person, and each household is uniquely identified is both essential for malaria researchers, and an impressive undertaking in the absence of street names and id cards, as it is common in rural areas in Zanzibar. In the two observed studies, an identification infrastructure was developed by the researchers from ground up. Sticky labels were printed according to the expected number of people, samples, and houses. These labels contained a unique ID and a matrix barcode. These IDs were not re-used for subsequent years and served only the identification aims of each study. Every year, data is produced from different studies and is stored in a variety of formats. One of the main sources of "data friction" [15], i.e. factors affecting the flow of data and information across organizations and disciplines, is that the data if formatted in a way that obeys the logic of different studies and reports made for scientific organizations or donors, such USAID as well as other international universities. At the beginning of each study, all relevant data is woven from a variety of sources. The past data constitutes the scaffolding for the new study: it is used to plan procedures for data collection, establish goals and later evaluate whether they are being met. It is based on this data that the field guides assigned by the sheha plan the routes and house visits for the enumerator teams. Targets are given for enumerators to reach, such as "10 households per day" for a given district. Yet, during the data collection in the field, it was often necessary to improvise around the planned routes, making use of the unwritten knowledge only people familiar with the village terrain have.

5.2 Scripting the work of enumerators

Let us return to the field. During field work, the IT consultant (first author) was mobile on foot or in a vehicle, often responding to calls for technical support from different enumerator teams that were spread out in different villages, each with a different field guide. A secondary role was to encourage data collection to be done through standard procedures. For example, part of the assigned tasks was to repeat the training procedures on how to use the digital questionnaires. After a few days in the field, it was apparent that the questionnaire was not always followed as it was

designed. As the enumerators gained experience with the questionnaire and became proficient in all the questions, they developed strategies to speed through different sections. For example, the procedure for individual questionnaire requires that everyone present in the household answers the questions concerning to them. Some questions were pertaining to the last time each person experienced malaria symptoms or traveling history. As an example, enumerators would sometimes ask in the beginning to everyone present if they have had malaria recently, or if they have left the district. Sometimes, the answers would be 'no' for both questions, for everyone present in the household. That would allow the enumerators to speed through those sections in the individual questionnaires without directly addressing each person.

Two technological components broke down daily throughout the field study: GPS collection, and matrix barcode scanning. Recording GPS position is a relatively novel way of mapping malaria, not possible before the introduction of mobile tools. Before, malaria prevalence was mapped at level of village or district. With GPS, researchers can map potential outbreak focus for households where blood was collected from and can also map factors collected through the questionnaire such as types of household and potential mosquito breeding sites. Another reason, discussed later, is to help in data cleaning, by disambiguating and matching households in the database where households have been visited on more than one occasion (see Figure 2). But every day, at each village, some tablets (2 or 3) would take hours to get a GPS fix. This was sometimes solved by connecting to a Wi-Fi hotspot in a mobile phone which would then download the almanac into the tablet. But lack of mobile coverage in most rural areas made this difficult. In addition, it was often not possible to get accuracy below 20 meters. Enumerators would sometimes walk around the households for a few minutes to obtain a GPS positioning, and often that would fail. For this reason, acquiring location was paradoxically made technically optional (it was possible to skip the question in the tablet), but daily encouraged to all enumerators.

Fig. 2. Map showing household IDs during data cleaning. Overlapping GPS coordinates of unmatched households could be used to re-identify households.

Matrix barcodes aimed at reducing data entry errors by the enumerators. There were two ways of recording matrix barcodes. The first way consisted on using the barcode scanner application in the tablets, and then inserting the number manually. The reason for it was the consent form was designed in a compact way so that several stickers were placed next to each other and there was a fear that a barcode pertaining to another person would be scanned. The second way was to insert the number manually twice. The intention was for enumerators to look at the labels twice and insert them. The system would then compare the two entries and give a warning if they were not the same. However, most times enumerators simply memorized the number and entered it twice,

not looking at the labels. So, neither enumerators nor barcode scanner were fully trusted by the researchers to input data correctly by themselves; instead one had to validate the other.

Before the field work starts, enumerators, field supervisors and researchers participate in a mandatory week of training. During this week, the overarching goals of the study are presented by the researchers and management, and standard operating procedures are discussed and finalized. It also includes technical training, provided by the IT consultant, on how to care for tablet computers and use the software. Training material for this effect was inspired by material available from the ODK website [2]. However, most enumerators did not require as extensive technical training as specified in the ODK training manuals since almost all were familiar with smart phones.

Many enumerators had difficulties scanning a barcode using the cameras in Android tablets. Some learned a technique that worked for them in the training week (Figure 3), and then found that it did not work in the field. Varying lighting either outdoors under strong sunlight or indoors in near darkness posed problems for scanning in the field. For these reasons, many enumerators chose manual input. The digital questionnaire allowed for double manual entry but the procedure highly unencouraged, with frequent reminders every morning to for teams that did not use the scanner. Deviating from the standard procedures could also be necessary due to how the population behaved during data collection. For example, in big households, especially when these were visited outside of school hours, children would enter or leave the house, requiring frequent amending of the data in the questionnaire, which would complicate the data validation scripts in tablets.

Fig. 3. Field supervisor teaching enumerators how to use the barcode scanner.

At the end of each day, the tablets were brought back to the office. Each tablet was then emptied of data and put to charge until the next day. Data from each tablet was downloaded to a computer, aggregated and made into a tabular format. Every day, one malaria researcher imported data into Stata and performed several statistical analyses, whose result influences how data is to be collected the next day. The first one was to count the number of households and compare them with the expected number, from previous data, as shown in the previous section. This was one measurement of performance of the current day but was not very actionable since previous data was generally not trusted.

When data is compiled, one common process is what is commonly called data cleaning. The definition of clean data depends on the uses envisioned for it, as well subjective filters and biases of the person who performs the data cleaning [8]. For these studies, there were two moments of cleaning: 1) during the fieldwork, 2) after the data collection was concluded. Unlike paper questionnaires, which could take months until they were manually transcribed and available in a format that could be manipulated, data collected digitally through tablets was made available

[2] https://opendatakit.org/help/training-guides/

daily. During the daily data cleaning, one researcher would go through the aggregated data and check if dates were correct, if there were any duplicate IDs, if the district and village names were correctly selected. According to malaria researchers, these were easier to correct when done daily, rather than at the end of the study. The researcher created daily statistics for each team, noting the number of times GPS coordinates were not taken, or if barcodes have been entered manually. A note written in paper was sent with the tablet for each team, and in the morning of the next day, the teams who did not follow the standard procedures were then advised by the field leaders to follow the procedures. But because procedures cannot always be followed, as noted earlier, these notes sometimes caused puzzlement and requests for technical support from the enumerators. The tension between standard procedures and the actual constraints of fieldwork was present throughout the two studies described in this paper. After the field work, there was one more moment of data cleaning, who did not directly influence ongoing data collection, but it influenced data collection recommendations for future studies. Based on statistics from input errors, new scripts are devised between study cycles as well. For example, new warning signs and data validation scripts in the digital questionnaire were added between the two phases of the 2016 study.

The 2016 study was done in two rounds, and the protocol included a revisit to each household. Personal IDs — random numbers generated by malaria researchers — had been left in the households in a copy of the consent form, as well as stickers with household IDs had been left in the doorways. At the end of the study, however, 1 in each 5 people in the second phase of the study could not be matched using the person ID to people in the first phase. These had to be matched manually. This was a process that relied on GPS coordinates when available (Figure 2), as well as name and age matching. These were also problematic since there are many ways of writing the same name in Swahili, and some people would state approximate ages. Sometimes children would be counted in different households since they moved during the day between relatives' households. This process of data cleaning often resulted in very few matches.

5.3 Mobilizing the population

One of the most used technologies during field work was the mobile phone, enabling coordination at a distance. The need for coordination in the field was well-known by ZAMEP from previous studies, and each enumerator received a phone voucher every day to enable communication in the field. In one occasion, field supervisors received several phone calls from enumerator teams noticing the lack of children in the village. The house visits were done often during school hours or after-school activities, such as madrasa. This prompted a string of phone calls from a ZAMEP team called Behavior Change Communication (BCC) to local authorities, and finally to teachers who were asked to excuse their students for the next day, so they could be home and participate in the study.

This example shows how quickly coordination between all actors involved in the study could achieve population-wide mobilizations. Zanzibar inhabitants are the most important actor when it comes to malaria studies. Malaria data is produced from them, and ultimately is used to inform government policies that affect them. They are only physically present during data collection. But mobilization of population begins months before the field work start. During training week, enumerators are given instructions on how to phrase the goals of the study in an understandable way, likely to generate high consent rates from the population. At this stage, the population cannot influence how the study is to be conducted. Population's only power is to not consent when the field work starts. However, the enumerators tend to speak for the population, based on their experience in participating in previous health studies, or on their own experience of being Zanzibar people and knowing the local customs.

Fig. 4. An enumerator takes blood samples from residents of a large household.

If there are added questions by KI researchers, they are in English, and are also to Swahili during training week. The enumerators discuss between themselves what would be the best translations that capture the essence of the question and serve the study goals. The language translations are part of translations in the Latourean sense [28]. The goals of the study are passed on to enumerators, who are enrolled as employees, hired on short term contracts to conduct field work. As such, they have little power to question the overarching study goals. However, each question in the questionnaire is discussed on an open table on the grounds of how it will likely be perceived by the population. Each question can be re-written several times until a consensus is achieved. For example, the question "is this person a permanent resident of this household" was a center of controversy, and several hours were spent in both 2015 and 2016 studies, discussing whether it was a valid question, and what would be the best way to phrase it or how to explain it to people. In case was the idea that in rural Swahili homes, people often tend to be more mobile than the study assumes. To ask if someone is permanent or temporary requires an understanding of how lives of people in rural areas are. One enumerator suggested that residents of one household could be identified by the people who eat from the same cooking pot, seeing that a building can host several families or households that, each, cook for themselves.

Translating interests of the government into the population does not stop with the enumerators. Before the study starts, the BCC team goes to each village and in cooperation with the sheha prepares the population for the visit of the malaria study workers by giving details about the study and asking members of all households to be at home when the enumerators visit. This includes community meetings where people attending can ask questions. During data collection, data cleaning was also made to identify areas in villages with a high number of refusals — i.e. where people did not consent to be part of the study. This prompted the BCC team to return to those areas of the village and attempt at convincing residents that the study was important, so that those areas could be re-visited. So, the practice of daily data cleaning played a role in mobilizing the population, in addition to mobilizing enumerators as described earlier, and obviously in addition to correcting errors in collected data.

6 DISCUSSION

As malaria shrinks in prevalence, the weight of each malaria case grows. In other words, as malaria endemic regions near malaria elimination status, governments are recommended to collect more data and dedicate more resources per malaria case [32]. It is known that mobile ICT tools are increasingly used in active surveillance in developing regions, and it is also not new that precise location technologies such as GPS are often used to make malaria more spatially defined. But with the ethnographic account presented here, we surface some of the coordination work that mobile

ICT tools enable during field work in order to precisely identify people and households, or to validify data while performing work efficiently. Below, we discuss how malaria, through practices enabled through the use mobile ICT tools, binds many objects together: more households are visited, asymptomatic individuals are made part of statistics, and more data is considered valid. Malaria is made into data through a complex socio-technical mix that involves WHO recommendations, institutional alignments, formalized procedures, and situated actions of data workers making use of mobile tools. Our descriptions of the complexity of the real-world data production practices are meant to enable not only for the design of better ICT tools, but also to imagine better practices of working with and through computers bringing together multiple perspectives from different actors, which a classic purpose of field studies in CSCW [6]. Implications for design are not immediately obvious from ethnographic work [14], but nevertheless we propose some directions and challenges for future work within public health surveillance.

6.1 Malaria datafication and the asymptomatic population

There are several disease surveillance systems and practices operating in parallel in Zanzibar, and they differ on how they impact on people not suffering from malaria symptoms. Let us consider the scenario of a patient, with malaria symptoms, going to a health facility seeking treatment. If they test positive for malaria, the case is added to the national register through weekly reporting, done using mobile technology on the system called MEEDS. An improvement on this system [13] introduces "active case response", through which the household of the malaria infected patient is visited: all household members are tested, GPS location recorded, and questions are asked about the household members. Finally, the population-wide studies and interventions described in this paper aim target the whole population in certain areas. With current sensitive diagnostic technologies [29], malaria does not need to be symptomatic to be visible. Asymptomatic individuals, which are still carriers of malaria parasites, are detected and treated, at a high speed, facilitated by ICT mobile tools.

The areas targeted by the studies described in this paper are "hot", just like the populations at risk. Malaria, transformed in and through data, becomes an entity to be managed and contained. This is an effect that echoes previous research in electronic surveillance [10]. In an electronic format, malaria is visible and actionable, not only to patients or point-of-care clinicians, but also to governmental decision makers. These are changes perhaps more felt by the people living daily with the disease. When a person is identified as part of a hotpop, or as living in a hotspot, she is likely to be targeted for future interventions. People living in hotspots are or course free to refuse being part of these studies, but they still have a disrupting effect in normal village life, due to their scale and importance given to them by the government.

Producing malaria data requires significant translation work to enroll a growing population. Mobile tools and real-time communication support the real-time coordination work necessary for these large-scale studies. A great deal of coordination work to enroll the population is done already before the studies start, offline, with BCC team and shehas in community meetings. But more importantly, as we have illustrated, in the field there is coordination work to enable people to be present at home when enumerators visit, by e.g. appealing to authorities to excuse children from going to school. This type of quick mobilization during field work is enabled by the introduction of mobile tools; they enable fast communication and real-time coordination, as well as enabling data from the previous day to be available for coordinating the next day.

Returning to the main point of the paper, it is then obvious that data production mobilizes organizations and individuals [38], making them orient themselves towards producing more data [26], but it is also clear that distributed coordination work is essential for this assemblage to produce data. In other words, rather than data being "raw", producing data on a population requires

mobilizing and coordinating, or change it. We have also shown that this assemblage of organizations and individuals tends to grow: malaria is transformed from an individually felt disease to a socio-technical entity in an electronic form, to be managed and eliminated on a country. This in turn re-arranges public health services around malaria and widens the range of stakeholders who must have an active role in malaria elimination, such as *all the people living in areas marked as hotspots*. They must now also perform work to eliminate malaria in the areas where they live, by regularly being available and volunteering data and blood samples. Malaria is not only a problem of individuals who may be sick and in need of treatment; whole districts now need to manage their malaria reservoir, akin to chronic patients in need to manage their illnesses.

Since current malaria eradication practices are not about reducing the amount of data produced about malaria or reducing the number of people implicated in data production — rather the opposite — the amount of coordination and translation work required is likely to increase. We do not have recommendations for facilitating the necessary translation and coordination work in increasingly larger-scale disease surveillance. Instead, our hope is that acknowledging the problem through the rich descriptions of these practices in this paper inspire technology developers, CSCW and public health researchers to think of ways together of improving large scale data production practices. We could however think of designing future mobile ICT tools in ways that transform data production while reducing the burden of these large-scale studies on populations. One could imagine leveraging existing mobile phones to collect some of the data without direct contact with the population. For example, technologies have already been proposed that make use of passively collected data from sources such as mobile phone usage data from residents of a country or a certain area in order to infer mobility patterns, for the purposes of disease surveillance [11, 40, 47]. This could reduce the burden of data collection on the daily life of the population, as some of this data can be inferred passively, but it would require additional privacy and ethical considerations [40].

6.2 Locating malaria through invisible work

Even if producing data always requires a shedding of context from where the data was produced, here we focus on the contextually rich socio-material practices performed by enumerators and field guides, to be able to locate houses in Zanzibar during field work. We do so to highlight and unpack the coordination work they perform, crucial for malaria data to be produced. The population-wide studies described herein require that data workers can find a significant sample of the households in a district. To do so, they rely on data from different sources within ZAMEP, produced in other studies and interventions, as well as lists of households (names of head of household) compiled by the shehas. Already these are coordination mechanisms [41] that support the field work. However, during field work some of these mechanisms are found to be incorrect or no longer valid. Achieving a desired sample number or, in the case of 2016 study, total household coverage, relies on knowledge of the terrain provided by the field guides. The knowledge of local guides, crucial to successfully conduct these studies, consists in what Scott [26, 42] calls, a "wide array of practical skills and acquired intelligence in responding to a constantly changing natural and human environment". This knowledge can be about the roads surrounding the villages, or the location of temporary houses. From the perspective of the researchers and enumerators, the temporary houses were "found", as they were not part of the household lists used to coordinate field work. Before the teams go on to "find new houses", they contact field supervisors using mobile phones, and these sometimes arrange for transportation and re-assign other teams to assist. This work of dealing with unexpected events is a type of second-order articulation work [35, 41]. It is not formalized, but it is essential for data production. This work is also "invisible" [23] because it is informal. The formal work of field guides — that they are paid to perform — is simply to guide enumerators to a previously agreed upon list of houses. However, as we have shown, field guides additionally informally participate in delineating

the geographical areas visited by the enumerators. Through this distributed work, *a greater number of households* can be included in statistics about malaria and, in turn, the geographical area where malaria is found widens.

It is known that ICT tools capable of geographical positioning change what location is, from being a contextual linguistic construct to a reified object through data [44]. In malaria research, it means that a new malaria case is no longer vaguely located in a district, but it can now be positioned exactly where it was detected. The idea is quite powerful, and it is at the center of WHO's malaria elimination guidelines. However, GPS coordinates are obtained through situated actions from the data workers, in this case the enumerators. When houses are found, enumerators collect geo-coordinates, among other types of data about the house. But positioning does not always work, possibly due to cloudy skies, or presence of trees. This creates a situation where the enumerators need to find an intersection between the physical space where the households are located and the spaces where it is possible to get GPS signal. Invisible coverage spaces have been shown to change how people orient themselves around physical spaces, by e.g. creating gathering areas in places with good mobile phone network coverage [19]. In this case, finding the intersection of GPS signal and physical space surrounding households is a largely unscripted type of work. Each enumerator acted situatedly, either by walking around the area near households, calling technical support or skipping the GPS recording entirely. How and where malaria ends up being located, as a data point with geographical coordinates, is a result of a mix of the materiality of digital tools, the terrain, and the situated actions of the data workers in the field.

An implication for the design of malaria studies is that it could be of value to surface the invisible work of the field guides and of enumerators when locating households. No matter how much work goes into planning malaria studies, locating malaria cases heavily depends on the situated coordination work of data workers while in the field, to find houses and then record coordinates. Neither finding or recording are trivial tasks. We therefore recommend a careful integration of the work of field guides into the planning of malaria studies, by allocating time for this additional work to be performed. In addition, one could also extend the training week procedures to include locating houses with field guides. These additional socio-technical procedures would support these re-occurring and crucial coordination mechanisms which are currently seen by malaria researchers as surprising.

6.3 Restricting actions to increase validity

Digital collection tools can act on enumerators in ways that paper questionnaires do not. They enable scripting actions of enumerators, and act on them during the act of data collection by restricting their input or giving warnings when data is thought not to be correct. These scripts are a result of first-order articulation work by malaria researchers and ZAMEP, where they saw as necessary to constrain enumerator actions to increase data validity. Technological tools can configure users [33], anticipating their actions and their failures, and shaping the range of possible actions [2]. However, it was necessary for enumerators to work around them repeatedly during field work, adapting to the changing circumstances of time limits, light conditions and so on. Workarounds were therefore the rule, rather than the exception, an observation that echoes previous CSCW work [35].

One implication for future studies is to acknowledge the constraints on the work of enumerators and allow for more explicit, predictable and artful [35] coordination mechanisms in the field. In the ODK training guides [3], it is almost always assumed that enumerators do not have any prior training in scientific studies, nor previous exposure to smartphones or tablets. In the cases described here, the enumerators hired by ZAMEP were health workers, most of them with extensive experience

[3] https://opendatakit.org/help/training-guides/

on data collection. For ODK, advanced features could be added enabling communication between enumerators and researchers, as well as between enumerators, allowing peers to find solutions to problems that may arise in the field, so that collectively agreed workarounds could be formalized around data collection methods. Meta-data is already currently generated by the malaria researchers at the end of each day, through the form of paper notes. Digital ICT tools enable data to be quickly analyzed and acted upon by researchers, making possible for data cleaning to be done at the end of each day. Currently, a result of this activity are coordination mechanisms in the form of paper notes sent daily to enumerator teams, with indicators of their performance and reminders and specific instructions to use barcode scanners and GPS. These notes could be transformed into digital notes, as part of a meta-data infrastructure. Meta-data could also be added detailing when procedures were followed by enumerators and when improvisation had to occur. These features could be implemented at the ODK application level, which is open-source, or as a separate system or recommended set of procedures, that would derive from a re-thinking of data collection procedures that treat enumerators as highly-qualified individuals, able to improvise around problems more effectively.

Returning to the main argument, with the aim of producing data which aligns correctly with previous data [38], ICT tools enable practices that aim at reducing the amount of data points which are discarded. Through barcode id matching, GPS, matching, and flexible coordination mechanisms that rely on daily data cleaning, *more data can be trusted by malaria researchers and government agencies*. This in turn has the consequence of increasing the power of malaria statistics.

7 CONCLUSION

The malaria parasite and the mosquitos who host it have been part of the life of humanity for millennia. Malaria has had a role to play in the rise and fall of cities and civilizations and has shaped our infrastructures, technologies and habits [43]. In this paper, we have shown how malaria is currently changing, as several countries make strides towards its elimination, from being a disease felt by individuals, to an electronic entity made by data. In an electronic format, malaria is visible and actionable, not only to patients or point-of-care clinicians, but also to decision makers. This change occurs through the work of data workers using ICT tools and modern diagnostics.

Malaria is not just out there to be recorded. Malaria data is produced, with mobile ICT tools, through a mix of sociomaterial practices that include situated coordination mechanisms such that 1) populations can be mobilized in real-time and more people can be surveyed, including individuals not suffering from malaria but possibly carrying the parasite, 2) more households can be included, and greater geographical areas can be covered, and 3) more validation mechanisms can be made and therefore more data can be trusted and included in malaria statistics. To be recorded as electronic data, malaria builds and mobilizes diverse human, organizational, and infrastructural worlds around it, who must now be dedicated to its production, management, and care.

We have also suggested some implications for design of future malaria studies, as well as design of data collection infrastructures to support them. Our suggestions are conservative since they predictably expand data production practices about malaria. Here we have suggested integrating passive forms of disease surveillance to reduce burden of data production on populations, surfacing invisible coordination work, and adding meta-data to acknowledge workarounds and formalized procedures. We are careful when doing these recommendations, as each would affect power relations and create new imbalances [45]. For example, making some types of work visible can incur in increased formalization and surveillance of work practices, which may be unwelcomed, or by recommending meta-data we are creating a new burden of accounting and creating new types of work. Each change must weigh costs and benefits and be informed by the perspectives of all actors involved. In addition, these are not the only suggestions that could arise from our observation and

analysis but are simply possible directions for data production practices in disease surveillance. We hope to engage CSCW researchers in the real-world contexts of disease surveillance and data production practices around the world. These contexts are likely to change dramatically in years to come: diseases change (as does our understanding of them), people and societies change and technologies change. How all these co-evolve, and which practices emerge from these complex assemblages, are likely to remain interesting and crucial topics of study for the CSCW community.

ACKNOWLEDGMENTS

We first and foremost thank all study participants, data workers and malaria researchers in ZAMEP and KI. We also thank our colleagues who read and commented on earlier drafts, and helped us frame the contributions and the format of the manuscript, especially Ulrika Morris and Airi Lampinen. Finally, we thank the anonymous reviewers for their generous and insightful comments. This work has been partly supported by the Knut and Alice Wallenberg Foundation (KAW) within the project "Engineering the Interconnected Society: Information, Control, Interaction" and Swedish Foundation for Strategic Research project RIT15-0046.

REFERENCES

[1] Mark S. Ackerman. 2000. The Intellectual Challenge of CSCW: The Gap Between Social Requirements and Technical Feasibility. *Hum.-Comput. Interact.* 15, 2 (9 2000), 179–203. https://doi.org/10.1207/S15327051HCI1523_5

[2] Madeleine Akrich. 1992. *The De-scription of Technical Objects.* MIT Press. https://halshs.archives-ouvertes.fr/halshs-00081744/ [Online; accessed 2017-04-12].

[3] Victor A. Alegana, Peter M. Atkinson, Christopher Lourenço, Nick W. Ruktanonchai, Claudio Bosco, zu Elisabeth Erbach-Schoenberg, Bradley Didier, Deepa Pindolia, Arnaud Le Menach, Stark Katokele, Petrina Uusiku, and Andrew J. Tatem. 2016. Advances in mapping malaria for elimination: fine resolution modelling of Plasmodium falciparum incidence. *Scientific Reports* 6 (13 7 2016), 29628. https://doi.org/10.1038/srep29628

[4] Mohamed Ali. 2015. Factors influencing performance of Malaria Early Epidemic Detection System (MEEDS) in Zanzibar: A case study of Zanzibar Malaria Elimination Programme. (2015). http://scholar.mzumbe.ac.tz/bitstream/handle/11192/908/MSc_HME_Mohamed_Haji_Ali_2015.pdf?sequence=1

[5] Jeremy P. Birnholtz and Matthew J. Bietz. 2003. Data at Work: Supporting Sharing in Science and Engineering. *Proceedings of the 2003 International ACM SIGGROUP Conference on Supporting Group Work*, 339–348. https://doi.org/10.1145/958160.958215

[6] Jeanette Blomberg and Helena Karasti. 2013. Reflections on 25 Years of Ethnography in CSCW. *Computer Supported Cooperative Work (CSCW)* 22, 4-6 (1 8 2013), 373–423. https://doi.org/10.1007/s10606-012-9183-1

[7] Geoffrey C. Bowker and Susan Leigh Star. 2000. *Sorting Things Out: Classification and Its Consequences* (n edition edition ed.). The MIT Press, Cambridge, Massachusetts London, England.

[8] danah boyd and Kate Crawford. 2012. Critical Questions for Big Data. *Information, Communication and Society* 15, 5 (1 6 2012), 662–679. https://doi.org/10.1080/1369118X.2012.678878

[9] Christian Bueger. 2015. Making Things Known: Epistemic Practices, the United Nations, and the Translation of Piracy. *International Political Sociology* 9, 1 (1 3 2015), 1–18. https://doi.org/10.1111/ips.12073

[10] Baki Cakici and Pedro Sanches. 2014. Detecting the Visible: The Discursive Construction of Health Threats in a Syndromic Surveillance System Design. *Societies* 4, 3 (28 7 2014), 399–413. https://doi.org/10.3390/soc4030399

[11] Guillaume Chabot-Couture, Vincent Y. Seaman, Jay Wenger, Bruno Moonen, and Alan Magill. 2015. Advancing digital methods in the fight against communicable diseases. *International Health* 7, 2 (1 3 2015), 79–81. https://doi.org/10.1093/inthealth/ihv008

[12] Jackie Cook, Berit Aydin-Schmidt, Iveth J. González, David Bell, Elin Edlund, Majda H. Nassor, Mwinyi Msellem, Abdullah Ali, Ali K. Abass, Andreas Mårtensson, and Anders Björkman. 2015. Loop-mediated isothermal amplification (LAMP) for point-of-care detection of asymptomatic low-density malaria parasite carriers in Zanzibar. *Malaria Journal* 14 (2015), 43. https://doi.org/10.1186/s12936-015-0573-y

[13] Gordon M. Cressman, Michael V. McKay, Abdul-wahid Al-Mafazy, Mahdi M. Ramsan, Abdullah S. Ali, Issa A. Garimo, Humphrey Mkali, and Jeremiah J. Ngondi. 2015. Using Mobile Technology to Help Eliminate Malaria in Zanzibar. *Online Journal of Public Health Informatics* 7, 1 (26 2 2015). https://doi.org/10.5210/ojphi.v7i1.5679 PMID: null PMCID: PMC4512391.

[14] Paul Dourish. 2006. Implications for Design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 541–550. https://doi.org/10.1145/1124772.1124855 [Online; accessed 2017-04-26].

[15] Paul N. Edwards. 2010. *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming.* MIT Press.
[16] Wendy Nelson Espeland and Mitchell L. Stevens. 2008. A sociology of quantification. *European Journal of Sociology/Archives Européennes de Sociologie* 49, 3 (2008), 401–436.
[17] Katayoun Farrahi, Rémi Emonet, and Manuel Cebrian. 2015. Predicting a Community's Flu Dynamics with Mobile Phone Data. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing*, 1214–1221. https://doi.org/10.1145/2675133.2675237
[18] Melanie Feinberg. 2017. A Design Perspective on Data. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2952–2963. https://doi.org/10.1145/3025453.3025837
[19] Pedro Ferreira, Pedro Sanches, and Alexandra Weilenmann. 2013. Awareness, Transience and Temporality: Design Opportunities from Rah Island. *Human-Computer Interaction – INTERACT 2013*, 696–713. https://doi.org/10.1007/978-3-642-40480-1_48 [Online; accessed 2018-04-12].
[20] Geraldine Fitzpatrick and Gunnar Ellingsen. 2013. A Review of 25 Years of CSCW Research in Healthcare: Contributions, Challenges and Future Agendas. *Comput. Supported Coop. Work* 22, 4-6 (8 2013), 609–665. https://doi.org/10.1007/s10606-012-9168-0
[21] Martin French. 2009. *Picturing public health surveillance: Tracing the material dimensions of information in ontario's public health system.* Ph.D. Dissertation.
[22] Carl Hartung, Adam Lerer, Yaw Anokwa, Clint Tseng, Waylon Brunette, and Gaetano Borriello. 2010. Open Data Kit: Tools to Build Information Services for Developing Regions. *Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development*, 18:1–18:12. https://doi.org/10.1145/2369220.2369236 [Online; accessed 2017-04-26].
[23] Erin Hatton. 2017. Mechanisms of invisibility: rethinking the concept of invisible work. *Work, Employment and Society* 31, 2 (1 4 2017), 336–351. https://doi.org/10.1177/0950017016674894
[24] Tobias Homan, Aurelio Di Pasquale, Ibrahim Kiche, Kelvin Onoka, Alexandra Hiscox, Collins Mweresa, Wolfgang R. Mukabana, Willem Takken, and Nicolas Maire. 2015. Innovative tools and OpenHDS for health and demographic surveillance on Rusinga Island, Kenya. *BMC Research Notes* 8 (2015), 397. https://doi.org/10.1186/s13104-015-1373-8
[25] Monika Alise Johansen, Jeremiah Scholl, Per Hasvold, Gunnar Ellingsen, and Johan Gustav Bellika. 2008. "Garbage in, Garbage out": Extracting Disease Surveillance Data from Epr Systems in Primary Care. *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, 525–534. https://doi.org/10.1145/1460563.1460646
[26] Aditya Johri and Janaki Srinivasan. 2014. The Role of Data in Aligning the 'Unique Identity' Infrastructure in India. *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work and Social Computing*, 697–709. https://doi.org/10.1145/2531602.2531687
[27] Elizabeth Kaziunas, Mark S. Ackerman, Silvia Lindtner, and Joyce M. Lee. 2017. Caring Through Data: Attending to the Social and Emotional Experiences of Health Datafication. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 2260–2272. https://doi.org/10.1145/2998181.2998303
[28] Bruno Latour. 1987. *Science in Action: How to Follow Scientists and Engineers Through Society.* Harvard University Press.
[29] Ulrika Morris. 2015. *Molecular methods in malaria control in the era of pre-elimination.* Inst för mikrobiologi, tumör- och cellbiologi / Dept of Microbiology, Tumor and Cell Biology. http://openarchive.ki.se/xmlui/handle/10616/44522 [Online; accessed 2018-04-05].
[30] Dawn Nafus. 2014. Stuck data, dead data, and disloyal data: the stops and starts in making numbers into social practices. *Distinktion: Journal of Social Theory* 15, 2 (4 5 2014), 208–222. https://doi.org/10.1080/1600910X.2014.920266
[31] World Health Organization. 2006. Communicable disease surveillance and response systems. *A guide to planning. Geneva: WHO* (2006).
[32] World Health Organization. 2014. *World Malaria Report 2014.* World Health Organization.
[33] Nelly Oudshoorn, Els Rommes, and Marcelle Stienstra. 2004. Configuring the User as Everybody: Gender and Design Cultures in Information and Communication Technologies. *Science, Technology and Human Values* 29, 1 (1 1 2004), 30–63. https://doi.org/10.1177/0162243903259190
[34] Kathleen H. Pine and Max Liboiron. 2015. The politics of measurement and action. *CHI 2015 - Proceedings of the 33rd Annual CHI Conference on Human Factors in Computing Systems: Crossings.* https://doi.org/10.1145/2702123.2702298 [Online; accessed 2018-02-22].
[35] Kathleen H. Pine and Melissa Mazmanian. 2017. Artful and contorted coordinating: The ramifications of imposing formal logics of task jurisdiction on situated practice. *Academy of Management Journal* 60, 2 (1 4 2017), 720–742. https://doi.org/10.5465/amj.2014.0315
[36] Theodore M. Porter. 2003. The management of society by numbers. *Companion to Science in the Twentieth Century* (2003), 97–110.
[37] Reid Priedhorsky, Dave Osthus, Ashlynn R. Daughton, Kelly R. Moran, Nicholas Generous, Geoffrey Fairchild, Alina Deshpande, and Sara Y. Del Valle. 2017. Measuring Global Disease with Wikipedia: Success, Failure, and a Research

Agenda. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 1812–1834. https://doi.org/10.1145/2998181.2998183

[38] David Ribes and Steven J Jackson. 2013. Data Bite Man: The Work of Sustaining a Long-Term Study. In *Raw Data is an Oxymoron*, Lisa Gitelman (Ed.). MIT Press. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6462156 [Online; accessed 2018-02-22].

[39] Pedro Sanches. 2016. Seeing mobility: how software engineers produce unequal representations. *Engineering Studies* 8, 1 (2 1 2016), 27–47. https://doi.org/10.1080/19378629.2016.1163572

[40] Pedro Sanches, Eric-Oluf Svee, Markus Bylund, Benjamin Hirsch, and Magnus Boman. 2013. Knowing Your Population: Privacy-Sensitive Mining of Massive Data. *Network and Communication Technologies* 2, 1 (26 4 2013). https://doi.org/10.5539/nct.v2n1p34 arXiv: 1412.2247.

[41] Kjeld Schmidt and Carla Simonee. 1996. Coordination mechanisms: Towards a conceptual foundation of CSCW systems design. *Computer Supported Cooperative Work (CSCW)* 5, 2-3 (1 6 1996), 155–200. https://doi.org/10.1007/BF00133655

[42] James C. Scott. 1998. *Seeing like a state: How certain schemes to improve the human condition have failed.* Yale University Press.

[43] Sonia Shah. 2010. *The fever: How malaria has ruled humankind for 500,000 years.* Macmillan. [Online; accessed 2017-04-26].

[44] Irina Shklovski, Emily Troshynski, and Paul Dourish. 2015. Mobile technologies and the spatiotemporal configurations of institutional practice. *Journal of the Association for Information Science and Technology* 66, 10 (1 10 2015), 2098–2115. https://doi.org/10.1002/asi.23321

[45] Susan Leigh Star and Anselm Strauss. 1999. Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work. *Computer Supported Cooperative Work (CSCW)* 8, 1-2 (1 3 1999), 9–30. https://doi.org/10.1023/A:1008651105359

[46] Lucy A. Suchman. 1987. *Plans and Situated Actions: The Problem of Human-machine Communication.* Cambridge University Press, New York, NY, USA.

[47] Andrew J. Tatem, Youliang Qiu, David L. Smith, Oliver Sabot, Abdullah S. Ali, and Bruno Moonen. 2009. The use of mobile phone data for the estimation of the travel patterns and imported Plasmodium falciparum rates among Zanzibar residents. *Malaria Journal* 8 (2009), 287. https://doi.org/10.1186/1475-2875-8-287

[48] Janet Vertesi and Paul Dourish. 2011. The Value of Data: Considering the Context of Production in Data Economies. *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*, 533–542. https://doi.org/10.1145/1958824.1958906 [Online; accessed 2017-04-10].

[49] Peter Wright and John McCarthy. 2008. Empathy and Experience in HCI. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 637–646. https://doi.org/10.1145/1357054.1357156 [Online; accessed 2017-04-26].

[50] Assembly Zanzibar. 1992. The local government (district and urban ammendment) act, 1992. http://www.zanzibarassembly.go.tz/act_1992/act_11.pdf [Online; accessed 2017-04-22].

Received April 2018; revised July 2018; accepted September 2018.