Developing a Data Dashboard for Population Health Surveillance: Widening Access to Clinical Trial Findings

David Concannon
Institution: The Bartlett Centre for Advanced Spatial Analysis, University College London
Address: Gower Street, London WC1E 6BT, United Kingdom
Email: david.concannon@ucl.ac.uk

Kobus Herbst
Institution: Africa Health Research Institute
Address: Africa Centre Building, Via R618 to Hlabisa, Somkhele, Mtubatuba, South Africa
Email: kherbst@ahri.org

Ed Manley [Corresponding Author]
Institution: The Bartlett Centre for Advanced Spatial Analysis, University College London
Address: Gower Street, London WC1E 6BT, United Kingdom
Email: ed.manley@ucl.ac.uk
Abstract

Background

Demographic surveillance platforms play a vital role in assessing the effects that disease has on a population, and the resulting impact of subsequent healthcare interventions. Population surveillance sites generate many datasets relevant to disease surveillance, however there is a risk that this data is under-utilised, misunderstood or not acted on in a timely manner due to the volume of data captured. Data visualisation offers stakeholders a means to quickly understand and interpret collected data.

Objective

This paper describes the development and evaluation of dashboard to visualise trial to staff and researchers at a demographic surveillance site and investigate the role that visualisation could play increasing the visibility and understanding of datasets produced therein.

Methods

This paper presents the development of a dashboard for visualising data generated within, a demographic surveillance platform at the Africa Health Research Institute (AHRI), in KwaZulu-Natal, South Africa. An evaluation study was undertaken to assess the effectiveness of the dashboards as a data dissemination tool. A mixed-methods approach combined benchmark task evaluation to assess usability with a questionnaire to study attitudes to the use of dashboards. The evaluation was recruited 20 participants drawn from scientific, operational nursing and community advisory staff working at AHRI.

Results

The results of the questionnaire showed that the majority of respondents felt that the dashboards provided a clear understanding of the results of the trial presented and would like to see the use of dashboards within the research centre to disseminate results. The evaluation demonstrated high usability for the dashboard across the study groups with scientific and operational staff having minimal issues in completing the tasks outlined. There were notable differences in the efficiency of task completion among different groups of respondents, indicating varying familiarity with data visualisation interfaces.

Conclusions

The paper has demonstrated the viability of data visualisation dashboards as a means of increasing the visibility and access to datasets at a population surveillance site. The usability differences between the different groups demonstrates the need for user-led design of dashboards in future, addressing varying computer and visualisation literacy’s present among user groups.

Keywords

Data visualisation; Data dashboards; Health and Demographic Surveillance; Sub-Saharan Africa; Treatment as Prevention; Clinical trials; Demographics; Real-time; Data literacy.
Introduction

Description of need for HDSS in sub-Saharan Africa

Demographic surveillance, the process of monitoring births, deaths, causes of deaths, and migration in a population over time, is one of the cornerstones of medical and public health research (Chandramohan et al., 2008; Byass et al., 2002). The availability of detailed data on these population statistics is essential to the planning, implementation and evaluation of any public health intervention (Baiden et al., 2006). In many developing countries there is a lack of comprehensive national vital statistics systems (Kahn et al., 2007). Without vital data on demographics and health outcomes the long term social, economic and demographic impact of major disease epidemics can only be estimated. (Cooper et al., 1998; Baiden et al., 2006). The lack of vital information presents major challenges for the use of an evidence-based decision-making process for public health interventions (Ye et al, 2012). This need for accurate data has led to the creation of Health and Demographic Surveillance Systems (HDSS) in areas were vital registration systems were poorly developed or non-existent (Byass et al., 2002; INDEPTH Network, 2002).

Definition, concepts of HDSS

A HDSS consists of the continuous monitoring of demographic and health characteristics of a population living in a well-defined geographic area. (INDEPTH Network, 2002). The goal of a HDSS is to generate high quality longitudinal datasets to capture the demographic and health changes of the set population (Tanser, 2007). HDSS collect information on key demographic and health events, additional information is often gathered on social, economic and behavioural data to gain greater understanding of the dynamics of the studied population. The goal of the monitoring system is to provide a platform for health-system innovations as well as social, economic and behavioural, interventions closely associated with research activities taking place at the site (INDEPTH Network, 2002). This way the HDSS can be used to generate evidence to guide health policymaking, allocate resources more efficiently and equitably, inform the development, implementation and evaluation of health interventions and strengthen the decision-making capacity of the health services in their location country (Chandramohan et al., 2008; INDEPTH Network 2002; Ye et al., 2012).

Information needs of a HDSS

The production of large and complex datasets required monitor disease burden poses many challenges on the public health community to explore, analyse and extract valuable information to make timely decisions (Caban & Gotz, 2015). The WHO has cited an accumulation of unanalysed data as a challenge facing HDSS sites (Baiden et al., 2006). Data collection is moving towards the use of electronic data capture. Employing mobile data capture methods reduces costs and increases the reliability of data produced (Mukasa et al., 2017; Mclean et al., 2017). The increased computerisation of health information at HDSS sites could lead to more efficient disease surveillance and data use. Yet for information to influence decisions it is essential to have a means to disseminate this information. AbouZahr (2005) suggests that due to the large quantities of data produced by health information systems the information overload at higher levels is such that data is in practice seldom used effectively for decision-making. Without the timely analysis of these datasets there is a decrease in their value as a tool to monitor disease trends (McNabb et al., 2006).

Advantages of data visualisation

The use of information visualisation offers a means to disseminate information in a timely manner and increase its visibility among stakeholders. Card, Mackinlay and Shneiderman (1999) defined data
visualisation as the use of visuals to “amplify cognition” to aid task completion. The use of visualisation acts as an intermediate step in converting data into information. Visualisation tools exploit human visual and spatial skills by using computer based interactive visual representations of data to amplify cognition. Visualisation can aid decision-making helping the user build accurate mental models that can leverage cognitive skills (Al-Hajj, 2013). Research in cognitive capacity demonstrates that humans can process more information presented graphically than in text (Miller, 1956; Tegarden, 1999). Keim (2008) states that the fundamental benefits provided by visuals are act as a frame or as a temporary storage area for human cognitive processes.

Larkin and Simon (1987) discuss how the use of data visualisation can support more efficient task performance by allowing substitution of rapid perceptual inferences for complex logical inferences and by reducing the search for information required to complete the task at hand. Visuals augment human memory to provide a larger working set for thinking and analysis and thus become external cognition aids. Intelligent use of visualisation exploit human’s visual perception capabilities to reduce information overload and build accurate mental models that can leverage human’s cognitive skills (Ware, 2012; Ware, 2010; Al-Hajj, 2013; Liu et al., 2010). Visualisation offers a way to shift cognitive load to the human perception system through the use of graphics (Lurie, 2007; Lohse, 1997). Visual representation can enhance problem solving capabilities by enabling the processing of more data without overloading the decision maker. Humans have evolved keen visual and spatial skills (Ware, 2010; Lurie, 2007). The exploitation of these abilities forms much of the basis behind data visualisations in use today (Few, 2006; Ware, 2010).

Data visualisations for HDSS

La Valle (2011) highlighted visualisation as a useful tool for gaining insight into large and complex datasets. Longitudinal datasets produced by HDSS is typically multivariate, complex and at different granularity levels (Ye et al., 2012) thus may represent suitable datasets for the use of visualisation tools. Datasets produced at HDSS sites are typically accessible only in formats that do not allow for the rapid extraction and analysis of salient information. Visualisation provides a means to access these datasets and rapidly extract useful information by augmenting human’s information processing skills (Keim, 2002). The role that surveillance data produced at HDSS can play in planning new health interventions can be diminished if the data is not communicated in a timely and understandable manner (Cheng, 2011).

Data visualisation dashboards

Data visualisation dashboards are a consolidated visual display of pertinent information, arranged in a manner such that the entire operation of an observed system can be monitored and understood at a glance (Few, 2006). In recent years dashboards have become a standard tool used by businesses and city managers to make better informed decisions in a timely and effective manner (Few, 2006). An effective dashboard can help policy makers and staff easily access and analyse important trends from indicators, supporting timely decision-making and quality improvement (Seow, 2014). Dashboards and other visualisation tools have seen little use at HDSS thus far but there is evidence of their effectiveness within other healthcare settings (West, 2015).

Case Study – Africa Health Research Institute

The setting of this study is the Africa Health Research Institute (AHRI) in KwaZulu-Natal, South Africa. AHRI’s field site operates as a HDSS, generating high quality longitudinal datasets to capture the demographic and health changes brought about by the HIV epidemic and to evaluate interventions to mitigate their impact (Tanser 2007). The objectives of the dashboard trial were to increase access to information for staff and community members to AHRI’s Somkhele field site. Giving a day to day picture of data collection operations and an overall picture of the progress of the treatment as prevention trial (TasP) that was taking place at the site during the development of the dashboard.
Methods
Stages of Data Dashboard Development

The development of a data dashboard within this context requires a number of specific implementation decisions. These factors involve elements of data design, user interaction, and the physical setting in which the dashboard will be installed, all of which contribute to the potential success of the dashboard. The development methodology outlined below is potentially applicable to the design and implementation of a dashboard within a similar setting:

- Dashboard Purpose and Concept: Specification of the purpose of the dashboard, including the context, target user group, and the expected objectives of users of the dashboard.
- Physical Interface: Specification of the physical environment in which the dashboard will be installed, including how users will gain access to the interface, how interaction with the dashboard will take place, presence of privacy and confidentiality concerns, and the connectivity between any devices and data sources.
- User Interaction and Flow: This specifies the intended process of user interaction with the dashboard interface, including the expected user ‘flow’ through the dashboard structure in achieving a user objective.
- Data Selection and Visualisation Design: This covers the design and structure of the data visualisation, including how the key attributes are selected, where they are located within the dashboard flow design, and their extraction from the underlying database.

The execution and design considerations involved in development of a data dashboard for the AHRI TasP trial context is outlined as follows.

Dashboard Purpose and Concept

The purpose of the data dashboard was to provide all staff at AHRI with the opportunity to learn more about the TasP trial, increasing the visibility and access to the datasets being produced updating as the trial progressed. The dashboard would give a clear and concise overview of the information collected in an understandable and accessible format while ensuring the anonymity of participants. The dashboards would be designed for a diverse range of users, anyone working at the centre could feel comfortable using the dashboards. For example, community members, fieldworkers, medical staff and resident scientists and visiting researchers. This design would enable users to either simply review overall progress, or to drill down into analysis.

To design the dashboard that met our required goals of data exploration tool to inform and engage users about the TasP trial, a number of design principles were established. The dashboard must include the functionality to

- Easily monitor key performance indicators
- Spatially interrogate datasets in relation to key performance indicators to identify trends
- Switch from a global to local view in relation to these indicators
- Drilldown to the local regions of the map to explore datasets in greater detail

Based on these design needs we developed the concept of a two-stage dashboard consisting of overview page and analysis page. The overview page would provide a global picture of the TasP trial and would feature a display of the key indicators relating to the progress of the trial and an interactive map to displaying these indicators to allow rapid identification of trends. Interaction with the overview page would allow the users to filter by indicator and act as a jumping off point to the analysis page. The analysis page would focus on a selected local trial region. Allowing the user to compare the performance of this region in relation to the global and augment their analysis with explanatory datasets.
Physical Interface

The intention of the dashboards was to provide AHRI with a public facing means of communicating the data produced at the demographic surveillance site. The dashboards would be on three large touchscreens within the research centre. The placing of the dashboards in public areas hoped to encourage collaborative use, visibility and discussion of datasets and could provide a means to explain the work of the institute field site to visitors.

User Interaction and Flow

The global performance of indicators is the starting point for the user’s interaction with dashboard, a set of anchors to guide their exploration of the datasets displayed and to compare performance on global and local level. The design of the dashboard was based on Shneiderman’s (1998) principles of visual information seeking mantra “overview first, zoom and filter and details on demand”. The overview component of the dashboard is the display of the global indicators combined with the spatial overview provided by the map. The user is able to filter data by changing the map display by indicator and the map provides a means to zoom to areas of interest based on spatial patterns of variance. Selecting a map region allows the user to access the details for that region.

Data Selection

A successful dashboard should provide an overall picture of the data presented with a key objective in mind. The choice of the key performance indictors (KPIs) is fundamental to the success of the dashboard. Given the diverse intended user base the indicators used would need be universally known among HDSS community. The indicators chosen were the UNAIDS 90/90/90 targets (Sidibe et al., 2016) alongside the rate of HIV prevalence over the study region. The 90/90/90 targets refer to the pathway, by which a person is tested, linked and retained in HIV care, and initiates and adheres to antiretroviral drugs (ARVs), they are indicators of progress in the overall reduction of HIV incidence. These targets state that by 2020, 90% of all people living with HIV will know their HIV status. By 2020, 90% of all people with diagnosed HIV infection will receive sustained antiretroviral therapy. By 2020, 90% of all people receiving antiretroviral therapy will have viral suppression. The table below gives a description of how each indicator was calculated. The use of the 90/90/90 targets gives a view of individual’s position in the HIV treatment cascade at any given moment. The indicators were recognised as common targets through the HDSS community at AHRI.

**Table 1 – Summary of selected KPIs and data types**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial participants</td>
<td>number</td>
<td>The number of individuals enrolled in the TasP trial</td>
</tr>
<tr>
<td>HIV prevalence rate</td>
<td>ratio</td>
<td>The percentage of individuals to known to be HIV positive</td>
</tr>
<tr>
<td>HIV positive individuals that know their status/ individuals known to be positive</td>
<td>ratio</td>
<td>The percentage of HIV positive individuals that know their status/the number of individuals known to be positive</td>
</tr>
<tr>
<td>HIV positive individuals on Antiretroviral Treatment (ART) / diagnosed individuals.</td>
<td>ratio</td>
<td>The percentage of HIV positive individuals on Antiretroviral Treatment (ART)</td>
</tr>
<tr>
<td>Individuals who are virally suppressed/individuals on Antiretroviral Treatment (ART).</td>
<td>ratio</td>
<td>The percentage of individuals who are virally suppressed</td>
</tr>
</tbody>
</table>


Framing data exploration in indicators allows the user to keep the global in mind while exploring detail. The dashboard’s exploratory hierarchy of datasets is the global view of KPIs then viewing the KPIs a local level. This allows the user to view the global and the local together, and make comparisons between both scales, gaining insight into local relative performance and identify emerging trends quickly and clearly (see Figure 1). Users could directly compare the local performance of a map region to all global indicators and interrogate factors that could influence the rate of the KPI’s with the help of explanatory datasets such as demographic or socio-economic data. The user is able to display explanatory datasets normalised by the regions total population, or normalised by population relating to the indicator (e.g. HIV positive individuals).

Figure 1
User interaction flow between global and local.
This image shows the interaction of KPIs at the global and local scales, allowing the user to explore the broad spatial trends in the indicator at the global scale, before drilling down into finer scale spatial variation and associations at the local.

Visualisation Design
The Map Interface

Design Justification

The landing page was designed to fulfil the function of a traditional dashboard by providing the user with access to the most salient information relating to the subject manner, while encouraging more detailed exploration of the datasets. The design intention was to provide the user a global overview of
the progress of the study in relation to the KPI’s, view a spatial representation of indicator performance and means to compare the global performance to the local. Once a region had been selected the user could move to the analysis page, that would focus on local performance with the aid of explanatory variables and the use of comparative visualisations. Following Shneiderman’s visual information seeking mantra the map page presents an overview of the datasets with the ability to filter datasets in relation to the KPI and then by local region performance details can be accessed by using these features to guide the user to the analysis page. The dashboard updates daily from the trial databases allowing the user to explore the shifting patterns of the study’s progress.

Interface Elements

The landing pages consists of 3 areas of information display (see Figure 2).

1) Map
2) Global performance of indicators
3) Local performance of indicators in a selected region.

The focus of the landing page was the map display (indicated with point 1 in Figure 2). The map provides a spatial view of the study performance through graduated symbols and acted as a vector for exploration of local regions. One of the weaknesses in the trials oversight protocol was lack of method to observe spatial variation in the progress of the trial without the generation of GIS maps, a time-consuming process. It was therefore decided that the map would be at the centre of the dashboard interface and would drive exploration of trial datasets. The map allows simultaneous overview of global trends and local variation in the observed attribute. The interface also allows the user to view trial results in relation to local infrastructure, e.g. roads and clinics.

Figure 2

Landing page of the dashboard

This image shows the interactive front page of the dashboard, displaying the global view of one indicator within the map (marked 1), measures of KPIs at the selected local region at the bottom of the screen (marked 2), and global KPI measures on the left (marked 3). Data is updated in real-time. The user is invited to manipulate the map and change selected regions through click or touch. The
Breakdown Page button leads to a more detailed exploration of KPI measures and related factors within the Analysis page (Figure 4).

Spatial data is displayed using a ‘binning’ technique that divides the trial area into a grid of uniform hexagons. Binning is a means of converting point data into a regular grid of polygons. Each element of the grid represents the aggregation of the points that fall within it. The result is a uniform representation of the trial region and allows simple clear visualisation of the datasets aided by quantification of the map using colour breaks. The most pressing concern was to ensure the privacy of the trials participants. This was the main driver of choosing a binning technique as participants were aggregated to a hex, resulting in the greatest granularity level of the data being the hex, with all local data displayed in relation to the performance of the hexagon region rather than the individual. The use of a grid acts a means of distortion that prevented users from directly identifying individual homesteads. A downside of this method is that the grid sections display geographically equal areas rather the population density (e.g. skewed towards urbanised areas).

The graduated colour map allows the user to identify areas of interest and to detect spatial autocorrelation within the trial region. The use of a binning technique also related to practical concerns to create regions to capture local performance and acts as a point of interaction to trigger events on the dashboard. When a hexagon is selected information relating to indicator performance and the number of trial participants residing in that region is displayed. A button on the information window allows the user to view this region in greater detail on the analysis page. Interacting with a map hexagon also triggers the display of KPIs for the selected region in the page footer.

The global performance of indicators was displayed in the left-hand sidebar (point 2 in Figure 2). The sidebar displayed the KPI in the form of donut charts with the percentage figure contained inside, when an indicator chart was interacted with the map display would update relating to that indicator while the bottom bar of the page would update to display explanatory information on the selected indicator.

The page footer (point 3 in Figure 2) allows direct comparison between the local indicators and the global in the left sidebar (See Figure 3). This design was to allow the user to consistently frame the global and local together without overloading them with information. Indicators are displayed using horizontal bar charts and contained an information button to display how the selected map indicator was calculated, and a button that moves the user to the analysis page to explore KPI’s in greater detail with the aid of explanatory datasets.
Points of comparison
This image highlights the points of interaction between global and local indicators, showing the points where a user is able to directly compare local and global measures of KPIs.

Local KPIs and Explanatory Variables

Design Justification

The analysis page provides a detailed view of the performance of indicators at a local scale, the “details on demand” of Sneiderman’s mantra (1998). The design intention was to view explanatory datasets in relation to the total population resident in each region and to add/remove population data in relation to the KPIs. The analysis page allows users to view data relating to HIV status, gender, age group, education, relationship and economic status. The analysis page facilitates exploration of datasets by allowing the user to explore the demographic differences of the trial participants relating to each indicator. For example, users could compare the education level of HIV positive individuals who have linked to care against those who have not linked to care. The analysis page allows for in-depth exploration and discussion of the datasets among users.

Interface Elements

The analysis page consisted of 4 areas of information (Figure 4)

1) Explanatory variables
2) Global and local comparison
3) Chart selection, mini map and global information display
4) Key numbers to relating to region

The focus of the analysis page is the explanatory variables section (point 1 in Figure 4). Within this window, charts are displayed in modules that can be added or removed through interaction with the interface. The user is presented with a set of charts displaying explanatory datasets relating to all trial participants in the local region. The charts are as follows;
1) A population pyramid, displaying participants in relation to age and gender
2) A normalised stacked bar chart displaying HIV rates by age group
3) A tree map displaying education levels
4) A tree map displaying relationship status
5) A tree map displaying employment levels

The user can add charts displaying population data in relation to each indicator to explore the differences between participants at different stages of the HIV treatment cascade, allowing the user to directly compare different outcomes. We used multiple visualisation methods to exploit their differing strengths to enhance the users’ data exploration and engagement. For example, the relative size differences of groups in the tree maps would allow the user to clearly and quickly learn the difference in employment levels between those on ART treatment and those yet to link to care.

The left-hand sidebar (point 2 in Figure 4) displays the KPIs in the local region to the global using a bar charts, as with the design of the map page allowing the user to keep global and local in their mind while exploring.

The right-hand sidebar (point 3 in Figure 4) acts as the control panel for the visualisation portal. The user can select charts to add to the centre display relating to an indicator, chart choices are represented by icons related to the dataset. The sidebar includes a mini-map highlighting the selected region and a bar chart display the global proportional relationships of education, relationship status and economic status within the trial.

This bar at the bottom of the page (point 4 in Figure 4) displays key figures relating to region such as gender breakdown, HIV prevalence and KPIs.

**Figure 4 - Analysis page**
This figure shows the Analysis page for a local region selected through the global map page. The page introduces a range of potential explanatory attributes that can be measured against KPIs to help develop hypotheses for future studies. Users are able to add and remove visualisations through interactive functionality.
Results

The data dashboard was evaluated by assessing its usability with 20 participants. Participants were selected from people working at AHRI drawn from different groups that could potentially use dashboards to gain insight from AHRI datasets. These groups chosen were scientific staff (primarily medics, healthcare specialists, and researchers), operational staff (those ensuring continued operation of AHRI), nursing staff, and Community Advisory Board (CAB) members. The CAB are members of the local community that are consulted and advise researchers on the implementations of studies. Five members of each group where randomly selected.

To assess the usability of the dashboards the participants performed a benchmark task evaluation consisting of five tasks (each incorporating two sub-tasks) of increasing difficulty requiring the participant to extract information using the dashboards. During benchmark task evaluations participants use visualisations to perform tasks to extract information, measuring targeted metrics. (Chen, 2000; North, 2006). The time taken and accuracy of the participants’ responses for the tasks were recorded and analysed to evaluate the dashboards usability. The use of a task evaluation can identify possible shortcomings in the representation of data within a visualization system (Amar & Stasko 2004). The task completion component of the evaluation was filmed with the permission of the participants to ascertain timing of task completion. The tasks undertaken by participants can be seen in Supplementary Material 1.

Alongside benchmarking, participants were also asked to self-evaluate their knowledge of the TasP trial and their level of computer literacy. Participants were also asked to complete a questionnaire using a Likert scale to assess their attitude towards the dashboard, covering the design of the interface, information presentation, and attitudes towards the future use of dashboards within AHRI. The questionnaire is provided in Supplementary Material 2.

The results of the evaluation can be broken down into three component parts – the results of the self-evaluation, the results of the task setting, and the results of the questionnaire.

Self-Evaluation

The self-evaluation asked the participants to give their level of computer literacy, awareness of the TasP trial and the dashboard project (see Table 2). Overall the CAB and nursing staff reported lower levels of computer literacy then the operational and scientific staff. While knowledge of the TasP trial and the dashboard projects varied within groups, with the nursing staff reporting the highest knowledge of the trial, which was to be expected as the nurses asked to participate where involved in the clinical component of the TasP trial.

Table 2 – Responses to self-evaluation from different participant groups

<table>
<thead>
<tr>
<th></th>
<th>Very Low</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
<th>Very High</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Community Advisory Board</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of computer literacy</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Knowledge of the TasP Trial</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Awareness of the dashboard project</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Nursing Staff</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of computer literacy</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Knowledge of the TasP trial</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Awareness of the dashboard project</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

**Operational Staff**

| Level of computer literacy | 0 | 0 | 2 | 2 | 1 |
| Knowledge of the TasP trial | 0 | 1 | 2 | 2 | 0 |
| Awareness of the dashboard project | 1 | 4 | 0 | 1 | 0 |

**Scientific Staff**

| Level of computer literacy | 0 | 0 | 0 | 5 | 0 |
| Knowledge of the TasP trial | 0 | 0 | 3 | 2 | 0 |
| Awareness of the dashboard project | 0 | 4 | 1 | 0 | 0 |

**Task Completion**

The results of the task component of the evaluation are displayed in the figures 5 and 6. The results of the task completion were for the most part positive. With the exception of the second part of the final task the majority of participants were able to complete the task successfully, and in the majority of tasks there was a greater than 70% correct completion rate (Figure 5).

![Figure 5](image_url)

**Figure 5**
Task completion rates, overall
This figure shows how all participants in the evaluation study performed on each set task.

The issue with the final task, which involved comparing data from multiple sources, may highlight a design weakness where data must be compared across separate charts rather than as different series within a single chart. Within these results, there was a great deal of variation within the sub-group performance, with the scientific and operational staff performing better in task completion than the CAB and nursing groups. Figure 6 outlines differences between participant groups.

Figure 6
Task completion rates, by group
This figure shows how different groups of participants performed during the evaluation study.

Overall the time taken to complete tasks increased with the increasing complexity of each task. Users on average spent more time on tasks where their answer was ultimately incorrect. Scientific and operational staff were generally quicker to answer questions, and more often correct, than CAB and nursing staff. Through one-way ANOVA, significant difference in time taken to complete tasks is found between scientific and nursing staff, but not between other groups (Figure 7). These results may be influenced by lower computer literacy on the part of the nursing and CAB participants, as self-reported. They indicate therefore that there are further steps to go to maximise accessibility across all interested groups.
Figure 7  
**Variance in time taken to complete tasks, by group**  
This image provides detail on the time taken by each group to complete tasks during the evaluation study. As can be seen, the CAB and nursing groups generally took longer to complete tasks than operational and scientific staff, indicating that more work is required around ensuring the dashboard is accessible to all user groups.

**User Questionnaire**  
The questionnaire consisted of eleven questions divided covering four topic areas, the insight provided by the dashboard, the dashboard interface, the future use of the dashboards at the AHRI and questions allowing participants to comment on the dashboards and the how they could be improved to provide greater clarity of information. Figure 8 shows the results of the questionnaire across all groups below.
Figure 8
Questionnaire results
This figure summarises participant responses to questions regarding their feelings about the dashboard and potential future use within AHRI.

Over 80% of all participants felt that dashboards provided them with a detailed understanding of HIV prevalence and treatment in the TasP trial region, with five participants strongly agreeing with the statement and eleven agreeing with the statement. In relation to the questions about the design of the dashboard here the results were mostly positive with the majority (60%) of participants commenting that the terms used in the dashboard aided in the understanding of the data presented, the dashboard was easy to navigate through and the charts where easy to understand. However, the CAB staff and operational staff had a more positive view of the design of the dashboards than the nursing or scientific staff. These differences can also be seen in the comments made relating to design and use of dashboards. A member of the nursing staff commented that:

“It must be used by scientists only, not everyone is familiar with use of graphs and percentages when analysing data” - Nursing staff member

A member of the scientific staff made a similar comment on the presentation of data relating to ability to understand data visualisation:

“For scientific staff the dashboard is user friendly, for the broader AHRI community, it might not be, particularly for TasP staff, the language and graphs used might not be easy to understand and follow” - Scientific staff member

95% of participants agreed that the dashboards were a useful tool for providing information on studies taking place at the centre, and that the dashboard should be used more widely as a tool for explaining results of the trials:

“Every staff member should be able to use the dashboard for information purposes” - Operational staff member
“I think there is enough information in this data dashboard because it is very expansive and could help us to understand easily any information that we need. And would help us to gain more information and if you don’t understand something it is easy for to go to the data dashboard to punch in that information and gain more knowledge” - CAB member

Discussion

In this paper, we outlined a framework for the development of a data dashboard for exploiting real-time data within the context of a population health surveillance site. Population health is similar to many other health disciplines in experiencing a large increase in the amount of data created at field sites. This is driven by new and more efficient data collection systems and techniques, such as mobile and online platforms and integration of datasets from different sites and studies. The potential of such data could be negated and restricted without a visualisation system to disseminate the data in manner suitable to audiences such as policymakers, researchers, and community members. These interfaces allow for the rapid detection of emerging health trends, highlighting of outliers and emerging clusters of activity. Once in place, fewer resources are needed to maintain a real-time monitoring system than would be required for creation of ad-hoc analyses.

The dashboard evaluation demonstrated the potential of the dashboards as a method to explain ongoing progress of research trials to staff and stakeholders at AHRI. The results of the questionnaire also demonstrated a very positive attitude towards their future use within AHRI across all groups studied. Demonstrating there is enthusiasm for not only the increased visibility of data, but also the increased use of data visualisation as a means to disseminate datasets.

The outlined framework demonstrates the potential for a visualisation platform to provide an exploratory interface for users to interrogate data, develop insights, and form new research hypotheses. The interface furthermore presents an opportunity for collaborations among researchers, through shared data exploration. Within the AHRI context particularly, it also allows stakeholders and community members to see how data collected in their community is being employed to further research and benefit the community at large. Increasing the visibility of data for community members increases transparency, and encourages active participants in ongoing studies. The evaluation demonstrated while the majority of all groups had positive attitudes towards the increased visibility of data, it is important to note variation among groups with respect to data needs.

The AHRI data dashboard can serve as a broad template for further development. However, further developments in the design and design process are future areas for exploration. The platform could be further developed to allow for increased flexibility in the types of data visualisation available, and expansion to other sites and contexts. However, it is clear that a more significant step lies in opening up the dashboard to a wider variety of users. Within the current design process, users were only consulted at later stages of the project, and there would be advantages in exploring user data needs in through a user-led process during later iterations. During the evaluation demonstrated that users spent more time when on a task when they did not complete the task. This indicates a research need to better understand the differences between how users interact with visualisations on the basis of their characteristics, and the need to understand data literacy in the development of platforms for diverse user groups.

Conclusion

The paper outlined the design and development of a framework for a data visualisation dashboard within a population health surveillance context. The framework outlines the designed user process, incorporating objectives concerning improving the visibility of data and providing access to users. The eventual design allows all users concerned with the surveillance trial to explore the data, develop
analyses and construct hypotheses with respect to the spatial relationships between attributes. A short evaluation of the dashboard yielded positive initial results and indicated that the dashboard was well received. Future work should explore further development within the existing framework structure, and improved accessibility to all users.

Declarations

Ethics approval and consent to participate

Ethics approval was granted for this study by the University of KwaZulu-Natal Biomedical Research Ethics Committee (BREC), reference BE497/16. Informed consent to participate in the study was obtained from all participants.

Availability of data and material

The datasets generated and analysed during this study are available from the corresponding author on reasonable request.

Funding

This project was funded under the i-sense (EPSRC IRC Early-Warning Sensing Systems for Infectious Disease) project (EP/K031953/1).

Authors' contributions

The project was conceived and designed by EM, KH and DC. The technical development of the data dashboard was undertaken by DC. The dashboard evaluation was designed and executed by DC, KH and EM. The evaluation analysis was carried out by DC. All authors read and approved the final manuscript.

Acknowledgements

The authors acknowledge the contributions of Siyabonga Nxumalo for running the evaluation at AHRI, Steven Gray for providing technical assistance on the development of the dashboards, and Dickman Gareta, Jaco Dreyer and the wider the research data management team at AHRI for all their assistance throughout the development process. The authors thank the Africa Health Research Institute for hosting the study, and particularly the participants who gave their time to take part in the evaluation. The authors further thank Professor Rachel McKendry for her continued generous support of this work.

References


**Figure 3**

Points of comparison

This image highlights the points of interaction between global and local indicators, showing the points where a user is able to directly compare local and global measures of KPIs.

**Figure 5**

Task completion rates, overall
This figure shows how all participants in the evaluation study performed on each set task.

**Figure 6**
**Task completion rates, by group**
This figure shows how different groups of participants performed during the evaluation study.

**Figure 7**
**Variance in time taken to complete tasks, by group**
This image provides detail on the time taken by each group to complete tasks during the evaluation study. As can be seen, the CAB and nursing groups generally took longer to complete tasks than operational and scientific staff, indicating that more work is required around ensuring the dashboard is accessible to all user groups.

**Figure 8**
**Questionnaire results**
This figure summarises participant responses to questions regarding their feelings about the dashboard and potential future use within AHRI.

**Table 1 – Summary of selected KPIs and data types**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial participants</td>
<td>number</td>
<td>The number of individuals enrolled in the TasP trial</td>
</tr>
<tr>
<td>HIV prevalence rate</td>
<td>ratio</td>
<td>The percentage of individuals known to be HIV positive</td>
</tr>
<tr>
<td>HIV positive individuals that know their status/individuals known to be positive</td>
<td>ratio</td>
<td>The percentage of HIV positive individuals that know their status/the number of individuals known to be positive</td>
</tr>
<tr>
<td>HIV positive individuals on Antiretroviral Treatment (ART)/diagnosed individuals</td>
<td>ratio</td>
<td>The percentage of HIV positive individuals on Antiretroviral Treatment (ART)</td>
</tr>
<tr>
<td>Individuals who are virally suppressed/individuals on Antiretroviral Treatment (ART)</td>
<td>ratio</td>
<td>The percentage of individuals who are virally suppressed</td>
</tr>
</tbody>
</table>

**Table 2 – Responses to self-evaluation from different participant groups**

<table>
<thead>
<tr>
<th></th>
<th>Very Low</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
<th>Very High</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Community Advisory Board</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of computer literacy</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Knowledge of the TasP Trial</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Awareness of the dashboard project</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Nursing Staff</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of computer literacy</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Knowledge of the TasP trial</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Awareness of the dashboard project</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>Operational Staff</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of computer literacy</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Knowledge of the TasP trial</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Awareness of the dashboard project</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Scientific Staff</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of computer literacy</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Knowledge of the TasP trial</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Awareness of the dashboard project</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>