HEALTHCARE DATA MANAGEMENT CHALLENGES IN DEVELOPING COUNTRIES: A SYSTEMATIC REVIEW

L. van der Merwe¹, W. Bam² and I. de Kock³
Department of Industrial Engineering
Stellenbosch University, South Africa

19072899@sun.ac.za¹, wouterb@sun.ac.za², imkedk@sun.ac.za³

ABSTRACT

There are many healthcare data management challenges in developing countries. These challenges have a major effect on the quality of healthcare service delivery. Therefore, it is very important to determine the most prominent healthcare data management challenges in developing countries to be able to address the right challenges effectively and efficiently. This paper discusses some of the prevalent healthcare data management challenges in developing countries through a systematic literature review. Scopus was used to retrieve literature on these healthcare data management challenges. The initial search was done on 16 April 2019 and yielded a total of 162 articles. The search terms included "data processing", "data management", "data administration", "data handling", "data control", "information management", "healthcare", "Health care", "south africa*", "developing countr*", "challenge*", "problem*" and "issue". After the exclusion process the eventual number of articles was 62. The structured literature review was used to scope all the different healthcare challenges. The challenges were categorised into different data management categories and data management challenges subcategories using Excel. The most prevalent healthcare data management challenges were identified from literature using the Excel scope of challenges. The literature review methodology is explained, then the study selection and characteristics of the results are described, followed by the analysis and discussion of the scope of healthcare data management challenges.

*Corresponding author
1 INTRODUCTION: HEALTHCARE DATA MANAGEMENT IN DEVELOPING COUNTRIES

Good data management is imperative for the effectiveness of any healthcare system. Despite this fact, developing countries struggle with the proper management of healthcare data. The poor management of healthcare data has an adverse effect on the effectiveness of healthcare systems, especially in developing countries.

Data management mainly entails the collection, storage, security and sharing of data gained from diverse sources [1]. According to Evariant*, an enterprise solution in healthcare, data management in healthcare is the basis that enables the holistic views of patients, personalisation of treatments, improved communication and enhancement of health outcomes. To achieve this, data must be aggregated and standardised [3]. After it is collected and aggregated, its accuracy, completeness and consistency must be ensured. Therefore, a data management plan is needed, coupled with the necessary platform to integrate data, manage its quality and utilise it productively.

Data is used in all aspects of the healthcare system. Data is collected, stored and used for patient record keeping, monitoring, diagnosis and treatment. It is also used in other parts of the healthcare system such as tracking medicinal stock levels and patient billing. Healthcare data is located everywhere: clinical and claims systems, Human Resources and financial applications and third-party sources [3]. Therefore, without the management of data, many of these healthcare system components will not be able to function properly.

There are some key healthcare data management challenges that cause the ineffectiveness of the healthcare system. These lacking data management components, discovered from literature, consist of data collection, storage, processing, transmission, retrieval, monitoring, reporting, utilization and analysis challenges. There are also many challenges with ensuring the security and quality of data. The integration of these various components is also very challenging. The various aspects causing the integration challenge are also discussed. Furthermore, the poor governance of data also contributes to the dysfunctional system. Effective and efficient data management is very dependent on technology and infrastructure, but there exist some challenges with technology and infrastructure that impede the effective healthcare data management in developing countries. Data management cost and the available financial support are also aspects contributing to the healthcare data management problem in developing countries. Humans also pose as a contributing factor to the dysfunctionality and lastly, the implementation of new data management systems are also very challenging, answering the question of why these challenges are not addressed sooner to improve data management and thereby improving the delivery of healthcare services.

This paper will discuss the different healthcare data management challenges in developing countries from literature. This enables the visualisation of the scope of challenges in healthcare data management challenges.

2 LITERATURE REVIEW METHODOLOGY

A structured literature review was conducted 16 April 2019. The search followed the process described in this section. This study followed the systematic review process as described by

* www.evariant.com/faq/why-is-healthcare-data-management-important
Pickering and Byrne [4]. Firstly, the electronic search platform and the search terms that were used are described in section 2.1. Secondly, in section 2.2, the exclusion process is described from the initial search results to the eventual number of search results. The required data will be extracted, as described in section 2.3, which will be analysed to draw conclusions from. Some variations to the process are made in section 2.3 as the steps in the process are inherent to how the study was carried out. The variation is explained in section 2.3.

2.1 Search Strategy

The first steps were to define the topic, formulate the research question, identify the keywords and identify and search databases [4]. The topic was defined as “Healthcare data management in developing countries”. The research question was “What are the scope of healthcare data management challenges in developing countries?” The electronic search platform used for this structured literature review is Scopus. It was decided to use only Scopus, because it is the largest abstract and citation database of peer-reviewed literature that includes scientific journals, books and conference proceedings. Scopus consists of an extensive database with good quality and diverse sources which ensures good worldwide research coverage. Scopus developers claim that it is the largest single abstract and indexing database ever built [5]. With Scopus it is possible to enter multiple search requirements with different search operators. This allows Scopus to search for results in a very specific field, while maintaining objectivity with regards to the inclusion of all results in that field. Scopus was developed by Elsevier and the characteristics of both PubMed and Web of Science were combined. This allows enhanced utility for medical literature research and academic needs [6].

The search terms were chosen to cover the research field of healthcare data management challenges in developing countries, and to specifically include South Africa. To ensure the comprehensiveness of the initial Scopus search, different keywords were used for the same concept. The variants of the term “data management” for this study include “data processing”, “data control” and “information management”. These terms were used in inverted commas to return only articles where the words of the terms are used in conjunction. Furthermore, healthcare was searched as one word and as two words enclosed in inverted commas to accommodate articles with the spelling difference. The search terms “developing countr” and “south africa” was used to narrow the search down to only developing countries and South Africa. These search terms were used with an asterisk to include articles where variants of these search terms are used. Lastly, synonyms of challenges were used with asterisks to ensure comprehensive coverage of articles that identifies healthcare data management challenges. These synonyms include problems and issues. Therefore, the relevant literature was retrieved with the following search keywords:

( TITLE-ABS-KEY ( "Data processing" OR "data management" OR "data administration" OR "data handling" OR "data control" OR "information management" ) AND TITLE-ABS-KEY ( healthcare OR "Health care" ) AND TITLE-ABS-KEY ( "south africa" OR "developing countr" ) AND TITLE-ABS-KEY ( challenge* OR problem* OR issue* ) )

2.2 Exclusion Criteria

The next step was to assess the publications to ascertain if it is relevant and whether it should be included [4]. The first exclusion criterion was to exclude articles that were published prior
to 2008. Data management has evolved tremendously over the years and including articles that are too old will result in including irrelevant challenges in this study. This study strives to address relevant healthcare data management challenges of developing countries.

The second exclusion criterion was by document type. Books were excluded from this review. Lastly, the abstracts of the remaining articles were read through to determine whether they are truly relevant to this study field. Articles were excluded during this phase based on whether they clearly mentioned any healthcare data management challenges in the abstract. Sometimes, articles were included in the initial search because the study used data management to address a totally different challenge than what this study addresses. Such articles, for example, were excluded during this phase.

This method ensured that only publications that are truly relevant to healthcare data management challenges in developing countries were included.

2.3 Data Extraction

After it was determined which articles to include in the study, the next step was to develop the structure for a personal database on the topic [4]. After the completion of the article selection process, data was extracted using MS Excel. Excel was used to develop a healthcare data management scope of challenges. The aim of the scope of challenges is to identify all the possible healthcare data management challenges of developing countries and to categorise these challenges into appropriate categories and subcategories. This scope of challenges enables the quantitative analysis of the extracted data. Other relevant data extracted from literature into the Excel sheet include:

- Year of publication
- Author(s)
- Article title
- Country or geographic area the study focused on

Whenever an article mentions a healthcare data management challenge of developing countries, an ‘x’ was marked under the category which that challenge belongs. Articles address multiple challenges and some challenges fall under multiple categories. The categories change and more categories were included as the data is extracted. This caused the scope of challenges to expand. At the end of the data extraction process, the categories were assessed to merge very similar challenges in the different categories. Challenges were classified under different data management components.

Due to the evolving nature of the scope of challenges, steps seven to 10 as described by Pickering and Byrne were automatically carried out in the reading of all the publications. As new challenges were found the structure of the database was updated until the final database was complete.

After the personal database was completed, the rest of the steps were carried out from analysing the findings, writing the discussion and evaluating the results and conclusions for the writing of this article.

2.4 Data Synthesis

The year of publication and the country the article focused on was included in the data extraction to map the literature studied. It gives an understanding of the landscape of the research done on healthcare data management challenges in developing countries. From the
year of publication, the distribution of when the articles on this topic was published, can be determined, and from the country the articles focused on it is possible to determine the distribution of where studies have been done on this topic.

The scope of challenges gives a holistic view of all the healthcare data management challenges that developing countries face. These challenges were categorised according to the different data management components. These data management components can be regarded as the main categories. Challenges were designated to subcategories under these categories. Other main categories are additional to data management components, but also have an impact on data management. The main categories that the challenges were designated to are:

- Poor governance
- Integration challenges
- Data collection challenges
- Data storage challenges
- Data processing challenges
- Data transmission challenges
- Data retrieval challenges
- Data utilization challenges
- Data monitoring challenges
- Data reporting challenges
- Data analysis challenges
- Data quality challenges
- Data security challenges
- Infrastructure and technology challenges
- Cost and financial support challenges
- Human factor challenges
- System implementation challenges

The subcategories under the categories were determined as they surfaced in the literature. When a new subcategory appeared in literature, it was added under the appropriate category. A subcategory, general challenges were, included under most categories if the challenge was mentioned without a specific regard of a subcategory.

3 RESULTS

This section of the paper focuses on the results yielded from the structured literature review. The results from the study selection is discussed and the characteristics of these results are presented.

3.1 Study Selection

The initial Scopus search, using the previously mentioned search terms, amounted to 162 articles. From there the search results were narrowed, excluding articles prior to 2008. This yielded 108 articles. Excluding books resulted in 106 articles. Furthermore, after the abstracts were read to determine the articles’ relevance to healthcare data management in developing countries, the eventual result amounted to 62 articles. Figure 1 illustrates the process of starting at the initial search protocol, to how the search results were narrowed down to the eventual number of search results through the different exclusion steps.
3.2 Study Characteristics

More than half of the articles relevant to this study were published from 2015 to 2019. This shows an increase in interest in this research over recent years. 2019, however, yielded only one article, but the reason for this can be attributed to the fact that the structured literature review was carried out early in 2019. 2015 and 2017 both yielded the most articles with a total of 12 articles each. During 2011 no articles relevant to this study was published. Figure 2 illustrates the number of articles published yearly. This figure indicates when the most and the least of the articles relevant to this study were published.
The countries or geographic area that the articles focus on were also captured. This indicates where the perspectives come from. This study focuses on developing countries, thus all countries included in this study are developing countries. Some articles did not state the country that it focuses on and addressed healthcare data management challenges for developing countries in general. When an article mentioned a country specifically, it was documented. Some articles did not mention a country, but referred to low- and middle-income countries, resource limited countries and sometimes it mentioned a larger geographical area such as Sub-Saharan Africa. As can be seen in Figure 3 the country that was focused on most, is South Africa with a total of 13 articles, followed by India with 10 articles. Four articles focused on Ethiopia and three on Bangladesh. There are many other countries like Tanzania, Uganda, Mexico, etc. included in the study that only one or two articles mentioned. It is interesting to note that 6 of the 13 developing countries mentioned is in Africa and that the country with the most articles, which is South Africa, is also on this continent. The other regions mentioned in articles can be described as:

- **Low- and middle-income countries**: these consider all countries with low or middle incomes.
- **Developing countries**: these consider countries around the globe that are not classified as developed.
- **Resource limited countries**: these are countries that have limited resources. This also includes all countries around the globe.
- **Southern Africa**: this includes Angola, Botswana, Lesotho, Malawi, Mozambique, Namibia, South Africa, Swaziland, Zambia, and Zimbabwe.
- **Sub-Saharan Africa**: the area of the continent of Africa that lies south of the Sahara. It consists of all African countries that are fully or partially located south of the Sahara.
- **African countries**: All countries in the African continent.
4 SCOPE OF HEALTHCARE DATA MANAGEMENT CHALLENGES IN DEVELOPING COUNTRIES

This section describes the healthcare data management challenges in developing countries. First, the healthcare data management challenges in developing countries are quantified through the data retrieved from the systematic literature review that was extracted in Excel on a scope of challenges spreadsheet. After that, the different healthcare data management challenges are explained according to their prevalence ranking.

4.1 Quantitative Analysis of Healthcare Data Management Challenges

The challenges found in the literature were categorised as stated earlier in the paper. The challenges were documented on an Excel spreadsheet as a scope of all the healthcare data management challenges identified from literature. An abstract of the spreadsheet can be seen in Figure 4 showing the article name, author(s), date of publication and how the subcategory challenges of the integration challenges category are scoped. Quantifying the number of times that challenges fall into each category gives an indication of which healthcare data
management challenges are the most prominent. The total number of occurrences recorded for all the categories combined is 1147.
For each category the challenges were further categorised into subcategories. This section will discuss the distribution of the challenges between the different categories and subcategories. Figure 5 illustrates the number of challenges that occurred in each category and Figure 6 illustrates the top 10 subcategories across the different categories.

As can be seen in Figure 5 the six categories with the most challenges are integration, data collection, human factors, data security, technology and infrastructure and data quality challenges. Integration challenges is the highest by far scoring 168 occurrences. Data collection challenges had the second-most occurrences with 127 occurrences. The category with the third most challenges was human factors with 124 occurrences, followed by data security with 107. Technology and infrastructure had 91 occurrences and data quality had 89 occurrences.

**Figure 5: Number of challenges occurred in different categories**

It is clear from Figure 6 that data not shared between different users is the subcategory with the most occurrences. This is followed by fragmented systems. Lack of digital data storage and missing or partial data is tied in third place. Lack of infrastructure is the next subcategory with the most occurrences. From there on it is no digital data capturing, patient privacy challenges, inaccurate data, general security challenges and network unavailability.

**Figure 6: Top 10 most occurring subcategory challenges**
It is interesting to note that the category data transmission did not have a high number of occurrences, but one of its subcategories, data not shared with different users, is the subcategory with the highest score. Out of the top ten data management subcategory challenges, categories data quality, data security and technology and infrastructure featured twice, while the other categories featured only once on the top 10. It is also interesting to note that integration challenges, the category with the most occurrences, also featured only once on the Top 10 subcategory list. All these findings convey the intricacy of the healthcare data management challenges in developing countries.

4.2 Discussion of Healthcare Data Management Challenges

This section describes the different categories challenges as found in literature. The discussion is presented according to the ranking of the highest scoring categories of healthcare data management challenges. The categories are presented in the order of highest scoring category to the lowest scoring category. Each category is introduced, followed by a discussion from literature of the different subcategory challenges of that category.

4.2.1 Integration challenges

In section 4.1 it was found that integration challenges was the category with the most occurrences. Challenges that contribute to the complication of integrating healthcare data management systems were designated to this category.

There are many factors contributing to integration challenges. There are many standalone information systems [7] and different systems have heterogeneous forms that make them even more difficult to integrate [8]-[10]. Often the necessary standardization and interoperability between systems are missing [11], [12]. Standardization and interoperability will enable different systems to share data among them.

Data is complex [13], and to manage data effectively it needs to be organized and aggregated. Aggregated data can help to make better decisions based on holistic views of data [10], [14].

Paper based data management systems make integration basically impossible [15], [16]. Incompatible technologies further complicate integration [15], [17].

Another issue is that even when it is possible to integrate systems and share data, new security issues arise [18]. Integration or sharing data should not compromise the security of the data.

As stated in section 4.1 it was found that fragmented systems is the second biggest challenge on the top 10. It is also the integration challenge with the most occurrences compared to other subcategories. It was found that in many developing countries fragmented systems was due to different subsystems operating separately and oftentimes even vertically [19]. Integration between governmental bodies or within the Ministry of Health is also missing [15]. Fragmented and heterogeneous systems traps data needed for proper decision making [20] and Sharifi et al. [21] stated that fragmented and inaccessible clinical information have a negative effect on healthcare quality and cost.

4.2.2 Data collection challenges

This category had the second-most challenges documented. The data collection challenges subcategories with the most occurrences are no digital data capturing, errors with collection, inefficient collection processes and methods, lack of proper entry forms, time constraint, duplication and different healthcare facilities collecting different data. It is also important to note that no digital data capturing was sixth on the top 10 subcategory list.
No digital data capturing mostly entails that data is captured on paper and documentation are done on paper as well [22], [23]. Sometimes there are systems that convert the paper-based data into digital format, but usually there is a lag before that is done [24]. Paper-based capturing makes it very difficult to exchange and access information and to monitor patient’s progress [7]. Masana and Muriithi [7] also stated that the lag before data is digitalized prevents real-time data accessibility.

Oftentimes handwritten data is illegible affecting the quality of the data [25]. Data capturing errors also affects the quality of data. These errors include partial data collection or entry mistakes by the capturer [9], [26], [27].

There is also a concern about the suitability of data entry forms [21]. Sometimes forms do not allow the entry of relevant data in the way that it is structured [28]. There exists a need to standardise these forms to ensure all relevant data is captured [25].

There are different data entry points. This causes the same data to be captured multiple times. This data is aggregated only later. Data is therefore duplicated by different data capturers [29], [30].

It has also been found that different healthcare facilities collect various types of data [31]. Different sub-districts and many healthcare facilities have different interpretations of how data should be captured and managed [32].

Another data collection challenge is the limited time available to collect data [15], [31], [33]. In developing countries, it takes long to collect data and often data capturers do not have enough time to collect the necessary data. Overworked healthcare personnel has multiple responsibilities and does not have the time to collect data [33], especially when paper-based data capturing is used which is very inefficient and not time effective [34].

4.2.3 Human factors

According to the analysis of section 4.1 human factors is the third biggest challenge category. It is interesting to note that none of its subcategories are in the top 10, but still it is the third biggest healthcare data management challenge category. All human related data management challenges were designated to this category. The five most prominent human related challenges are the low skill level of staff, the lack of training, the lack of staff, digital illiteracy and lack of participation.

Many healthcare workers do not have the required skill to do the data management tasks they need to do. They make mistakes with data collection, do not know how to use the routine health information system, and lack data security, data management and data analysis skills [35].

There also exists a lack of training [21]. A lack of training means that healthcare workers are not trained before they must do their work and it also means that their skills will not improve in the future, because there are no training opportunities. Even in cases where there were training, there are no ongoing data management training available [27].

Digital literacy is another major challenge [18], [19], [35]. For healthcare workers to manage data, they need to be able to use the necessary information and communication technologies. Healthcare workers do not have the expertise to operate equipment such as tablets, computers, and smart phones [36].
Not only is the skill level or the digital literacy of the staff low, there is also a major shortage of staff [37]. This includes a shortage of data capturers [25]. The shortage of staff influences data quality [32].

Another factor is that humans do not participate in the data management system as they should. Braa and Sahay [30] found that sometimes health managers lack faith in the health management information system. A lack of collaboration of doctors and health personnel can lead to poor data quality [15].

4.2.4 Data security challenges

Data security had the fourth most challenges identified. Any challenge regarding the protection of data or protection against data breaches is included in this category. The subcategory with the most occurrences is patient privacy. General data management challenges also has just as many occurrences and both these subcategories shares the seventh place on the top 10 subcategory challenges list. Authorization, confidentiality and data integrity is the other major data security challenges.

Medical systems are more connected and networked. It is necessary to have identifiable health data in health data repositories across these systems for data accessibility and sharing, but the increase in connectivity and identifiable health data causes an increase in patient privacy risks and security breaches [9], [20]. Electronic documents have many privacy and security risks because they are accessed and transferred easily [15]. Unmanaged servers endanger data, bandwidth and other devices on the network [38].

Preservation of anonymity and security of patient records are major concerns, but in some developing countries there are no privacy laws that protect identifiable data [12], [36]. There also exists a need for well-coordinated regulatory frameworks for proper governance of the privacy and security of patient health information [19].

Another major concern is unauthorized access [18]. Authorization services include policy management, role management and role based access control [39], but sometimes even authorized users can use data maliciously [10].

4.2.5 Technology and infrastructure challenges

This category is the fifth biggest challenge category. It involves all data management challenges related to the technologies and infrastructures needed for data management in healthcare. The highest scoring subcategories were infrastructure challenges, network availability, software issues and power supply problems.

The infrastructure required for data management is lacking [11], [28], [31], [40]. Some of these infrastructures needed for data management include information technology infrastructure [39], communication infrastructure [17], and technological infrastructure to allow appropriate information storage and sharing [12]. New technologies cannot be introduced to healthcare data management systems because the current infrastructure is unable to support it [14].

Network availability is also a big problem. Ganiga et al. [39] found that there is a great need for high speed internet connectivity, but often the internet connection is unreliable. The unreliability of internet forces lower level subsystems to be completely paper-based [41]. Internet signal strength is often weak [34] and bandwidth costs are high, inhibiting the health systems’ effectiveness [19].
Power supply has been found to be epileptic [18]. In many resource limited areas healthcare facilities do not have a reliable electricity supply to support data management activities such as electronic data storage [14], [38].

4.2.6 Data quality

Data quality is the sixth biggest data management challenge category and it has two subcategories in the top 10 subcategory challenges list. Missing data is tied in third place and inaccurate data shares the seventh place. These are the two main data quality subcategories, but poor data quality in general was also often mentioned [11], [15], [42]. Other data quality issues were data duplication, unstructured data and data discrepancy or inconsistency.

Quality data is important for many healthcare functions. Missing and inaccurate data impedes functions such as record linking, proper diagnosis and analysis, patient monitoring and clinical and public health decision-making [43].

Causes of missing or inaccurate data are nurses that simply do not write out all patient data [21], doctors that do not write diagnosis, but only symptoms and prescriptions [14] and errors made by less qualified staff [20].

4.2.7 Data storage challenges

This challenge category has the seventh most challenges documented. Challenges is designated to this category if it posed problems to the storage of data. The highest data storage subcategory, lack of digital data storage, is ranked second on the top 10 list for subcategory challenges. Loss of data is the second biggest data storage concern. Other data storage subcategories include unstructured storage, inefficient storage, reliability of storage and the lack of storage infrastructure.

Many healthcare systems in developing countries are still managing data in traditional paper-based systems [29], [39], [44]. This makes it difficult to access patient information [45] and gives rise to problems in data aggregation, transmission and analysis [29]. These paper-based stored healthcare data is also located at different geographical locations. This limits the adoption of a system-wide approach to healthcare management [15]. For some manual paper-based systems health statistics that are recorded in log books are sent to regional offices for data capturing of metrics into a centralized database, but these log books are sent infrequently [46].

Loss of data is also a common data storage challenge. Data that is needed for patient care or program management gets lost [47]. In paper-based data storage the loss of health record books are common [48] and in some paper-based systems the record room only maintains five years of patients. This results in the loss of continuous patient data [39]. Chimbari [49] found that datasets are collected at high costs, but are not analysed and gets lost over time. Digital data storage is prone to loss of data in the case of hardware or software failure. Efficient back-up is needed to prevent loss of data [38].

4.2.8 Cost and financial support challenges

This category is the eighth biggest healthcare data management challenge category. All the different cost aspects of healthcare data management are categorised under this category. Implementation cost was mentioned the most in literature. Other data management cost aspects mentioned include infrastructure costs, technology costs, systems costs, data storage costs and training costs.
Cost for implementing healthcare data management systems is the biggest cost challenge and is the main reason why developing countries struggle to adopt digital healthcare data storage [47]. Cost is also the reason why existing digital medical records are not integrated into Information Technology [10]. Additional costs of acquiring, installing and maintaining equipment is needed for integration [12].

Technology expensiveness is also a great concern. Healthcare companies have developed technologies to improve healthcare data management, but the problem is they are too expensive for developing countries [50]. Healthcare data management software and hardware are expensive [21], as well as equipment to monitor patients at healthcare facilities [34].

### 4.2.9 Implementation of systems challenges

Implementation of systems is ranked ninth from the different challenge categories. Challenges regarding the implementation of systems is designated to this category. Its main subcategories are resistance to change, training, technical infrastructure available to support implementation and the existing culture. The cost of implementing systems is the greatest implementation challenge, but that is categorised under the category cost and financial support.

Healthcare systems users sometimes have negative perceptions about the usefulness and the associated threats regarding new systems. This causes a resistance to adopt the new systems [15]. Threats regarding privacy and security make organizations reluctant to adopt new systems [19]. One example of resistance to change due to perceived usefulness is that reports from radiology is handwritten. Staff objected to typing it in electronically, because it takes too long and workloads are too high [14]. Patients also resist change. Warkulwiz et al. [36] found that patients did not want their data to be entered electronically, and preferred paper-based systems instead.

The availability of technical infrastructure is very important for systems implementation [11], but poor infrastructure makes this very difficult [38]. To ensure adoption of widespread implementation, current infrastructure available should be used to support the implementation of systems, rather than having to develop new and sophisticated technologies to support implementation [17].

Another factor is that the need for training makes implementation difficult [25]. When new systems are implemented, it involves training which is very expensive [21]. After implementation, ongoing training of personnel is also needed [27].

### 4.2.10 Governance challenges

According to section 4.1, this category is ranked tenth. All challenges that relates to the governance of healthcare data management is categorised in this category. Its main subcategories are policies, legislation and standards. Other subcategories include regulations, frameworks and leadership.

Policies are needed at a national level to ensure coherence [12], but there exists a lack of adequate policies and procedures [31]. Some common policies that are missing are policies regarding security [15], Information and Communication Technologies [51] and well-defined access policies to ensure authentic users can access data [45]. Policies for the use of information systems are also important, for it grants access to the systems to authorized users anywhere and at any time [16]. Kaposhi et al. [32] recommended changes in knowledge translation, data verification, programme management and standardization policies.
Having the necessary standards are also very important to have a system-wide approach to patient healthcare management [15]. It is found that many developing countries do not have informational and care standards and have limited existing regulations [14]. Digital health record standards have also been adopted only recently [52].

It was also found that there is an absence of legislation regarding data management [12]. Health data is collected from public and private hospitals, but because the required legislation is not in place, this data cannot be shared or used [15]. Turan and Palvia [15] also found that there are no legislation safeguarding personal health information.

4.2.11 Data transmission challenges

Data transmission was quite a low scoring challenge category, but one of its subcategories, data sharing with different users, had the highest number of occurrences of all the subcategories. Challenges regarding the transmission of data from one place to another by whatever means were included in this category. Other data transmission challenges are difficulties with disseminating data to patients, data transfer latency, unreliable network availability for transmission and transmission errors. In some cases, data was transferred in paper form.

A key challenge in developing countries is to make healthcare data accessible from rural to urban [39]. Paper based systems only allows data to be accessed from one place [53]. Therefore, there exists a need to digitalise data for data accessibility and data sharing, but if digital systems are not linked or integrated, data accessibility and exchange to different users will still pose as a major challenge [7].

Some clinical information technologies do not allow data sharing between clinicians, labs, hospitals, pharmacies and patients [54] and data is not shared between different levels of healthcare either [37]. Different healthcare facilities use their own systems with their own localized data networks. This inhibits data sharing and the adoption to a system wide approach to patient healthcare management [15]. Data is isolated in silos which impedes data sharing between care providers [10].

Legislative frameworks and an unwillingness of companies to share data are yet some other factors contributing to the lack of data sharing [19].

4.2.12 Data retrieval challenges

Data retrieval ranked twelfth. Its subcategories consist of challenges regarding the retrieval and accessibility of data. The main subcategories are historical data retrieval, real-time data retrieval, the lack of remote data accessibility and timely data accessibility.

Paper-based and manually filed medical records are difficult to access and impedes good service to patients [55]. Electronic databases can make retrieval of stored data easier [47]. Data retrieval is important to make accurate clinical decisions [10] and inaccessible clinical information has a negative effect on quality if healthcare [21].

Many developing countries use systems that do not allow timely data accessibility. For instance, healthcare systems that are divided into different categories and operating on different system makes access to patient data very difficult [18]. Other example is processes that are error-prone and time consuming causes a delay in data accessibility [53].
Recently, there has been an increase in the need to access data from remote locations [22]. This is needed so that doctors can view patient health records from anywhere and give advice and treatment [45].

5 CONCLUSION

This structured literature review gives a comprehensive overview of the scope of healthcare data management challenges in developing countries. It also quantifies which challenges occurred the most in literature. The quantification of challenges gives an indication of which challenges are the most prominent. Addressing the most prominent healthcare data management challenges will improve healthcare service delivery the most.

It is found that there has been an increase in the interest of healthcare data management challenges over recent years as more than half of the articles are from 2015 onwards. The country that articles focused on most is South Africa followed by India.

The six major healthcare data management challenges categories identified by this study is integration challenges, data collection challenges, human factors, data security challenges, technology and infrastructure challenges and data quality challenges.

The top 10 healthcare data management challenges subcategories are data not shared with different data users, fragmented systems, the lack of digital data storage, missing or partial data, the lack of infrastructure, no digital data collection, patient privacy issues, inaccurate data, general security challenges and network availability challenges.

This paper gives an idea to the scope and complexity of the healthcare data management challenges in developing countries. To solve healthcare data management challenges in developing countries, the challenges identified in this study are the major challenges that should be addressed for effective and efficient improvement of healthcare service delivery.

6 REFERENCES


