

Submitted: May 17<sup>th</sup>, 2017

Accepted: September 22<sup>nd</sup>, 2018

## Towards Geodata Maintenance: A Case of DHIS2 GIS Implementation in Malawi

Patrick Albert Chikumba<sup>1,2</sup> and Gloria Chisakasa<sup>2</sup>

<sup>1</sup>Department of Informatics, University of Oslo, Norway.

<sup>2</sup>Faculty of Applied Sciences, The Polytechnic, University of Malawi.

### Abstract:

**Background:** As noted in literature of geodata maintenance, the emphasis is on the geographic database update. However, geodata maintenance is beyond the geographic database update. Thus, this paper proposes administrative and technical actions for geodata maintenance in health in which GIS is becoming popular.

**Methods:** This case study was conducted in health management in Malawi. An interpretive qualitative research method was adopted. Data was collected through participant observation, as the principal method, which was supplemented by semi-structured interviews and artifact examination. Participant observation was done in one and half years during the deployment of DHIS2 GIS. Thematic analysis was employed as the data analysis approach.

**Findings:** This study has identified six necessary actions for geodata maintenance: (1) identify the need, (2) communicate the need, (3) assess the need, (4) edit the model, (5) acquire the geodata, and (6) edit the dataset. The first two actions are administrative and the last three actions are technical while the third one is both administrative and technical.

**Conclusion:** The geographic database update involves two actions: edit the model and edit the geodata. Other four actions are required for acquiring geodata to make a complete set of geodata maintenance operations. However, to have a sustainable geodata maintenance, there is a need to incorporate geodata maintenance into operations of health management information system in general and GIS implementation process in particular. The involvement of stakeholders at all levels is also essential.

**Keywords:** Geodata; Geodata Completeness; Geodata Maintenance; Geographic Database Update.

### 1. Introduction

In literature of Geographic Information System (GIS), it has been emphasized that data is fundamental. Without adequate and reliable data, GIS is not useful [1]. The adequacy and reliability of data can be achieved through data maintenance. The data element of GIS includes both spatial data (also referred to as geodata) and non-spatial data. In this paper, the interest is on geodata, which models user defined geographic entities or spatial features. Geodata is geographically referenced in some consistent manner using, for example, latitudes and longitudes, national coordinate grids, or postal codes [2].

In health sector, as observed in the studies of Chikumba [3], Msiska [4] and Saugene and Sahay [5], geodata suffers from lack of maintenance. Geodata needs to be continuously maintained to a certain

level of usefulness so that its quality is not diminished. In addition to accessibility and development of geodata as fundamental to GIS, maintenance of geodata is also of paramount importance. Data maintenance is a process of continual improvements and regular checks for keeping the high quality data.

In literature of geodata maintenance, the focus has been on update of geographic databases [1, 6]. Geodata maintenance is beyond the geographic database update. Data maintenance incorporates key components of data management process and involves organized administrative and technical actions. This paper intends to contribute to geodata maintenance in health sector in developing countries, in which GIS is becoming popular. The case used in this study is DHIS2 GIS implementation in Malawi. The aim of this study was threefold, firstly to summarize the current state of knowledge with respect to geodata maintenance, secondly to use interpretive and qualitative methods to investigate readiness for geodata maintenance in health management in Malawi and lastly to propose administrative and technical actions for geodata maintenance in health management.

## **2. Geodata Maintenance and Completeness**

### *2.1 Geodata Maintenance*

Geodata has properties that need to be well defined and continuously maintained: *what* (thematic elements, i.e. attributes), *where* (spatial elements, i.e. space or location), *when* (temporal elements, i.e. time and events) [7] and *relationship* between spatial features [8]. Although space is the dominant member of these components, time is critical to understanding phenomena as events that appear and disappear in space over time; and without theme there is only geometry [9]. The attributes represent elements that are not geometrical and are used, for example, in queries, analyses and visualization of spatial data. The degree of importance of relationships between spatial features depends on an application being used in GIS. Ulubay and Altan [8] argue that an effective geodata management requires that location and attribute data be variable independent of one another i.e. attributes can change character but retain the same location and vice-versa.

Geodata is held in a geographic database, which is constantly updated as new information arrives in order to continuously support user requirements. Longley, Goodchild, Maguire and Rhind [1] define a *geographic database update* as any change of geometry, attributes, or database schema. Required tasks include edits of geometry, attributes, indexes, topology, and even importing and exporting data. There are other required actions to be performed apart from the database update. According to British Standards Institution [10], *maintenance* is the combination of technical and administrative actions aimed at retaining an item in a state in which it can perform a required function. From the data perspective, “maintenance of (spatial) data can be defined as the combined activities to keep the data set up-to-date and as supportive as possible to the user community” [6]. Huisman and de By [6] explain that data maintenance involves acquisition and capturing of new data into the system, possibly replacing outdated data; and the need for such actions stems from requirements that the users impose or due to changes in

the real world. Hence, in this paper, *geodata maintenance* is taken as a set of organized technical and administrative actions that are carried out in order to keep geodata valuable in a particular GIS application.

Geodata, as a resource, has a value that is to be maximized. In this paper, the value of geodata is in form of a solution to a GIS problem or satisfaction of a GIS need. Thus, the value of geodata needs to be maintained in order to continuously meet GIS needs and solve GIS problems. The maintenance of geodata in an organisation leads to improved and new usage of the data thereby leveraging its value. However, maintenance of data requires a well-specified management regime [11], which means that during geodata maintenance right decisions should be made at certain points of time.

## 2.2 Geodata Completeness

Geodata requires being complete and well matched with existing data so that they can be properly integrated. If a new dataset is placed correctly into the context of other available data, it is utilized as expected [8]. It is necessary to make different datasets compatible so that they are reasonably displayed on the same map for sensible analysis [12]. In a successful integration of different datasets, the completeness is one of the concerns to be addressed. *Geodata completeness* refers to the exhaustiveness of set of spatial features and their attributes in the geographic database in relation to an abstract universe or model world [9, 13]. The abstract universe is a representation of the reality with a desired level of simplification or generalization [9]. For instance, in most GIS applications, health facilities, such as hospitals, are usually represented as points in geographic databases.

The completeness of geodata is generally assessed at two levels: model and data completeness [9, 13]. When the model and/or digital data are not complete, geodata maintenance is probably required. *Model completeness* is the commission or omission relationship between the spatial features in the model world and those in the reality whereas *data completeness* refers to the commission or omission relationship between datasets and their attributes defined in the model world and those available in the digital data [13]. Basically, according to Yang [13], *omission* means that some required spatial features and/or their attributes are not included in the geographic database while *commission* means that extra spatial features and/or their attributes, which are not necessarily required, are included in the geographic database.

Since different people use GIS for different purposes and the phenomena they study have different characteristics, right decisions should be made about what features should be modeled and how they should be represented in GIS [1]. This model completeness is application dependent in the sense that the model world is created for a particular application [13] and therefore, it is an aspect of the fitness-for-use [14]. This implies that this type of completeness is achieved if the context of GIS is properly analysed by defining problems to be solved. As Ramasubramanian [15] points out, the clarity in problem definition is critical in GIS implementation. Once a spatial feature has been modeled, it is expected that required elements (i.e. spatial, temporal, and thematic) are collected and captured into the geographic

database. At this stage, the data completeness is assessed. Data completeness is application independent and hence geodata can be shared and re-used in various contexts.

### **3. Research Context**

This study was conducted in the public health sector in Malawi from June 2015 to December 2016. Malawi is a landlocked country in southeast Africa and borders with Tanzania to the northeast, Zambia to the northwest, and Mozambique to the east, south and west. The health system has five levels of management (nation, zone, district, facility, and community). There are twenty-nine health districts, which are grouped into five health zones. In every health district, there is a district health office (DHO), which is responsible for providing all necessary support to health facilities in its health district; including planning, coordination, management, monitoring and evaluation. The health district is further divided into health facility catchment areas within which there are communities and villages.

Since 2002, there have been various initiatives in order to strengthen Malawi national HMIS and different technologies have been implemented including DHIS2 in 2012. DHIS2 ([www.dhis2.org](http://www.dhis2.org)) is a tool for collecting, validating, analysing, and presenting aggregate and patient-based statistical data, tailored (but not limited) to integrated health information management activities. From 2015, the MoH has been implementing DHIS2 GIS in order to enhance data analysis, integration and presentation. During the time of our study, CMED was carrying out a major reform on DHIS2; that is, the reconfiguration of DHIS2 in which the deployment of GIS was one of its milestones.

This study involved participants from the national and district levels. At the national level, participants were from Central Monitoring and Evaluation Division (CMED) and DHIS2 team. To strengthen its national health management information system (HMIS), the MoH established CMED that is involved in coordination, data management, advocate and facilitation of the information use in various activities such as policy formulation, planning and program implementation at all levels. Since CMED does not have adequate technical capacity for implementing and supporting DHIS2 and other technologies, in 2009, the MoH established a technical team, referred to as DHIS2 team, which is composed of information technology (IT) experts from some of its collaborating partners. CMED coordinates the activities of DHIS2 team. At the district level, participants were from Blantyre health district, which is one of the largest districts with the population of slightly over 1 million and has both city (having a central/referral hospital) and rural settings. In addition, it was close to where the authors live. Thus, it was relatively cheap to visit in terms of logistics.

### **4. Research Methods**

In this single case study, interpretive qualitative research methods were adopted to understand social and natural settings and practices on the building and maintenance of geodata. People have different ways of looking at the reality [16] and it is important to study them in their natural settings

aiming at producing factual descriptions [17]. Participant observation was the principal data collection method, which was supplemented with semi-structured interviews and artifact examination.

#### *4.1 Participant Observation*

Participant observation allowed us to observe and participate at the same time. In this research, we participated as GIS implementers in the DHIS2 GIS deployment exercise. We were part of DHIS2 team. As the implementers, we participated in various activities starting from acquisition of spatial data up to demonstration of the live DHIS2 GIS. First, geodata were acquired from various sources such as CMED, UNICEF Malawi, Jhpiego and National Statistical Office (NSO). Second, the geodata completeness were assessed in which missing geodata were identified. Third, the missing geodata were internally developed. Fourth, geodata were preprocessed to make it suitable for uploading into DHIS2 GIS. Fifth, setting up and testing of DHIS2 GIS were done concurrently. The last activity was to demonstrate the live DHIS2 GIS to some users for their feedback; three CMED managers at the national level and two HMIS officers and two health program coordinators in Blantyre health district. All demonstrations were done at the participants' workplaces.

#### *4.2 Semi-structured Interviews*

As the participant observers, interviews enabled us to step back and examine the interpretations of our fellow participants in some details [18]. Semi-structured interviews were conducted at different times and in some cases, we had to interview one participant more than once, which provided us opportunity to confirm, verify and even build on information from previous interviews [19]. At the national level, eight interviews were conducted with two CMED managers and three members of DHIS2 team. These interviewees interact with DHIS2 almost on daily basis. At the district level, five interviews were conducted with two HMIS officers and three health program coordinators in Blantyre health district. HMIS officers are the data managers and provide support to users of health information systems in their respective health districts. They work closely with health program coordinators. All interviews were conducted at the individual participant's workplace. During interviewing, we used the note-taking technique to record responses in which rough but extensive notes were made. Immediately after the individual interview, we wrote up our notes in full and sent to the respective participant through email for verification and feedback.

#### *4.3 Artifact Examination*

The examination of artifacts provided us with the avenue of gathering data apart from the participant observation and interviews. Norum [20] refers to artifacts as "things that societies and cultures make for their own use" including written texts (e.g. documents, diaries, memos, meeting minutes, letters), archival records, and those in the form of film, television, and music. Various documents were analysed which involved minutes, emails, reports, policies, strategies, manuals, and

forms. The main objectives were to understand (a) policies, procedures and strategies that govern the implementation of various technologies and data management; (b) various roles in the implementation of technologies; and (c) various ways of presenting data and information. Apart from documents, geodata of administrative boundaries and health facilities were analysed for completeness and relationships.

#### *4.4 Data Analysis*

This case study employed thematic analysis approach, which is a qualitative analytic method for identifying, analysing and reporting patterns (themes) within data. The notions of model and data completeness guided the data analysis. The concept of completeness was applied to understand required spatial features and datasets for GIS application in the health management, and what kinds of maintenance practices are expected on geodata.

### **5. Research Findings**

From the perspective of geodata completeness, three aspects have been applied in this study: the reality, the model world, and the digital data. The research findings have been presented in respect of these aspects. The aspect of reality has helped in the definition of an application domain of DHIS2 GIS in Malawi through identifying necessary spatial features and their relationships. The aspect of model world has assisted in understanding thematic and spatial elements of the spatial features for DHIS2 GIS database. The aspect of digital data has defined the geodata for DHIS2 GIS implementation. In this section, we have assessed geodata being used in DHIS2 GIS in order to understand its completeness and this has led to the identification and definition of administrative and technical actions on geodata maintenance.

#### *5.1 Institutional Arrangement – The Reality*

To determine spatial features for DHIS2 GIS in Malawi, we briefly describe its context from the institutional arrangement perspective. In this paper, *institutional arrangement* is taken as a formal governance structure that is established to manage human interactions and its rules and regulations which often form the basis for guiding activities of the organisation [21]. In the context of health management, this arrangement involves networks of entities and organisations in planning, provision, management and monitoring of health programs and services. As shown in Table 1, the focus is on management levels, healthcare services, health facilities, and CMED personnel.

**Table 1:** Management levels, healthcare services, health facilities and CMED personnel

Management Levels	Healthcare Services	Health Facilities	CMED Personnel
<b>Nation</b>	Tertiary	Central Hospitals	Health Economists, Statisticians, DHIS2 Team, M&E Officers
<b>Zone</b>			
<b>District</b>	Secondary	District Hospitals	HMIS Officers
<b>Facility</b>	Primary	Rural Hospitals down to Health Posts	Facility In-charge
<b>Community</b>		Village & Outreach Clinics	Community Health Workers

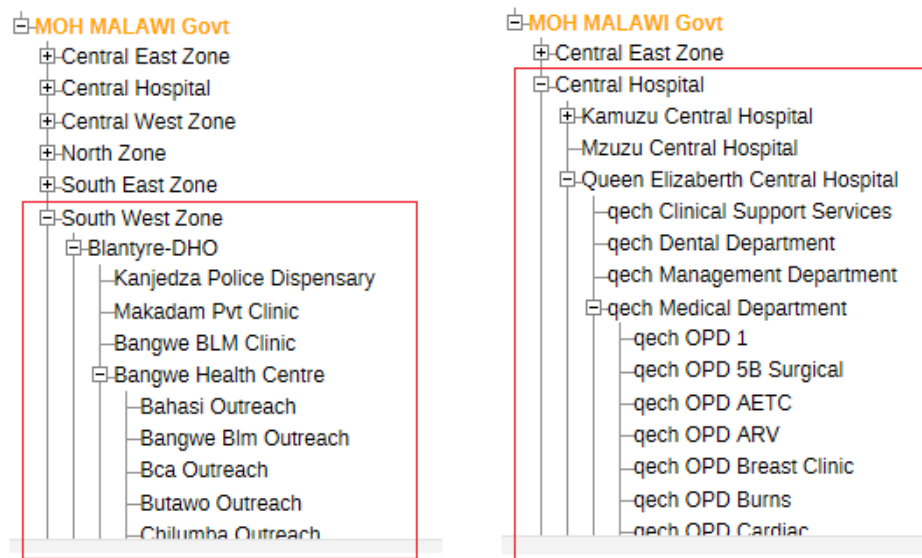
When we talk of GIS application in health, health facilities are very critical because “central to a fully operational Health Information Systems (HIS) is a basic inventory of all functioning health facilities and the services they provide” [22]. We take a health facility as any place where people can go and get required health services which are provided by a health care agency. In Malawi, all health facilities from central hospitals down to health posts have permanent structures. Village and outreach clinics have temporary structures such as houses of community health workers or small summer huts constructed by communities. In some cases, outreach clinics can be serviced at a community facility (e.g. school). The health facilities are grouped into three: primary, secondary and tertiary (see Table 1). Primary healthcare is the first contact for healthcare services which are usually provided by community/rural hospitals down to village and outreach clinics. Secondary healthcare services are provided by district hospitals. Tertiary healthcare services are provided by central hospitals.

Health facilities provide various health services to the population in their respective catchment areas. The World Health Organisation defines the catchment area as a geographic area served by a health program or institution and it is delineated on the basis of such factors as population distribution, natural geographic boundaries and transportation accessibility [23]. In Malawi’s health system, the critical catchment areas are health districts. Health facilities, except the central hospitals, are grouped by the health districts and they report to their respective District Health Offices (DHOs). The central hospitals provide specialized services to the nation at large and they report directly to the national level.

Population distribution is one of the requirements in the GIS implementation for health management in Malawi, which refers to as the arrangement or spread of people living in a given catchment area usually according to variables such as age, gender, and economic status. For example, the MoH and other stakeholders need population distribution data for estimating who may have access to the health facilities within the recommended distance of 8 kilometres. Data on population distribution in the catchment areas can help in planning, health service delivery, monitoring and evaluation, among others.

In DHIS2, the hierarchical structure has been implemented with respect to the management levels; having organisation units which represent the catchment areas and health facilities. Data analysis at the district level is usually by health facilities while at the national level it is by health districts and in some cases by health zones. DHOs need to know various events occurring in their respective health

facilities. Central hospitals have their own structures of wards, departments, and hospital. One participant pointed out that a central hospital is not really a single facility but “a group of hospitals which are represented by specialized departments.” In DHIS2, central hospitals form their own ‘virtual’ zone and have departments and wards. Figure 1 illustrates the hierarchical difference between a health district (e.g. Blantyre) and a central hospital (e.g. Queen Elizabeth).



**Figure 1:** Part of Blantyre DHO and Queen Elizabeth Central Hospital

One of the responsibilities of CMED is to implement and manage technologies, including DHIS2 GIS, with support of its collaborating partners such as universities, non-governmental organisations (NGOs), development partners, and other government agencies. This is due to inadequate technical capacity in CMED. CMED has no ICT personnel; IT Unit of the MoH is mainly responsible for hardware and software maintenance, and the payroll system. As a long-term strategy, CMED has been investing on HMIS officers at the district level in terms of GIS capacity building. One participant said: “As custodians of GIS in their respective health districts, we bought GPS for all HMIS officers and then trained them.”

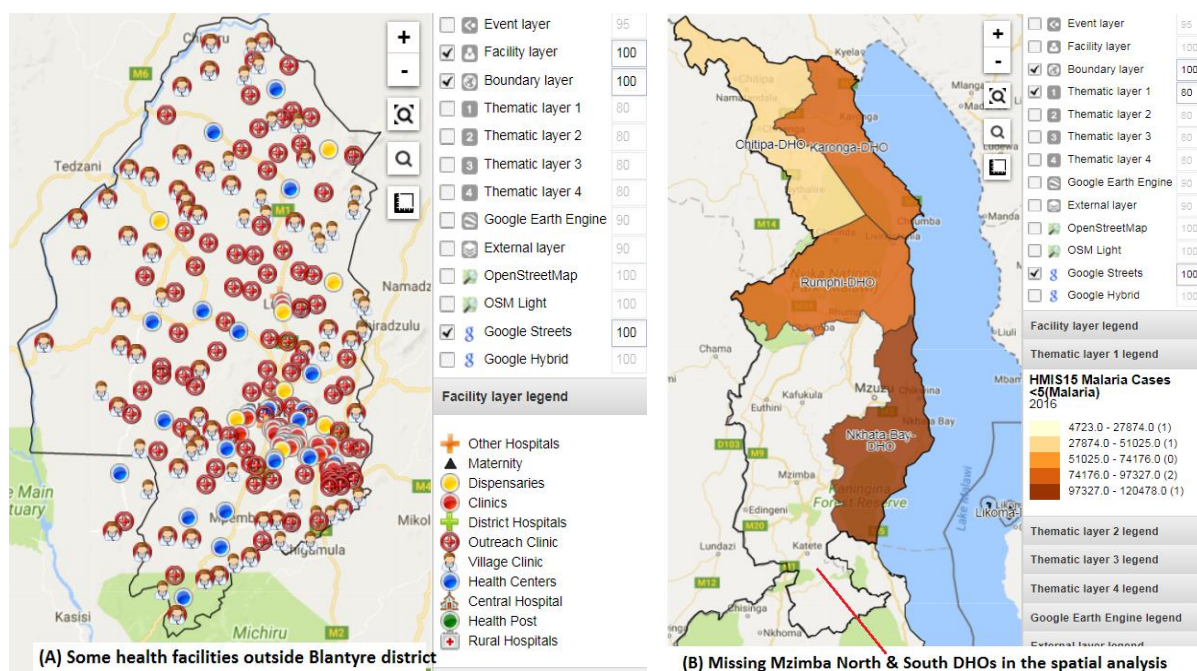
## 5.2 Model Completeness – The Model World

Since geodata models geographic entities for a particular GIS application, it is important to identify which spatial features are to be modelled. This information may help to determine relevant sources of data and build required spatial datasets. A good model contains relevant spatial features and their attributes [9, 14] so that the acquired geodata answers questions of ‘what’, ‘where’ and ‘when’ [7]. If relevant spatial features are not identified properly, this can affect the acquisition of data and even its maintenance. In this study, we have identified three basic spatial features namely health facilities, catchment areas and population distribution, which are needed by health programs and services in

DHIS2. We have assessed the model completeness of DHIS2 GIS from the perspectives of these three spatial features.

In DHIS2 GIS, health facilities are represented as points. However, the main concern is the central hospital as a reporting health facility. In practice, departments and wards usually report data. In the current status of DHIS2 GIS, spatial analysis is possible at the hospital level but not at department and ward levels. As illustrated in Figure 1, the departments and wards are included in DHIS2 but they are not effective in GIS module because they do not have spatial elements. One participant said: “Usually at this central hospital, health managers and other stakeholders demand the analysis by department or ward.” By considering departments and wards as part of DHIS2 GIS, we have noted that issues concerning map scale and level of details in the visualization should carefully be considered. For instance, if departments and wards are represented as points, the central hospital itself may be represented as a polygon. However, in DHIS2 GIS, polygons are usually used to represent health catchment areas.

Geodata of health districts, as catchment areas, is not commonly available. Instead, administrative boundaries are used. This has raised some concerns. First, in some cases, administrative boundaries mismatch with the structure of health systems. For instance, in some situations, health facilities are outside the administrative district (see Figure 2(A)). However, it is possible to have health facilities outside the administrative district “because of how the district health catchment area has been defined”- one participant commented. Second, as illustrated in Figure 2(B), there exist two health districts (i.e. Mzimba South and Mzimba North) in one administrative district (i.e. Mzimba) which has resulted in not being included in the spatial analysis.



**Figure 2:** (A) Health Facilities in Blantyre and (B) Spatial Analysis of North Zone

In version 2.21 of DHIS2, the one being used by the time of this study, population distribution has not been modeled. The version 2.24 or later of DHIS2 supports map of layers from Google Earth Engine including the population layer. Hence, to include the population distribution is just a matter of migrating to the latest version of DHIS2. The Google Earth Engine layer allows people to display satellite imagery and geospatial datasets with planetary-scale analysis capabilities. However, participants in this study mentioned age groups (e.g. under 5 children) as one specific characteristic of the population. This requirement is difficult to be met in the population layer from the Google Earth Engine. Overhead imagery has its own limitations; for example, it “may reveal a building and may even indicate that it is a factory but usually cannot detect the products made” [24].

### *5.3 Data Completeness – The Digital Data*

In Malawi, the existing spatial datasets in DHIS2 GIS include health facilities, health zones, and health districts. These datasets have spatial elements – that are represented by longitudes and latitudes – and thematic elements for describing the spatial features. Geodata of health facilities were sourced from CMED and UNICEF and that of districts were sourced from National Statistical Office (NSO) and Jhpiego. Geodata of health zones were generated internally. One important observation is that regardless of its source, every dataset requires to be verified and preprocessed. In the preprocess exercise, spatial data is converted into a form suitable for entry into GIS application [25]. In DHIS2 GIS, data sets are converted into GML (Geography Markup Language) format before being imported into the geographic database. GML is the XML grammar that serves as a modeling language for geographic systems and as an open interchange format for geographic transactions on the Internet.

The verification on geodata was performed in consultation with HMIS officers and CMED management to check for the digital data completeness. This exercise involved three main activities. First, we checked if organisation units in DHIS2 had corresponding spatial data elements and added health facilities that were not available in DHIS2. Second, we renamed organisation units which had the similar names, particularly village and outreach clinics. Third, we checked if health facilities in DHIS2 were assigned to the right facility type; otherwise affected health facilities would not correctly be analysed and visualized in GIS application.

Geodata of health facilities from CMED has few attributes (see Figure 3) as compared to that from UNICEF (see Figure 4). Attributes of geodata from CMED include facility name, type, ownership (controlling agency), and administrative location (e.g. district and region). Geodata from UNICEF has additional attributes of contact of facility in-charge (name, email, and phone), cluster, and status of facility (functional or non-functional).

YEAR	FACILITY NAME	DISTRICT	REGION	FACILITY TYPE	OWNERSHIP	SOURCE	LATTITUDE	LONGITUDE	DATUM
2013	Amitofo Care Centre	Blantyre	South	Clinic	NGO	GPS	-15,74768	35,107470	WGS84
2013	Aquid Lifeline Project	Blantyre	South	Clinic	NGO	GPS	-15,622170	35,056040	WGS84
2013	Auction Holdings Clinic	Blantyre	South	Clinic	Company	GPS	-15,811950	35,064170	WGS84
2013	Ayali Private Clinic	Blantyre	South	Clinic	Private for profit	GPS	-15,659290	35,021140	WGS84
2013	Bangwe Health Centre	Blantyre	South	Health Centre	Government/public	GPS	-15,82382	35,08046	WGS84
2013	Beit trust Cure International Hospital	Blantyre	South	Other Hospital	Mission/ Faith-based (other than CHAM)	GPS	-15,80354	35,01753	WGS84
2013	Belton Kappepye Memorial Clinic, Kappepye	Blantyre	South	Clinic	Private for profit	GPS	-15,80869	34,99434	WGS84
2013	Bernald Kachale Private Clinic	Blantyre	South	Clinic	Private for profit	GPS	-15,78018	35,07352	WGS84
2013	Blantyre Adventist Hospital	Blantyre	South	Other Hospital	Christian Health Association of Malawi (CHAM)	GPS	-15,7824	35,0031	WGS84
2013	Blantyre City Assembly Clinic	Blantyre	South	Clinic	Government/public	GPS	-15,80333	35,03895	WGS84
2013	Blantyre Dream Project Clinic	Blantyre	South	Health Centre	NGO	GPS	-15,79182	35,01273	WGS84

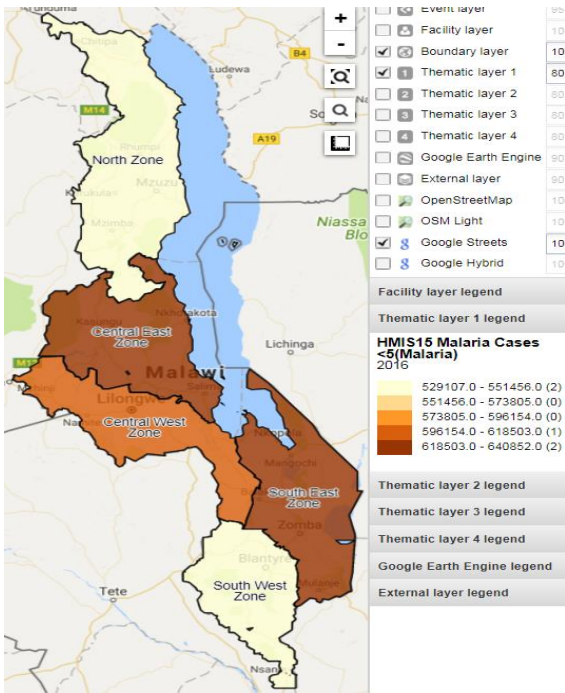
Figure 3: Part of geodata from MoH collected in 2013

Health Facility	Village	GVH	TA	District	Eastings	Northing	Type of Facility	Controlling Agency	Name	Phone	Email	Cluster	Status	Year
BAHASI	Wisaki	Mwamadi	Machinjiri	Blantyre	35,100392	-15,833177	Outreach	CBO	Brenda Katunga	881061577	N/A	Bangwe	Functional	2015
BALAZA	Balaza	Chibweya	Chigaru	Blantyre	34,905817	-15,419463	Outreach	MoH	Masauko Jembe	993901196	N/A	Mdeka	Functional	2015
BALUTI	Baluti	Chimire	Somba	Blantyre	34,991079	-15,826820	Outreach	MoH	Alick Phiri	882658056	N/A	Zingwanga	Functional	2015
BANGWE	Namatapwa	Mwamadi	Machinjiri	Blantyre	35,080477	-15,823832	Health Centre	MoH	Chastopher Mxunga	88371775	cmxunga@yahoo.com	Bangwe	Functional	2015
BANGWE BLM	Mwamadi	Mwamadi	Machinjiri	Blantyre	35,080416	-15,830897	Outreach	MoH	Lizzie Mwambazi	881558969	N/A	Bangwe	Functional	2015
BCA	Chipagala	Mtenje	Machinjiri	Blantyre	35,076611	-15,842377	Outreach	MoH	Lydia Maulidi	882919019	N/A	Bangwe	Functional	2015
BLANTYRE ADVENTIS	Blantyre City	Blantyre City	N/A	Blantyre	35,003089	-15,782349	Hospital	CHAM	Kirby Kasinja	1820488	bah.hosp@yahoo.com	Blantyre Advre	Functional	2015
BONDO	Mlomba	Likomba	Machinjiri	Blantyre	35,080881	-15,667137	Village Clinic	MoH	Simon Dzimbiri	884280169	N/A	South Lunzu	Functional	2015
BOTA	Bota	Lunguzi	Lundu	Blantyre	35,04368	-15,495089	Village Clinic	MoH	Rabson Jeremiah	888498872	N/A	Lundu	Functional	2015
BUTAWO	Butawo	Mwamadi	Machinjiri	Blantyre	35,109931	-15,808739	Outreach	MoH	Evon Ng'omba	884262347	N/A	Bangwe	Functional	2015
CHABVALA	Chabvala	Chabvala	Kassisi	Blantyre	34,834342	-15,833894	Health Centre	MoH	Calisto Ngwira	990918838	N/A	Chabvala	Functional	2015
CHAGWIRAUTA	Chagwirauta	Imbwa	Kunthembwe	Blantyre	34,858251	-15,807708	Village Clinic	MoH	Chiyangano Chalombo	881033389	N/A	Chabvala	Functional	2015
CHAKWIYA	Chakwiya	Chakwiya	Makata	Blantyre	35,102671	-15,551672	Village Clinic	MoH	Mercy Muhaniwa	996029252	N/A	Makata	Functional	2015
CHALUMBA	Makochera	Chilipa	Makata	Blantyre	35,10367	-15,662377	Village Clinic	MoH	Grey Kazembe	999218699	N/A	Makata	Functional	2015

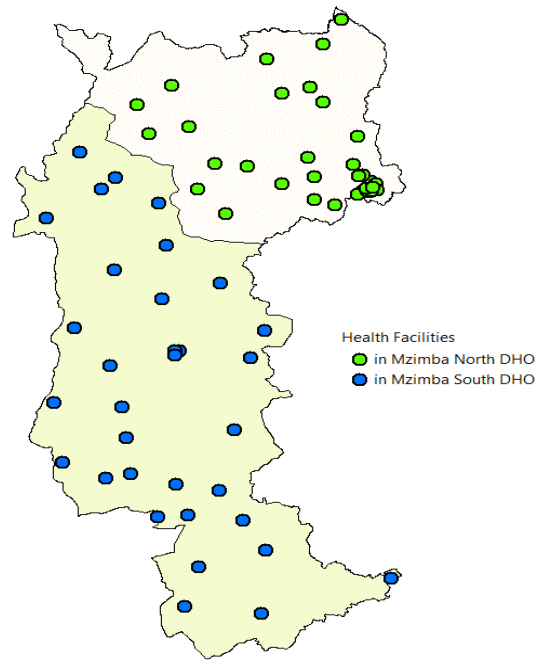
Figure 4: Part of geodata from UNICEF collected in 2015

During the service provision assessment survey in 2013, out of 1060 only 977 health facilities were visited [26], which means that 83 health facilities were not available in the dataset including Zomba mental hospital (the referral hospital). During the deployment of DHIS2 GIS, out of 8632 village and outreach clinics, 209 clinics were not added in DHIS2 due to unavailability of their respective clusters. In this context, the cluster is a health facility which a village or outreach clinic should report to. We reported this to HMIS officers in affected health districts, CMED and UNICEF for necessary actions. One important response was that some clusters had not yet been added in DHIS2 as health facilities or their names had changed. The DHIS2 GIS implementation team and HMIS officers took necessary actions collaboratively.

As mentioned earlier, administrative boundaries are used to model health districts and zones. Two main activities were done to make sure that the geodata is complete for DHIS2 GIS. First, geodata of health zones had to be created by dissolving administrative districts. In GIS, *dissolve* is a process of creating a new feature by merging adjacent polygons or lines that have a common value in a specified attribute. We applied the 'dissolve' tool in ArcGIS 10.2 to generate the new dataset of zones. As shown in Figure 5, it is possible to have the spatial analysis by health zones. Second, in the case of Mzimba, there was a need to split it into two health districts. This idea was presented to some participants in this study and the CMED management, and all agreed to it. However, both technically and legally, it is not easy to demarcate administrative area. The administrative area is legally documented and if it is to be changed, the legal protocols need to be followed. For the sake of this study and to be used in DHIS2 GIS, Mzimba was split into Mzimba South and Mzimba North by dissolving health facility catchment areas that were created in 2003. Figure 6 illustrates the demarcation of Mzimba into the two health districts.



**Figure 5:** Spatial Analysis by Health Zone



**Figure 6:** Mzimba North & South

## 6. Analysis and Discussions: Towards Geodata Maintenance

In the definition of geographic database update [1], the concept of change is emphasized. Any change to geometry, attributes or database schema is determined by a certain problem to be solved or a need to be met. Since a geographic database is always updated when new data exists [1], our observation is that before the database is updated, certain actions should be performed to generate such data. This case study has revealed such actions. Thus, this section discusses those proposed actions and the context in which they may be performed.

### 6.1 Proposed Actions for Geodata Maintenance

From the findings presented above, we have identified actions that are necessary in maintaining geodata to a certain level of usefulness. The process of maintaining geodata is initiated by a need that is imposed by a data user or due to a real world change. This need is to be communicated to an agency that assesses the need for its feasibility. Then, the agency provides feedback to the reporting agency whether it is possible to implement the change or not. If it is possible, the geographic database update is expected to take place and the reporting agency and other stakeholders are communicated accordingly. In this context, we have proposed six actions as (1) identify the need, (2) communicate the need, (3) assess the need, (4) edit the model, (5) acquire the geodata, and (6) edit the dataset.

#### 6.1.1 Identify the Need

In information systems, users always have new requirements in order to be in line with changes in the context or beyond. These new demands influence some changes in the system including data. In GIS,

geodata is among the key resources in meeting user needs. We have observed that a change in geodata can be influenced by demands from within or outside the health management system. In over 20 years ago, the political system in Malawi demanded the split of some administrative districts. It took the government of Malawi a long time to update geodata of affected districts. For instance, administrative boundaries, used in studies of Chikumba [3] and Saugene and Sahay [5], were not updated to reflect the change. The Ministry of Health (MoH) has no direct control over such external demands. Hence, in this paper, our interest is on demands imposed by GIS users within the health management system.

We have observed that demands from the national level affect two or more health districts, programs or services. For example, the decision on building geodata for Mzimba North and South health districts was from the national level. Even including the population distribution in DHIS2 GIS was the decision made at the national level. Demands at district and lower levels are particularly those affecting health facilities. HMIS officers are responsible for making sure that details of health facilities in their respective health districts are always up-to-date in DHIS2. At the community level, village and outreach clinics are under the management of community health workers (CHWs) who are probably to identify any changes affecting geodata of those health facilities.

Demands can also come from a service or program. For instance, in planning and monitoring service, DHIS2 users need to use GIS in assessing the accessibility of health facilities, which may lead to determination of locations of new health facilities. The nutrition program expects GIS to enhance the monitoring and tracking services by providing spatial visualization of locations of its clients and associated logistics resources.

#### 6.1.2 Communicate the Need

All health-related issues concerning the nation are managed at the national level and those concerning individual districts are managed at the district level. The data management in HMIS follows this type of governance. CMED officers manage data at the national level while HMIS officers are data managers at the district level. We have observed that issues concerning geodata of health facilities are reported to HMIS officers and those concerning the geographic database model, like the issues of Mzimba and population distribution, are reported to CMED at the national level. Even issues being raised by health program coordinators, such as patient tracking in the nutrition program, are reported to CMED at the national level because it involves many health districts. These demonstrate that demands should be communicated to the right authority or office for further assessment.

#### 6.1.3 Assess the Need

We recommend that a change on geodata should carefully be analysed in order to determine its complexity and required resources for its implementation. The nature of change can differ depending on what part of geodata is likely to be affected. For instance, changes on the database schema require the technical know-how on database management and GIS, including programming in some cases, which

may result in consuming more technical resources than those changes on attributes. Some concerns in the assessment of the need are:

- Level of the change – it can be on the database model, spatial or thematic elements.
- When to implement the change – a change can take time (e.g. boundary data); or it needs urgent attention (e.g. a new health facility); or it can wait; or even not to be implemented.
- Responsible personnel and required competence – a change can be done at the district level or national level by the internal personnel or the external expertise.
- An expected effect on GIS technology and related systems – sometimes, a change needs the upgrade of technology through, for example, migrating to the latest version, developing an additional application, or reconfiguring some components.

#### 6.1.4 Edit the Model

According to Longley, Goodchild, Maguire and Rhind [1], *editing* is the process of making changes to a geographic database, which involves adding new objects or replacing existing ones. Generally, the type of change can be either a new spatial feature or new data entry. If a new spatial feature is demanded, the model may be modified. In this process, the new spatial feature is to be integrated with existing ones in the geographic database so that it can easily be analysed, presented and visualized [8, 12]. The focus is on defining relationships between spatial features, which is the application-dependent.

#### 6.1.5 Acquire the Geodata

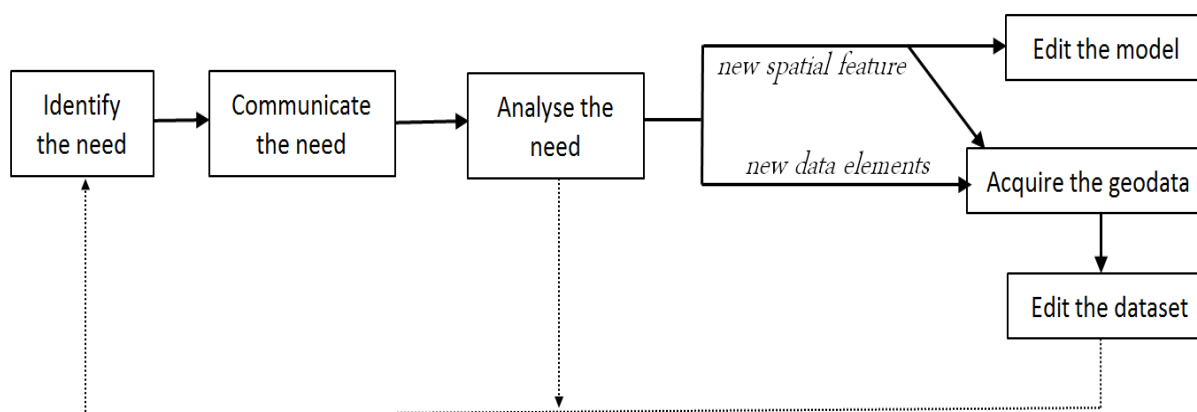
Geodata can be acquired through various ways such as being sharing, internally developing or capturing in the field. “Geodata may be collected by both government organisations as well as private agencies ... be shared and re-used by different users and applications” [2]. This is achievable if there are no critical omissions or commissions and the dataset is in the suitable format for maximum efficacy [27]. The organisation is forced to capture geodata in the field or develop internally if there is no other source to get such geodata. In collaboration with other organisations, the MoH captured the geodata of health facilities.

#### 6.1.6 Edit the Dataset

The new data entry involves changes on spatial and/or thematic elements; add new contents or replacing outdated ones. Mzimba South and North health districts were added as new objects in the health district dataset, in which spatial and thematic elements were entered into the geographic database. The health zones, as catchment areas, were already modeled in DHIS2 and what remained was to acquire and update their spatial elements.

## 6.2 Proposed Model for Geodata Maintenance

From the discussions above, in Figure 7, we illustrate how the proposed actions for the geodata maintenance are related. As shown in Table 2, they are categorized as technical and administrative actions. In this paper, we take both administrative and technical actions as any actions or decisions made by stakeholders of a particular GIS application that affect the maintenance of geodata. Technical actions require the practical knowledge of GIS while administrative actions may not really need such knowledge.



**Figure 7:** Proposed Actions for Geodata Maintenance

**Table 2:** Categories of Proposed Actions

Action	Category	Description
<b>Identify the Need</b>	Administrative	A stakeholder in the GIS application domain identifies a need for change in the system.
<b>Communicate the Need</b>	Administrative	The need is communicated to a responsible agency for further actions. The reported need is to be documented for sharing and referencing.
<b>Assess the Need</b>	Administrative/ Technical	This involves both administrative and technical decisions. If the reported need is complex and external support is required, collaborating partners can be consulted for assistance in terms, for example, finances, expertise or materials. Basic knowledge of GIS and geodata is required.
<b>Edit the Model</b>	Technical	This is the responsibility of technical people, who have the practical knowledge of GIS including database management. In most cases, these are developers or implementers of GIS.
<b>Acquire the Geodata</b>	Technical	This requires geodata sharing and practical knowledge of geodata collection using at least GPS. It also requires basic knowledge of coordinate systems e.g. longitudes and latitudes.
<b>Edit the Dataset</b>	Technical	This is the responsibility of technical support teams (e.g. DHIS2 team at the national level and HMIS officers at the district level), who have practical knowledge of GIS and geodata.

In the context of government, geodata can generally be in three categories namely physical data (physical features on the Earth); political data (artificial designations that define an area as part of political entity); and socioeconomic data (e.g. population, economics, and social patterns) [24]. We treat

health facilities as physical data; health districts and zones as political data; and population distribution as socioeconomic data. In terms of acquisition and management of these types of data, each category has unique characteristics. We have observed that health facilities require frequent changes, particularly village and outreach clinics. Due to their temporary structures, village and outreach clinics are likely to change locations over a short time or even become non-functional, and new clinics are established. Even thematic elements, such as contact details of a facility in-charge (see Figure 4), are expected to change frequently due to the mobility of health personnel.

Although the users of DHIS2 GIS are at the national and district levels, community health workers (CHWs) at the community level may play a great role in the maintenance of geodata, particularly that of village and outreach clinics. Since geodata is a new phenomenon to CHWs, it is important to have a conducive environment for promoting their involvement in geodata maintenance. They can be involved in *identify the need* and *communicate the need* processes. As observed in some studies – for example Moyo, Nkhonjera and Kaasbøll [28] – both technical and social systems should be considered in the implementation of technical phenomena at the facility and community level; including management commitment and individual attitudes towards a phenomenon.

Geodata maintenance requires resources such as people, hardware, software, procedures, reliable communication systems and finances among others. Collaboration has enabled the MoH to acquire necessary geodata for DHIS2 GIS. We argue that collaboration is also useful in the geodata maintenance, mainly in *edit the model* and *acquire the geodata* processes. Since it lacks the IT capacity, CMED is unlikely to carry out its geodata maintenance activities without being assisted by other organisations. In the health sector, collaboration is a key to success in the delivery of health program or service, and implementation and maintenance of technologies due to the complexity of health problems. Collaboration can accomplish complex and far-reaching tasks more effectively than individual institution doing alone. CMED can utilize existing collaborations to complement its resources in geodata maintenance as it has done in other GIS initiatives.

## **7. Conclusion**

In literature of data maintenance in GIS, much has been discussed on the update of geographic databases, which involves changes on spatial and thematic elements, and database schema (model). From the discussions above, we relate *edit to the model* to changes on the database schema and *edit the dataset* to those on geometry and attributes (i.e. spatial and thematic elements). This study has demonstrated that, additional four actions are required to have a complete set of geodata maintenance operations involving both administrative and technical actions. The four actions are *identify the need*, *communicate the need*, *assess the need* and *acquire the geodata*. As Longley, Goodchild, Maguire and Rhind [1] and Huisman and de By [6] point out, geographic database is always updated when new data arrives or is acquired. These four additional actions are necessary to make new data available for the

geographic database update. However, these proposed actions are based on our observations during the implementation of DHIS2 GIS in Malawi. Further studies are required.

From the definition of GIS implementation as an on-going process of decision-making through which users become aware of, adopt, and use GIS [15], we take geodata maintenance as part of this process. All involved agencies need to be aware of and adopt good practices on geodata maintenance in order to build a culture towards the sustainability of DHIS2 GIS. Geodata maintenance as an innovation needs to be incorporated into operations of HMIS. In order to promote the involvement of various stakeholders at different levels, there is a need, for example, to have reliable management support and ongoing communication among those stakeholders.

## References

1. Longley PA, Goodchild MF, Maguire DJ, Rhind DW. *Geographic information systems and science*. 2nd ed. West Sussex: John Wiley & Sons Ltd; 2005.
2. Awange JL, Kiema JBK. Environmental geoinformatics: monitoring and management. *Environmental Science and Engineering*. 2013.
3. Chikumba PA. Application of geographic information system (GIS) in drug logistics management information system (LMIS) at district level in Malawi: opportunities and challenges [Thesis]. Norway: *University of Oslo*; 2009.
4. Msiska B. *Challenges and opportunities in using GIS for monitoring and management of HIV/AIDS: a case study from Malawi* [Thesis]. Norway: University of Oslo; 2009.
5. Saugene ZB, Sahay S. The challenge of customizing global open source software to local country contexts: the case of DHIS GIS for health management. *11th International Conference on Social Implications of Computers in Developing Countries*; Kathmandu, Nepal; 2011.
6. Huisman O, de By RA. Principles of geographic information systems: an introductory textbook. Enschede: *The International Institute for Geo-Information Science and Earth Observation (ITC)* ; 2009.
7. Peuquet DJ. *Representations of space and time*. New York: The Guilford Press; 2002.
8. Ulubay A, Altan MO. A different approach to the spatial data integration. Symposium on Geospatial Theory, *Processing and Applications*; 2002; Ottawa, Canada; 2002.
9. Veregin H. Data quality parameters. In: Longley PA, Goodchild MF, Maguire DJ, Rhind DW, editors. *Geographical information systems*. 2nd ed. New York: *Wiley*; 1999. p. 177-189.
10. British Standards Institution. *Glossary of terms used in terotechnology*. London: BSI; 1993.
11. IGGI Working Group. *The principles of good data management*. 2nd ed London: ODPM Publications; 2005.
12. Flowerdew R. Spatial data integration. In: Maguire DJ, Goodchild MF, Rhind DW, editors. *Geographic information systems*. London: *Longman*; 1992. p. 337-358.
13. Yang T. Visualisation of spatial data quality for distributed GIS [Thesis]. Australia: *The University of New South Wales*; 2007.
14. Joksic D, Bajat B. Elements of spatial data quality as information technology support for sustainable development planning. SPATIUM No 11, *Institute of Architecture and Urban & Spatial Planning of Serbia*. 2004; 11:77-83.
15. Ramasubramanian L. GIS implementation in developing countries: learning from organisational theory and reflective practice. *Transactions in GIS*. 1999; 3(4):359-380.
16. Hancock B, Ockleford E, Windridge K. An introduction to qualitative research. London: The NIHR RDS EM/YH; 2007.

17. Ebrahim GJ, Sullivan KR. *Mother and child health - research methods*. London: Book Aid; 1995.
18. Walsham G. Interpretive case studies in IS research: nature and method. *European Journal of Information Systems*. 1995; 4:74-81.
19. Bolderston A. Conducting a research interview. *Journal of Medical Imaging and Radiation Sciences*. 2012;43(1):66-76.
20. Norum KE. Artifacts. In: Given LM, editor. *The SAGE encyclopedia of qualitative research methods*. London: *SAGE Publications*; 2008. p. 25-26.
21. Nemarundwe N. Institutional collaboration and shared learning for forest management in Chivi district, Zimbabwe. In: Wollenberg E, Edmunds D, Buck L, Fox J, Brodt S, editors. *Social learning in community forestry management: linking concepts and practices*. Bogor: *Center for International Forestry Research*; 2002. 85 - 108
22. Noor AM, Alegana VA, Gething PW, Snow RW. A spatial national health facility database for public health sector planning in Kenya in 2008. *International Journal of Health Geographics* 2009;8:13.
23. WHO. Ageing and health technical report vol.5: a glossary of terms for community health care and services for older persons. WHO/WKC/Tech.Ser./04.2; 2004.
24. Mennecke BE, West Jr. LA. Geographic information systems in developing countries: issues in data collection, implementation and management. *Journal of Global Information Management*. 2001; 9(4):45-55.
25. Barrington MR, Johnson DE, Harris NR, Krueger KM. A digital elevation model and associated GIS data bases for the Northern Great Basin experimental range. Corvallis, Oregon: Oregon Agricultural Experiment Station; 1994.
26. Ministry of Health, ICF International. Malawi service provision assessment (MSPA) 2013-14. Lilongwe, Malawi, and Rockville, Maryland, USA: *MoH and ICF International*; 2014.
27. BGS International. *Geodata for development: a practical approach*. Esource book. Nottingham: International Geoscience Services Ltd; 2012.
28. Moyo C, Nkhonjera T, Kaasbøll J. Using mobile technology to improve timeliness of data reporting: a case study from Malawi. In: Cunningham P, Cunningham M, editors. *IST-Africa conference week 2015 6-8 May 2015; Lilongwe, Malawi*; 2015.