<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Starting with small health data opportunities for mHealth in Africa</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Heavin, Ciara; O'Connor, Yvonne</td>
</tr>
<tr>
<td><strong>Publication date</strong></td>
<td>2016</td>
</tr>
<tr>
<td><strong>Type of publication</strong></td>
<td>Article (peer-reviewed)</td>
</tr>
</tbody>
</table>
| **Link to publisher's version** | https://www.witpress.com/elibrary/dne-volumes/11/3/1205  
http://dx.doi.org/10.2495/DNE-V11-N3-338-351 |
| **Item downloaded from** | http://hdl.handle.net/10468/6443                             |

Downloaded on 2019-04-20T18:10:00Z
STARTING WITH SMALL HEALTH DATA OPPORTUNITIES FOR MHEALTH IN AFRICA

CIARA HEAVIN & YVONNE O’CONNOR
Health Information Systems Research Centre, University College Cork, Ireland.

ABSTRACT
The need to obtain data to understand effective and available child mortality-reducing control measures in rural areas of developing countries is great. Evidence shows that this challenge can potentially be overcome with the increased availability of Information and Communication Technology (ICT) to support the data/information/knowledge needs of healthcare delivery services in low resource settings. Recognising the benefits of ICT and the need for improvements in the Nigerian health sector, this paper outlines the plans for the technical feasibility assessment of the IMPACT (using Mobile Phones for Assessing, Classifying and Treating sick children) smartphone application to capture, store and analyse child health assessment data. IMPACT is a secure, scalable, user friendly mobile health (mHealth) innovation that is being developed to support ‘small data’ capabilities within the context of healthcare in the community in Enugu State, Nigeria, Africa. Notwithstanding the heightened focus on ‘big data’ in health, this research is interested in investigating the opportunities associated with doing ‘small healthcare data’ well, with the long term view of building to the big data scenario for healthcare in the community in Enugu. This paper outlines the plan for the IMPACT project considering the implications for health data, knowledge management in healthcare and the big data opportunities to support disease surveillance, healthcare delivery and resourcing and healthcare practitioner education.

Keywords: data, IMPACT application, Information and Communication Technology, knowledge management and big data, mobile health (mHealth).

1 INTRODUCTION
In developing countries, millions of children under the age of five die due to inadequate healthcare services in remote, rural areas. In Nigeria, the under-five mortality rate is 124 per 1,000 births. According to the Nigerian Health Profile published by the World Health Organisation [1], malaria was the leading cause of death, closely followed by acute respiratory infections. While the number of child mortalities in Nigeria has reduced in the last decade, the rate is still high when compared with the global average and falls short of achieving Millennium Development Goals (specifically, Millennium Development Goal 4).

To address this situation, the World Health Organisation (WHO) and UNICEF have developed clinical guidelines for community/primary health workers to deliver healthcare services to children under the age of five in remote, rural areas of Low-to-Middle Income Countries (LMIC). These guidelines are known as integrated Community Case Management (iCCM) and are adapted to individual countries basis based on the National Child Health Index. iCCM is an algorithmic approach, which enables primary healthcare (PHC) workers to assess, classify and treat seriously ill children. Employing iCCM PHC workers capture patient-related personal data (e.g. name, caregiver name and address) and clinical assessment information (e.g. the presence/absence and duration of symptoms), to classify illness and identify appropriate treatment recommendations for sick children. The
underlying aim of iCCM is to reduce death, illness and disability while promoting improved growth and development among children under five years of age [2].

Numerous LMIC employ iCCM guidelines [3, 4]. In Nigeria, iCCM is currently being rolled out in two states (i.e. Niger and Abia) with future rollouts planned in an effort to provide basic healthcare services to over 300,000 children [5, 6]. However, research has documented shortcomings of the paper-based iCCM approach with reports of poor PHC worker adherence to the guidelines, leading to misdiagnosis and mistreatment of seriously ill children [7].

With this in mind, the purpose of this paper is to outline the ‘small data’ opportunity and the prospective ‘big data’ mHealth scenario following the introduction of an ICT-enabled mHealth artefact amongst PHC workers in Nigeria. While our investigation is focused on supporting the capture, storage and analysis of child health data at an individual level, which will potentially enable improved availability of patient health information in the short term. We are keen to understand the ‘big data’ possibilities and the opportunity for creating new knowledge and better insights for paediatric health data in the community in Enugu.

This paper is organised as follows, the next section considers the opportunities for mHealth in a developing country such as Nigeria. Following this, the concepts of data, information and knowledge are positioned within the context of Health Information Systems (HIS). The notion of big data in mHealth is considered. The proposed research approach for the study is outlined and is followed by discussion on the IMPACT project and the potential knowledge creation and big data opportunities. Finally, the conclusions are presented.

2 MOBILE HEALTH (MHEALTH) IN DEVELOPING COUNTRIES

The substantial growth in mobile technologies has revolutionised how many industries conduct their work, primarily in a developing world context. Since the establishment of the Millennium Development Goals (MDG) in 2000 and the Sustainability Development Goals (SDG) in more recent times the healthcare sector in developing countries has witnessed the introduction of various mHealth technologies [8, 9]. Numerous mHealth technologies were introduced (cf. [10, 11]) in an attempt to improve the livelihood of civilians in resource-poor settings whether at a local, regional and/or national level [12, 13]. Given the geographic dispersion of healthcare services in developing countries, the interaction of many healthcare providers and the need to balance the complex sequence of events required to care for each patient requires substantial application of mobile technologies [14].

Existing mHealth studies acknowledge that the point-of-care delivery is vital for the success of any application in the clinical healthcare environment [15–17] as healthcare providers perform clinical work on various individuals in multiple locations [18]. Davis [19] posits that achieving a dramatic increase in access to data and computing requires the utilisation of mobile technological devices. One of the integral driving forces, therefore, behind these mHealth innovations is the need to obtain health-related data, information and knowledge [17] in an attempt to galvanise global attempts at providing a better lifestyle for people in developing countries.

3 HEALTH DATA, INFORMATION AND KNOWLEDGE

Establishing definitions for data, information and knowledge within the context of the healthcare domain is essential. Indeed, ‘people imagine data in different ways’ [20]). To overcome this, a common meaning or ontology among healthcare providers, patients and other stakeholder around key medical concepts is required. An ontology, according to Staab and Studer [21], is a description (like a formal specification of a program) of concepts and relationships that can exist for an agent or a community of agents. The concept is important in terms of building a common language [22] for the purpose of enabling knowledge sharing and reuse, which is integral for a standardised delivery of
healthcare [23]. In particular, it is noted that many Information Systems (IS) researchers (for example: Dennis, Earl, El Sawy and Huber) have used the terms information and knowledge interchangeably. Figure 1 represents data, information and knowledge as a continuum. Figure 1 indicates that the extremes of each phenomenon are distinct however there is significant overlap between data/information and information/knowledge.

According to Davenport and Prusak [24] ‘the distinction between knowledge and information is seen as more of a continuum than a sharp dichotomy. Most projects that focus on internal knowledge [repository] deal with the middle of the continuum-information that represents knowledge to certain user’. The healthcare domain has transformed into a business sector that is driven by information resources and complex knowledge with an increasing volume of transactional health data and a need for high value knowledge to support decision making needs of healthcare providers [26]. However, the healthcare sector is criticised for the proliferation of terminologies across medical specialities, locales, healthcare facilities and clinical content [27]. Overcoming this challenge is reported to have numerous benefits for the healthcare sector. For example, the development and use of widely understood and accepted ontologies can build more powerful and more interoperable IS in healthcare [28] by facilitating the easy transfer and (re)use of patient-related data [29].

There exists vast amounts of information about patient history, symptoms, functions and lifestyle; information about diseases, diagnostic, drugs and treatment methods [30]. Working in information-intensive institutions, such as the healthcare sector [31], it is important for consistent ontologies to exist so that healthcare practitioners can create knowledge from all the available information if they are to adhere to today’s key aspect of medical practice; namely “evidence based medicine” [32, 33]. Evidence-based medicine practice integrates individual clinical expertise and the best evidence found in research [34]. This style of medical practice provides doctors with conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients [35]. Clinical/medical knowledge is of critical importance to patient care delivery [36] and handheld mobile devices can provide healthcare providers with accurate and real-time information when needed [37].

Medical knowledge refers to information about health and disease [38] that is understood to be generalisable to the care of all patients [39]. According to [40] the patient treatment process can be improved by providing medical knowledge. However, medical knowledge changes rapidly [41]; therefore, it is important that medical professionals constantly create knowledge to ensure that patient safety is not compromised. Sufficient medical knowledge may help prevent doctors from
misdiagnosing and prescribing medicine that cause adverse drug events resulting in potential problems [42]. El Morr and Supercaze [34] argue that knowledge created by one healthcare practitioner is of utmost importance to the community of healthcare practitioners in order to deliver quality of care. However, to make use of knowledge, Yang et al. [43] posit that it is important to understand the way knowledge is created.

But what constitutes the term ‘knowledge’? Generating numerous debates, the definition of knowledge has been of interest to many scholars as knowledge is seen as a versatile concept with multi-layered meanings [44]. From a review of the literature Alavi and Leidner [45] found several perspectives of knowledge including (1) a state of mind, (2) an object, (3) a process, (4) a condition of having access to information or (5) a capability (see Table 1). Knowledge, in this study, is viewed as a condition of access to information [46], a process of applying expertise [47] and the potential to influence action [48].

Knowledge has been differentiated into two main modes: explicit and tacit [44]. The former is defined by Grant [49] and Berman et al. [50] as knowledge that is based on facts and theories that can be codified, replicated and transmitted to others easily in formal and systematic language. The latter is knowledge that is largely embodied or personal knowledge [44, 51]. That is, the implicit knowledge that people have developed [52], which often evolves after experience [53].

Knowledge is of extreme importance to healthcare as the healthcare industry has become a knowledge-based community that depends critically on knowledge management activities to improve the quality of care [54]. The goal of knowledge management initiatives is to make readily available the right knowledge to the right processes (human or computer) at the right times in the right presentation for the right costs [55]. Within knowledge management, Allard [52] posits that the creation of knowledge (sometimes referred in the literature as knowledge creation or generation) is viewed as essential to an organisation. Table 2 provides the definition for knowledge creation as being described as a process, output or outcome as proposed by Mitchell and Boyle [56].

According to Mitchell and Boyle [56], there is a lack of evidence however as to the existence of a comprehensive definition of knowledge creation. This lack of consensus has often led to the creators of knowledge being unaware of the different ways in which knowledge is created, and as a result, Persson and Stirna [57] posit that valuable knowledge may be lost. To prevent this, the authors argue that a consistent and integrated approach to knowledge creation is required. One possible way is to initially gather small datasets moving along the continuum to larger dataset. The subsequent promotion of a ‘big data’ environment could then be used to analyse trends/patterns in large data sets, thereby generating new insights in the medical domain [58, 59].

<table>
<thead>
<tr>
<th>Knowledge perspective</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge vis-à-vis data and information</td>
<td>Data is facts, raw numbers. Information is processed/interpreted data. Knowledge is personalized information.</td>
</tr>
<tr>
<td>State of mind</td>
<td>Knowledge is the state of knowing and understanding.</td>
</tr>
<tr>
<td>Object</td>
<td>Knowledge is an object to be stored and manipulated.</td>
</tr>
<tr>
<td>Process</td>
<td>Knowledge is a process of applying expertise.</td>
</tr>
<tr>
<td>Access to information</td>
<td>Knowledge is a condition of access to information.</td>
</tr>
<tr>
<td>Capability</td>
<td>Knowledge is the potential to influence action.</td>
</tr>
</tbody>
</table>
'BIG DATA' SCENARIOS IN MHEALTH

Big data describes a dataset that is so large and complex that they ‘require advanced and unique data storage, management, analysis and visualisation technologies’ [60]. A vast amount of industry, company, product and customer data across businesses scenarios, including retail, automotive, search engine, can be gathered from many external and internet sources [61, 62]. The three key attributes of big data are volume, velocity and variety (also known as the 3 ‘V’s). These attributes characterise big data in terms of: (1) the availability of large volumes of data and the benefits from having more data to develop better models, (2) significantly increased data processing speed, thereby better facilitating decision making and action taking and (3) the nature of data. Data are messy and complex due to the many sources of the data and the many formats of the data with more than 90% of big data being unstructured [63] and inconsistent [64]. Some researchers include ‘value’ as a fourth ‘V’, indicating that top-performing organisations cite Big Data and Business Analytics as a key differentiator [65, 66] to guide both future strategies and day-to-day operations [64].

Big data refers to a huge amount of data that conventional data warehouse technologies cannot store, manage or analyse, but is required to generate improved insights when evaluating new business opportunities and to enhance organisational decision making [67]. Existing research recognises the affordances of ‘big data’ in the health domain, extolling big data as an opportunity to leverage patient and practitioner data as a means of improving the quality and efficiency of healthcare systems [63]. Big data in healthcare denotes large and complex electronic health data-sets, which are perceived to be difficult (or impossible) to manage with traditional data management tools [68].

According to Raghupathi and Raghupathi [69], ‘big data’ in a healthcare context amalgamates clinical data from Computer Physician Order Entry (CPOE) systems and Clinical Decision Support Systems (CDSS) (physician’s written notes and prescriptions, medical imaging, laboratory, pharmacy, insurance and other administrative data); patient data in Electronic Health Records (EHR); machine generated/sensor data (e.g. vital sign monitoring), social media posts (e.g. Twitter) and web pages and less patient-specific information, including emergency care data, news feeds and articles.

Table 2: Representative uses of the term ‘knowledge creation in the literature (adapted from Mitchell and Boyle 2010).

<table>
<thead>
<tr>
<th>Knowledge creation component</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process</td>
<td>Refers to the initiatives and activities undertaken towards the generation of new objects. i.e. Method or means through which knowledge is created and can be differentiated from the end result</td>
</tr>
<tr>
<td>Output</td>
<td>Defined in terms of an immediate product of the knowledge creation process, such as the representation of an idea, and can be differentiated from its impact on the organisational system. i.e. Refers to the development of new ideas that reflect a significant elaboration or enrichment of existing knowing</td>
</tr>
<tr>
<td>Outcome</td>
<td>Defined in terms of a value-adding object. Refers to new knowledge diffused, adopted and embedded as new products, services and systems.</td>
</tr>
</tbody>
</table>

4 ‘BIG DATA’ SCENARIOS IN MHEALTH

Big data describes a dataset that is so large and complex that they ‘require advanced and unique data storage, management, analysis and visualisation technologies’ [60]. A vast amount of industry, company, product and customer data across businesses scenarios, including retail, automotive, search engine, can be gathered from many external and internet sources [61, 62]. The three key attributes of big data are volume, velocity and variety (also known as the 3 ‘V’s). These attributes characterise big data in terms of: (1) the availability of large volumes of data and the benefits from having more data to develop better models, (2) significantly increased data processing speed, thereby better facilitating decision making and action taking and (3) the nature of data. Data are messy and complex due to the many sources of the data and the many formats of the data with more than 90% of big data being unstructured [63] and inconsistent [64]. Some researchers include ‘value’ as a fourth ‘V’, indicating that top-performing organisations cite Big Data and Business Analytics as a key differentiator [65, 66] to guide both future strategies and day-to-day operations [64].

Big data refers to a huge amount of data that conventional data warehouse technologies cannot store, manage or analyse, but is required to generate improved insights when evaluating new business opportunities and to enhance organisational decision making [67]. Existing research recognises the affordances of ‘big data’ in the health domain, extolling big data as an opportunity to leverage patient and practitioner data as a means of improving the quality and efficiency of healthcare systems [63]. Big data in healthcare denotes large and complex electronic health data-sets, which are perceived to be difficult (or impossible) to manage with traditional data management tools [68].

According to Raghupathi and Raghupathi [69], ‘big data’ in a healthcare context amalgamates clinical data from Computer Physician Order Entry (CPOE) systems and Clinical Decision Support Systems (CDSS) (physician’s written notes and prescriptions, medical imaging, laboratory, pharmacy, insurance and other administrative data); patient data in Electronic Health Records (EHR); machine generated/sensor data (e.g. vital sign monitoring), social media posts (e.g. Twitter) and web pages and less patient-specific information, including emergency care data, news feeds and articles.
in medical journals. Therefore, mHealth technologies with consistent ontologies will enable the creation of ‘big data’ for healthcare scenarios, thus enabling the discovery of associations and understanding of patterns and trends within the datasets for diagnosis, prognosis and treatment of illness and enhanced understanding of healthcare resource distribution [70]. Table 3 presents examples of the 4 ‘V’s associated with big data. It is evident from this table that ‘small data’ (whether structured or unstructured) is initially required to ensure that ‘big data’ opportunities are feasible.

In order to achieve this big data scenario, we recognise the need to leverage mHealth as part of this ‘small healthcare data’ opportunity to support the provision of child healthcare services at the community level, with the potential positive impact on paediatric health and reduction in mortality at a local level in Enugu. With additional technical infrastructure and new ways of leveraging existing channels to support the capture, storage and dissemination of health-related data, we aspire to achieving the big data mHealth scenario characterised in Table 3. The next section outlines the research approach employed as part of this study.

### Table 3: Characterisation of big data mHealth scenario.

<table>
<thead>
<tr>
<th>Knowledge type</th>
<th>Volume</th>
<th>Velocity</th>
<th>Variety</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health</td>
<td>Patient Data</td>
<td>Direct interactions with patients</td>
<td>Structure, semi-structured and unstructured data</td>
<td>High value</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>· Trends – disease surveillance</td>
<td>Patient data</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>· Demographic data</td>
<td>Population data</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Location-based data</td>
</tr>
</tbody>
</table>

5 RESEARCH APPROACH

A three-phased approach will be employed as part of this study. Commencing in March 2016, the first four months of the IMPACT (usIng Mobile Phones for Assessing, Classifying and Treating sick children) research programme will be allocated to gathering the necessary data from secondary data sources and key stakeholders (via frequent contact, interviews and focus groups). In parallel, the ‘IMPACT’ smartphone application is being developed in accordance with the results obtained from this fact finding phase coupled with the iCCM guidelines for Nigeria (see Fig. 2).

Following this, primary data collection will be conducted in two parts. Firstly, an ‘As-Is’ assessment of PHC workers and their role in the assessment of young children. Data will be collected using interviews and observational techniques. We will employ positional and reputational method [71] to identify participants. During this trip we will also conduct some preliminary testing of the IMPACT app to identify any potential challenges (these may be technical, human, cultural) to executing the feasibility study. The data gathered during this trip will underpin the baseline data for the study.

Secondly, a two week feasibility study (visit 2) will be conducted in Enugu State Nigeria in early 2017. We will work with our Nigerian partners in Enugu State University of Science and Technology (ESUT) to run a technical feasibility study of the IMPACT application. Prior to conducting the feasibility study, the team will conduct a one-day training session with the 50 PHC workers involved in the study. The IMPACT project team will work on the ground with the PHC workers training them in the use of the IMPACT app; this will include role play scenarios involving the community to simulate the app working in a real life scenario. The PHC workers will use the IMPACT app and the
smartphone will remain with them for 10 days. PHC workers using the ‘IMPACT’ mobile app will be asked to fill in a diary whereby they are free to express their experiences using the application during that time. Participation will be voluntary and participants will be free to leave the study at any time. Participants who agree to be included in the study will be asked to complete a signed consent form and will be assigned a unique identification number thereafter to ensure data anonymity. Google analytics will be used to capture and analyse the application’s usage during the technical feasibility assessment. This usability data will extend this big data opportunity offering new insights into how direct users (i.e. PHC workers) engage with the technical intervention.

### 6 DISCUSSION

At the outset the focus of the project is to develop an ICT artefact; namely IMPACT, an mHealth iCCM application suitable for use as part of a medical assessment of young children in Nigeria. The iCCM guidelines for Nigeria are currently being reviewed, codified and customised as part of the ‘IMPACT’ application. ‘IMPACT’ is being developed (Figs 3 and 4 illustrate the splash screen and the home screen for the IMPACT app, respectively) using a clinical guideline decision support rule engine, which uses classification and treatment rules for assessing a sick child defined in XML (eXtensible Markup Language). This will enable the team to initially generate small datasets on patient-related data (more specifically, paediatric data). The use of XML to codify the CCM rules for Nigeria supports the scalability of this rule base, as well as the flexibility to extend the IMPACT app capabilities by developing other electronic guidelines based on new WHO/UNICEF directives.

The app will be user tested with clinical experts in Enugu State University of Science and Technology to ensure that ‘IMPACT’ can be used on the ground in Nigeria without impeding on the current delivery of healthcare services to children under five years of age. The underlying rationale for conducting this testing is to capture both the explicit and tacit knowledge of PHC workers in
Nigeria to improve (if required) the IMPACT application. Further input will be obtained from key stakeholders within the Nigerian healthcare sector (include examples) to identify how health data is currently disseminated between different healthcare bodies. This mHealth innovation is a cost-effective solution as it uses low cost but secure mobile technologies, which enable the communication and dissemination of critical datasets to necessary stakeholders who currently have very limited or no access to such health data and information.

The IMPACT app, as illustrated in Figs 3 and 4, is being developed as a secure, cloud-based solution. Using SQLCipher for Advanced Encryption Standard (AES) full database encryption and Amazon Web Services (AWS) for cloud data storage, the IMPACT app will be easily scaled to support the big data scenario outlined in this paper.

The proposed technical feasibility study will provide an outstanding opportunity to test and evaluate the security, robustness and efficacy of the IMPACT application in an African country where technical infrastructure can be a challenge. This feasibility study will act as a proving ground for the capture and storage of complete, accurate, real-time child health records (i.e. ‘small data’). In the future, this aggregated data (i.e. ‘big data’) could be used to generate automated, timely and complete reports on
specific disease outbreaks (i.e. disease surveillance) at the level of the state or indeed at a country level. Figure 5 illustrates the prospective small data to big data opportunity outlined in this study.

Leveraging the data/information/knowledge and big data lessons from the complexities of new product development [62], Figure 5 highlights the integral elements for consideration as part of this mHealth project. This prospective big data scenario is underpinned by the capture, storage and availability of small patient data, individual patient knowledge and subsequent creation of new health knowledge. In order to achieve better insights, it is imperative that the technical infrastructure and efficient and effective processes are in place to support the sophisticated analysis of large volumes of complex health data.

7 CONCLUSION

A number of key outputs are anticipated as a result of this research project. The technical feasibility of a low cost mHealth intervention, the IMPACT app, will be assessed to understand this opportunity for this artefact in terms of better supporting the control of malaria and/or neglected infectious diseases amongst children in resource poor country(s). The project will offer an effective and accessible mobile solution to better support PHC workers in their roles in the Nigerian community. In doing this, IMPACT will provide an ideal and safe test-bed to develop PHC workers medical knowledge and expertise in the use of mHealth technology.

Although small health datasets have been established in low resource setting such as this, we argue that mHealth ‘pilotitis’ is hindering the establishment of big datasets for addressing child mortality and morbidity in developing countries. As outlined in this study, mHealth devices must be scaled up and sustainable to facilitate for mHealth big data scenario. The IMPACT application and complementary project work packages underpin the long term plan to create and leverage an mHealth ‘big data’ scenario in a low resource setting.
Knowing more about the collective, in terms of insights into population health, will consequently support the knowledge needs of healthcare practitioners in terms of the prospective diagnosis and treatment of an individual. If correctly implemented, big data could be used to guarantee public health, determine and implement appropriate treatment paths for patients, support clinical improvement, monitor the safety of healthcare systems, assure managerial control, and promote health system accountability to the public. After all, ‘Data-driven decisions are better decisions— it’s as simple as that; [63].

ACKNOWLEDGEMENT
This work is funded by the Irish Research Council grant no. ‘New Horizons REPRO/2015/116’.

REFERENCES
http://dx.doi.org/10.1089/tmj.2012.0328


[38] Verhoeven, A.A.H., Information-seeking by general practitioners. 1999.


