Comparing strategies to integrate health information systems following a data warehouse approach in four countries

Edem K. Kossi ¹, Johan Ivar Sæbo ¹, Ola H. Titlestad ¹, Romain R. Tohour ¹, Jørn Braa ¹
¹Department of Informatics, University of Oslo
Address for correspondence:
Department of Informatics
P.O. Box 1080, Blindern
0316 Oslo
Norway
Email: ekossi@gmail.com

Abstract:
This paper addresses one of the major obstacles of reaching the Millennium Development Goals: inefficient and unreliable information systems. Leading international organizations have called for integrated data warehouses as one of the solutions, but this remains hard to achieve. This paper presents four country cases of standardizing and integrating health data which are all following what is here termed a data warehouse approach; data from across different health programs are organized in one database framework – or data warehouse. In all countries, fragmentation of health information in different partly overlapping sub-systems run by different vertical health programs represented a major problem for the efficient use of health information. While South Africa developed a new integrated system in addition to the existing fragmented sub-systems, Zanzibar, Sierra Leone and Botswana, all aimed to encompass all or most of the data from the existing systems. The three latter countries all followed slightly different approaches; more or less incremental in the approach to standardizing health data, and more or less strict in whether to include “all” data, and whether to solve all inconsistencies between the various data sets included early on. The four cases demonstrate that integration is as much, and maybe more, about aligning organisational-political actors as it is about technical solutions. The technical solutions are however important in aligning these actors and in enabling integration. We argue that “attractors”, technical solutions or standards that achieve a certain level of success and enable the building of momentum, are important in aligning the various political actors. In turn, these attractors need to evolve within the changing context of a growing health information system in order to achieve the scale needed to address the Millennium Development Goals with full force.

Index Terms—Integration, standardization, health information systems, data warehouse
1. INTRODUCTION

In 2010 ten years have passed since world leaders agreed to eight Millennium Development Goals (MDG), that are to be achieved by 2015. Three of these goals are health-related, namely goal 4; reduce child mortality, goal 5; improve maternal health, and goal 6; combat HIV/AIDS, malaria, and other diseases. To reach these goals, managers at all levels need to base their decisions on reliable, timely information for the resources spent to have the most impact. International organizations need to monitor progress and adjust policies and funding priorities, national health administrations need to strengthen the health services by identifying bottlenecks and emerging trends, and distribute health personnel and resources, and local managers need to respond to the needs of their communities. All this depends on well-functioning Health Information Systems (HIS), one of the six pillars of a health system (WHO 2005, 2007). At the same time, there has been a realization that many of the problems with HIS in developing countries stem from the practice of installing specific, narrow information subsystems, often covering limited information needs. Without a comprehensive strategy, these subsystems evolve into a range of fragmented, expensive, often overlapping and inefficient systems, in turn leading to less use of information for health improvement. With the number of global health initiatives focusing on specific health priorities still rising (IHP+, 2010), this fragmentation of efforts, and their HIS, seriously undermines the work towards the MDGs.

This has led to a several international organizations that specifically address these challenges with new strategies. The International Health Partnership (IHP+, established in 2007), aims to harmonize aid agencies, and align international efforts with national systems. The Partnership in Statistics for Development in the 21st Century (PARIS21, established in 1999) helps countries adapt evidence-based policy making and implementation to reach the MDGs. The Health Metrics Network (HMN, established in 2005) was created to strengthen country HIS and information use, through integration of subsystems and data sources. Other international donors, the initial source of many of the fragmented systems, have as a result adapted a strategy of improving integrated national HIS rather than implementing new, narrow, specific subsystems.

This process of integration, however, is less than trivial. Integration takes place at several levels at the same time, from the purely technical to the organizational, and the interplay of these levels throughout the process very much determines the appropriate strategy. A central part of the HMN technical strategy to overcome the problems of fragmentation and lack of data standards is the establishing of data warehouses, or data repositories, for the storage of essential data and indicators from multiple data sources. This data repository will then provide shared access to information across the various health programs and user groups (HMN 2008). In this article we describe and discuss different varieties of such a “data warehouse approach” as they are applied in four countries. Botswana, Zanzibar and Sierra Leone represent relatively recent developments (since 2005/07), whereas in South Africa the process has been ongoing since 1995. While in
Botswana, Zanzibar and Sierra Leone they have followed various strategies to include all data routinely collected from the health services by different health programs in one data warehouse system, in South Africa they started out with a minimum essential data set strategy (i.e. only a sub-set of data reported), which has gradually evolved to include more data through a longitudinal standardisation process.

Integration as demonstrated in the four cases, however, is not primarily a technical process, but rather a political one. Nevertheless, the technical and the non technical dimensions sway one another. In all cases the overarching problem of fragmentation was caused by different organizational actors working independently. Providing a technical solution to integrate two different systems is relatively easy, but the politics of negotiating this among the actors involved is key to attain any use, and hence sustainability. However, in turn, the potential of the technical solution to get actors understand it and perceive their interest is not to be neglected. Our cases show that mindfully using ICTs can align different actors towards a common goal; the development of strong and use-oriented health information systems so necessary to reach the MDGs.

This article will proceed as follows. First we present the methods applied. Then we present relevant literature on the issues of integration and standardization in developing countries, before presenting the four cases, the discussion and finally the conclusion.

2. METHODOLOGY

2.1 Action research in four countries
This article draws on case material from direct involvement and experiences in the development and assessment of health information systems in South Africa, Zanzibar, Sierra Leone and Botswana. The authors are all involved in the one or more of the four countries, as well in the broader network of the Health Information Systems Programme (HISP) (Braa et al. 2004; see also www.hisp.info). HISP started in South Africa in the 90’s as a conscious effort to adapt and apply central features of the Scandinavian action research tradition in IS development, such as user participation, evolutionary approaches and prototyping, to the context of Africa (Braa, Hedberg, 2002). Action research aims at generating new knowledge through taking part in the full cycle of design, development, implementation, use, and evaluation and analysis together with all involved parties before the interventions are adjusted accordingly, and the cycle starts all over again (Susman, Evered, 1978). In each of the four countries the authors have been engaged in repeated prototyping cycles of the DHIS software, system development such as revision of data sets (i.e. what data to collect and process) and capacity building, and exploration of ways to access, present and analyse the information in the system.

The authors have through this immersion in the system reform processes played various roles. The work spans from the purely technical domains (software development) to the political (strategy meetings with directors, ministers), but it has been the in the role as interpreters and mediators between these various domains that we have been able to address the very complex task of integration from many angles. In all countries, one or more of the authors have been involved over several years, both through many on-site
visits (length of stay varies from a week to a year, with total length of over 6 months in all countries), and also through continuous communication via email and telephone.

The selection of countries came naturally as South Africa was the place the HISP started, and the subsequent involvement in other countries came about due to specific projects funded by external partners, partly based on the proven success in South Africa. Following the increased focus on health information systems globally, increasingly international donors and agencies play a role in the circulation of best practices, also in the countries included in this study; the European Union funded the project in Botswana, DANIDA (the Danish aid agency) initiated the one in Zanzibar, and WHO and Health Metrics Network initiated the project in Sierra Leone. The authors have also been involved in similar processes in other countries, but the four presented here were found to best highlight the distinct approaches to integration. They are also countries were the authors have been the most involved, and would thus provide richer material to analyse.

2.2 Linking global and local efforts
The research and lessons in South Africa, where HISP was initiated in 1995, forms the basis for the research and interventions in the three other countries. The overall HIS development approach first used in South Africa, including the focus on the local district level, data standardisation based on essential information needs, and local use of own data, has been tried adapted to the other country contexts. As the context differs considerably between the countries, a major research challenge has been to ensure a framework for cross country comparative and collaborative learning and intervention. The methodological approach to address this cross country challenge is based on a “network of action” strategy (Braa et al. 2004), where, first, a mutual beneficial “win-win” collaboration is sought established between action projects, which then is fostered through, second, the circulation of people, best practices and lessons, and software and other materials. The PhD program at the University of Oslo is including 15-20 African students and is facilitating comparative action research within and between countries and thus enables the needed “circulation of people” and sharing of experience. Three of the authors of this article are part of this program. Shared development and application of the DHIS software, which is used in all four countries, make up an important generator of “win-win” collaboration between the countries, while at the same time providing a concrete vehicle of sharing material and experiences.

2.3 Data analysis process
The projects described in this paper have been ongoing for several years. Through the active involvement in the work in these (and other) countries, we have naturally gone through a maturation process in our approaches to integration. The development of the four models of integration discussed in this paper has come gradually, and coincides also with a change in climate towards integrated approaches from international donors, often a key source of funding for health systems in poor countries. Lastly, the continuous development of the software used has led to more flexibility in terms of leveraging the software for the different integration strategies. The data analysis stems from the repeated evaluation cycles of action research, namely asking the questions what worked, why, or why not, which subsequently lead to a different approach in the "next" country. Over the
years, this led to the distinct strategies explained in this paper, and the confirmation that there is no "one size fits all".

3. CONCEPTUAL FRAMEWORK

3.1 Heath Information Systems
National Health Information Systems (HIS) in developing countries are often plagued by fragmentation, where vertical, disease-specific programs run their own information-gathering systems (Stansfield et al. 2008). A typical situation in a developing country is that; 1) there is a national HIS run by the Ministry of Health aiming at covering data from across the various health services and disease specific health programs (e.g. HIV/AIDS, Tuberculosis, Maternal & Child Health, Disease Surveillance, Vaccination Programme (called EPI)); 2) the general health services and the national HIS are relatively weak with insufficient resources; 3) while many of the vertical programmes are comparatively rich as they are funded from international sources (e.g. HIV/AIDS), and these programs need to provide their funders with quality information on their activities and achievement in order to maintain funding. As a consequence, 4) not being satisfied with the data from the national HIS, the vertical programs develop their own sub-HIS with international funding (Braa et al, 2007).

There is often lack of data harmonisation across these sub-systems which both results in duplication of data collection and reporting requirements among health workers at community and facility-level, as well as extremely difficult conditions for health managers at all levels to get the full picture of the health status of their population. This can be illustrated by the challenging task for a manager to monitor and make evidence-based interventions on MDG 4, to reduce child mortality, which would involve the need to get data from many different programs and sub-systems of the HIS; the reproductive health program for still births and neonatal deaths, the immunisation program during the child's first year, from the nutrition program, and for deaths due to major diseases they would need to get data from Malaria, Tuberculosis and HIV programs, and most likely from the program for Integrated Management of Childhood Illness (IMCI). To add to the complexity these data will normally be available in a range of different formats, from paper reports to a mix of different electronic outputs, and another key challenge for the managers is the struggle to get data from the private facilities as well as from the community, in order to get a more complete picture of all the relevant health events in their area.

The involvement of multiple funders and organisational actors make integration highly political. HIS integration is thus a non-trivial process, involving more than just integration at a technical level. Sahay, Monteiro and Aanestad (2008) bring the attention to the political aspects of integration, and claim that, with a few notable exceptions, literature on the topic seems to be too positive and too technology-oriented. For instance, integration would also include aligning the efforts of several ministries (Sæbø, Braa, & Chandna, 2007), a challenge of institutional nature.

3.2 Integration and interoperability
Despite its common use and perceived attractiveness, the qualifier ‘integration’ or ‘integrated’ in the context of information systems remains notoriously slippery as “individuals often have a different understanding of the meaning of the word” (Gulledge, 2006, p. 5), or that it can “mean all things to all people” (Wainwright, Waring, 2004, p 229). During a one week online forum organised by the RHINO network on “integrated HIS” in July 2009, it became evident that also in the global health HIS domain there are different understandings of the term integration. While the more technical oriented participants focusing on electronic medical record (EMR) systems, argued that integration is not possible because it means monolithic rigid systems and that open standards and interoperability is the way to go, the public health participants tended to argue that integration is best understood as efforts to harmonise and bring together the various vertical disease specific reporting systems and sub-HIS. The HMN framework (HMN 2008) is in fact combining these two perspectives by arguing for a central data repository that integrates the different data sources to one harmonised system, and the key mechanism for doing so is to apply interoperable solutions between the various sub-systems and the integrated data repository itself.

The dominant perspective on integration in information systems research has largely been prescriptive in character and technical in orientation (Grimson et al., 2000). Strategies and approaches to integration vary (Hasselbring, 2000) and include technical solutions at different levels and forms of integration frameworks or mechanisms are identified such as data base schema integration, middleware software platforms (Corba, COM2), application or service architectures (service oriented architecture, SOA, .NET).

More recently, however, the complexity facing integration within and across organisations have been brought more into the foreground. Singletary (2004) surveyed practitioners’ perceptions of the downsides to integration including lock-in with vendors, costs, and project risks. Wainwright and Waring (2004) have identified four domains of integration, namely technical, systems, strategy, and organizational, and argue the importance of including the organisational domain in the integration strategy, which “is rarely acknowledged in the IS/IT literature” (ibid. page 342).

3.3 Understanding the integration process - mechanisms for harmonisation and alignment
We have pointed out above how the literature distinguishes integration as belonging to various levels, ranging from the technical to the organisational, or the political. What is missing, however, is a deeper understanding of the relationship and interaction between the different domains or levels of integration. In order to address this issue we will adapt an integrative framework developed for managing knowledge across boundaries by Carlile (2004). The framework includes three progressively complex borders; syntactic, semantic and pragmatic, and three correspondingly progressively complex processes; transfer, translation and transformation. Communication, sharing and agreeing across these increasingly complex borders are about developing appropriate boundary objects that match the needs of the borders (Carlile 2002, Star and Griesemer 1989), an adequate common lexicon, or standards, at the syntactic border, shared meaning at the semantic
border and aligning interests at the pragmatic, political or organisational border. A forth element of the proposed framework is the iterative cycling through the levels in order to step by step develop common understanding and alignment of interests. Repetitions are needed because consequences cannot be seen and addressed with one try, and new agreements and changes are developed through an iterative and evolutionary process. We draw on this concept in trying to understand the standardisation processes needed to achieve HIS integration, and standards tend to be evolving through negotiations between actors facilitated through ongoing efforts to develop standards in practice, rather than through early agreements (Braa et al, 2007, Egyedi, 2002, Hanseth, Monteiro & Hatling, 1996).

In the form of shared methods or technological artefacts the boundary objects play an important role in providing the capacity to negotiate interests (Star and Griesemer 1989), and may "facilitate the transforming of specialized knowledge into novel jointly produced knowledge that transcends each community’s local interests" (Barret et al 2007). Barret et al (2007) go further in trying to understand how the boundary object (in this case a technological artefact) not only helps to translate meaning, but also plays an important role in reshaping boundary relations between occupational groups in an organisational setting both positively and negatively, and draws on the notion of boundaries as dynamic and relational (Abbott 1995).

Eoyang's (1996) work from complexity science on the concept of "attractors" is helpful in understanding the role of technological artefact in enabling organisational change in our case studies. Eoyang argues that “attractors”, e.g. successful pieces of software, play a key role in adapting the various actors to common practices within a complex system by creating momentum through shared practices. Building on this concept again, Braa et al (2007), argue that attractors are crucial in the evolution of new standards generally, and in the complex health system context particularly. An important part of HIS integration at country level is to get the various actors to agree on shared standards. Attractors, such as successful examples of software or data standards, may be important in convincing, or aligning through practice, the various actors in health care (ibid.).

Having worked extensively with national HIS in South Africa, Braa and Hedberg argues that inscribing flexible standards in the software used for an HIS could allow a “hierarchy of standards” (Braa & Hedberg, 2002). For health management, there are different information needs depending on the hierarchical distance to the patient. Doctors would need detailed information on patients, districts would need community information to supplement their core health indicators, while national levels would only need a few health indicators to prepare a budget and take strategic decisions. Typically, the amount of information required would diminish as one went higher in the hierarchy, that is, closer to the national level. Scaling of national standards to lower levels had to be complemented with the ability of lower levels to define their own, additional, information standards.

Drawing on the concept of inscriptions (Hanseth & Monteiro, 1997), we can see that certain solutions prescribe certain behaviour, and that inscribing software with incentives
to integrate can be a fruitful approach. If attractors; successful standards or artefacts that offer a gravitational field among seemingly disparate actors (Braa et al. 2007) embedding such inscriptions can be created, a powerful tool for integration has been developed.

4. CASE STUDIES
In this section we present a series of four case studies from South Africa, Botswana, Zanzibar, and Sierra Leone, all related to the global Health Information Systems Programme (HISP), and more specifically to the implementation of the DHIS software. First we provide a general background on the DHIS’ role in supporting an integrated HIS approach.

4.1. The DHIS software - flexible design to enable integration
Very often a computerised HIS is a direct computerisation of the various forms and reporting structures that exists in the paper-based HIS. A result of such an approach is often lack of flexibility to modify the systems when the paper-based equivalent changes or to incorporate new needs like integrating additional programmes or additional indicators. Such approaches, which strengthen the existing structures of fragmentation, are typically found in the various health programmes where focus is often on reporting huge amounts of detailed data upwards in the hierarchy to the national level and to donor agencies.

To integrate information at the district level and to provide flexibility to meet the rapidly changing requirements of the health care domain, the DHIS design strategy is based on a more flexible model that goes beyond the typical “computerise the forms”-approach. In stead of computerising the form as a whole, the DHIS is computerising every data item or data element in the form independently. By breaking up the form into more fine-grained building blocks of data, the software also breaks up the fragmentation the forms represent and provides the possibility of manipulating and presenting data across health programs and forms, a prerequisite for integrated data analysis. This fine-grained or atomic design enables the user to define calculated indicators (combining data elements into formulas), custom reports, tables and graphs for analysis that are completely independent of the collection forms and can in stead be understood as user-defined assemblies or reconstructions of the data (elements) that was captured using the forms. This flexibility makes DHIS a more suited tool for public health management than the more typical computerised reporting tools that exist in the various fragmented departments of the health system. The DHIS allows e.g. a district manager to combine all relevant data on child mortality (MDG4), originally from multiple sources and sub-systems, into one common report that can be used for routine monitoring of interventions related to this specific target.

4.2. South Africa
In South Africa the HIS reform process started in the 90’s after the abolishment of apartheid and was part of the ANC Reconstruction and Development Plan (ANC 1994a, Braa & Hedberg 2002). During apartheid the health services were extremely fragmented and inequitable and health services and geographical areas were separated according to
race and the system of “homelands”. Until May 1994 there were 14 departments of health at the central level; the general National one, separate “white”, Coloured” and “Asians” administrations, and 10 “black” homeland administrations. As a result, there were no health data standards and a multitude of forms for reporting data were in use. The new national health plan (ANC 1994b) made it very clear that the health system needed to be reconstructed based on equity in health services provision to ensure the inclusion of those who had suffered during apartheid. In order to monitor to what extent this goal was properly addressed and to pinpoint communities in particular need, access to good quality essential health data, integrated across services and geographical areas, was seen as an absolute requirement.

HISP started in three pilot districts in Western Cape Province in 1995 with the aim of developing a district-based integrated information system. It turned out to be difficult to get all the actors to agree on common standards for reporting. After an “endless” array of workshops, the breakthrough came with the development of the DHIS in the conjunction with the testing of a first version of a minimum integrated data set in the pilot districts in 1997/98. The Province agreed to test it further and the success was rather instant; for the first time health managers had data available at their own desktop computer without having to wait for annual reports or enquire for data at various head offices. Having seen the success in Western Cape, the neighbouring Eastern Cape province also wanted to go for the DHIS software. In Eastern Cape, a project funded by USAID had developed and implemented a minimum data set in all the health facilities in the province, but had serious problems with the software they were using. They contracted HISP and implemented the DHIS successfully at the end of 1998 and could provide full coverage data for the province early 1999. Seeing the success in two provinces, other provinces also wanted to join the process and by 2001 the DHIS and data set approaches were endorsed by the National Health Information committee of South Africa (NHISSA) as a national standard. Since 2001 South Africa has established national standard essential data and indicator sets which all provinces are required to report. In addition to this, and given the federal institution of South Africa, each province is free to collect and manage the data and indicators they want. The principle of hierarchy of standards, enabled by the DHIS, was important in creating agreement among the provinces and other actors. According to this principle, each organisational unit in the health hierarchy is free to collect the data they want and thereby pursue their own interests, as long as they adhere to the standards of the level above. Each province could then continue to collect their own data as long as they reported the required standards to the national level. There are large differences between the data sets collected by the different provinces, but they all include the core national standards. The national core standards have gradually been expanded to more use areas since 2001.

This standard reporting format was in the beginning an addition and “on top” of all other reporting forms. Through its relatively instant success in terms of data quality and easy and immediate access, more provinces joined, but also, as the momentum grew, gradually vertical programs included their basic data requirements within the national minimum data set, which later changed name to “essential data sets”. The immunization program, for example, realized that the quality of their data that were included in the minimum
report was higher than in their own data reporting system, and subsequently “gave in” and included their vaccine reporting in the now national essential data set.

There are several electronic record systems in South Africa, or computerized transaction systems as seen related to the data warehouse debate. For the Tuberculosis electronic register and from some Anti retroviral Treatment electronic registries for AIDS patient, “extract, transform and load” functionality is established with the DHIS, the standard more technically advanced “Western” data warehouse model is gradually being adopted.

4.3. Botswana

The HIS in Botswana is extremely fragmented, with several ministries involved in the collection, compilation, and use of health data. In 2005, a centralized initiative to integrate the health information systems was established through the Health Statistics Unit (HSU), which was to collect one form containing the essential data from a range of health programs. However, the main reason for collecting this data was for the compilation of an annual book on health statistics, not for supporting local health information use. At the national level, the compilation of the health statistics book was two years delayed (Sæbø, Braa, & Chandna, 2007). As a response, various health programs still relied on their own systems, which were designed and run by each individual health program. At the district level, however, most of these information systems were handled by one or two persons in one office, representing a form of human data warehouse. Some of the richer health programs, most notably the HIV program of Prevention of Mother To Child Transmission (PMTCT), set up their own infrastructures to collect timely data by hiring their own people and equipping them with computers and software solutions.

The approach in Botswana was to make a direct mapping of the paper-based system into the DHIS version 1.4, meaning that each cell in the paper forms represented a unique definition in the software. Even if the paper forms had overlaps and duplications, the duplications were transferred to DHIS, without standardization of, in essence, equal data definitions. This approach would not reduce the amount of data to be entered, it would not solve the duplications (which turned out to sometimes contain different figures in different forms), but it did allow for a fast implementation by avoiding “political” obstacles in the set-up of the database. As long as no paper form was changed as it was copied to an electronic format, a database could easily be agreed upon by the various stakeholders. This strategy opted for a quick solution in an environment where the capacity to negotiate an integrated solution was not present at the time. The semi-integrated data collection tool, the DHIS, was then to be used as a platform for further integration. The nature of the project organization posed severe limitations in this direction.

As part of an EU funded research project (BEANISH^1), the Ministry of Science,  

^1 Building Europe Africa collaborative Network for applying IST in Health care sector, see http://208.76.222.114/confluence/display/BEANISH/Home
Technology, and Communication oversaw the initiation of the data warehouse implementation. Assigning the administration of the implementation to their IT-unit seconded to the Ministry of Health, the project was distanced from the health domain from the start. While the IT-unit physically worked at the Ministry of Health building, its main activities here consisted of setting up the network, assigning e-mail addresses, procuring hardware, and the like. The placement of the project implementation with the IT-unit meant that the project came to be seen as a “computer project” by the various health-affiliated stakeholders. This division between health and IT became stronger over time, and the project failed to enlist “champions” in the health programs. To complicate further, public health services at the district and facility levels were run by yet another actor, the Ministry of Local Government. Unable to tie the necessary links with the Ministry of Health-run national health programs and the Ministry of Local Government, the IT-unit, formally under the Ministry of Science, Technology, and Communication, became increasingly focused on just the computer-technical side of the implementation.

A spin-off of the data warehouse project in 2007 is worth mentioning. Looking for a software to ease data collection for their CRIS analytic software, UNAIDS adopted DHIS. Driven by local champions both at national and district level, UNAIDS rolled-out this new software bundle nationwide over the following year. Representing so far a new, independent, vertical reporting structure, the network UNAIDS has established of national and local actors might also manage to integrate the other health information systems in place, from which UNAIDS retrieves their HIV/AIDS specific data.

### 4.4. Zanzibar

In November 2004 a joint survey to study the HIS status was conducted by the Danish International Development Agency (DANIDA), World Health Organization (WHO), Ministry of Health and Social Welfare (MOHSW) and University of Oslo. The survey revealed poor health information system characterised by fragmented data collection and reporting due to many subsystems in healthcare service provision, and little use of information. As a result of this, in 2005 HISP started working with the newly established Health Management Information System Unit (HMIS Unit) at the Ministry, to implement a computerised HIS. Key objectives of the project were to 1) strengthen information use at the district level, and 2) set up an integrated data warehouse at the national level to facilitate access to information by all health programmes and stakeholders. Ministry stakeholders, health programme managers and district medical officers, and HISP consultants conducted a series of meetings to agree on a set of standardized data collection forms (datasets) that would form the basis of the integrated data warehouse.

To implement the same integrated data repository approach throughout the country (10 districts and a population of 1 million), a combined paper and computer strategy was implemented. The glue between all levels were the standardised essential datasets including essential data from all health programs taking part in the integrated approach (Reproductive and Child Health, Immunisation, HIV/AIDS, and Disease Surveillance). These standard datasets were both made available on paper forms to be used for data reporting between the health facilities and the district level, and electronic forms as part
of the DHIS software used to register, validate, analyse and report data at district and national levels.

In 2005 the customisation and implementation of DHIS started by using the new version 1.4, which was actually under development. This caused some technical problems during the first 1-2 years, as Zanzibar acted as a test site and bugs documenter for the new DHIS version being developed in South Africa. The approach to visually break up the harmonised database into multiple essential datasets to accommodate health program needs was different from the South African way of using only one essential dataset for all the data collected, and this was an important feedback to the global DHIS development process. Communication of requirements and testing of new functionality from Zanzibar became an important process in supporting the development of a more flexible and globally relevant DHIS software.

Every year since the first version of the unified data collection forms there has been a revision process (a 3-5 days workshop) whereby the data sets are revised and updated based on negotiations among key stakeholders (health programs and other departments at the MOHSW), and as a result the paper and electronic forms are updated on an annual basis to adapt to these revisions. While the standardisation process in Zanzibar towards one harmonised system has been relatively successful due to solid political backing the other key objective from the reform process, to strengthen information use at the district level, has been a slower process. Changing the culture of dealing with health data from a hierarchical reporting perspective to one of information for local management has proven a long term process and still requires lot of emphasis on capacity and awareness building.

4.5. Sierra Leone

Sierra Leone, a relatively small country in West Africa, is one of the poorest countries in the world and was ravaged by civil war that had lasted for ten year before it was officially declared over in January 2002. The public health system, which suffered from a huge loss of both personnel and infrastructure during the war, is slowly rebuilding the capacity to improve the service provision across the country. This effort is led by the government and supported by many international agencies, and aims at achieving the health millennium development goals. The rapid growth of various health initiatives has created a situation of fragmented information systems, common also in other developing countries (Braa et al, 2007, Sæbø et al, 2007). Figure 1 depicts the fragmented situation characterized by overlapping data collection tools and data elements. The overlapping rate between two different forms can vary from 0 to 50%. In 2008, each facility reports about 17 forms.
From the fragmented HIS described above, the main issue was to provide meaningful and relevant information for decision making and to diminish the workload of staff who are collecting and reporting the data. The suggested way forward was to use DHIS to integrate the various data flows and data sources and thereby to provide an integrated framework for M&E and data management. To achieve this, attention was devoted to the data warehouse as a starting point. All data elements were identified and sorted out in order to identify and avoid a) duplication from overlapping data forms b) overlapping data elements – in term of their definition. As a result a coherent integrated data warehouse was built, where one data element in the database can be related to a field in several data collection forms. Figure 2 shows how multiple, duplicate data elements from the existing paper system were integrated “behind the scenes” in the data warehouse.

In January 2008, this integration approach was implemented in 4 of the 13 districts in Sierra Leone, and 6 months later in 3 more districts. Intensive training was carried out, each district captured their data in the DHIS and exported the data to the national DHIS by the use of memory sticks. An extensive process to capture and import backlog data from the various electronic systems from all districts was put in place, and a rather extensive national data set was available for analysis during the second half of 2008. All stakeholders were made part of the process which convincingly documented the problems...
with the current system such as overlapping data collection forms, inconsistent data definitions, poor data quality in terms of both correctness and completeness. At the same time, through actually doing it, it was also documented that shared common data sets in a national repository was possible to achieve. This learning process sparked an increased interest to revise the current collection forms, and during 2009 a series of meetings took place among the key stakeholders to negotiate a new set of harmonised data collection forms. As a result, since January 2010 a new set of completely rationalised and harmonised forms have been in use, very different from those of the previous years.

Coupled with extensive training in the analysis and use of the resulting information, districts are starting to show improving health indicators. With the availability of data from across the country, one district medical officer was appalled at his district's comparatively low institutional delivery rate, an indicator directly linked to MDG Target 5a (reduce maternal mortality rate), but also MDG 4 (reduce child mortality rate). Involving health staff and civil society in discussing approaches to improve the situation, they managed to significantly increase the rate both in urban and rural areas over the course of 2009.

4.6. Summary of the case studies

Figure 2 illustrates the different approaches to HIS integration in the four countries. The fine-grained data element design of DHIS allows for breaking up the forms into smaller atomic data elements which may or may not (in the case of Botswana) be harmonised in the data warehouse. From left to right we see increased harmonisation of data both in the forms and in the data warehouse. Table 1 summarises the four case studies on key topics related to the integration process.
Figure 2 Integration strategies and degree of integration

Table 1: Summarising the cases

<table>
<thead>
<tr>
<th>Data warehouse standardisation approach</th>
<th>South Africa</th>
<th>Botswana</th>
<th>Zanzibar</th>
<th>Sierra Leone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Essential data sets. Specific paper reporting forms collected uniformly across country, eventually in addition to other forms</td>
<td>Essential data sets. Specific paper reporting forms collected uniformly across country, eventually in addition to other forms</td>
<td>Including all existing report forms – little revision of overlap and inconsistencies between forms and data elements. Fragmented db.</td>
<td>Full revision of all report forms – no overlap between forms and data elements. Annual revision process</td>
<td>All existing report forms – no revision of forms, overlaps solved in database, inconsistencies identified and addressed. Harmonised db.</td>
</tr>
<tr>
<td>Why this approach was selected</td>
<td>Extreme fragmentation caused effort to establish new unified reporting forms in addition to all other reporting routines</td>
<td>Capacity to lead integration not available. Data warehouse to highlight to stakeholders the need for integration.</td>
<td>All stakeholders agreed to revise and coordinate their reporting forms and solve problems of overlaps</td>
<td>Previous efforts failed because not all programs were included. Now include all data reporting and stakeholders. Show overlaps and inconsistencies.</td>
</tr>
<tr>
<td>Main challenges with the approach</td>
<td>Keep pace with new developments and technologies and remain in the forefront; e.g. on top of HIV/AIDS reporting development</td>
<td>To direct the early momentum gained into real integration. Avoid entrenchment of fragmented system by computerizing it.</td>
<td>Integrated data warehouse does not automatically lead to more local use of information. More capacity in the use of information</td>
<td>Align all actors in the integration process. Local capacity building. Supporting infrastructure. Harmonization of collection tools</td>
</tr>
</tbody>
</table>
5. DISCUSSION

As we have seen from the cases, there are both similarities and differences between the countries that highlight certain lessons. The similarities are perhaps the easiest to pinpoint; an initial situation of fragmentation and independent information structures, quite obstructive to efficient management to improve the health situation, including reaching the MDGs (not to forget, the poor information systems also influence other health areas than those spelled out in the MDGs). The differences are more diverse, and stem from the socio-political entity of the wider health system, including power structures, skill levels, pressure to see quick results due to funder’s priorities, and entrenchedness of the former systems. In the following discussion, we try to categorize the main findings according to concepts introduced in chapter 3.

5.1 Contrasting the cases - Data warehousing and integration

South Africa and the three other countries represent distinct integration strategies. While in South Africa the new system was built bottom-up and incrementally in addition to an existing array of partly inconsistent sub-systems, the other three countries tried to encompass all the existing data and systems, each following different strategies. South Africa has gradually extended the scope of the data sets by convincing and enrolling new service areas and programs through providing good results and thus creating a momentum (Braa et al, 2007). Zanzibar represents the opposite approach: to a large extent, the old reporting systems were revised and combined into a newly integrated system. This was possible through strong political and financial support, and because they met the basic requirements of the important stakeholders, not met by the old systems. In Sierra Leone the aim was similar, but they went through an intermediate process of including all the existing reporting formats and data sets in the data warehouse while at the same time solving the inconsistencies in the internal data structure. Through this process problems with the existing fragmented system and the advantage of integration was documented and the stakeholders have been convinced to revise and integrate the entire system of data collection, and new integrated data standards have been decided upon. Solid political support from the Ministry of Health and the Health Metrics Network was important. In Botswana also the aim of the intermediate process was to document and convince stakeholders, but the results were less convincing. Lack of
political support and commitment from within the Ministry of Health were major obstacles, and the results from the pilots did not create sufficient interest.

At the technical database level the different strategies may be outlined as follows:
South Africa is different from the others in that they followed a data for decision making approach, only including the most important data from each area. As the momentum grew, gradually more areas and data were added, which was easily accommodated through the flexible database structure. In this way national standards for essential data were developed “on top of” other existing systems. The three other countries, Botswana, Sierra Leone and Zanzibar, all followed different all-encompassing strategies; 1) Botswana: include all data “as they are”, without solving inconsistencies between them, 2) Zanzibar: start with revising the data collection tools and solving the inconsistencies before setting up the data warehouse, 3) Sierra Leone: two phases; a) accept that there are different overlapping data collection forms in use, but solving the inconsistencies in the meta data structure of the database, b) use this approach to demonstrate that “full” integration similar to the approach in Zanzibar is possible.

Ideally, the best approach to standardizing health data and integrating different vertical data reporting systems would be to make all actors agree upon and implement a revised and harmonized set of standards. This is to some extent what was achieved in Zanzibar. However, due to the wide variety of vertical donor initiatives, which all tend to have their own separate reporting requirements, and overall changes in information requirements within and between the various vertical health programs, standardization of health data is a complicated process. Standards tend to be evolving through negotiations between actors facilitated through ongoing efforts to develop standards in practice and by good examples working as “attractors” that are creating momentum rather than through early agreements (Braa et al. 2007, Eoyang 1996, Egyedi 2002, Hanseth et al. 1996). This is illustrated well by the case of South Africa where the first limited data set and database became successful, first in one, then two provinces, before later being adopted across the country as a result of the momentum created by successful use. The first phase in Sierra Leone also managed to create an attractor as the usefulness of the data warehouse approach was demonstrated in such a way that the involved actors agreed to go for full integration in a phase 2.

In Botswana, however, the problems of fragmented health and lack of standards were addressed mainly through a technical solution, which turned out not to be very useful; the problem of fragmentation was addressed by including all data in one data warehouse. By “dumping” all data from different health programs in one data warehouse, without solving the overlaps and inconsistencies between the data sets, analysis and use of data across the data sets were still difficult. Overlap of data will still produce different and confusing figures for the same phenomenon. The number of pregnant women receiving antenatal care, for example, is reported through both the Mother and Child Health program and the PMTCT, resulting in different figures for the same number of pregnant women. It is therefore obvious that integration and standardisation of data between these
two health programs need to be negotiated between the two organisations, and they will have to agree on shared data definitions at the semantic level. There are no technical short-cuts.

It may be added that maintaining a heterogeneous data warehouse as in Botswana will require more skilled data managers interacting with advanced data warehousing tools to produce an abstracted, harmonized level on top of the heterogeneous data sources, than what is required for a strictly standardised one. Consequently, this approach is not recommended for developing countries where human resources make up the bottleneck.

Contrary to the case of Botswana, we see that in the three more successful cases, the processes of developing the system and standardising the health data reporting have taken place at the organisational level. To what extent the solution supports integrated information use is the key criterion for success.

5.2 Spanning boundaries through flexibility

Carlile's integrative framework is helpful in highlighting how the four countries have followed different paths in their quest for integrated health information systems, and in understanding how software and implementation strategies collectively developed in the global network of HISP have played the role as boundary objects in this longitudinal collection of case studies.

The integration process in Sierra Leone followed Carlile's processes of 1) transfer – adoption of the technical solution to share data and identify indicators and data elements, while collection forms were kept as before. 2) translation – a new common lexicon of harmonised data elements was created in the back-end of the data warehouse, removing overlaps and duplicates that existed in the collection forms. Examples of integrated data analysis were produced and shared among stakeholders. 3) transformation – in 2009 after using the integrated data warehouse for more than one year, stakeholders acknowledged problems of data quality with the existing fragmented collection forms and were convinced of the need for data harmonisation. A revision process was started and in February 2010 the new harmonised and standardised data collection forms were implemented countrywide. And 4) iteration - during the harmonisation process, stakeholders agreed to meet on regular basis to review and renegotiate the shared and harmonised standards. This iterative process, similar to the development of the gradually expanded essential dataset in South Africa and the annual revision of forms in Zanzibar, allows for integration of new and emerging initiatives and information needs, such as the country-wide launch of the free health care initiative in Sierra Leone, April 2010.

Although we have demonstrated that the political and organisational dimensions of integration are critical, there is also a role to play for the technology. Furthermore, the cases of South Africa and Sierra Leone show that the DHIS data warehouse was instrumental and worked as a boundary object in providing the capacity to negotiate interests between the actors and to transform knowledge and practices. In the first phase
In Sierra Leone the flexible data element design of DHIS allowed for a harmonised data warehouse at the back-end while keeping the clear borders of the health programs' collection forms intact at the front-end. The inherent flexibility of the technology could this way show the benefits of integrated data analysis and attract interest (Braa et al 2007), while at the same time delaying the political battle to harmonise the collection forms until the stakeholders had been convinced. This maturation process was followed by a second phase where the data standards and collection forms were harmonised in a negotiation process between the health programs and the national HIS unit. In this process the relational borders between these sub-systems were reshaped into a mutually beneficial integrated information system (Barret et al).

In Botswana the DHIS failed to play the role as a successfully attractive boundary object and did not enable any change to the relational borders between health programs, departments and ministries, and as result little integration took place. While the data warehouse in Sierra Leone was harmonised at the back-end and therefore could communicate the benefits of integration, in Botswana the clear borders of the health programs' forms were mirrored in the technology (see Figure 2). This way the DHIS was not able to demonstrate integrated information use and consequently failed to bridge the technical and the organisational levels of the HIS in Botswana.

5.3 Maintaining attractiveness in a changing environment
To explain the difference in approaches, especially the contrasting cases of Sierra Leone and Botswana outlined above, it is important to look at how the boundary objects and attractors have evolved over time. As noted in chapter 2, this study spans many years, in which changes have occurred at different levels. The implementers (and authors) have been part of an evolving network, in which feedback and ideas from countries and research institutions have shaped the understanding of health information systems and approaches towards integration.

The DHIS application has gone through many development cycles since its first use in 1996. The needed flexibility in the design as well as the necessary experience of the DHIS implementers were not available in 2005 in Botswana, but had been developed in time for a more successful data warehouse design in Sierra Leone in 2007. Important lessons were learned in Botswana and Zanzibar, especially in dealing with multiple datasets and accommodating collection forms across health programs, which was very different from the one essential dataset approach of South Africa. These new experiences and emerging requirements were fed back into the global HISP network, which again benefitted other countries, including Sierra Leone.

Furthermore, the global community of health has changed, with global partnerships like HMN and IHP+ campaigning for harmonization and integration, and giving legitimacy to formerly more marginalized groups advocating this at the country level. The strong backing of HMN in Sierra Leone played an important role in pushing the agenda for change and strengthened the effect of the DHIS as a boundary object, in contrast to the
more difficult political context in Botswana two years earlier.

On top of all this, the projects themselves changed in nature as they scaled; from pilots to established systems which a lot of people depended on for their daily work. What was useful for bringing two initial groups together, might not be sufficient to keep growing, reaching the crucial scale necessary to offer a comprehensive system supporting the MDGs.

In South Africa, more than ten years ago, a relative simple but flexible software application supporting a minimum essential data set created an attractor, which was both addressing immediate needs for many users but also demonstrating how “similar” problems could be solved. The easiness with which to define and implement new data sets integrated with other data sets using the DHIS led to innovations and new use areas, and it led to the spread of the DHIS and HISP approaches to new countries. The strategy that has led to success in South Africa and to some extent also in Sierra Leone and Zanzibar, and generally in the DHIS and HISP networks, has been to create an attractor by initially providing easy to make, but useful solutions, which may later evolve and become more comprehensive as new actors and use areas are included (Braa et al. 2007).

Through the changing context brought by external and internal factors, as well as scaling of the initiatives themselves, the forces of gravity emanating from the attractors grow and wane in strength. Maintaining attractiveness, then, becomes a key goal. The first efforts in South Africa started before the MDGs were formulated. The projects in Zanzibar and Botswana were initiated as HMN was just starting to build global legitimacy for an integrated approach. By the time Sierra Leone reformed their HIS, an integrated system supporting the MDGs was very much a core requirement. The DHIS software has over the years closely followed this trend. Flexibility to integrate various data sets and to customize data collection forms and reports have ensured that health programs could opt for a quick win; joining the “integrated” data warehouse without having to reform their data collection tools and to use the data warehouse as an active instrument in the continuous process of standardisation and integration, as we have seen in Sierra Leone.

The reason why DHIS is still used to create attractors in new countries and use areas is that it has been under continuous development and has managed to evolve with the requirements of the ever changing needs of global health, and has thus managed to stay attractive over time. The fact that DHIS is still an attractor, however, also demonstrates that the problems of fragmentation and challenges of integration are not yet solved.

6. CONCLUSION

With ten years spent pursuing the MDGs, there is still a lack of adequate health information systems for management and monitoring capacity in many countries. In response to this, and the sorry overall state of health information systems and use in developing countries in general, the international community has advocated integration to reduce the inefficiencies of the fragmented legacy systems of decades of uncoordinated
development interventions. This paper has addressed the challenges to such integration by investigating four country cases, all of which had different approaches to integration.

We have argued that integrated information systems and quality information are needed in order to achieve the MDGs. In order to, for example, substantially reduce maternal and child mortality and to successfully combat HIV/AIDS, both targeted action and coordination of health services will be needed. This again will require the use of quality data for both local level decision making and services delivery, as well as integration of information across health programs.

We have described and discussed four cases of applying a shared data warehouse – as a tool to facilitate practical integration of information across different organisational structures. The lack of success in Botswana and the relative success in the other three countries, clearly indicate that although following a technically termed “data warehouse approach”, in order to be successful, integration needs to be rooted at the institutional and organisational-political level. The data warehouse approach, when applied successfully, has turned out to work as an important mediator and boundary object facilitating the negotiation between the organisational actors.

We argue that “attractors”, technical solutions or standards that achieve a certain level of success and enable the building of momentum, are important in aligning the various political actors. The networks of action approach followed in this longitudinal research project enables circulation of methods and software between country implementations, and facilitates learning and sharing over time. These networking mechanisms are important in making sure that the attractors evolve within the changing context of a growing health information system, in order to achieve the scale needed to address the Millennium Development Goals with full force.

7. REFERENCES


