Mobile-based and open-source case detection and infectious disease outbreak management systems: a review [version 1; peer review: 1 approved, 1 approved with reservations]


1 Institute of Health Informatics, University College London, London, NW1 2DA, UK
2 Institute for Global Health, University College London, London, WC1E 6JB, UK
3 Centre for Clinical Microbiology, University College London, London, NW3 2PF, UK
4 Fondation Congolaise pour la Recherche Médicale, Brazzaville, Congo
5 Institut für Tropenmedizin, Universitätsklinikum Tübingen, Tübingen, Germany
6 Ifakara Health Institute, Dar es Salaam, Tanzania

Abstract
In this paper we perform a rapid review of existing mobile-based, open-source systems for infectious disease outbreak data collection and management. Our inclusion criteria were designed to match the PANDORA-ID-NET consortium’s goals for capacity building in sub-Saharan Africa, and to reflect the lessons learned from the 2014–16 West African Ebola outbreak. We found eight candidate systems that satisfy some or most of these criteria, but only one (SORMAS) fulfils all of them. In addition, we outline a number of desirable features that are not currently present in most outbreak management systems.

Keywords
outbreak management, outbreak response, infectious disease surveillance, mobile health, mobile health software, open source, mhealth,
**Introduction**

In an infectious disease outbreak, gathering, sharing and analysing accurate, real-time data about persons with the disease and their contacts is crucial for effective, targeted interventions. This was particularly evident in the 2014–2016 Western Africa Ebola epidemic, during the initial stages of which, poor data management practices significantly contributed to the difficulties in containing the outbreak in Sierra Leone.\(^1\)

The PANDORA-ID-NET consortium\(^1\) aims to build capacity for effective outbreak response in sub-Saharan Africa, supported by adequate research and training. As part of this mission, we endeavour to develop a real-time data sharing platform for disease outbreaks that leverages centralised data management and uses mobile technologies for data gathering and feedback. In line with our capacity-building goals, we have committed to accomplishing this using open-source technologies, so that the resulting software packages can be easily deployed on regional IT infrastructure and maintained and further developed by local staff. In addition, this will help ensure that all data collected through an outbreak response system can be stored and processed in the region of origin. The purpose of this paper is to review existing “best practice” open-source systems for outbreak data collection, collation, sharing and analysis, which are used in low- and middle-income settings and could serve as a foundation for the PANDORA-ID-NET consortium data sharing platform.

The rest of the paper is organised as follows. First, to outline what is required from an outbreak management system, we briefly describe the key data-related components of an Ebola outbreak response as identified by the World Health Organization (WHO).\(^2\) We do so under the assumption that other infectious disease outbreaks will share many essential features with those of Ebola virus disease (EVD).\(^1\) Secondly, to understand the possible challenges in the implementation of these data-related components, and to formulate our inclusion criteria for this review, we summarise the existing academic literature on the shortcomings of data management in the 2014–2016 West Africa Ebola outbreak response. Finally, we identify and review a number of publicly available, mobile-based (usable on a smartphone or tablet), open-source outbreak management systems that satisfy our inclusion criteria. Some of these systems were developed in response to the above EVD outbreak, while the rest were designed in other circumstances. We performed this review by leveraging an existing systematic review by Tom-Aba et al.,\(^3\) by conducting our own supplementary rapid literature search and by contacting global experts in field epidemiology.

**Key data-related components of outbreak management**

The WHO’s Ebola Virus Disease Consolidated Preparedness Checklist\(^1\) outlines 11 key components of an outbreak response (see Box 1 for the full list). Of these components, case management, surveillance, contact tracing and laboratory results require active maintenance and sharing of accurate data amongst outbreak response professionals, and have a high degree of dependence on one another. We give a brief description of each component below.

<table>
<thead>
<tr>
<th>Box 1. WHO Ebola outbreak management components</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Coordination – to enhance national public health emergency preparedness and response plans, and national command and coordination structures; to minimise duplication of efforts and ensure maximum impact from the available resources.</td>
</tr>
<tr>
<td>• Rapid response team – to be able to act immediately once the first suspected case is reported, irrespective of its geographic location.</td>
</tr>
<tr>
<td>• Public awareness and community engagement – to harness public trust, counteract false rumours and enhance behaviour to reduce the risk of EVD exposure.</td>
</tr>
<tr>
<td>• Infection prevention and control – to minimize the risk of transmission of EVD in health-care settings and in the community.</td>
</tr>
<tr>
<td>• Case management – to strengthen clinical care whilst minimising the risk of transmission to others, including health workers, and to eliminate additional risks such as unsafe burials.</td>
</tr>
<tr>
<td>• Surveillance – to detect and report any persons with an illness compatible with EVD, or any other unusual health events possibly associated with EVD.</td>
</tr>
<tr>
<td>• Laboratory results – to produce rapid confirmations of cases, which is crucial for containing an outbreak, tracing contacts and providing emergency healthcare.</td>
</tr>
<tr>
<td>• Capacities at points of entry – to prevent cross border transportation of EVD cases by implementing standard operational procedures at international airports, seaports and major land crossings.</td>
</tr>
<tr>
<td>• Budget – to ensure that both preparedness and response activities are costed in a coordinated and planned manner and sufficient resources are identified to enable their rapid implementation.</td>
</tr>
<tr>
<td>• Logistics – to ensure that the logistical capacities (such as supply chain management and staffing) required to implement the above components are in place.</td>
</tr>
</tbody>
</table>

**Case management**

Ebola case management\(^6\) involves i) suspected case identification (either through routine surveillance or following community-based surveillance alerts), ii) determining whether suspected cases are probable cases by following case definition guidelines, iii) establishing whether probable cases are confirmed cases by obtaining laboratory results, and iv) clinical management of probable and confirmed cases.\(^7\) In ii), the data is recorded using a case investigation form, which must be comprehensive and standardised. Additionally, each case must be uniquely linked to the original surveillance alert, the case investigation form data and the subsequent laboratory results. As we describe later, failure to link this data appropriately can lead to case duplication, which in turn affects the accuracy of surveillance data.

---

\(^1\)Note that we exclude systems that deal exclusively with vector-borne outbreaks, as such systems require comprehensive vector control features that are beyond the scope of this review.
Contact tracing
According to the WHO guidelines, all potential contacts of Ebola cases should be identified and closely observed for 21 days from the last day of exposure; contacts that develop illness should be immediately isolated to prevent further transmission of infection. Contact tracing can be broken down into three stages: contact identification, contact listing and contact follow-up.

Contact identification
Contact identification is performed for all cases meeting the EVD case definition, whether alive or deceased. This is done by asking about the activities of the case (the patient) and the activities and roles of the persons around the patient since onset of illness. Persons who lived with or visited the patient since the onset of illness should be identified, as well as the persons, places and healthcare facilities that the patient had visited after they became ill. The contact information collected for each case should be verified for consistency and completeness via repeated interviews conducted during later visits.

Contact listing
At this stage, contact information is entered into the contact listing form. The recorded data for each contact should include their name, address, phone number, sex, age, relation to the case, date of last contact with the case, type of contact and whether they are a healthcare worker. This information is used to assess each contact’s level of risk of EVD infection, and emphasis is placed on reaching out to high-risk contacts first.

Contact follow-up
Contact follow-up involves monitoring each identified contact daily for 3 weeks, recording whether they develop symptoms indicative of EVD and referring them to hospital if they do (at which point they also become a new case). If a suspected case subsequently tests negative for EVD, their contacts no longer need to be followed up.

Laboratory results
According to the WHO guidelines, all probable and suspected Ebola cases should have their blood specimens collected by trained medical staff at designated healthcare facilities. A number of tests are used to determine the Ebola status of a suspected or probable case, including viral RNA, viral antigen and immunoglobulin M antibody tests.

Surveillance
The EVD consolidated preparedness checklist states that an event-based surveillance system must be in place to “enable timely follow-up of information/rumours from all sources including the community, media, etc.”.

The WHO Ebola contact tracing manual recommends asking people who are being interviewed whether they are aware of any other suspected cases. This is known as “active case finding” and is usually carried out during home visits while performing contact tracing and contact follow-up.

In countries that have had no reported cases of Ebola, alert systems are advised for major land border crossings with already affected countries and for airports, seaports, and healthcare facilities in capital cities. These alert systems should report sick persons who meet the case definition criteria and who have arrived from countries with suspected EVD cases.

Challenges in the management of the 2014–2016 Ebola outbreak
During the 2014–2016 West African Ebola outbreak, the government of Sierra Leone suffered considerable setbacks in its attempts to curb the epidemic during its initial stages, which contributed to the persistence of the overall outbreak in the region. Some of these setbacks arose as a result of the country’s political challenges and its post-war environment. However, problems with data management were also a significant factor. Specifically, Oyada et al. state that inadequate management and integration of multiple data sources meant that at the peak of the outbreak, “reliable epidemiological statistics to determine the actual number of confirmed cases and deaths and to effectively monitor the outbreak could not be obtained.” In the same paper, the authors further highlight individual data management problems that contributed to these inaccuracies:

- Late arrivals of case investigation forms at district response centres – as a result, databases were not being updated in time to provide daily summaries and laboratory results had to be used instead for this purpose.
- Missing variables on case investigation forms and in database entries – many records did not have information on date of illness onset, sex, age, or residence. Additionally, entries were often inconsistent across multiple databases. These factors limited the ability to measure critical indicators, such as the case fatality rate and incidence rate, and to pursue effective contact tracing.
- Misclassification of Ebola as other diseases endemic to the region – one example is malaria, which shares common clinical features with Ebola and may have contributed to delays in recognising some EVD cases. Such confusion was made more likely by missing data such as laboratory results.
- Difficulties in detecting and merging duplicate case entries – arising from the use of separate databases for case management and laboratory results, and a lack of database user training.
- Lack of a consolidated database that captured and linked all data sources in a structured way – inconsistencies in Excel database formats across different regions complicated national level case data aggregation.

Overall, inaccurate data ingestion and poor data consolidation were the two central data management issues faced during this outbreak by the communicable disease control teams. Therefore, in the next section, we aim to build on these lessons and restrict our review criteria to only include outbreak management software systems that implement all four data components described in our first section in an integrated manner. Furthermore, we choose to focus only on systems that...
collect information in real-time via mobile applications running on smartphones and tablets, so as to mitigate the problem of late data arrival highlighted above.

Methods
To identify mobile-based and open sourced outbreak management systems that could meet the challenges that occurred during the management of the 2014–2016 Ebola outbreak, we built upon the systematic review by Tom-Aba et al. that considered systems developed in response to the 2014–2016 West Africa Ebola outbreak. It is reasonable to focus on this outbreak given its scale and impact, and the search strategy employed by the authors is likely to have captured a significant part of development activity in the field of mobile-based outbreak management. However, in the review the authors state that "while it would have been valuable to conduct this review beyond the application of EVD and haemorrhagic fevers and beyond 2015, removing these selection criteria from the search strategy would have resulted in an unmanageably large output with an extremely low positive predictive value". For this reason, we felt it necessary to carry out our own, supplementary rapid literature review to ensure that we have not missed major developments that occurred after the 2014–2016 epidemic.

This review was conducted by a single reviewer and only considered publications written in the English language. We used Google Scholar as the primary tool for our literature searches. Our preliminary inclusion criteria were "outbreak management systems that include at least two of the four key data components described in the ‘key data-related components of outbreak management’ section and have a mobile component". Note that these are not specific to EVD and are less stringent than our ultimate inclusion criteria that necessitate all four data management components being present. The preliminary criteria act as a filter to increase the positive predictive value of the candidate outbreak management systems without significantly reducing the sensitivity of the search results.

Table 2 shows the search queries we used to create a pool of candidate publications. Note that this involves searching for all pairwise combinations of keywords corresponding to the key data components. Additionally, we used the keywords “outbreak” and “android”. We settled on the latter as a catch-all term for “mobile applications”, since the latest data shows that Android has a market share of 83.1% in all of Africa. We also carried out a PubMed search using the following MeSH query:

(((Software[mh:noexp]) OR Mobile Applications [mh:noexp]) OR Cell Phone[mh]) OR Telemedicine[mh:noexp]) AND ((Disease Outbreaks[mh]) AND Epidemiologic Methods[mh]).

To increase the coverage of our review further, we contacted international experts that participated in the second International Conference on Reemerging Infectious Diseases (ICREID). The conference took place in March 2019 in Addis Ababa, Ethiopia. Conference organisers contacted every participant via email and asked them to tell us about outbreak management systems they had used.

Results
Table 3 summarises the results of our literature search, executed in February 2019.

The Google Scholar search strategy yielded 1016 unique publications; only nine described systems satisfying the preliminary criteria, of which six were not listed in the systematic review by Tom-Aba et al. The PubMed search returned 231 results but did not increase the number of publications matching the above criteria. Via personal communication with expert field epidemiologists, we were informed of three commercial and proprietary outbreak management systems. Finally of 475 ICREID 2019 conference attendees, 38 told us about their experiences and indicated that they had taken part in at least one outbreak investigation. In total, 16 reported using an electronic outbreak management system, and three additional systems were discovered this way.

The combined search strategy detailed above yielded 1022 unique citations, of which 1007 were excluded, and 15 were retrieved. A PRISMA flow chart providing a visual summary of our screening process can be found in Figure 1.
Table 1. Comparison of the eight identified outbreak management systems meeting all four key data management criteria. Note that none of the above systems offer clinical trial data management. *Planned feature.

<table>
<thead>
<tr>
<th>Systems</th>
<th>Surveillance</th>
<th>Case mgmt.</th>
<th>Contact tracing</th>
<th>Lab results</th>
<th>Open source</th>
<th>One health</th>
<th>Passive data</th>
<th>Training materials</th>
<th>Multiple diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td>SORMAS</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>CommCare Ebola</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sense Ebola Followup</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>AfyaData</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>HPZone</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>EpilInfo</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Conduent Maven</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Altalas WorldCare</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 2. Google scholar queries for the rapid review of additional systems not identified in Tom-Aba et al.4. There were 1016 unique publications after pooling all query results.

<table>
<thead>
<tr>
<th>Query</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>(&quot;contact tracing&quot; OR &quot;contact tracking&quot;) &quot;surveillance&quot; &quot;android&quot; &quot;outbreak&quot;</td>
<td>60</td>
</tr>
<tr>
<td>(&quot;contact tracing&quot; OR &quot;contact tracking&quot;) &quot;case management&quot; &quot;android&quot; &quot;outbreak&quot;</td>
<td>22</td>
</tr>
<tr>
<td>(&quot;contact tracing&quot; OR &quot;contact tracking&quot;) &quot;laboratory&quot; &quot;android&quot; &quot;outbreak&quot;</td>
<td>51</td>
</tr>
<tr>
<td>&quot;surveillance&quot; &quot;case management&quot; &quot;android&quot; &quot;outbreak&quot;</td>
<td>86</td>
</tr>
<tr>
<td>&quot;surveillance&quot; &quot;laboratory&quot; &quot;android&quot; &quot;outbreak&quot;</td>
<td>960</td>
</tr>
<tr>
<td>&quot;case management&quot; &quot;laboratory&quot; &quot;android&quot; &quot;outbreak&quot;</td>
<td>71</td>
</tr>
</tbody>
</table>
### Table 3. Outbreak management systems with at least two key data components.

<table>
<thead>
<tr>
<th>Systems</th>
<th>Surveillance</th>
<th>Contact tracing</th>
<th>Case mgmt.</th>
<th>Lab results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Systems identified in Tom-Aba et al.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BioCaster Portal</td>
<td>✓</td>
<td>?</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Bio-Sense 2.0</td>
<td>✓</td>
<td>?</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Collaborative Overarching Multi-feed Biosurveillance</td>
<td>✓</td>
<td>?</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>CommCare Contact Tracing</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sense Ebola Followup</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Ebola Tracks</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mHealth real-time infectious disease interface</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>NNDSS</td>
<td>✓</td>
<td>?</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Open Data Kit</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sentinel surveillance system (SSS)</td>
<td>✓</td>
<td>?</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>“Smartphone-based contact tracing system”</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SORMAS</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>“Telemedicine”</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Systems identified through our rapid review strategy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AfyaData</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPZone</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Conduent Maven</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Atlas WorldCare</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>EbolaDiag[16]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Coconut Plus[17]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Ebola telemonitoring system[18]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Smartphone-based molecular diagnostics system[19]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>OpenMRS-Ebola EHR[20]</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

In partnership with the field workers and epidemiologists at the Ebola Emergency Operations Centre in Nigeria, building on the latter’s experience of successfully containing the country’s Ebola outbreak that was concomitant with the West African epidemic[27]. The authors identified different “personas”, or user roles, involved in the process of EVD outbreak management (e.g. informants, rumour officers, surveillance/case/contact officers and supervisors). For each role, the authors established the relevant activities, information items and interactions with other roles, and used these to define the information flow between the different user types. To give one example, “a rumour officer conducts an initial triage on all incoming rumours concerning possible cases, and uses the rumour information checklist to report the results of this triage to the surveillance supervisor”. SORMAS was then designed to facilitate and capture this information flow via the corresponding system modules. One reportedly unique feature of this design was that the flow of information was bi-directional, such that supervisors could instruct officers on the follow-up steps that needed to be taken[27].

The latest (2018) update of the system incorporates epidemiological data capture, outbreak management user roles and information flows for several additional diseases, such as Lassa fever, monkeypox, bird flu, dengue, yellow fever, measles, meningococcal infections, Yersinia pestis and cholera[33]. Data aggregation and analysis are implemented as a web-based multi-user overview dashboard. It includes live case data, epidemiological curves, case maps displayed via an embedded GIS visualisation module and transmission chain network diagrams constructed from contact tracing data (see Figure 2–Figure 5, taken from 33). These features were implemented as a result of several “design thinking” focus group workshops with Nigerian target users, carried out in Nigeria and Germany between 2015 and 2017[33,34]. As of 2018, the Nigerian Centre for Disease Control recommends that health
Figure 2. National level dashboard of SORMAS: monkeypox, December 2017.

Figure 3. National level dashboard of SORMAS: monkeypox, cumulative incidence by Local Government Authority, December 2017.
Figure 4. National-level dashboard of SORMAS: monkeypox, GIS-location of cases (red) and contacts (yellow), December 2017.

Figure 5. SORMAS: contacts and transmission chain dashboard.
facilities and local governments across all of the country’s 36 states deploy SORMAS\textsuperscript{39}. An online demo of the SORMAS dashboard is publicly available\textsuperscript{40}.

CommCare Ebola Response
CommCare is a cloud-based software development platform that enables users to design open source mobile applications for data collection and to distribute these applications to Java-based and high-end Android phones, with the collected data being sent back to the CommCare HQ data analysis portal\textsuperscript{37}. A special EVD management system was set-up using this platform, comprising of case management, surveillance, contact tracing and laboratory result tracking, but with the caveat that the gathered data is hosted at CommCare\textsuperscript{38}. As a result, this platform does not satisfy our capacity-building requirement of enabling affected countries to host the outbreak management system entirely on local IT infrastructure. Additionally, while the system has been built using open-source technologies, its source code is not available for public download. This means that it cannot be enhanced independently of the original vendor, which further invalidates the use of CommCare Ebola Response under our criteria.

Sense Ebola Follow-up
Sense Ebola Follow-up was originally an Android-based, real-time data capture mobile application that guided health workers through the process of registering a contact and performing a follow-up. It was developed at the peak of the West African Ebola outbreak in 2014. Since then it has been extended to work with both Android and iOS devices, and provides case investigation, contact listing and contact follow-up forms, as well as functionality to make a laboratory investigation request. A server-side data integration and analysis dashboard is also provided\textsuperscript{1,29}. However, as with CommCare Ebola Response, the source code is not available for public download, which makes Sense Ebola Follow-up unsuitable for capacity building purposes.

AfyaData
AfyaData is an open-source infectious disease surveillance system developed at the Southern African Centre for Infectious Disease Surveillance (SACIDS) in Tanzania\textsuperscript{15}. It consists of an Android-based mobile application and a web-based application acting as a server. The mobile app is used for collecting and submitting surveillance data, and for receiving and tracking feedback while the server component is responsible for data storage and management. The following activities are supported by the system: community-based participatory disease surveillance (for timely detection and reporting of disease events at the community level), official surveillance strategy (for timely collection and submission of disease data at the health facility level) and two-way communication feedback to individuals who report disease events at community and health facility levels. Notably, contact tracing and laboratory result management are absent from the set of features listed by Karimuribo et al.\textsuperscript{15}. Despite this, we have included AfyaData as an exception because development work on this system is ongoing\textsuperscript{13}. The system has been trialled during the latest Ebola outbreak in the Democratic Republic of Congo\textsuperscript{19}, where it was used for syndromic surveillance.

Commercial and proprietary systems
During the 2014–2016 West African Ebola outbreak, eleven people were treated for EVD in the United States\textsuperscript{41}. This created a risk of an EVD outbreak in the country and two commercial software vendors responded to this threat by offering outbreak management tools with specific Ebola-related features. Maven by Conduent\textsuperscript{21} and WorldCare by Atlas Software\textsuperscript{2,41} both purportedly provide case management, surveillance, contact tracing and laboratory result management functionalities. However, no public documentation is available for these systems (one has to request a private demo to see how they operate) and there is currently no prospect of either Maven or WorldCare becoming free of charge and open source.

HPZone is Public Health England’s disease surveillance and outbreak management system\textsuperscript{23}. It provides case management, surveillance, contact tracing and laboratory result management for up to 100 different diseases but is neither publicly available nor open source.

Our emails to participants at ICREID identified four additional systems: EpiInfo\textsuperscript{24} (used by four participants), Open Data Kit\textsuperscript{18}, HPZone\textsuperscript{37} and KoBoCollect\textsuperscript{25} (each used by one participant). The other eight respondents did not provide the names of the systems they used, and one participant reported using an Excel spreadsheet. In total, six respondents used their outbreak management system on a tablet or a smartphone. Of the named systems, EpiInfo, KoBoCollect and HPZone were reported to have all four key data components of outbreak management. However, EpiInfo is a desktop-based application and neither it nor HPZone are open source. Meanwhile, KoBoCollect is similar to Open Data Kit, in that it is a mobile-based data collection platform that requires users to design their own data entry forms instead of having built-in outbreak management functionality. As a result, none of the systems mentioned by the survey respondents satisfy our inclusion criteria.

Discussion
Our review found numerous existing mobile-based outbreak management systems, but only SORMAS satisfies the capacity-building requirements of PANDORA-ID-NET (open source code and the ability to store gathered data in the region of origin) and fully integrates data from case management, contact tracing, laboratory work and surveillance components, in line with the lessons learned during the 2014–16 West African Ebola outbreak.

The PANDORA-ID-NET consortium aims to take a multi-faceted approach to building outbreak response capacity. In addition to standard preparedness measures, this includes strengthening the surveillance of zoonotic diseases, improving the quality of collected syndromic and laboratory data and facilitating clinical trials of relevant interventions. We describe below the corresponding features that would be complementary to the four key data components of an integrated outbreak management system and outline why these features are desirable.

Zoonotic disease surveillance
Many infectious disease outbreaks are due to emerging infectious diseases (EIDs). The latter are defined as infections
that have appeared in a population for the first time or have previously been present but are rapidly increasing in incidence or geographic range. It is estimated that approximately 60% of such diseases have zoonotic origins and 72% of these have originated in wildlife. Surveillance and management of zoonotic EIDs has multiple components:

- **Domestic and livestock disease control.** This requires animal vaccination, vector control, test and treat, cull and quarantine programs, as well as the participation of local populations in animal disease surveillance and reporting.

- **Wildlife disease control.** This is achieved through regular examinations of animals in the wild for signs of disease using macroscopic, histopathological, bacteriological and parasitological analyses.

- **Ecological surveillance.** There is mounting evidence that human-led ecological disruption is contributing to emerging (and re-emerging) zoonotic diseases. It has been hypothesised that changes in land use in the Guinea Forest Region, such as intensified agriculture and clear-cut logging, played a significant part in the ecotypic shift of the Ebola virus before the 2014–2016 West Africa outbreak. This shift turned a self-limiting rural infection into a proto-pandemic disease, capable of spreading to densely populated urban areas. The goal of ecological surveillance is to monitor the environment for changes, e.g. in land use or wildlife population levels, and establish how these changes alter the risk of zoonotic infection re-emergence. One example of this type of surveillance is predicting the spread of zoonotic monkey malaria infection in areas with sparse incidence data, using deforestation patterns and other environmental information.

The domestic, livestock and wild animal disease control components are collectively referred to as “one health” in the context of epidemiological surveillance. Out of the systems we identified, only AfyaData offers animal disease surveillance features and none have ecological surveillance capabilities.

**Passive data collection**

There has been recent progress in rapid Ebola testing using mobile and point-of-care devices. Outbreak management systems would benefit from automatically collecting data from such devices to increase the data collection speed and reduce data entry errors. Of the reviewed systems, only CommCare Ebola offers such functionality.

**Clinical trial data management**

At the peak of the 2014–2016 Ebola epidemic, the WHO convened a panel to discuss the ethical permissibility of using unregistered treatments during such outbreaks. It concluded that “it would be acceptable on both ethical and evidential grounds to use as potential treatments or for prevention unregistered interventions that have shown promising results in the laboratory and in animal models but have not yet been evaluated for safety and efficacy in humans, provided that certain conditions are met.” However, this recommendation came with the caveat that “researchers have a moral duty to evaluate these interventions (for treatment or prevention) in clinical trials that are of the best possible design in the current exceptional circumstances of the West African Ebola outbreak, in order to establish the safety and efficacy of the interventions or to provide evidence to stop their use. Continuous evaluation should guide future interventions.” In light of these recommendations, clinical trial data management will become essential in future outbreak management scenarios. However, none of the reviewed systems support this functionality.

To summarise the above, zoonotic disease surveillance, passive data collection and clinical trial data management – features that are important for strengthening outbreak response preparedness – are not currently present in most of these systems, including SORMAS.

Our review of systems was rapid, since it was conducted by a single reviewer and only considered publications written in the English language. Additionally, the grey literature search scope was limited. However, this review was also rigorous, as we leveraged the results of an existing systematic review of this field by Tom-Aba et al., conducted our own supplementary review to increase coverage, and contacted experts working in this area to identify systems they have used that were not published in the grey or grey literature that we searched. Although the scope of Tom-Aba et al.’s review is limited to the 2014-16 Ebola outbreak, our broader supplementary search did not uncover additional systems that satisfy our own inclusion criteria. As a result, we believe that our review comprehensively covers the field of mobile-based outbreak management systems (apart from those that deal exclusively with vector-borne diseases). However, this is notwithstanding the fact that there are proprietary systems for internal use without sufficient public exposure that we could not include in this review. Humanitarian agencies may have their own in-house systems for outbreak management without any associated publications. One example is Médecins Sans Frontières’ Health Surveillance Programme (HSP), which uses mobile devices in conjunction with the commercially licensed Dharma software platform for data collection and management via a dashboard interface. In 2015, HSP was deployed in Najaf Governorate, Iraq for real-time identification of scabies outbreaks. Although MSF report significant data collection speed improvements and cost savings, very little information is available on the actual functionality of this system. Systematic evaluation of outbreak management system performance remains an outstanding research question. SORMAS boasts a formidable level of functionality, has an impressive regional deployment coverage and has been subjected to several field-user evaluation exercises that have returned satisfactory results. However, disease outbreaks are unforeseen events by
their nature, making it challenging to design a consistent set of evaluation criteria for objectively rating and comparing how successful these systems are in reducing the numbers of new cases in an ongoing epidemic.

The implication of our review is that SORMAS is a mature system that is ready to be disseminated and rolled out in regions with high infectious disease burden, and, owing to the system’s open-source design, can be further adapted to meet specific public health needs of individual countries. Future areas of research and development for SORMAS could include zoonotic disease surveillance, passive data collection and clinical trial data management. A thorough systematic review for each of these areas would help to establish the existing best clinical trial data management. A thorough systematic review that is ready to be disseminated and rolled out in regions.

Conclusion
Mobile-based digital outbreak management is an emerging field, with new tools being continuously developed. Out of the identified outbreak management systems, only SORMAS satisfies all of the inclusion criteria. Additionally, we have outlined a number of desirable features that are not currently present in most outbreak management systems, including SORMAS. Our paper provides a timely review of the state of the art mobile-based and open source case detection and infectious disease outbreak management systems that we hope others can use as a source for identifying tools for their research and communicable disease control efforts.

Data availability
All data underlying the results are available as part of the article and no additional source data are required.

References

14. GlobalStats StatsCounter; Mobile operating system market share in Africa (January 2019). Retrieved February 1, 2019. Reference Source
22. Atlas Medical Software; WorldCare. Reference Source
27. Fähnnich C, Denecke K, Adoeye OÖ, et al.; Surveillance and Outbreak Response


32. SORMAS: GitHub repository. Retrieved February 1, 2019.


38. Personal communication with Dr. Leonard Mboera.


41. Atlas Medical Software: Surveillance and outbreak management. Retrieved February 1, 2019


52. Dharma platform website.

Page 13 of 17
Open Peer Review

Current Peer Review Status: ✔️ ❓

Version 1

Reviewer Report 29 June 2020

https://doi.org/10.21956/wellcomeopenres.17236.r38726

© 2020 Fraser H. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Hamish S. F. Fraser
Brown Center for Biomedical Informatics, Brown University, Providence, RI, USA

This article is a focused systematic review of disease surveillance systems for the management of Ebola based on the experience and reports of systems deployed to tackle the West African Ebola epidemic in 2014-2016. This is a highly important topic even before the Sars-Cov-2 pandemic, with many early miss-steps and lessons learned.

As a result of the lack of well documented and appropriately designed systems at that time considerable work has been carried out to build and validate new systems. The criteria used by the authors are generally appropriate, focusing on essential functionality, evidence of successful use and open source software. The later is important in ensuring accessibility to the code in the least developed countries, local control of the code and clinical data, and the ability to carry out capacity building of local programming teams to support and improve systems. The use of the earlier review by Tom-Aba et al is appropriate.

Overall this is a well written article of immediate relevance to the global health community well beyond the low income settings these tools were generally developed for. However there several limitations particularly in the interpretations of the results in the broader context of global health informatics and clinical data sharing.

Suggested revisions:
1. The search strategies are documented in the article but additional data on grey literature searched and from interviews is less clear. Clarification of what was done would be helpful.

2. The underlying strategy for disease surveillance systems in the article appears to be that they are separate tools used for a specific outbreak and running in parallel to any existing health information systems – whether paper or electronic. This is very problematic as it has the potential to divert attention away from the management of the many other diseases and health needs in vulnerable populations during outbreaks. In between outbreaks this strategy is likely to markedly weaken the surveillance functionality. What is needed is for a system like SORMAS or CommCare to run in the background receiving high quality, timely
data from routine health care contacts in health facilities and the community. This will then support syndromic surveillance, notifiable disease reporting and tracking of lab results that allow early detection of new outbreaks.

Therefore an EHR system such as OpenMRS does not and arguably should not try to manage community case finding and contact tracing. It should share a common data set with such systems and allow seamless interoperability so that there is one view of all patients suspected, tested and confirmed to have the disease. Therefore fulfilling the 4 functions on the list can be accomplished by two or more systems that have very good interoperability. This is the goal of Digital Square, OpenHIE (ohie.org) and OpenMRS (openmrs.org) for example with interoperability between OpenMRS and CommCare well established and with SORMAS made possible recently. This interoperability work is being rapidly accelerated at present due to COVID-19. HL7 FHIR is now being used to develop standardized interoperability profiles and interfaces for these systems.

3. Notifiable disease reporting and syndromic surveillance data for example should come from routine health systems, such as EHRs with primary care functionality. Emphasis needs to be placed on data collection quality, completeness and timeliness. Lack of direct links to primary care data was a critical fail point in the eIDSR system in Tanzania due to the poor quality and management of paper registers (unpublished study).

4. Central curation of a core data set is a major advantage in this work allowing a wide range of data collection tools to feed into the surveillance system even if they do not support full interoperability. This was a key finding in the OpenMRS-Ebola and Buendia (MSF/Google) systems in the 2014-2016 Ebola outbreak and is central to work on data collection and interoperability for COVID-19 data. It has been accomplished for OpenMRS and OpenHIE with the CIEL concept dictionary.

5. Some mention of the significance of the work for the current COVID-19 pandemic would be helpful.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions drawn adequately supported by the results?
Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Global Health Informatics, EHR systems, open source software in health, evaluation of health information systems

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 11 June 2020

https://doi.org/10.21956/wellcomeopenres.17236.r38827

© 2020 Rweyemamu M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mark M. Rweyemamu

SACIDS Foundation for One Health, Sokoine University of Agriculture, Morogoro, Tanzania

The authors have reviewed the features of their own SORMAS system and seven others against the criteria set by WHO for Ebolavirus disease case management.

The review seems to focus particularly on information/data flow from diverse sources within a health system. As a review on case management, the emphasis is more on expert management support than efficiency of individual components.

The combined experience of the Ebola epidemic in West Africa and the current COVID-19 pandemic are likely to lead to an increasing role mobile systems in disease surveillance and case management system. This review is timely.

The SORMAS was the only one that met all the criteria used for the review. It is of note that the review found only two systems that were based on Opensource programming and only one that was One Health based.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Infectious diseases; One Health

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.