USE OF COMPUTERIZED TEMPLATES AND MOBILE TECHNOLOGY IN PUTTING EPIDEMIOLOGICAL SURVEYS INTO OPERATION IN THE FIELD.

Onicio Batista Leal Neto[1]; Elainne Christine de Souza Gomes[1]; Cesar Martins de Albuquerque[2]; Jones Albuquerque[2,3] & Constança Simões Barbosa[1].

[1]Aggeu Magalhães Research Center – Fundação Oswaldo Cruz, Av. Professor Moraes Rego, s/n – Cidade Universitária. Cep: 50.670-420 – Recife, Pernambuco, Brazil.

[2]C.E.S.A.R. edu - Rua do Bione, Porto Digital – Recife, Pernambuco

[3]UFRPE – Departamento de Estatística e Informática – Recife, Pernambuco

Author for correspondence: Onicio B. Leal Neto – onicio@cpqam.fiocruz.br

Abstract: This paper presents the methodology for constructing computerized templates serving as an instrument of understanding of infectious disease epidemiology scenarios, providing subsidies to intervention. Additionally being developed an application for the Android platform that will help field investigative processes for endemic diseases, performing as an instrument of great value for epidemiological studies of parasitic diseases.

Key-words: Public Health, Epidemiology, Geoprocessing, Android Platform.

Background:

To understand endemic processes and their influence on the health of populations, it is indispensable to conduct studies that promote such understanding. In epidemiological surveys in the field, when the aim is to detect the occurrences, distribution and magnitude of a given health condition, a conception of the survey area as a space where individuals live together becomes necessary (Barcellos et al, 2002).

In Brazil, some healthcare services still manually construct templates for the localities where epidemiological surveys will be carried out, following the geographical recognition manual of the National Health Foundation (FUNASA, 1996) to the letter. This is a laborious activity without operational practicality. It is expensive and slow, given that the field technicians have to draw diagrams manually, guided by the sun and cardinal points. The sketches produced are often imprecise and liable to lead to interpretational error.

With the advent of geoprocessing and remote sensing technologies that use the Global Positioning System (GPS), it has become more practical to put field data into their special locations. Through using GPS, it is possible to construct tracks of localities (paths followed by individuals who were using GPS devices) and thematic networks containing computer-manipulated georeferenced elements, with emphasis on point records of disease cases or vector foci. The data gathered is used to generate thematic maps that make it possible to visualize the local epidemiological landscape. These maps constitute a valuable analytical instrument for health surveillance, since they highlight the spatial distribution of events and improve the way in which they can be seen, thereby contributing towards the etiological exploration of health events (Gleiser, 2010). All these advances and technologies are at the disposal of healthcare services at a relatively low cost, when the speed of obtaining and manipulating the health information is taken into consideration, along with the optimization of human and material resources in developing these activities.

With the evolution of mobile technology and smartphones, it has been possible to develop open source platforms, i.e. open-access programming code in which users with software programming knowledge can develop applications that are specific for the healthcare field and make them available online for download (Keller et al, 2009). Use of the Android platform (a smartphone system developed by Google) within epidemiology is a new alternative for computerization of epidemiological studies that facilitates the field technicians' work process through reducing errors and creating a means of instantaneous remote communication between information gathered in the field and an online virtual database (Freifeld, 2010).

The aim of the present study was to highlight the use, efficacy and practicality of these instruments in epidemiological studies in the field, taking the example of data recently gathered in surveys conducted by the research group of the Schistosomiasis Reference Service and Laboratory of the Aggeu Magalhães Research Center, Fiocruz, Pernambuco. Construction of computerized templates became a practical necessity in order to attain the level of precision and detailing that studies demand, while also giving priority to speed and accuracy in the operational processes in the field.

Methodological procedures:

1. Acquiring satellite signals

GPS devices need to acquire satellite signals in order to refine their precision, given that there is a margin of error within the process that ranges from 1 to 10 meters. This can be minimized if the GPS device is initialized 15 minutes before starting the data gathering. The signal level display on the device needs to be checked to ensure that it is visible and that at least three satellite intensity bars are complete. To create a database within the Geographical Information System (GIS), the geodesic system was chosen with the aim of achieving the best correlation between the geoid and the reference ellipsoid. These are the elements that make it possible to understand the real shape of the terrestrial globe and its assumed mathematical surface, respectively. Since the geoid is irregular, there is no standard ellipsoid, and so each country uses the one that is most appropriate for its geographical reality. In Brazil, the geodesic system most used is the South American Geodesic System (SAD-69). The coordinate projection format that was used as the reference standard for constructing the templates was the Universal Transverse Mercator (UTM) system, since this is the system that has been used by the Brazilian institutions responsible for systematic mapping of the country since 1955, and this has been used in a variety of topographic mapping studies (Burrough, 1998).

2. Operating the GPS device in the field

The GPS device selected needed to have track and waypoint construction tools. The main purpose of track construction is to be able to virtually trace out the path that the technician followed in the field. In the geographical recognition process for constructing the template for the locality of Porto de Galinhas, for example, this option was used for producing virtual drawings of the street blocks at this locality by georeferencing the block *in loco* at the same time as each building received a specific registration number. The numbering was done in increasing order, from 1 to infinity (in an S shape), going from left to right and returning from right to left along the second line of blocks, in a sequence such that each block would be closest to the subsequent one, thereby avoiding large shifts from one to another, as can be seen in Figure 1 (FUNASA, 1996).
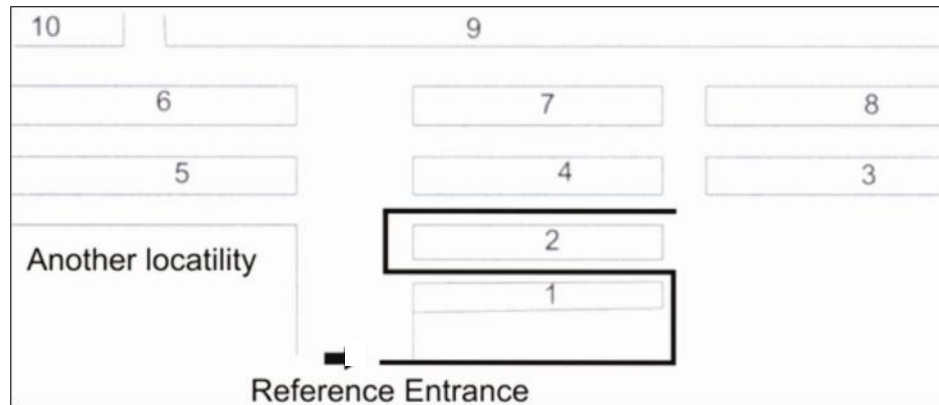
Figure 1: Demonstration of the path to be followed (FUNASA, 1996).

3. Construction of the template and the spatially organized thematic maps

The georeferenced data from the field were transferred from the GPS device to a computer via a Universal Serial Bus (USB) port, and the raw data were modeled using the GPS TrackMaker Pro software (Geo Studio) to adjust the precision of the tracks, identify water accumulations, breeding sites, foci and homes with individuals parasitized by *S. mansoni*. Since the epidemiological studies that had been done required greater detailing, new analysis tools that would enable construction of risk maps using the tool of spatial analysis were required. In this manner, the environmental variables gathered could be correlated with the epidemiological variables diagnosed. For this type of analysis, the data were transferred to the TerraView version 3.5 free spatial analysis software (www.inpe.br). This made it possible to make descriptive and spatial statistical inferences and to refine the data and include new information on the thematic maps.

Once the template had been finalized, a variety of possibilities for data analysis and/or consultation opened up. Among the procedures for estimating event density, the kernel estimate is the best known and most used procedure, and this plays an important role within the epidemiological context (Gatrell *et al.*, 1996) for identifying the concentrations of given health conditions. Kernel estimates consist of a statistical technique for exploratory nonparametric interpolation that generates a density surface for visual identification of hot spots, which are formed by concentrations of indicator events that are agglomerated in a spatial distribution. The distribution of points or events is transformed into a continuous graduated surface that then expresses occurrences of risks inherent to the health condition under examination. This procedure

makes it possible to filter the variability of a dataset, without altering the essence of its local characteristics (Bailey and Gatrell, 1995).

4. Use of smartphones in the field.

The research group of the Schistosomiasis Reference Service tested out a smartphone device for use in mapping the locality of Porto de Galinhas in November 2010, using an application for the Android 2.1 platform that was developed by the study team. The ANKOS application (A New Kind Of Simulator) was considered, given the need to develop a portable computing environment for simulations that would be easy to access and handle, and would hold a database created to register the families in the study. In addition, this instrument made it possible to georeferenced events such as vector foci and breeding sites and positive human cases of *S. mansoni*. It also allowed the image files to be correlated with the data on the forms (Lillesand *et al*., 2004). After the construction stage of this model, the Android platform (the operating system for the device in question) was adapted for use in the smartphones. The research group also created a homepage (www.epischisto.org), which functioned as a remote-access database for the researchers and students in the group, thus ensuring that there was an instantaneous backup for the data generated in the field. The hardware used for the application was the Motorola model MB 502, which was shown to be a robust device. It has a QWERTY keyboard, which facilitated data input in the field. It has an integrated GPS receiver, thus promoting the georeferencing of the information acquired, with data storage in the device through a flash memory and with synchronized data transfer through the 3G mobile network to the website www.epischisto.org, where storage in the database took place.


Figure 2: Tracking produced by Android in the locality of Porto de Galinhas.

4.1 ANKOS

A layered pattern was used for the system solution because of the characteristics and operational requirements of the ANKOS system, as a future simulation environment. This pattern enabled decoupling between data presentation, business rules and data persistence, taking the characteristics presented above into consideration.

4.1.1. Layers

The model used was divided into three layers, as described in the following:
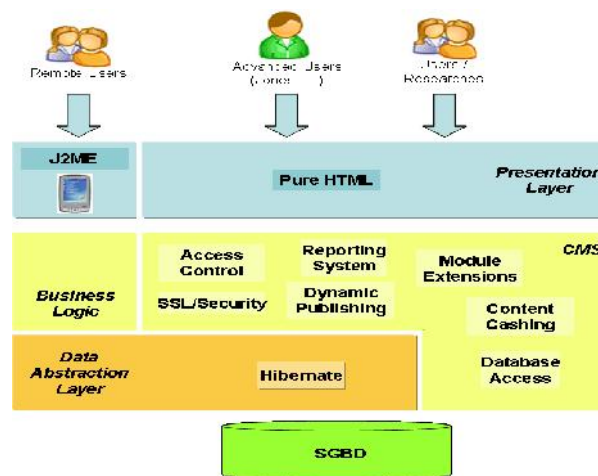
- **DBMS (Database Management System)**
  This is a relational database that stores the data in a consistent filing system. It organizes the data from field surveys, entered by health agents, into tables. The maps obtained are also stored here.
- **Data Abstraction Layer (DAL)**
  The data abstraction layer provides a set of interfaces with the layer immediately above it, with the aim of abstracting a non-volatile storage system. To translate business classes into relational tables, the Hibernate technology is used.
- **Mobile Business Logic**
  This layer concentrates the business rules of the ANKOS system, for interactions through the mobile-ANDROID interface.
- **Presentation Layer**
  This is responsible for supplying an access interface to the system developed by the different types of user and for the different forms of interaction. For the web system, HTML enables data entry.
- **CMS**
  In addition to the facilities mentioned earlier, CMS promotes extension of its functionality through modules. For the ANKOS system, a specific module was developed in Plone/Zope. This module concentrates the business rules for the web interface, as well as forms, consultation screens and a reporting system.

4.1.2. Technology

The technologies that were used in developing the system were the following:

- ANDROID platform
- Programming language: Java
- Language for displaying dynamic content in HTML pages: JSP
- CMS: Plone/Zope
- Database: MySQL
- Others: Hibernate, XML

Results

In the field survey carried out in Porto de Galinhas, use of GPS made it possible, over a 60-day period, to map out 4300 buildings, 162 street blocks and 38 breeding sites for vector molluscs, covering an area of approximately 10.400 km². In easily accessible streets, the path could be followed and georeferenced with the surveyor using the GPS device inside the field vehicle, while the other streets were surveyed on foot. The data that were gathered and imported to the GPS TrackMaker Pro software resulted in the template for the locality (Figure 3). In the same way, all water accumulations in which the mollusc that is the intermediate host of *S. mansoni* was present were georeferenced.

Figure 3: Template for the locality of Porto de Galinhas, Pernambuco.

Figures 4 and 5 show the evolution of the template into thematic maps with a variety of possibilities for analysis on the gathered data, such as construction of an intensity estimator (kernel), definition of areas of influence for the health condition (buffer) or aggregation of events that occurred in nearby areas (clusters).

The GPS integrated with this device georeferenced the information acquired, which was stored in the device in a flash memory, with synchronized transfer through

the 3G mobile network to the website www.epischisto.org, where the research group could immediately access the information and thus evaluate it and redirect the survey stages.

Discussion

For epidemiological analysis on data gathered in population-based studies, it becomes essential to use spatial analysis tools for health data, in preparing risk maps showing locations of events and the conditions associated with such occurrences. For this type of analysis, a unified georeferenced cartographic base has to be available, for use in representing the set of thematic networks of a given locality. In Pernambuco, these unified bases have only been produced for some cities in the metropolitan region and a few other isolated municipalities. Therefore, to conduct studies in places where these geodata are unavailable, it becomes necessary to construct georeferenced templates as a cartographic base, in order to enter information generated through epidemiological surveys.

Figures 2, 3, 4 and 5 are results generated through the practicality, speed and precision of using computerized tools and resources, which also show the epidemiological information graphically. These investigative and data analysis tools for generating health mapping have applicability that should go beyond academic studies: they are shown here to be useful instruments for planning health service actions, with regard to primary healthcare activities and health surveillance.
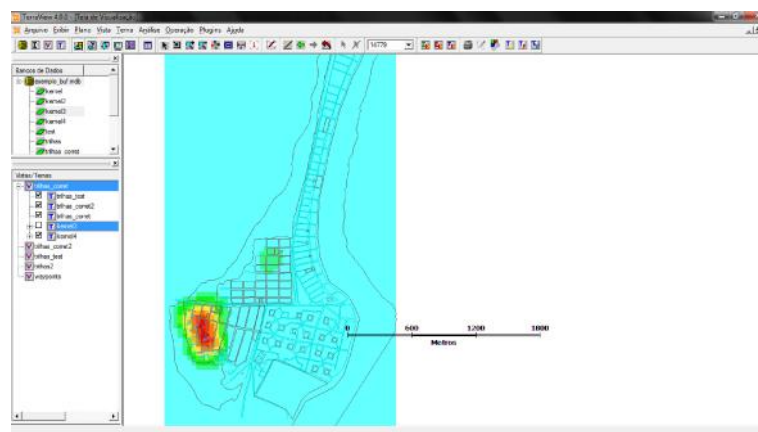

Figure 4: Thematic map with representation of the kernel estimator for breeding sites of *Biomphalaria glabrata*.

The evolution from a simple template (Figure 3) showing the exact locations of cases and foci of the molluscs that are vectors for the disease, to the thematic maps (Figures 4 and 5) improved the analyses on the gathered data and provided greater quantities and quality of information. This in turn provided greater support for correlations regarding the causal nexus of this disease in this location. In Figure 4, construction of an intensity estimator (kernel) made it possible to calculate the extent of the risk of becoming ill due to schistosomiasis. In Figure 5, use of the buffer technique made it possible to show the areas that are subject to influence for occurrences of the health condition, through aggregation of events happening in nearby areas (clusters).
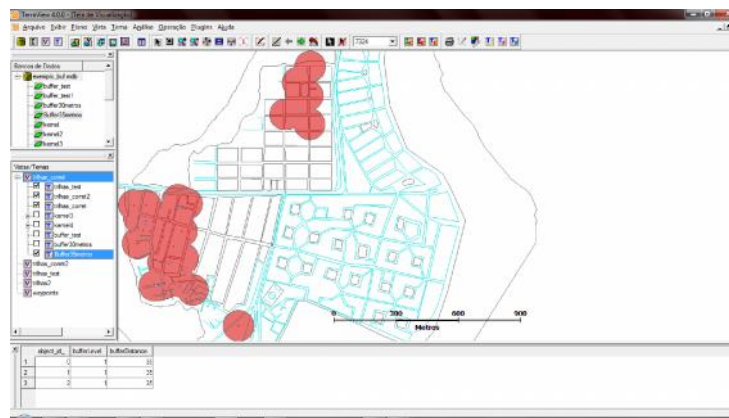

Figure 5: 35-meter buffer for the breeding sites of *Biomphalaria glabrata*.

With a view to future perspectives for aiding in the field stages of epidemiological studies, the Epi Schisto Risk Modeling research group (www.epischisto.org) is testing an application to be used in mobile devices (smartphones) that will replace the traditional method of data gathering, i.e. an exclusive device for GPS and the patients' registration files. Starting from the concepts of *Infodemiology* and *Inforveillance* described by Eysenbach (2009), it is now possible to manipulate epidemiological data via the internet, such that the data are gathered and analyzed in tempo real. With this logic as the starting point, the group has been developing studies applied to the realities of Pernambuco: the Android platform for promoting integration of data gathered in the field has been constructed and used, which has enabled immediate remote analysis for generating thematic maps, thereby providing rapid data interpretation and epidemiological diagnoses for the events.

The ANKOS application has demonstrated other possibilities, going beyond speeding up the analysis processes and spatially organizing the georeferenced data, in

terms of enabling storage of images from the study area and viewing them in real time both on the smartphone and on the group's homepage. This enables instantaneous viewing of the study location, even if the research group is far away. The study that has been developed at Porto de Galinhas has been the pilot study for validating the software and these new tools, which can be used in new surveys for data gathering and analysis, thus making it possible to generate and disseminate knowledge on the epidemiology of schistosomiasis more rapidly.

References

Bailey, T.C.; Gatrell, A.C. **Interactive spatial analysis.** London: Longman Scientific and Technical, 1995. 413 p.

Barcellos, C. C.; Sabroza, P.C.; Peiter, P.; Rojas, L.I.: Organização Espacial, Saúde e Qualidade de Vida: Análise Espacial e Uso de Indicadores na Avaliação de Situações de Saúde. Informe Epidemiológico do SUS 2002; 11(3) : 129 - 138.

Burrough, P. A. , R. McDonnell. **Principles of Geographical Information Systems** . Oxford University Press, 1998. 333 p.

Eysenbach G (2009) Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 11: e11.

Freifeld, C.C.; Chunara, R.; Mekaru, S.R.; Chan, E.H.; Kass-Hout, T.; Iacucci, A.A.; Brownstein, J.S.: Participatory Epidemiology: Use of Mobile Phones for Community-Based Health Reporting. PloS Medicine, Volume 7, Issue 12, December 2010.

Fundação Nacional de Saúde. Manual de Reconhecimento Geográfico/Fundação Nacional de Saúde – FUNASA, 1996.

Gatrell, A. C.; Bailey, T. C.; Diggle, P. J.; Rowlingson, B. S. Spatial point pattern analysis and its application in geographical epidemiology. **Transactions, Institute of British Geographers**, v. 21, p. 256-274.

Gleiser, R.M & Zalazar L.P.: Distribution of mosquitoes in relation to urban landscape characteristics. Bulletin of Entomological Research (2010), 100: 153-158.

Keller, M.; Blench, M.; Tolentino, H.; Freifled, C.C.; Mandi, K.D.; Mawudeku, A.; Eysenbach, G. and Brownstein, J.

Lillesand, T. M.; R. W. Kiefer, and J. W. Chipman. Remote Sensing and Image Interpretation. Jhon Wiley & Sons, 2004, 5a Edição.