

# **Fighting a moving target: Leapfrogging to new information systems for malaria vector monitoring and control**

Mauro Braganca, Bruno de Sousa, Jacques Derek Charlwood

In order to adapt control efforts to the moving target of malaria, new ways of rapidly processing and using data are required to produce usable information. Portable electronic devices with software applications, collectively known as mHealth, are increasingly used to assist health services and manage patient information. Here we describe the development of an mHealth system, using mobile phones, for data collection and analysis in resource-constrained environments. The system overcomes many of the difficulties presented by other mHealth systems. An asynchronous, Internet connection-independent data collection system, similar to web-based architecture, is proposed. Reporting and advanced statistical and epidemiological analysis, with external data integration, such as environmental datasets, is demonstrated by a pilot assessment of the system in Mozambique. We argue that this technology can provide the entomological and epidemiological information needed for sensible decision-making processes and the development of public health policies regarding malaria control.

# Leapfrogging to New Information Systems for Malaria Vector Monitoring and Control

Mauro Braganca<sup>1,2</sup>, Bruno de Sousa<sup>2</sup> and J. Derek Charlwood<sup>3,5</sup>

1: Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal.

2: Centro de Malária e Doenças Tropicais, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Lisbon, Portugal.

3: DBL Centre for Health Research & Development, 57 Thorvaldensvej, Fredriksberg, Denmark.

4: Faculdade de Psicologia e de Ciências da Educação, Universidade de Coimbra, Coimbra, Portugal.

5: Entomology, London School of Hygiene and Tropical Medicine, London, United Kingdom.

Mauro F. Bragança - [mbraganca@gmail.com](mailto:mbraganca@gmail.com)

Bruno de Sousa - [bruno.desousa@fpce.uc.pt](mailto:bruno.desousa@fpce.uc.pt)

Jacques Derek Charlwood - [jdcharlwood@gmail.com](mailto:jdcharlwood@gmail.com)

## Abstract

In order to adapt control efforts to the moving target of malaria, new ways of rapidly processing and using data are required to produce usable information. Portable electronic devices with software applications, collectively known as mHealth, are increasingly used to assist health services and manage patient information. Here we describe the development of an mHealth system, using mobile phones, for data collection and analysis in resource-constrained environments.

The system overcomes many of the difficulties presented by other mHealth systems. An asynchronous, Internet connection-independent data collection system, similar to web-based architecture, is proposed. Reporting and advanced statistical and epidemiological analysis, with external data integration, such as environmental datasets, is demonstrated by a pilot assessment of the system in Mozambique.

We argue that this technology can provide the entomological and epidemiological information needed for sensible decision-making processes and the development of public health policies regarding malaria control.

**Keywords:** mHealth; Malaria

## 29 1. Introduction

30 Whilst advances in insecticides, the discovery of new drugs, or even the development of a vaccine will help towards  
31 the ultimate goal of malaria elimination, it is the effective collection and use of epidemiological information that will  
32 enable them to be deployed in the most cost-effective manner. The information obtained from new tools for the  
33 monitoring and control of vectors and cases (such as novel traps and diagnostic tests), will only be useful if it  
34 informs strategy. Given that malaria represents a nexus of relationships between the host, the parasite and the vector,  
35 and that the epidemiology of the disease depends to a very great extent on local conditions, this strategy may, in  
36 future, be based at the level of a village or even a household. Old ideas such as species sanitation, in which  
37 individual vector species are the target of control, may be rediscovered and used more effectively. There is no reason,  
38 given the advances in information technology, why this should not be so.

39 In 1975, Djukanovich & Mach wrote that 'The deprofessionalization of the health services is the single most  
40 important step in raising health standards in the Third World' and in 1984 Charlwood wrote that 'the way to improve  
41 health standards in the world is to make people aware that they can look after themselves and to give them the  
42 wherewithal for doing so' (Charlwood 1984). We still believe this to be the case and, with the recent advances in  
43 information technology, especially the World Wide Web and mobile devices, this is now surely possible. We argue  
44 that our best weapon against the moving target of malaria is information, and not just the standard one-way flow of  
45 information, but a two-way dialogue designed to tackle the disease at its source. In the article quoted above it was  
46 argued that the future role of Malaria Control Organizations should be to provide information, help, advice and  
47 feedback to communities affected by the disease, and that such a role would require collaboration with individual  
48 community health workers (CHW's). It would necessitate active monitoring systems, using computers, to test the  
49 effect and efficacy of the various control methods adopted. Establishing this service on a central computer may  
50 however present problems. Main-frame computers are highly dependent on existing infrastructures and require  
51 maintenance and related human resources. In developing countries these factors, in conjunction with a variety of  
52 social-economic constraints, have prevented the modernization of health care services, especially in rural settings.  
53 However, mobile phones are good solutions to these problems. They are reasonably cheap, require little in the way of  
54 maintenance and are relatively energy-independent.

55 Developing countries can 'leapfrog' to new and up-to-date technologies in areas where previous versions of such  
56 technologies have not been deployed ([Davison et al., 2000](#)). Although the benefits of leapfrogging to more cutting-  
57 edge information and communication technologies (ICT) are controversial, they can accelerate development, by  
58 skipping less efficient, more costly and polluting technologies.

59 Mobile health (mHealth) fits into the category of leapfrogging technologies. mHealth, *sensu lato*, refers to the use of  
60 portable electronic devices with software applications to assist health services and manage patient information  
61 ([Källander et al., 2013](#)). The scope of mHealth is, therefore, inherently associated with the use of information and  
62 communication technologies for health information and services: eHealth ([Wu & Raghupathi, 2012](#))([Schweitzer &](#)  
63 [Synowiec, 2012](#)). We argue that this technology can provide the entomological and epidemiological information

64 such as maps, risk assessments and statistical analysis needed for sensible decision-making processes and the  
65 development of public health policies including malaria control.

## 66 **mHealth: the opportunity is here**

67 Recent studies suggest that, in addition to the effects of increasing prosperity, ICT factors substantially improve a  
68 country's public health delivery. This has considerable implications for health policy strategy in developing countries  
69 ([Wu & Raghupathi, 2012](#)). The hardware and software costs of such technologies, particularly mobile phones, are  
70 diminishing and they present the opportunity to overcome traditional obstacles to deliver health services to the poor  
71 in low and middle-income countries ([Fraser & Blaya, 2010](#); [Schweitzer & Synowiec, 2012](#)). In much of sub-Saharan  
72 Africa, there are more mobile phones than fixed-phone landlines ([Kaplan, 2006](#)). The continuous growth of mobile  
73 network coverage and unprecedented penetration of mobile devices in such countries is creating a perfect system for  
74 leapfrogging ICTs such as mHealth ([Ngabo et al., 2012](#)) contributing towards the achievement of the Millennium  
75 Development Goals (MDGs). The existence of a so called "digital divide" along the socio-economic gradient is less  
76 pronounced in mobile phone ownership than for other ICTs, making it suitable to deliver health services in under-  
77 served rural locations at lower costs ([Kaplan, 2006](#))([Budiu & Raluca, 2012](#)).

## 78 **Scope of mHealth and related problems**

79 In the last five years major mobile technology pilot projects have been attempted in low- and middle-income  
80 countries to monitor and control a variety of diseases ([Ngabo et al., 2012](#)). Coinciding with the MDGs, these  
81 countries are striving to improve community-based programs that reduce child mortality ([Zambrano & Seward,  
82 2012](#)). Health reforms represent an opportunity to demonstrate the value of mHealth and to integrate the concept  
83 with existing health information systems ([Whittaker, 2012](#)). Integrated community case management of malaria,  
84 pneumonia, and diarrhea delivered by CHW's are now being implemented at scale in several African and Asian  
85 countries, including Uganda, Mozambique and Cambodia ([Källander et al., 2013](#)). Developing countries are seizing  
86 potential partnerships between governments, technologists and industry as well as academia and non-governmental  
87 organizations to improve health services ([Källander et al., 2013](#)). mHealth and eHealth have become the new trend in  
88 health information systems in developing economies. Nevertheless, despite their contribution, recent literature  
89 suggests that modern technologies have yet to reach full operational levels and they have so far provided modest  
90 benefits in rural settings. Failure to up-scale to regional or national levels remains a common denominator in  
91 developing countries ([Källander et al., 2013](#)).

92 There is a myriad of factors that play a role when up-scaling. Firstly, guidance to implement mHealth technologies at  
93 scale is almost non-existent. There is also a necessity for guidelines on the rights of data usage, and storage, and  
94 sufficient qualitative data to explain the potential information collected by mHealth alongside close program  
95 monitoring ([Källander et al., 2013](#)). Additionally language and socio-cultural barriers present operational difficulties  
96 that most existing systems fail to deal with ([Leon et al., 2012](#)). In developing countries, mHealth interventions are,  
97 with a few exceptions, based in non-governmental organizations (NGOs) and are not integrated into the mainstream  
98 public health services ([Leon et al., 2012](#)). Sustainability of mHealth/eHealth, on a countrywide level, however,

depends on their incorporation, within national health care programs ([Källander et al., 2013](#)). At the CHW level, possible disadvantages of using mobile phones include the risk of inaccurate data input, difficulties in reading for those with poor vision and literacy problems ([de Jongh et al., 2012](#))([Leon et al., 2012](#)).

## **Benefits of mHealth and potential applications**

The various applications of mHealth rely on the rapid collection, transmission, storage and transformation of data into information, as compared with paper-based systems. These applications allow timely access to data, real time monitoring, rapid analysis and auto-generated reports ([Leon et al., 2012](#)). In fact, mHealth can play a role at every step from healthcare providers to patients ([Ngabo et al., 2012](#)). mHealth can be used by health authorities to support disease surveillance, early warning and prevention, policy and decision-making and can provide assistance in other operations such as data collection, communication, reporting and strategic planning ([Leon et al., 2012](#)). Innovations in mHealth can conceivably change how data are used in health programs, leading to faster and decentralized decision-making and better reallocation of resources ([Handel, 2011](#)). Increased efficiency in disease monitoring and evaluation by community health workers (CHW) are also reported in resource-constrained environments ([Wu & Raghupathi, 2012](#))([Leon et al., 2012](#))([Rajput et al., 2012](#)). At this level mHealth can help health systems to promote equality and reduce disparities. Clinical service delivery can also benefit from dissemination of clinical updates and learning materials, tools for clinical decision-making and patient record management ([Ngabo et al., 2012](#)). At the patient level, mHealth can improve access to health services at lower costs, enhance patient involvement and self-management capabilities. It can be used to monitor a patients medication compliance (in chronic diseases), or be used to remind patients of follow-up consultations ([de Jongh et al., 2012](#))([Nilsen et al., 2012](#))([Handel, 2011](#)). At the research level, substantial reductions in time, from data collection to data analysis, is possible ([Leon et al., 2012](#)).

Cost-effectiveness studies, at different levels of intervention, consistently suggest a saving per unit of spending over traditional manual data collection and transmission ([Handel, 2011](#); [Rajput et al., 2012](#)). Nevertheless, despite the advantages over paper-based data-collection systems, mHealth costs are still substantial ([Rajput et al., 2012](#)) and the decrease in overall costs, due to eHealth, is debatable ([Schweitzer & Synowiec, 2012](#)). There is little research into the economic impact of such investments in lower and middle-income countries ([Schweitzer & Synowiec, 2012](#)). Many authors suggest potential benefits, but few take it beyond speculation, by measuring the direct outcomes of such investments ([Schweitzer & Synowiec, 2012](#)). Healthcare providers must also not misinterpret integration and implementation as an endpoint to their responsibilities ([de Jongh et al., 2012](#)).

The mHealth approach is incrementally gaining ground in developing countries, especially in settings where there is no power grid or infrastructures other than cell phone network coverage ([Fraser & Blaya, 2010](#)). On an operational level costs include mobile equipment acquisition, replacement or maintenance and charging ([Källander et al., 2013](#)), although some authors suggest the use of CHW personal phones to minimize initial investments ([Ngabo et al., 2012](#)). The central level costs are dependent on server side maintenance and programming, human resources, initial training and costs related to the development of metrics to measure system performance ([Schweitzer & Synowiec, 2012](#)). In general mHealth models rely on network coverage and communication for data centralization, therefore Short Message Service (SMS) or Internet usage present additional costs.

Despite the wide availability of open-source solutions most mHealth solutions are based on proprietary architectures. Device or operating system driven architectures have, in fact, slowed the development pace and constitute a barrier to development ([Chen et al., 2012](#)) ([Whittaker, 2012](#)). Programmers are often faced with device native languages that are very different from widely used programming languages such as web languages ([Fraser & Blaya, 2010](#); [Ngabo et al., 2012](#)). As previously seen in other areas of information technology development the use of proprietary software on devices can limit their flexibility to incorporate functionality in data collection ([Rajput et al., 2012](#), Whittaker, 2012).

Most mobile health systems, described in the literature, are patient center models ([Handel, 2011](#)). The main problems of such models are that they do not necessarily support existing health information systems. In fact, the majority lack an evidence base to improve health in patients ([Bastawrous & Armstrong, 2013](#)). Additionally, the developmental pace of these technologies is far superior to our capacity to evaluate their validity and efficacy ([Budi & Raluca, 2012](#); [de Jongh et al., 2012](#)). Results of some interventions suggest that access to health services or information via mHealth still remains unclear ([Budi & Raluca, 2012](#); [de Jongh et al., 2012](#); [Whittaker, 2012](#)).

Typical mHealth implementations are designed using a one-way communication pathway system, in which information is diffused using SMS. Two-way (or multiple) communications exist although they do not generally occur in real-time ([Källander et al., 2013](#)). Kallander et al (2013), in a review, reported that lack of immediate response via SMS and privacy concerns about SMS content were important. Inconsistencies in SMS formatting are often a problem with reporting systems and have an error of up to circa 10% of collected data in rural settings ([Gurman et al., 2012](#)).

Some mHealth interventions suggest improvements in community-based strategies to tackle disease in rural settings. Health decisions and policy making, in such cases, do not merely rely on health post data and CHWs actions, but also on population feedback. This two-way information pathway is fragile, mostly because it is based on the principle of community feedback without fully considering equipment necessities, costs or mobile phone penetration ([de Jongh et al., 2012](#)).

Given the nature of mobile phones, with their data connectivity and intrinsic sensors it is now possible to collect unprecedented amounts of data. Unfortunately, there is a lack of modular tools and techniques for drawing meaning and scientifically valid inferences from much of this data ([Chen et al., 2012](#)). In fact, data are frequently full of bias, noise, variability and gaps. In remote sites, increased storage of valuable data raises the potential for data loss ([Fraser & Blaya, 2010](#)). More sophisticated and effective tools for data visualization, analysis and synchronization are required ([Chen et al., 2012](#); [Fraser & Blaya, 2010](#)).

Data management and analysis often relies on data centralization, which implies a fully operational communication tier. The conceptual downside of this scheme is the bandwidth requirement and inherent costs, subsequent data validation and information production. Some authors contemplate alternatives to remote access to servers such as



offline data entry for sites with unreliable Internet access, such as the Haiti, Peru and Kenya mHealth interventions (Fraser & Blaya, 2010). It is suggested that the ideal solution relies on the use of local servers that automatically synchronize the data with a central site (Fraser & Blaya, 2010).

Few mHealth projects specifically target CHW or local health policy makers (Källander *et al.*, 2013). A common cause is the inability of systems to generate relevant or meaningful statistics (Ngabo *et al.*, 2012). Gurman, in a systematic review, describes methodological flaws and adequate sample sizes for statistical inference. A small number of published health interventions discuss the theoretical basis of interventions (Whittaker, 2012). Present mHealth systems largely remain confined to official channels in part because the architecture of the system is not a 'one size fits all', thus excluding many potential users. In fact developers are often absorbed in promoting their preferred technologies instead of working with end-users to develop the most useful solutions. The lack of data sharing, open standards and absence of tools to make sense of health data are, therefore, critical bottlenecks limiting the impact of mHealth to improve health outcomes and inter-operability (Chen *et al.*, 2012). Naturally, this is the source of inefficient results with little impact on health strategies. A different approach is required to iterate, adapt and improve what others have done.

Medical information, in developing countries, traditionally uses aggregated data collections and only recently are individual records being collected in health facilities (Fraser & Blaya, 2010). The problem of aggregated data is that they can only provide inherently aggregated statistics with little impact at an individual level (Ngabo *et al.*, 2012). Disease control strategies depend on the information provided from the field. Alternative strategies for individual records must rely on a capable local database that can generate reports that are submitted to higher levels in the health system but have an impact at a local level (Fraser & Blaya, 2010).

Information used for decision making often travels a long and winding road. Accurate data collection is an essential initial procedure. Failure to ensure the quality of information at source often raises concerns in terms of internal and external validity of studies, where dealing with compromised data will consequently lead to misguided decisions based on biased information (or wrongly assumed inferences). In other words: garbage in garbage out (GIGO). An example of this might be malaria case reporting. Most official agencies from governmental to international ones like the World Health Organization (WHO) rely on statistics from health centers and hospitals to obtain data on malaria incidence. Yet at the grassroots level an unknown number of patients (and lost to follow-up), seek treatment outside official channels. Many of those who can afford it may go to private clinics (that are widely available) while others go to a local pharmacy, and yet others may buy medicine from ambulatory salespersons. This compromises the usefulness of official data and may lead to inappropriate decisions. For example, agencies like the Presidents Malaria Initiative (PMI) rely on Government statistics to determine where to conduct IRS campaigns. Should the use of government clinics or rate of reporting vary between potential target areas, then intervention campaigns may miss the areas that most need them. Data integration from different sources seems to be the only way to achieve effective control.

The World Wide Web has increasingly become more important in our daily life and most operations that would require face-to-face interaction can now be done using on-line forms. Internet browsers provide human readable web-pages by interpreting an HTML (Hypertext Markup Language) ([15445:2000\(E\), 2000](#); [Boehm, 2010](#)). Although HTML is becoming progressively more powerful, other languages have improved user interaction. Javascript or server side languages such as PHP or ASP are relevant examples. Conceptually all these languages are used to create source code files that are stored on web servers and made accessible, through the Internet, using HTTP (Hypertext Markup Transfer Protocol) ([Rosen & Rich, 2009](#)). They provide an interface between users and databases. An important aspect of this is that local web servers can operate off-line and serve other client computers provided that they are in the same physical network. Using this characteristic it is possible to establish a local web server and benefit from Internet technologies without being online.

It is here that the benefits of mHealth, using real-time continuous biological and environmental data collection, can greatly improve understanding of the underlying causes of disease, particularly malaria, while dealing with the inherent lack of infrastructures and Internet in developing countries ([Collins, 2012](#)).

Our main technical objective was to create a low-cost mHealth solution (“Leapfrog”) able to tackle the difficulties noted by other mHealth solutions specifically in rural settings, using open source technologies and easy to create, maintain and integrate computer languages. Secondary objectives included significant improvements in data collection and integration, of real-time environmental data sources. We wanted to generate information for CHW’s and central level authorities using CHW data inputs.

## 2. System architecture and Preliminary results

Although the use of mobile phones for mHealth in malaria generally involves patient records and treatment, there is ample scope for it to be used for the monitoring of mosquito vectors. An open source system has been developed with this in mind ([Bragança \*et al. in prep\*](#)). The conceptual architecture of this system, ‘Leapfrog’, is to have a mobile phone mimic an Internet web server. By harvesting the power of web server technologies and putting them into a phone, complex features become possible rendering forms, validating user entries, reporting, integrating with external data sources and advanced inferential statistics and modeling.

An important characteristic of this architecture is that it does not require a permanent Internet connection because data are registered on the phone and not online. In fact, Internet connections are reserved for synchronization and external data integration processes only. [Figure 1](#) elucidates the differences between this model and other eHealth or mHealth designs.

Another specificity of this system is the use of common web-based languages for form and other interface design. Languages, recognizable by any web developer, such as PHP, HTML, JavaScript and Ajax and SQL were used to create the system.



To accomplish this architecture, two freely available source code implementation of an HTTP web server and a database were used: Apache, an HTTP web server, and MySQL, a multi-querying, multi-user and robust database. PAMP a specifically designed open source port operating in a Symbian Series60 environment, over Open C (a set of industry-standard POSIX and middleware C libraries) was used although the same architecture is also available to more modern smartphones. Data can be registered directly into the device available forms. Using the phone wireless technology other Wi-Fi capable devices can view the forms, and other interfaces, available on the phone system, thus making data introduction, visualization or analysis possible, and eventually easier, from a laptop, computer or other mobile phones or PDAs within certain range.

Reports, data viewing and specifically engineered statistics for available data are also produced. The major difference is that information can be readily produced at a local level empowering CHW, while still feeding data to higher levels authorities. Synchronization to a remote centralized database or higher level is simplified since data is already in a structured query language (SQL). During synchronization, external environmental datasets, regarding, for example, precipitation (or other mosquito ecology related variables), are automatically downloaded and appended or merged, by date, with the phone local database. Forms can also be updated remotely during synchronization processes, an especially useful feature in up-scaling.

## Field Testing

System field tests were conducted between 5<sup>th</sup> May and 11<sup>th</sup> July 2011 on the 1.9 x 6.2 km peninsula of Linga-Linga (23°43'S, 34°18'E), in Morrumbene District, Inhambane Province, Mozambique. The data are presented to demonstrate the systems possibilities rather than to provide definitive analyses. Linga-Linga has been described by Charlwood *et al.*, (2013) and is located 8 km to the west of Morrumbene, the district capital, and 10Km to the north of Inhambane, the provincial capital. Access to Linga-Linga is either by boat or via a sand road from the mainland. Linga-Linga has no running water or electricity and literacy levels are low. Despite this the majority of villagers own mobile phones.

The Mozambican-Danish Rural Malaria Initiative (Mozdan), has conducted studies in Linga-Linga since 2006 (under the ethical clearance 123/CNBS/06 issued by the Mozambican Health Ministry). The objectives of this initiative are to control malaria on the peninsula. Among other community actions, Mozdan established a health post and base camp for entomological studies (Charlwood *et al.*, 2013). The project received ethical clearance from the National Bioethics Committee of Mozambique (reference 123/CNBS/06) on the 2nd of August 2006.

To improve the monitoring of mosquitoes an information system using Leapfrog was developed. For field trials, the Nokia E63 model was chosen based on reliability, screen size, energy efficiency, user comfort, and device price (around 110USD in 2011). The phone used a 369 MHz ARM 11 processor with 128MB of RAM and a Li-Po 1500mAh (BP-4L) battery, running up to 10-18days (longer if using GPRS only). Local Mozdan operators following a set of instructions provided remotely by Bragança performed system installation.

## The malaria vector population

Entomological data was collected for four consecutive months during routine surveys in Mozdan field site and entered on the phone with weekly uploads to the internet. A battery operated CDC miniature light-trap, model 512, was used throughout this period to collect host-seeking mosquitoes inside bedrooms in the village. Female mosquitoes were identified to species according to morphological characters and were subdivided into 'unfed', 'part-fed', 'fed', 'semi-gravid' or 'gravid' categories according to the visual aspect of their abdomen as described by Detinova (1962). *Culex quinquefasciatus* was the most common mosquito collected (Figure 2).

## Generic external data integration

With the basic entomological data in hand, automatic and semi-automatic methods and procedures were used to integrate variables from remote sensing and public domain environmental datasets, to match the spatial resolution of Linga-Linga and time frame of the mosquito collection. These variables were chosen based on the measures of association (odds ratio, relative risk, risk ratios) from previous studies in Inhambane province and the Mozdan expert opinions. The identification of reliable sources that provide systematic and up-to-date data was a crucial step to automatize data feeds that could be integrated in the system. Data to compute vegetation indices (NDVI and EVI) were obtained from an online data repository at the NASA Land Processes Distributed Active Archive Center (LP DAAC). Wind speed and direction were acquired from the U.S Air Resource Laboratory (ARL) integrated in the National Oceanic and Atmospheric Administration (NOAA).

Precipitation, humidity, mean sea level pressure and temperature data were obtained from the Mozambican Regional Bureau of Water Administração Regional de Águas (ARA) do Sul and NOAA. Information of coastline tides based on harbor measurement was available at the Portuguese Hydrographic Institute. These measurements were computed based on the harmonic analysis of serial observations of variable durations. Mozdan supplied other environmental settings such as distance to water bodies and possible larval breeding sites.

## Example of data source extraction, automation and integration

The Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI), based on the principle that this vegetation absorbs radiation in the visible region of the electromagnetic spectrum, were used to measure the amount of green vegetation on the peninsula – an important feature in vector ecology. To acquire these two indicators, from Moderate Resolution Imaging Spectroradiometer (MODIS) of the Terra (MOD13Q1) and Aqua (MYD13Q1) satellites, we developed a routine script, to request and download, matching data from the LP DAAC archive. MODIS Level-3 16-day composite Vegetation Indices are standard MODIS products available at 250m pixel size (provided in tiles of 10 by 10 degrees and distributed in the Sinusoidal projection). GDAL (a translator library for raster and vector geospatial data formats) and a Python script were used to extract raster values of vegetation indices for a nine point grid (Figure 3). A principal component analysis (PCA) was used to obtain a more expressive vector of the EVI / NDVI variation. The dataset was systematically appended to phone database during synchronization operations.

## Example of data analysis and results

Results were obtained by a system-generated set of instructions, in R (2.13.0) (a statistics software package),

according to a previously defined (but customizable) analytical plan using data stored in the phones MySQL database. R generated code can either be rendered in R Studio environment or on the local phone using R2HTML or Knitr (R packages) exports. Although R syntax scripts for such functions as Neural Networks or advanced statistical modeling (Linear Mixed Effect analyses (LME), Generalized Estimating Equations (GEE) and other Generalized Linear Models (GLM)) are complex, they can be rendered into easily understandable outputs or plots visible on the phone (Figure 4 and 5). For example, in the most prevalent mosquito species, *Cx. Quinquefasciatus*, the GEE system analysis suggests that:

1) for every degree (°C) increase in average daily temperature there is a 31.7% (C. I. 7.9%-60.6% based on Wald 95% Confidence interval for Exp(B)) increase in the number of unfed *Culex quinquefasciatus* collected in light-traps ( $p = 0,007$ ).

2) for every unit increase in tide measurements there is a 5% (Confidence Interval [1%-9%] based on Wald 95% Confidence interval for Exp(B)) change in the numbers of unfed *Culex quinquefasciatus* in light-traps ( $p = 0,019$ ).

3) for every unit increase in the NDVI PCA component there is a 35.1% (Confidence Interval [9.6%-66%] based on Wald 95% Confidence interval for Exp(B)) increase in the numbers of unfed *Culex quinquefasciatus* in light-traps ( $p = 0,005$ ).

### Server side rendering

Other operations require processing and rendering on the server-side. The result of these operations can later be incorporated on the phone during synchronization. For example, high mosquito counts can flag actions at a central level, such as estimating the source or the possible dispersion of mosquitoes. In our case the possibility of airborne transport of mosquitoes across the Inhambane bay was evaluated using a hybrid single-particle Lagrangian integrated trajectory (HYSPLIT) particle dispersal model (Figure 6). Backward wind trajectories were simulated using GDAS (Global Data Assimilation System) to run a mixed approach of Lagrangian (transport) and Eulerian (dispersion) methods to compute simple air parcel trajectories at different heights (measured in meters above mean sea level) as has recently been applied for long-distance *Culicoides* modeling (Eagles, 2013; Garcia-Lastra, 2012).

### Conclusion

Although the Leapfrog system can be used for the complete variety of applications that other mHealth systems are designed for, here we have attempted to show its potential versatility by using it to both register and analyze mosquito data collected over a relatively short period from an isolated area in Mozambique. We present the data and analysis to show the systems capabilities rather than to produce definitive statements about the ecology or control of the mosquitoes.

Updated environmental data, like the one used in this system can, contribute to a practical understanding of the ecology of malaria vectors and might provide evidence-based solutions for future integrated control actions when fighting the moving target of malaria. For example, as the wet season progresses, the kind of water used by malaria

vectors such as *Anopheles gambiae* or *An. funestus* can change. Integrating rainfall, evaporation rates and local topography might help villagers decide where or when or if to apply anti-larval measures.

To tackle inefficiencies inherent to the up-scaling process such as device cost, difficulties in maintenance, system development and form updates, we chose open source implementations of existing web server technologies that can be installed in low and mid-price devices and render popular languages such as HTML without using the Internet. The use of web languages opens the possibility for local programmers to work with health authorities making the system development and integration with existing informative systems much easier. Most importantly, the system is independent of third party developers or NGO support. Another aspect of web languages is that they allow complex data validation, for example, our system uses the very same technologies of an online car insurance simulation, flight booking or a tax return form. They are responsible for major improvements in data quality and therefore the quality of the information produced.

Even though there are many innovations in this particular mHealth system, the production of information understandable to non-statisticians, such as the ordinary health decision maker, makes it an improvement over other such systems. By flagging only statistically significant findings and make it comprehensible to CHW's we empower not only central level decision and policy makers but also local agents. The visualization of information can motivate users to keep entering data and the two-way system of information exchange allows health care providers to anticipate needs, such as drugs, and may help in deciding best care practices, or control measures, in local settings. Another important feature of this system is that anyone with a reasonable mobile phone can use the system, and can therefore be integrated into the data collection and dissemination process. Thus, in addition to health centres and CHWs, people offering services in the private sector can use it. This would improve the reliability and usability of statistics on, for example, incidence data from local pharmacies.

Of course, while many people may own a phone not many own or run light-traps! Many Districts in countries like Tanzania, however, have one or perhaps even two, people responsible for malaria. Their job is primarily control, and often vector control. Decisions that they make are often based on little or no data. This is not because the data cannot be collected (setting up light-traps is not rocket science) but because people rarely have the wherewithal to analyze it. Having district personnel collect data that once introduced on the phone automatically provides feedback on such things as the best method to use for control, would help in decision making. By feeding local data into a central database there is the possibility of improving the precision of estimates of transmission, which results in more fine tuned possibilities for control. As malaria transmission (and hence immunity) is reduced then, in the absence of a vaccine, more areas will become prone to epidemics. Mosquito data is particularly important in epidemic prone areas, which means that most of the effort in mosquito monitoring is the setting of the traps rather than the counting of the mosquitoes collected (since there will be many traps without any mosquitoes at all). A system such as 'Leapfrog' will become increasingly necessary. One way of predicting the likelihood of an epidemic months before it occurs is to determine anopheline density inside houses (Linblade et al, 2000), and to use suitable 'stopping rules' that may include environmental factors such as temperature or dew point (Nygreen et al., 2014) as co-variables. Whilst the collection of the data is not difficult (an aspirator, a torch and a decent pair of eyes can suffice), the

integration of the data and the prediction requires a more developed capability. The proposed system includes a neural network that learns from collected data, and using forecast data integration, can help choose the best settings (for example, mosquito presence) for certain interventions such as to spray or not to spray – which can also be done at the local level.

The myriad of constraints present in the control of a moving target demands a flexible, robust, intricate but feasible solution. We don't argue that 'Leapfrog' is the ideal solution but the features presented might inspire other mHealth systems to integrate external data and to provide information other than basic summaries and, most importantly, empower health workers.

## Acknowledgements

We thank Virgilio de Rosario for helpful comments during the development of Leapfrog, Dominique Shoham for correcting the English and the Bill and Melinda Gates Foundation for financial help in setting up the pilot project in Linga Linga (under the Grand Challenge Grant 'Turning Houses into Mosquito Traps').

## 386 References

- 387 15445:2000(E). (2000). Information technology - Document description and processing languages - HyperText  
388 Markup Language (HTML). *ISO/IEC - Information Technology Task Force*.
- 389 Bastawrous, A., & Armstrong, M. J. (2013). Mobile health use in low-and high-income countries: an overview of the  
390 peer-reviewed literature. *Journal of the Royal Society of Medicine*, 106(4), 130-142.
- 391 Boehm, A. (2010). *Murach's HTML, XHTML, and CSS*: Mike Murach & Associates.
- 392 Budiu, J. N., & Raluca. (2012). *Mobile Usability*: New Riders.
- 393 Chen, C., Haddad, D., Selsky, J., Hoffman, J. E., Kravitz, R. L., Estrin, D. E., et al. (2012). Making sense of mobile  
394 health data: an open architecture to improve individual-and population-level health. *Journal of medical Internet*  
395 *research*, 14(4).
- 396 Collins, F. (2012). How to Fulfill the True Promise of “mHealth”. *Scientific American*, 307(1), 16-16.
- 397 Davison, R., Vogel, D., Harris, R., & Jones, N. (2000). Technology Leapfrogging in Developing Countries-An  
398 Inevitable Luxury? *The Electronic Journal of Information Systems in Developing Countries*, 1.
- 399 de Jongh, T., Gurol-Urganci, I., Vodopivec-Jamsek, V., Car, J., & Atun, R. (2012). Mobile phone messaging for  
400 facilitating self-management of long-term illnesses. *Cochrane Database Syst Rev*, 12.
- 401 Eysenbach, G., & Group, C.-E. (2011). CONSORT-EHEALTH: improving and standardizing evaluation reports of  
402 Web-based and mobile health interventions. *Journal of medical Internet research*, 13(4).
- 403 Fraser, H. S., & Blaya, J. (2010). *Implementing medical information systems in developing countries, what works*  
404 *and what doesn't*. Paper presented at the AMIA Annual Symposium Proceedings.
- 405 Gurman, T. A., Rubin, S. E., & Roess, A. A. (2012). Effectiveness of mHealth behavior change communication  
406 interventions in developing countries: a systematic review of the literature. *Journal of health communication*,  
407 17(sup1), 82-104.
- 408 Handel, M. J. (2011). mHealth (Mobile Health)—Using Apps for Health and Wellness. *EXPLORE: The Journal of*  
409 *Science and Healing*, 7(4), 256-261.



- 410 Källander, K., Tibenderana, J. K., Akpogheneta, O. J., Strachan, D. L., Hill, Z., ten Asbroek, A. H., et al. (2013).  
411 Mobile health (mHealth) approaches and lessons for increased performance and retention of community health  
412 workers in low-and middle-income countries: a review. *Journal of medical Internet research*, 15(1).
- 413 Kaplan, W. A. (2006). Can the ubiquitous power of mobile phones be used to improve health outcomes in developing  
414 countries. *Global Health*, 2(9), 1-14.
- 415 Leon, N., Schneider, H., & Daviaud, E. (2012). Applying a framework for assessing the health system challenges to  
416 scaling up mHealth in South Africa. *BMC medical informatics and decision making*, 12(1), 123.
- 417 Lindblade, K.A., Walker, E.D. and Wilson, M.L., 2000. Early warning of malaria epidemics in African highlands  
418 using Anopheles (Diptera: Culicidae) indoor resting density. *Journal of Medical Entomology*, 37 (5), 664-674  
419 Ngabo, F., Nguimfack, J., Nwaigwe, F., Mugeni, C., Muhoza, D., Wilson, D. R., et al. (2012). Designing and  
420 Implementing an Innovative SMS-based alert system (RapidSMS-MCH) to monitor pregnancy and reduce maternal  
421 and child deaths in Rwanda. *The Pan African Medical Journal*, 13.
- 422 Nilsen, W., Kumar, S., Shar, A., Varoquiers, C., Wiley, T., Riley, W. T., et al. (2012). Advancing the science of  
423 mHealth. *Journal of health communication*, 17(sup1), 5-10.
- 424 Nygreen, D., Stoyanov C., Lewold C., Mansson F, Miller J., Kamanga A. & Shiff C.J. (2014) Remotely-sensed,  
425 nocturnal, dew point correlates with malaria transmission in Southern Province, Zambia: a time-series study. *Malaria*  
426 *Journal* 13:231 doi:10.1186/1475-2875-13-231.
- 427 Rajput, Z. A., Mbugua, S., Amadi, D., Chepng'eno, V., Saleem, J. J., Anokwa, Y., et al. (2012). Evaluation of an  
428 Android-based mHealth system for population surveillance in developing countries. *Journal of the American*  
429 *Medical Informatics Association*, amiajnl-2011-000476.
- 430 Rosen, L. S., & Rich. (2009). *Web Application Architecture: Principles, Protocols and Practices*: Wiley.
- 431 Schweitzer, J., & Synowiec, C. (2012). The economics of eHealth and mHealth. *Journal of health communication*,  
432 17(sup1), 73-81.
- 433 Whittaker, R. (2012). Issues in mHealth: findings from key informant interviews. *Journal of medical Internet*  
434 *research*, 14(5).
- 435 Wu, S. J., & Raghupathi, W. (2012). A panel analysis of the strategic association between information and  
436 communication technology and public health delivery. *Journal of medical Internet research*, 14(5).
- 437 Zambrano, R., & Seward, R. (2012). Mobile technologies and empowerment: Enhancing human development  
438 through participation and innovation.



## Conceptual differences between Leapfrog and other mHealth architectures

Figure 1 - Conceptual differences between conventional mHealth applications (mHealth blue square), web (Server blue square) and LeapFrog (Mobile Phone turned into Server blue square)

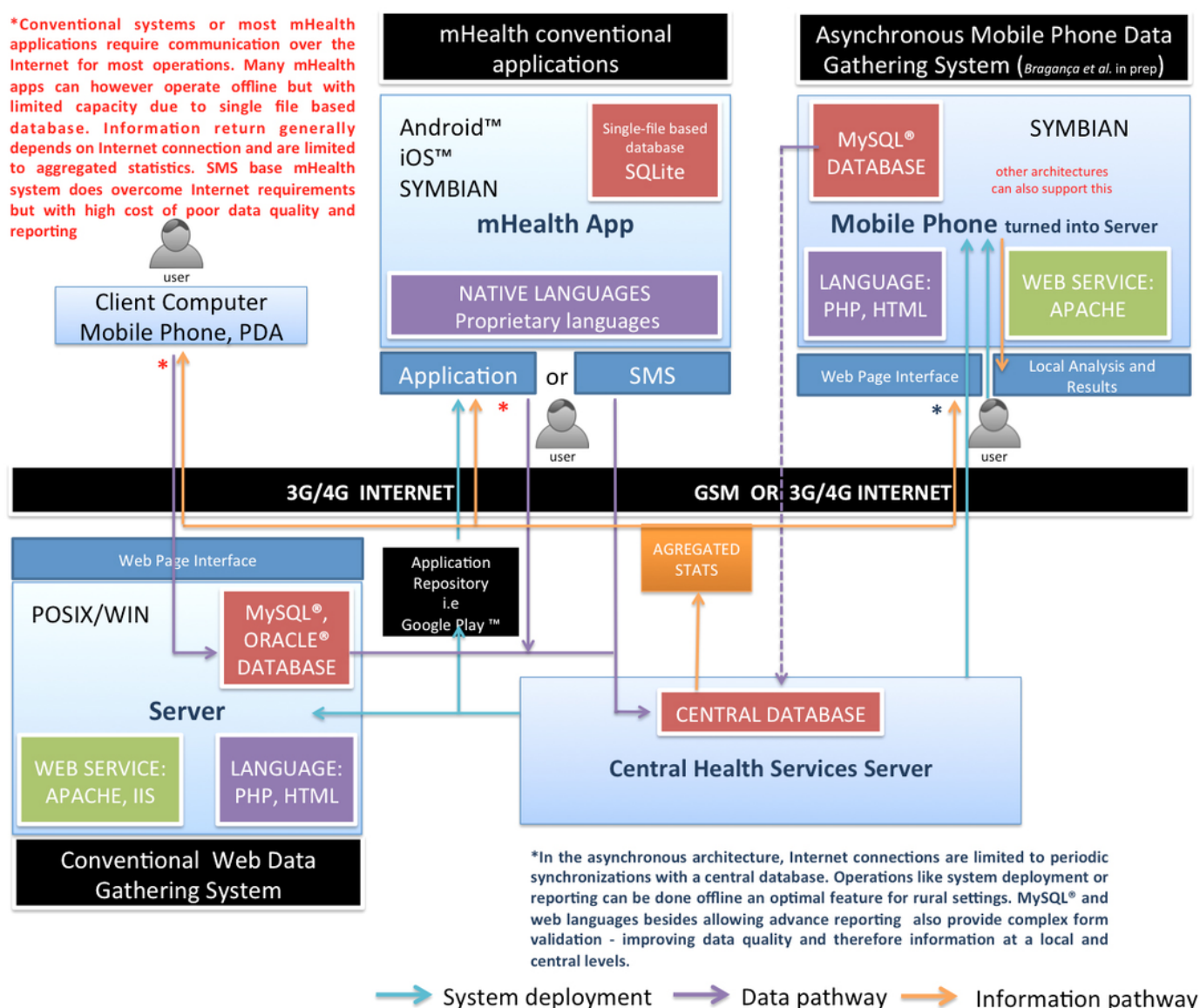
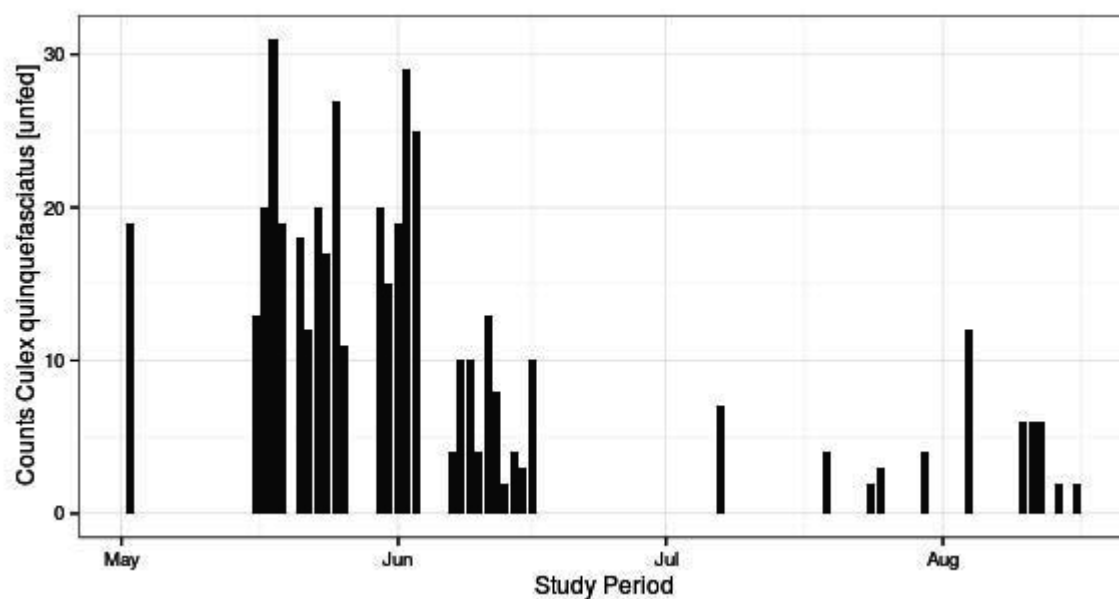


Figure 2 -Numbers of *Culex quinquefasciatus* unfed collected from the light-trap during the system testing period, Linga Linga, Mozambique.



## Normalized Vegetation Index for Linga-Linga

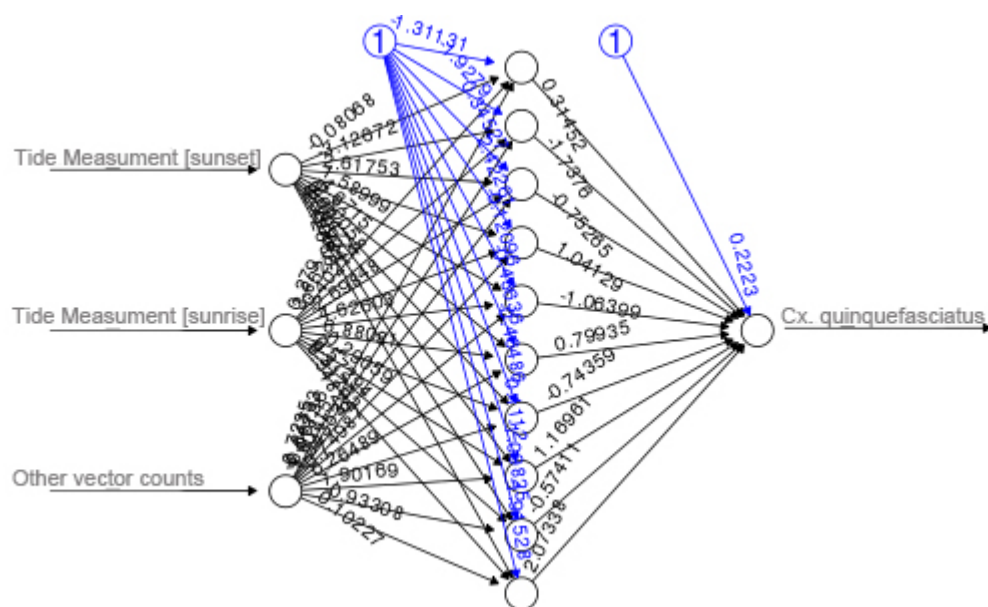
Figure 3 - Map of Linga Linga Vegetation Index (NDVI) for Linga-Linga obtained via LP DAAC (NASA) using automation processes and the nine-point reference grid used extract the NDVI/EVI values. Resulting data was integrated into the local phone database to match location and time frame during synchronization processes.





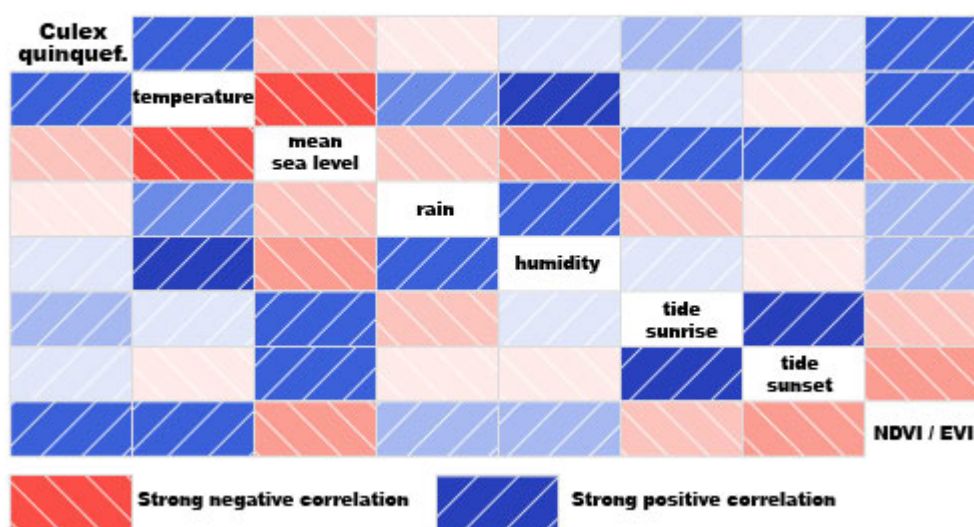
## Example of system outputs

Figure 4 - Example of a neural network output - using tide patterns and other mosquitoes to predict mosquito counts of *Cx. quinquefasciatus* - worst prediction has 35% of error.



Another example of system outputs

Figure 5 - Multiple visual correlation of counts of unfed *Cx. quinquefasciatus* and average daily temperature, mean sea level and atmospheric pressure, precipitation, humidity, tide close to sunset and sunrise, and NDVI. A dark blue color indicates a strong positive correlation whilst a dark red one a strong negative correlation.



# 6

Example of Server side rendering

Figure 6 - Example of a wind backward trajectory using HYSPLIT – useful to understand the possibility of mosquito dispersal from Inhambane to Linga Linga.

